System parameter and state estimation

Eykhoff, Pieter

Published: 01/01/1972

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):
SYSTEM PARAMETER AND
STATE ESTIMATION

P. Eykhoff

And when all is said, he that publishes a book runs a very great hazard, since nothing can be more impossible than to compose one that may secure the approbation of every reader.

Miguel de Cervantes Saavedra

Don Quixote, vol. 2, ch. 3.
SYSTEM PARAMETER AND STATE ESTIMATION

Preface

In recent years many aspects of system parameter and system state estimation have been discussed in a number of papers, at many conferences and in several university courses. The interest in this subject apparently has widely differing roots:

- the desire of practicing engineers in process industries to obtain better knowledge of specific plants in order to obtain improved control and thus lower operating cost or increased efficiency.

- the study of high performance aerodynamic and space vehicles and not incidentally the study of more down to earth systems such as railway carriages and hydrofoils.

- the study of the human being in tracking action situations and in other types of control or learning processes.

- the investigation of biological systems, e.g., neuromuscular systems such as the eye pupil control system, extremity (arm, leg) control, heart rate control system, etc.

This interest in estimation theory has not arisen solely from the need for increased system performance (or increased knowledge of system dynamics) but it is also a reflection of the dramatically changed possibilities of applying estimation theory brought about by the advent of modern computer hardware and software. The need for such computing facilities will become apparent when we come to discuss specific estimation algorithms.

More or less independently of the interest in estimation theory from the above sources under the general heading "engineering sciences", there has also been considerable interest and activity from econometricians and sta-
ticians who have been working on dynamic models (econometric models) of either some portion or the whole of social activity. This latter work has lead to occasional cross fertilization with the engineering sciences and is now culminating in the development of a general mathematical theory of systems applicable to systems of whatever type, be they social or physical.

The net result of this development is a large number of publications, either accentuating a particular type of approach or describing a certain case study. In this book we have tried to cast the estimation problem in a coherent form from a rather broad frame of reference. It is written with regard for the many different types of scientists and engineers concerned with dynamic systems - be they production processes, aircraft, biological objects, economic and other types of models.

Due to the nature of the subject, this book has to deal with statistics, system dynamics, stochastic signals, analogue, digital and hybrid simulation, and optimization. The limited treatment of some subjects (brought about by the constraint of a finite number of pages of text) may be supplemented by reference to the extensive bibliography of pertinent literature.

A few remarks on the organization of the book are in order here. When contemplating the theme, one may come up with a variety of legitimate distinctions to be used for the scheme of organization, e.g.,

- type of application
- statistical basis: least squares, Markov, maximum likelihood, Bayes' estimates
- linearity and nonlinearity of the model with respect to parameters (note that this is not the same as linearity with respect to dynamic behaviour)
- signal representation: sampled and non-sampled
- implementation: analogue, digital, hybrid
- type of approach: open loop and closed loop estimation methods

In order to provide the reader with convenient means of reference and to accentuate the coherence of the field the following approach has been chosen:

The chapters Introduction and Basic approaches serve as an explanation of the most important approaches towards the solution of estimation problems.

The part Prolegomena summarizes the knowledge that is needed; the reader will find some condensation of knowledge about signals, dynamic systems (processes) and topics from probability theory, statistics and optimization theory.
The part Estimation of parameters is organized along the following lines:

```
<table>
<thead>
<tr>
<th>types of estimates</th>
<th>types of signals</th>
<th>types of approach</th>
<th>types of models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>open loop</td>
<td>linear in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sampled</td>
<td>or explicit</td>
<td>methods in</td>
</tr>
<tr>
<td></td>
<td>signals</td>
<td>parameters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>closed loop</td>
<td>linear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>or implicit</td>
<td>nonlinear in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>methods</td>
<td>parameters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>open loop</td>
<td>linear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>or explicit</td>
<td>methods</td>
<td></td>
</tr>
<tr>
<td></td>
<td>continuous</td>
<td>linear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>signals</td>
<td>or implicit</td>
<td></td>
</tr>
<tr>
<td>least squares</td>
<td>closed loop</td>
<td>nonlinear in</td>
<td></td>
</tr>
<tr>
<td>and Markov</td>
<td>or implicit</td>
<td>parameters</td>
<td></td>
</tr>
<tr>
<td>estimates</td>
<td>methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parameter</td>
<td>periodic or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimation</td>
<td>large signals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>maximum likelihood</td>
<td>estimates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayes estimation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

This way of presentation leads from the simple to the more complicated, from the particular to the more general case. In spite of the reversed trend in many research papers we feel that, for the first approach, the engineering student and researcher is more open to inductive reasoning, even where he recognizes the great beauty and merits of the deductive approach. By chapter 5 is introduced already the relation between the different types of estimates.
Spectrum analysis techniques are included in those chapters where transformation to the frequency domain is practical.

The estimation assignment is more complex, if not only parameters, but also (state) time functions in the system under study have to be determined.

The description of system behaviour in terms of "state variables" is reviewed in chapter 4.

The part **Simultaneous estimation of parameters and states** reviews first the state estimation problem. Here the deductive approach is chosen as it is expected that the reader has been exposed to this problem before or that, from the parallelism with the parameter estimation problem, he can readily find his way. The next chapter is devoted to the combined parameter and state estimation techniques.

The last part of the book, **Applications**, refers to specific estimation cases that have been discussed in the literature, e.g. in control, in automatic industrial measurements, in biological systems and in learning or pattern recognition problems.

The book is written with potential users in mind, e.g., a graduate engineer or research worker, a postgraduate student or a college lecturer.

With respect to the background needed for reading this text, the following pre-requisites will be sufficient:

- graduate level mathematics (advanced calculus, differential equations, complex-variable theory, some acquaintance with linear algebra and matrices)
- some basic knowledge of feedback control theory.
- some knowledge of probability theory and elementary statistics.
- some knowledge of digital- and analogue computer techniques.

For easy reference, part of the necessary background knowledge is summarized in appendices.
List of Contents

Preface and list of contents

Chapter 1 SOME CONCEPTS; model building, parameter and state estimation

1.1 Models
1.2 Model building
1.3 Structure, parameters and states
1.4 Representation of the problem
1.5 Normal operating conditions
1.6 Some fields of application
   - estimates for "diagnostic" applications
   - estimates for control applications
   - estimates for automatic (industrial) measurements
   - estimates for automatic (industrial) decisions
   - estimates for automatic (industrial) adjustments
   - pattern recognition
1.7 A wider perspective

Chapter 2 BASIC APPROACHES TO THE PROBLEM; statistical and engineering

2.1 General aspects of the problem
   - process input signals
   - a priori knowledge about the process
   - estimation scheme
2.2 Classes of implementation
   - use of explicit mathematical relations ("open loop")
   - use of implicit or model-adjustment technique ("closed loop")
2.3 Classes of statistical approaches
   - least-squares estimation
   - Markov or generalized least-squares estimation
   - maximum likelihood estimation
   - minimum risk estimation
2.4 Model-adjustment techniques
   - "odd-function" and "even-function" error
   - techniques for obtaining the partial derivatives or gradients
2.5 Parameter and state estimation

PROLEGOMENA

Chapter 3 SIGNALS; deterministic and stochastic

3.1 Classes of signals
3.2 Orthogonal functions
3.3 Description of deterministic signals
3.4 Description of stochastic signals

Chapter 4 PROCESS MODELS; linear, varylinear, nonlinear

4.1 Classes of process models
4.2 Linear models
   - deterministic signals; non-sampled
   - deterministic signals; sampled
   - stochastic signals
4.3 Varylinear models
4.4 Nonlinear models
   - Volterra representation
   - other representations
4.5 Controllability, observability, identifiability
4.6 Models based on these descriptions
   - linearity-of-the-dynamics versus linearity-in-the-parameters
- linear input-output models
- linear state models; canonical forms
- nonlinear models

Chapter 5  SOME PROBABILISTIC NOTIONS; estimation theory, convergence schemes, stochastic approximation

5.1 Estimation theory
- characteristics of estimators
- Bayes estimators
- maximum likelihood estimators
- relations between the different estimators
- some other aspects
5.2 Deterministic convergence schemes
5.3 Stochastic approximation

Chapter 6  SAMPLED SIGNALS; explicit methods

6.1 Regression analysis
- regression curves and planes
- estimation from a finite number of observations
- linear unbiased estimators
- least squares estimators
- Markov or generalized least-squares estimators
6.2 Implementation of "open loop" estimation schemes
- general considerations
6.3 Accuracy; some error causes
- noise error
- truncation error
- state-correspondence error
- errors due to simplification of the implementation
- sampling error
6.4 Residuals, noise properties and model order
6.5 Extension to generalized models and processes with feedback
- the bias problem
- processes with feedback
- repeated least squares
- instrumental variables
- tally principle
- Levin's method
6.6 Some computer procedures

Chapter 7  SAMPLED SIGNALS; implicit or model-adjustment techniques

7.1 Models, linear-in-the-parameters
- corrections proportional to the gradient
- least squares approach to strings of observations
- least squares approach to single (pairs of) observations
- stochastic approximation
- contraction mapping
7.2 Extension to generalized and state-space models
- successive linear regression and filtering
- generalized-least-squares estimation
- extended matrix method
- state space models
7.3 Some computer procedures and results
7.4 Models, nonlinear-in-the-parameters.
Chapter 8  CONTINUOUS SIGNALS; explicit methods

8.1 Operations on "analog" signals
- general properties of correlation measurements
- the use of orthogonal filters
- statistical errors of correlation measurements

8.2 Operations on quantized signals
- amplitude quantizing
- types of correlators

8.3 Operations on binary signals
- the use of auxiliary signals

8.4 Differential approximation method

8.5 Higher order correlation functions

Chapter 9  CONTINUOUS SIGNALS; implicit or model-adjustment techniques

9.1 Models, linear in the parameters
- generalized error
- the choice of the error criterion and the type of information processing
- the choice of the dynamic operators

9.2 Minimization of the instantaneous error
- basic equations
- convergence properties
- influence of additive noise
- examples

9.3 Minimization of a time-averaged error
- basic equations
- convergence properties

9.4 Parameter sensitivity functions
- parameter influence (sensitivity) techniques
- sensitivity points methods
- continuous model-adjustment
- intermittent model-adjustment
- examples
- intermittent model-adjustment

9.5 Simultaneous use of two models or repetitive use of one model

9.6 Model with parameter-perturbation

Chapter 10 TEST SIGNALS; periodic or large ones

10.1 Impulse and step signals
- the influence of additive noise
- the use of correlating filters

10.2 Sinusoidal test signals
- the influence of additive noise
- related procedures

10.3 Binary signals
- m-sequences
- errors when using an m-sequence test signal

10.4 Other approaches
Chapter 11 BAYES' AND MAXIMUM LIKELIHOOD ESTIMATION

11.1 Bayes' estimation
11.2 Maximum likelihood estimation (M.L.E.)
   - the achievable accuracy
   - some of the most important properties
11.3 Some implementation schemes
11.4 The effect of the process-input signal

SIMULTANEOUS ESTIMATION OF PARAMETERS AND STATES

Chapter 12 PROCESS STATE ESTIMATION; a review

12.1 The Wiener filtering problem
12.2 The Kalman-Bucy filter
   - one-stage filter
   - multi-stage filter
12.3 Further aspects, applications and examples

Chapter 13 PARAMETER AND STATE ESTIMATION

13.1 Nonlinear nature of the problem
13.2 Least squares versus Bayes' approach
   - Bayes' estimation
   - maximum likelihood estimation
   - least squares estimation
13.3 Two point boundary value problem
13.4 Gradient methods
13.5 Quasi linearization
13.6 Invariant imbedding

APPLICATIONS

Chapter 14 SELECTED APPLICATIONS FROM DIFFERENT FIELDS

14.1 Automatic control
   - optimal control; the separation hypothesis
   - self-optimizing and self-adaptive
14.2 Automatic industrial measurements and adjustments
14.3 Estimation of biological parameters
14.4 Learning systems; pattern recognition
14.5 Economics

APPENDICES

A. List of symbols used
B. Summary of notions from probability theory and stochastic signals
C. Summary of matrix operations
D. Multi-dimensional Laplace transform
E. Transients in statistical average and variance of correlation measurements
Chapter 1 SOME CONCEPTS;
Model building, parameter and state estimation.

1.1. Models

At the heart of the natural sciences and engineering there are the general notions of observation and measurement. Based on observation, the scientist builds a physical insight into the problem being studied and from that insight he formulates a theory by trial and error. This theory is a proposed concept of that aspect of nature which he is studying. Guided by this concept he designs new experiments. The observation of the results of these experiments either confirms the theory, dictates a change of the theory or rejects it completely. Although a concept may be beautiful and appealing to the mind of the scientist, the factual results nevertheless are dominant in this interaction of theory and experiment. From this point of view it is legitimate to say that in the fields of natural science and engineering experiments and observations (measurements) are most fundamental.

Of almost equal importance is the idea of model building (Rosenblueth and Wiener, 1945). It can in fact hardly be separated from the observations and experiments mentioned before. The formulation of a theory (as a proposed concept of the aspect of nature being studied) may rightly be called "model building"; the theory stands as a verbal or mathematical "model" of reality. For our purposes, we will define a model as a representation of the essential aspects of an existing system (or a system to be constructed) which presents knowledge of that system in a usable form.

In this definition, "system" may be interpreted in a rather general way, e.g. as "a collection of objects arranged in an ordered form, which is in some sense purpose or goal directed". Everything not belonging to the "system" is part of the "environment". Depending on the borderline drawn between the system and its environment, one may characterize the system by input/output
relations. What constitutes a system depends on the point of view of the observer. The "system" may be e.g., an amplifier consisting of electronic components, or a control loop including perhaps that amplifier as one of its many parts, or a chemical processing unit having many such loops, or a plant consisting of a number of units, or a number of plants operating together as a system in the environment of a global economy.

The scope of the definition of model may be clarified by some remarks.

- The model does not need to be a description of the actual mechanism of the system. It may simulate or "mimic" the system behaviour. One may construct a controlled arm prosthesis without understanding how the human being is able to utilize his limbs. It is the skill of the model maker to mask unexplained influences by more or less alchemistic properties such as "crackability", "spinning factor" etc.

- Models may be conceptual, physical or mathematical (also called in that order: phenomenological, empirical and analytical) depending on: 1) the judgment of what is the essential aspect for the particular purpose, 2) the model building techniques that can be used and 3) the amount and quality of knowledge available. A clear and well known example of the development of a model can be found in celestial mechanics. The Ptolemaic model of cycloidal motions stood as quite an adequate conceptual model for the observations of planetary motions. The Copernican model, which explained the observations as a result of both the motions of the earth (observer) and of the observed planet, can stand as a physical model. Kepler's laws provide a mathematical model that offers the possibility of numerical predictions. From his own experience the reader will be able to give examples of other models in the different areas cited.

Another distinction between models may be in terms of "hardware" (instru-
ments, apparatus) and of "software" (mathematical expressions, computer programs).

The object for which a model has to be made may be either a planned or an existing system. For engineers the importance of a model in the design stage of a process needs no further emphasis. For research scientists the model may provide a guide to further experiments.

- The knowledge should be presented in a usable form. This is an essential aspect since the model must provide the background for further decisions. If the model is too complex its usefulness becomes questionable. The relative simplicity is a dominant feature of model construction. A model is a representation of reality with a reduced complexity. In many cases, the model complexity has to have a definite relation to the complexity of reality in order to be useful (e.g., for living systems).

Models can be found in many fields of human observation. These fields include physics, biology, astronomy, engineering, economics, sociology and psychology.

In this text the attention will be focused mainly on the engineering fields, although the estimation techniques presented may be and are being used in the other fields mentioned as well. In engineering, one may distinguish among the following purposes of, or uses for models:

- Research. One wants to give an interpretation of knowledge or measurements; the immediate use does not need to be clear beforehand but the gain in comprehension and consequently a concise formulation is of paramount importance (diagnostic situation). The model has to provide clues for further investigation.

- Design. The knowledge of the components or subsystems to be used has to be expressed in a model, compatible with the design criteria (stability, error criterion, economic yield, intrinsic safety, etc.). In this context, the adaptation of an existing system due to higher demands or because a more economical
performance is required has to be considered as part of a design too. Design essentially is a feedback action with the designer in the loop.

- Control. The control actions on a system depend on the knowledge about that system. In this case one can distinguish:
  
a) normal operating conditions, e.g., feedback or feedforward control, types of static optimization, dynamic optimization, adaptive control, batch control. Apparently there is a close relation between the usability of a "difficult" dynamic system and control-means available.

b) emergency situation, such as partial breakdown, when the control action depends on information about the type and degree of failure.

c) start-up and shut-down situation, where certain steps of a programmed schedule may depend on the values of the system variables or parameters.

For the (engineering) scientist there is a real challenge in the attempt to analyze the common feature behind the immense number of applications of (automatic) control systems.

In the first place there is the wide spectrum of man-made control systems in engineering, ranging from the simplest forms of on-off control to complicated computer control in process industries. In addition to these, there are also the control mechanisms in society (e.g., the enforcement of the law) and the abundant number of control loops in living "systems".

For such a wide variety of "systems" only a rather loose notion can be expected to indicate the common factor behind the applications of control. Such a notion may be designated as "uncertainty".

Since the celebrated application of Watt's regulator it has become apparent in innumerable cases that the introduction of some feedback mechanism can be used as an effective expedient for combatting uncertainties. These uncertainties may be the result of unpredictable influences (disturbances) of the environment on a system, or they may originate inside the system under
consideration (e.g., wear, aging, catalyst poisoning, etc.). Simple feedback can diminish the influence of uncertainty, or at least shift the adverse effect of uncertainty from an important system quantity to an unimportant quantity. Of course, this is no panacea and there are limitations to the application of this remedy. Such limitations are met in those cases where very large parameter variations occur; such a situation may lead to the use of adaptive control principles. Another limitation may be found in the requirements for optimizing some economic criterion, leading to optimal or self-optimizing systems.

In designing ordinary control the required knowledge of the behaviour of the process is limited to (possibly) rather inaccurate data needed for stability considerations and to the data needed for evaluation of possible improvement by the application of feedback, i.e. the nature of loads and disturbances, constraints on the control variables, etc. Throughout the history of control theory it has been known that the knowledge of a system and its environment, required to design a system fully, is seldom available a priori. Even if the equations governing a system are known in principle, it often happens that knowledge of particular parameters is missing; it is not uncommon that the models which are available are much too complex. Such situations naturally occur in many other fields. There are, however, two facts which are always present in identification problems occurring in automatic control:

- It is often possible to perform experiments on the system in order to obtain the knowledge which is lacking.
- The purpose of the identification is to design a control strategy.

One of the factors which undoubtedly contributed very much to the great success of frequency response techniques in "classical" control theory was the fact that the design methods were accompanied by a very powerful technique for systems identification, i.e., frequency analysis. This technique made it
possible to determine the transfer functions accurately, and this is precisely
what is needed to apply the synthesis methods based on logarithmic diagrams.
The models used in "modern" control theory are with a few exceptions parametric
models in terms of state equations. The desire to determine such models from
experimental data has naturally renewed the interests of control engineers
in parameter estimation and related techniques.
For systems with large parameter variations (leading to adaptive control)
and systems with strict "economic" criteria (leading to optimal or self-
optimizing control) there often is a definite need for more accurate and
timely information on the systems, both with respect to parameters and state
variables. Consequently, the necessity for process identification, modeling
and parameter estimation is apparent. One would approach such a process as
being a celebrated black box, i.e., a (sub)system with one or more "inputs"
and "outputs" but of unknown structure. In most engineering situations,
however, the black box approach is not a very realistic one. The experimenter,
in many cases, has derived some a priori knowledge from physical insight
into the process under consideration. This may give information on the
structure of a conceptual model for that process, and probably even an
approximate knowledge of the values of the coefficients (parameters) in
that model. Consequently, the box is more or less "grey" or translucent.
Even the fact that we know we want to consider a distillation column and not
a lunar vehicle implies a prior knowledge (Pavlik).
Some types of problem characterization is provided by the answers to the following questions:

- Does one want a model for static behaviour (e.g., for static process optimization) or for dynamic behaviour (e.g., for dynamic process optimization or for aircraft dynamics control) of a complete non-linear or a linearized model? The answer to this question may imply clues for the desired accuracy, the sophistication of the model, the mathematical or statistical approach, etc.

- Should the model be assembled off-line, with paper and pencil from basic laws and some isolated experiments, or can we work on-line and are we allowed to experiment with the existing process?

- What budget considerations have to be taken in account?

From this point on the number of questions increase rapidly and become more and more awkward. We shall mention just a few:

- How is the quality of the model to be judged?

- How is all the relevant knowledge to be incorporated?

- What is the optimum strategy to obtain the missing knowledge?

- How are nonlinearities to be handled?

- How can a complex system be approximated by a simple model?

The answers to these questions are strongly object oriented. Some general considerations are given in chapter 4.

1.2 Model building

In many cases model building starts with the application of basic physical laws (Newton's laws; Maxwell's laws; Kirchhoff's laws; mass-balance; energy/heat-balance; impulse balance; entropy balance) to the process being studied, e.g. a mechanical, an electrical or a thermodynamic process. From these laws, a number of relations between the variables at hand follow, e.g., ordinary
or partial differential equations, difference equations. These are relations between "forces" (intensive quantities) and "currents" (extensive quantities), e.g.,

<table>
<thead>
<tr>
<th>&quot;forces&quot;</th>
<th>&quot;currents&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>function of time only</td>
<td></td>
</tr>
<tr>
<td>temperature difference</td>
<td>heat flow</td>
</tr>
<tr>
<td>concentration &quot;</td>
<td>mass &quot;</td>
</tr>
<tr>
<td>electric potential &quot;</td>
<td>electric current flow</td>
</tr>
<tr>
<td>function of time and place</td>
<td></td>
</tr>
<tr>
<td>temperature gradient</td>
<td>(vector) heat flow density</td>
</tr>
<tr>
<td>concentration &quot;</td>
<td>&quot; mass flow &quot;</td>
</tr>
<tr>
<td>electric potential &quot;</td>
<td>&quot; electric current flow density</td>
</tr>
</tbody>
</table>

If all external and internal conditions of our system are known quantitatively and if our physical knowledge about the system is complete then, at least in principle, the numerical value of all coefficients (parameters) in those relations can be determined. This situation may occur for example in connection with space-vehicle dynamics. Such a situation is, however, quite rare in down-to-earth (engineering) applications. Our a priori knowledge is restricted because of uncertainty as regards the environment, and because of an incomplete knowledge of the physical aspects inside the process.

Even if our knowledge of the physical effects in each "elementary cell" of the process were perfect, then the devil of dimensionality would probably still force us to use a simplified description of the process as a whole. This may imply that the parameters and states of our model are less amenable to a priori calculation. For that reason one has to rely on (process) identification and estimation. Consequently, model building consists of the following steps:
- selection of a model structure based on physical knowledge;
- fitting of parameters to available data (estimation);
- verification and testing of the model (diagnostic check);
- application of the model to its given purpose.

In view of the wide field of disciplines and the great variety of applications no general rules can be given for these steps. The model structure is selected on the basis of initial (a priori) knowledge about the system and its purpose. Actually finding an appropriate model may be a very difficult problem even in a narrow field of application.

The main emphasis of this text is on the estimation aspect. The use of the word "estimation" is due to the fact that in almost all realistic situations the observations made on the system under study are contaminated with random influences (disturbances; errors). Consequently one has to call upon statistical methods to obtain a "best" result from our measurements by "filtering out" the influence of the disturbances.

The verification and the testing of the model is strongly related to the estimation; as a follow up of the estimating procedure one should investigate to what extent the model really "explains" the behaviour of our system. Conceptually this can be visualized by driving the system and the model by the same "input" signal and by studying the character and magnitude of the difference of both "output" signals (errors or residuals).

The application of the model to its given purpose again is strongly object oriented.

For some examples of this procedure in process industry the reader is referred to Eykhoff, Van der Grinten, Kwakernaak and Veltman (1966). Examples in other fields can be found in the literature mentioned in the chapters 13 through 17. For another concise illustration of the (iterative) sequence of model-structure selection, fitting, diagnostic check and application (e.g., forecasting) cf. Box and Jenkins (1969).
1.3 Structure, parameters and states

Identification is defined by Zadeh (1962) as the "determination on the basis of input and output, of a system within a specified class of systems, to which the system under test is equivalent". Using Zadeh's formulation it is necessary to specify a class of systems, \( S = \{S\} \), a class of input signals, \( U \), and the meaning of "equivalent". In the following we will call "the system under test" simply the process and the elements of \( S \) will be called models. Equivalence is often defined in terms of a criterion, an error or a loss function, which is a functional of the process output \( y \) and the model output \( y_M \), i.e.

\[
E = E(y, y_M)
\]

Two models \( M_1 \) and \( M_2 \) are then said to be equivalent if the value of the loss function is the same for both models i.e.

\[
E(y, y_{M_1}) = E(y, y_{M_2})
\]

There is a large freedom in the problem formulation; this freedom is reflected in literature on identification problems. The selection of the class of models, \( S \), the class of input signals, \( U \), and the criterion is influenced largely by the a priori knowledge of the process as well as by the purpose of the identification. It has to be recognized that a model represents three types of knowledge, viz.:
structure, expressed in terms of mathematical identities, block diagrams, flow diagrams or graphs, connection matrices.

- parameter values, i.e., quantities that show no dependence on the input or independent variables.

- values of the dependent variables (state) at a certain instant or as a function of time.

The importance of the structure cannot be over-emphasized. Its choice is guided by the type of application of our model; the choice may be decisive for the ultimate success or failure of the estimation scheme.

Often identification has the connotation of starting from scratch without any prior knowledge about the process. As mentioned before, in the majority of engineering situations and in a number of cases in biology as well, this is not a realistic assumption; from the structure of the "process" and at least a partial understanding of its operation, a certain amount of a priori knowledge will be available, e.g., the model structure. In such cases the knowledge to be derived is reduced to numerical values of a number of parameters (coefficients of differential equations governing the dynamics of the process, coefficients of a linear or nonlinear "model" of the process, etc.) and/or numerical values of the state variables. Consequently the identification problem is reduced to that of (process) parameter and/or state estimation.

Parameter estimation is defined as the experimental determination of values of parameters that govern the dynamic and/or nonlinear behaviour, assuming that the structure of the process model is known. It should be pointed out
that the distinction between knowledge of structure and parameter is not as straightforward as it may appear at first sight. The change from a parameter value $\beta \neq 0$ to $\beta = 0$ may represent a simplification in the structure as that "branch" of the model may be deleted. On the other hand it is clear that in this way no "additions" to the model can be made; the potential capabilities of the model are limited by the choice of the model. Little work on systematic improvement of models by adding new structural elements seems to have been done. At this state of the art, the human intelligence and intuition appear to be still the limiting factor.

In some cases, one is interested in a more detailed knowledge, e.g. a continuous information about the state of the process; this leads to state estimation. Roughly speaking the "state" of a process is a variable (e.g. a vector) which, together with the subsequent input to the process, completely determines the subsequent behaviour. Consequently, if the process is left alone, i.e., if there is no input signal, then knowledge about the state at a certain time (together with knowledge about process structure and parameters) is sufficient to predict its future behaviour. As an example one may consider the initial conditions of a differential equation as a state vector; cf. chapter 4. The assignment of determining the best estimate of the state can occur in modern control applications as e.g., optimal control. If there is in addition an uncertainty about a number of parameters then we have the combined problem of parameter and state estimation. From this description it is clear that in the majority of cases, state estimation is required to be done under normal operating conditions. For parameter estimation the determination can be based either on (active) experiments or (passive) observation of the process under study.

1.4 Representation of the problem

Some simple representations of the class of problems at hand are given in fig. 1.1.
The quantities \( u(t) \) and \( y(t) \) are the input and the output of the process respectively. The differences among figs. 1.1a, b and c lies in the observability of the input signal, viz.:
- observable, - observable but contaminated with noise, - unobservable; c.f. Kerr (1965). In all cases the output \( y \) is contaminated with noise. From the observable signals and possible a priori knowledge about probability density functions the information on parameters and/or state has to be derived. In feedback and in multiple-input/multiple-output systems, combinations of these models may appear. The consequences of these different situations will be studied in chapter II.

The term "process" can be used in the broadest sense of the word and may refer to an engineering system (e.g. aircraft control system, chemical plant, ionospheric transmission paths or other transmission systems), to a biological system (e.g., human tracking action, medical diagnosis), and perhaps also to an economic and sociological system.

Of course, the representation can be complicated by several factors. First of all there may be the multiple-input/multiple-output situation where each input may affect a number of outputs. In this case the input and output may be represented by vector quantities \( u(t) \) and \( y(t) \) respectively, where each component of the vector is a time function corresponding to an input or output signal. Another type of complication is found in that situation where the noise is not of the additive type but where it is multiplicative (e.g. a changing gain or attenuation that is not a parameter of primary interest).

In Section 1.1 we emphasized the importance of models. The estimation problem can conceptually be stated in the following way. The process and a model are subjected to the same input signal. The process output, contaminated with noise, and the model output are compared. Then it is necessary to find an adjustment of the model that is optimal in some predefined sense.
The parameter values are not directly observable. Consequently the optimum has to be defined using a criterion with respect to the output signal(s), or a criterion related to the expected error of the estimated parameter value; cf. sect. 2.3. The output signal is related to the parameter values by a functional relationship. This causes the complications of the process-parameter estimation and, generally speaking, makes the introduction of computer elements necessary. It was mentioned before that the criterion is often a minimization of a scalar loss function. In such a case it is natural to ask several questions:

- Is the minimum achieved?
- Is there a unique solution?
- Is the uniqueness of the solution influenced by the choice of input signals?
- If the solution is not unique, what is the character of the models which give the same value of the loss function and how should $S$ be restricted in order to ensure uniqueness?

Answers to some of these problems have been given for a simple class of linear systems arising in biomedical applications by Bellman and Åström (1969). The loss function is chosen ad hoc when the identification problem is formulated as an optimization problem and it is a consequence of other assumptions when the problem is formulated as an estimation problem.

Often the criterion is expressed as a functional of an error e.g.,

$$E(y, y_M) = \int_{0}^{T} e^2(t) dt$$

where $y$ is the process output, $y_M$ the model output and $e$ the error; $y$, $y_M$ and $e$ are considered as functions defined on $(0, T)$. Notice that the criterion can be interpreted as a least squares criterion for the error $e$. 
The case
\[ e = y - y_M = y - M(u) \]
where \( M(u) \) denotes the output of the model if its input is \( u \), is referred to as the \textit{output error}. It is the natural definition when the only disturbances are white noise errors in the measurement of the output; fig. 1.2a.

The case
\[ e = u - u_M = u - M^{-1}(y) \]
where \( u_M \) denotes the input of the model which produces the output \( y \), is called the \textit{input error}; fig. 1.2b. The notation \( M^{-1} \) implies the assumption that the model is invertible, roughly speaking that it is always possible to find a unique input which produces a given output. ¶

¶ Rigorous definitions of the concept of invertibility are discussed by Brockett and Mesarovic (1965), Silverman (1969), Sain and Massey (1969).
From the point of view of estimation theory, the criterion with the error defined as the input error would be the natural criterion if the disturbances are white noise entering at the system input.

In a more general case the error can be defined as

$$e = M_2^{-1}(y) - M_1(u)$$

where $M_2$ represents an invertible model; fig. 1.2c. This type of model and error are referred to as generalized model and generalized error; Eykhoff (1963).

A special case of the generalized error is the "equation error" introduced by Potts, Ornstein and Clymer (1961).

Another type of identification problem is obtained by imbedding the problem in a probabilistic framework. If $S$ is defined as a parametric class, $S = \{S_\beta\}$, where $\beta$ is a parameter, the identification problem then reduces to a parameter estimation problem. Such a formulation makes it possible to exploit the tools of estimation and decision theory. In particular it is possible to use special estimation methods e.g., the maximum likelihood method, Bayes' method, or the min-max method. It is possible to assign accuracies to the parameter estimates and to test various hypotheses.

However, in many probabilistic situations it turns out that the estimation problem can be reduced to an optimization problem. In such a case, the loss function is, however, given by the probabilistic assumptions. Conversely,
to a given loss function it is often possible to find a probabilistic interpretation.

The estimation problem is closely related to the statistical theory of optimal systems (Pugachev, 1963). The optimum filtering problem (fig. 1.3a) belongs to the same category as does the prediction problem (fig. 1.3b and c) and combinations of these problems. It is also related to adaptation and learning problems (Tsypkin, 1966, 1967, 1968, 1970).

The problem becomes still more complicated if the "process" being studied is part of a closed loop. Fel'dbaum (1960, part 1) considers the situation where the control signal is used for a twofold purpose:
- to study or learn the characteristics of the process or the ways of controlling it.
- to implement the regulation, to direct the process to the required state.

For the combination of these investigating and directing actions he coins the word dual-control. He points out that these two functions are sometimes separated in frequency range (frequency sharing operation) or in time intervals (time sharing operation). This separation, however, is not necessary and a general theory on these dual control systems is given by Fel'dbaum (1960, 1961). The essential problem is indicated by: "In dual control systems, there is a conflict between the two sides of the controlling process, the investigational and the directional. An efficient control can only be effected by a well-timed action on the object. A delayed action weakens the control process. But the control can only be effective when the properties of the object are sufficiently well known; one needs, however, more time to become familiar with them. A too 'hasty' controlling member will carry out the operational movement without making proper use of the results of trial investigations performed on the object. A too 'cautious' system will bide its time unnecessarily long and process the received information without directing the object to its required state at the right time. In each
case, the control process may not prove the best one and may not even prove to be up to the mark. Our problem is to find out, one way or another, which combination of these two sides of the regulation would prove to be most suitable. The operations must be so selected as to maximize a criterion of the control's quality.”

"Two factors should be taken into account by the controlling member which decides the specific amount of regulating movement at any given moment of time:

a. The loss occurring in the value of the quality criterion due to the fact that the outcome of the operation at a given moment and at subsequent moments of time, will cause a deviation of the object either from the required state or from the best attainable one. The average value of the loss shall be called the **action risk**.

b. The loss occurring in the value of the quality criterion due to the magnitude of the controlling action has not proved the best to obtain information on the characteristics of the object; in view of this, the subsequent actions will not be the best possible ones either. The average value of this loss shall be called the **investigation risk**.

Another point of general interest is the relation between process parameters and signal parameters. Many types of deterministic signals can be generated, at least conceptually, as an impulse response or step response of a linear or non-linear filter. Many types of stochastic signals can be obtained by the filtering of white noise. This implies that in such cases signal characteristics can be expressed in filter (process) parameters and vice versa. Consequently there is a close tie between process parameter and signal parameter estimation.
1.5 Normal operating conditions

In many cases, the performance of an experiment implies that a number of important environmental conditions are preset or under control. This means that the experiment can be repeated. In other cases, e.g. celestial mechanics (before Sputnik) one had to be satisfied by making passive observations only.

There are cases in which the duration of the experiment is not an essential issue. In terms of accuracy, the amount of information that can be derived from the experiment is not restricted by time or by the number of trial runs. Restrictions may be induced by economic considerations or by other practical limitations. Of course, there also do exist fundamental limitations such as the uncertainty principle of Heisenberg, which states for example, that it is impossible to specify or determine simultaneously both the position and velocity of a particle as accurately as may be wished.

The situations where process parameter estimation is applicable are often unlike the experimental situation indicated before. The observation time is restricted, either through economic necessity or for fundamental engineering reasons. Included in the latter are process characteristics that change in time and yet have to be determined. Then, parameter estimation has to be performed under normal operating conditions.

In some cases this has to be achieved through the use of (random) signals that may be present already; in other cases test signals of a special class with a sufficient level may be introduced.

It may be noted that in the previous lines the practical situation is clearly over-simplified. Many actual experiments do have aspects of the experimental as well as the normal operating conditions. A whole range of
situations can be imagined. At one end of this range are the astronomical observations of celestial mechanics; this process is clearly under normal operating conditions. An example at the other end of the scale is the determination of transfer characteristics of a network in the laboratory using sinewave generators, etc.

Most cases, e.g., industrial-process measurements, will be in between these extremes because only part of the environment will be under control. It will often be necessary to make some kind of experiment, observe the process while using perturbations as input signals and observe the corresponding changes in process variables. In practice there are, however, often severe limitations on the experiments that can be performed. In order to get realistic models it is often necessary to carry out the experiments during normal operation. This means that if the system is perturbed, the perturbations must be small so that the production is hardly disturbed. It might be necessary to have several regulators in operation during the experiment in order to keep the process fluctuations within acceptable limits. This may have an important influence on the estimation-results.

When carrying out identification experiments of this type there are many questions which arise naturally:

- How should the experiment be planned? Should a sequential design be used, i.e. plan an experiment using the available a priori information, perform that experiment, plan a new experiment based on the results obtained, etc. When should the experimentation stop?

- What kind of analysis should be applied to the results of the experiment in order to arrive at control strategies with desired properties? What confidence can be given to the results?
- What type of perturbation signal should be used to obtain as good results as possible within the limits given by the experimental conditions?

- If a digital computer is used, what is a suitable choice of the sampling interval?

In spite of the large amount of work that has been carried out in the area of system identification, we have at present practically no general answers to the problems raised above. In practice, most of these general problems are therefore answered in an ad hoc manner, leaving the analysis to more specific problems. It is worthwhile to recognize the fact, that the final purpose of identification is often the design of a control system, because this simple observation may resolve many of the ambiguities of an identification problem. A typical example is the discussion of whether the accuracy of an identification should be judged on the basis of deviations in the model parameters or in the time-response. If the ultimate purpose is to design control systems, then it seems logical that the accuracy of an identification should be judged on the basis of the performance of the control system designed from the results of the identification.

In any case, it is of vital importance that

- the input signal should excite all modes of the process, i.e., the process has to be controllable, cf. chapter 4, and the signals have to be of a class called persistently exciting, cf. section 11.4

- the output signal contains sufficient information on "what is going on in the process", cf. chapter 4: observability and identifiability.

Generally speaking, the time interval needed for the estimation has to be as short as possible in view of the need for actual information on the (slowly
changing) parameters. On the other hand, the minimal time interval is bounded by the statistical effects (inherent variance of the measurements, influence of additive noise). In a sense, the time interval needed for the estimation can be related to the instrumentation costs. This can be explained by consideration of fig. 1.4.

Assume that only one process parameter $b$ has to be determined. One could then use $n$ models, each with the same structure and a different parameter value $\beta_i$ ($i=1, \ldots, n$). The output of the process is compared with that of each of the models. The difference or error is weighted by the boxes marked C according to some criterion, e.g., by squaring and integrating over a time-interval $T$. Now the value of $b$ is known after one such "measuring time-interval" $T$, for we know that $b \approx \beta_i$ if $E_i < E_j$ for each $j \neq i$.

Under stationary conditions, the same knowledge can be obtained using one model and setting in the $j$-th measuring time-interval $\beta = \beta_j$. Consequently in this case the desired information is obtained only after an interval of length $nT$; the instrumentation costs, however, are much less than in the previous case. For this reason one has to be careful in comparing different parameter-estimating schemes. For more parameters, say $m$, the total number of models or the total number of measurement intervals becomes extremely large, viz. $n^m$.

Although it was mentioned before that there exists a continuous range of cases, the situation indicated by "under normal operating conditions" differs from the measurements made under test-conditions. This difference consists of:

- the information stored in the process (history);
- the time requirements with respect to measurements and evaluation;
- the choice of test signals to be used.

These differences necessitate a detailed study of the efficiency of the
information-processing scheme that one wants to use for the estimation under consideration. The results of such an analysis have to be given in terms of:

- the speed of convergence towards the asymptotic value;
- the bias of the asymptotic value compared to the true parameter value;
- the variance of the resulting parameter information as a function of time;
- the influence of additive noise and of non-linearities in the process; etc.

These aspects have to be compared to the speed with which process parameters may change.

Let us consider briefly the complications mentioned above.

The information stored in the process depends on the history of the input(s). For the elementary case of a linear time-invariant process, this effect can be seen from the convolution form of the input-output relation

\[ y(t) = \int_{-\infty}^{t} u(\theta) h(t-\theta) d\theta \]

where \( u \) is the process input, \( y \) is the undisturbed process output and \( h(t) \) is the process impulse response. Starting from a given time \( t=T \) we wish to observe the input and output in order to obtain information on \( h(t) \). The output \( y(t) \) for \( t>T \), is determined also by \( u(t) \) for \( t<T \). This contribution to the output can be written as:

\[ y(t) = \int_{-\infty}^{T} u(\theta) h(t-\theta) d\theta \]

It acts as a disturbance to the output, since it is not related to the part of the input signal \( u(t) \) for \( t>T \) which we are observing. This difficulty can also be described by using state-space techniques, cf. chapter 4. Roughly speaking, a state space of dimension \( n \) can be assigned to each process described by a differential equation of order \( n \). The coordinates of this space represent \( n \) independent variables, e.g. the values of the derivatives of order 0 to \( n-1 \). At each time \( t \) the "state" of the system is given by a point in the state space called the state vector \( x(t) \).
This point moves under the influence of the equilibrium-restoring forces (dissipation of energy) and the forcing functions (input signals).

At time $T$ the state $x(T)$ contains all information of the past history of the system. Under force free conditions (the input signals disconnected) the state vector of a stable system will in time tend to an (some) equilibrium state position $x_0$. This component of motion of the state vector, the free response, is added to that caused by the action of the input signals.

Thus the free response acts as a disturbance signal if the input and output signals are observed say on a time interval starting at $t=T$. An extended form of the estimation problem is then: estimate the unknown parameters and the unknown state at $t = T$.

The time requirements for the measurement and evaluation are clearly dependent upon the fact that the parameter estimation has to use signal characteristics which cannot be obtained until after that signal has been observed over some interval of time. In general, there will be a close relationship between the length of the interval and the accuracy of the information obtained. The choice of test signals is also pertinent to the previous remark.

In the experimental situation in the laboratory, large or periodic test signals can often be used. Under normal operating conditions, this is sometimes the case as well. If test signals are not allowed, then we have to make the best possible use of the available disturbances. If only a small signal may be added to the input then we have to consider the signal to noise ratio, the best choice of the predetermined characteristics of the test signal, the influence of the type of signal on the information processing operation, etc.

When the process parameters are changing, then there is a fundamental uncertainty in the estimation results.
1.6 Some fields of application

The following diagram indicates the theories that are relevant for parameter and state estimation. It also indicates some fields of application.

Theories

- measurement theory
- control theory
- system theory; linear and nonlinear models
- signal theory; deterministic and stochastic
- information theory
- statistics
- stochastic approximation

Parameter and state estimation

Ways of application

- "diagnostic" applications
- control applications
- automatic (industrial) measurements
- automatic (industrial) decisions
- automatic (industrial) adjustments
- pattern recognition

Fields of application

- communication engineering
- power engineering
- mechanical engineering
- aeronautical engineering
- chemical engineering
- physics
- geology
- economics
- biology/medicine

Related problems

- signal detection
- optimal filtering
- prediction
- learning
Through the stochastic-signal theory, stochastic approximation, statistics and information theory a considerable use is made of probabilistic notions and results. The basic aspects needed, and assumed well-known, are summarized in appendix B. The theoretical aspects of signal and model description are discussed in chapters 3, 4 and 5.

The list of applications suggest that the potential uses of parameter estimation extend beyond those in control engineering. In each application, the goal of the estimation procedure has to be clear beforehand. This goal determines the accuracy needed and the effort permitted.

**Estimates for "diagnostic" applications**

In this category belong the measurements of transfer characteristics of a process, e.g., the measurements of a Bode plot, the determination of the impulse response, the determination of poles and zeros of a transfer function, the estimation of nonlinear effects. The celebrated Fourier filtering method, i.e., the instrumentation of the equations:

$$
\text{Re} \left[ H(j\omega) \right] = \lim_{k \to \infty} \frac{1}{kT} \int_{0}^{kT} y(t) u \sin \omega t \, dt
$$

$$
\text{Im} \left[ H(j\omega) \right] = \lim_{k \to \infty} \frac{1}{kT} \int_{0}^{kT} y(t) u \cos \omega t \, dt
$$

where:

- $u(t)$ is the process input signal.
- $y(t)$ is the process output signal.

belong to this class. Apart from applications in electrical and electromechanical systems one also finds applications in:

- chemical engineering, e.g., the determination of catalyst aging or poisoning; the determination of rate coefficients in chemical reactions.
- industrial processes, e.g., the determination of heat transfer coefficients.
- physical measurements, e.g., the measurement of the electrical impedance of a mercury electrode (i.e., an electrode that consists of a mercury drop that grows at the end of a tube. After this reaches a certain volume
then it falls and a new drop starts growing etc. With such an arrange-
ment there is always a "clean" electrode surface available).

- mechanical engineering, e.g., the measurement of the transfer characteristics
  of railroad wagons, textile machines.

- nuclear reactors, e.g., the determination of reactivity coefficients.

- aircraft and spacecraft engineering, e.g., the determination of the airplane
dynamics during flight.

- biology and medicine, e.g., the determination of the human characteristics
  when performing a "tracking action"; a simple arrangement is shown in fig.
1.5. where an optical signal is given and where the assignment is to track
the spot on the cathode ray tube as accurately as possible. Studies can be
made of changes in the characteristics, e.g., due to "learning", fatigue,
environment and drugs.

Other examples from a large variety of possibilities: the pupil reaction
on incident light; the determination of the parameters for a blood-circu-
lation model.

- economy.

In the following categories the accent lies on the type of application for
which the estimate is to be used.

**Estimates for control applications**

If the purpose of the identification is to design a control system, the cha-
racter of the problem might vary widely depending on the nature of the con-
trol problem. A few examples:

- Design a stable regulator.

- Design a control program for optimal transition from one state to another.

- Design a regulator which minimizes the variations in process variables due
to disturbances.
In the first case it might be sufficient to have a fairly crude model of the system dynamics. The second control problem might require a fairly accurate model of the system dynamics. In the third problem it is also necessary to have a model of the environment of the system.

The danger of identifying systems under closed loop control also deserves to be emphasized. Consider the classical example of fig. 1.6.

An attempt to identify $H_p$ from measurements of $u$ and $y$ will give

$$\hat{H}_p = \frac{1}{H_R}$$

i.e. the inverse of the transfer function of the feedback. In industrial applications the feedback can enter in very subtle ways e.g. through the action of an operator who makes occasional adjustments. Fisher (1965) has shown the interesting result that the process may be identified if the feedback is made nonlinear.

**Optimal control systems.** Here the optimality is based mostly on an economic criterion. In many cases the quality of the optimal control appears to be strongly dependent on the information that is available concerning the dynamical characteristics of the process that has to be controlled. If these characteristics change in time we have to determine them continuously. For optimal control the state of the process is also of great importance. This leads to a combination of parameter and state estimation. As pointed out by Aoki (1967), only for a
limited class of systems (viz. a linear process and using a quadratic cost function) can the optimal control scheme be divided into an optimal state estimator followed by an optimal controller; cf. chapter 14. For other types of systems, this division into two separate problems leads to an approximation of the optimal control. Yet it may be an important aid to make the control problem tractable. As examples we can use the approach of and the landing of an airplane, the control of batch processes, the control of continuous industrial processes, where the optimal setpoints of the controllers are determined by use of a (partial) model of the process. Interesting examples are described in the case studies (Peschon and others, 1968):
- on optimum control of electric power systems
- of Kalman filtering in the C-5 aircraft navigation system
- in the paper-making industry
- on the optimal control of a flexible launch vehicle.

For other examples the reader is referred to chapter 14.

Self-optimizing and self-adapting systems. For this class, the same holds as for the previous category. Now we want to adjust the parameters of the control system, based on the knowledge of the parameters of the process, in such a way that the transfer characteristics are optimal. An illuminating example of this is the airplane which, as a result of the difference in air density close to the earth surface and at 30,000 ft. altitude, shows a great difference in dynamic behaviour. Another illustration of adaptive control by process-parameter estimation is found in driving an automobile. Due to the weather conditions, there is a chance that the road is slippery. This may be detected by the introduction of small perturbations with the steering wheel. Evaluation of the results of these perturbations gives the information that is used in controlling the speed.
Estimates for **automatic (industrial) measurements**

On account of the economic aspects in this class, the speed of measurement is of great importance. As an example of a new approach to an old problem the automatic measurements of impedance can be used; cf. fig. 1.7. The impedance can be described by the differential equation

\[
C \frac{d^2 u}{dt^2} + \frac{1}{R} u - i = 0
\]

The model satisfies the equation

\[
y \frac{d^2 u}{dt^2} + \rho u - i = e
\]

Due to the instrumentation shown in the diagram the model will be adjusted in such a way that the error signal approaches zero, that is to say its mean squared value will be minimized. This principle can be extended in several directions, among which the determination of more complicated impedances of nonlinear devices e.g. quartz-crystals, electromechanical transducers.

**Estimates for automatic (industrial) decisions**

In these applications the accent also lies on the speed of information processing, particularly if this speed is so great that the human being is not able to do the job.

As an example one may consider the decision of whether or not to switch off a generator in an electrical power station if there appears to be a short circuit in the distribution system. Here the criterion is whether or not the generator will fall out of step. This will be governed by the state variables of the generator and the power distribution system.

**Estimates for automatic (industrial) adjustments**

An automatic adjustment can be considered as a series of repeated decisions on the continuation and the direction of an adjustment. The adjustment of band-pass filters during the manufacturing of radio and television sets can be interpreted as the adjustment of a model (the bandpass filter to be adjusted) according to a standard or reference circuit.
Another application is the compensation of a communication channel with regard to the inter-symbol interference. These properties change if the cable configuration changes by interconnections. Using a filter that consists of a delay line and potentiometers, one can obtain an automatic compensation (adaptive filter) that adjusts itself automatically with respect to the minimal interference.

Pattern recognition

This can be considered as the determination of a number of "parameters" of the object under study together with the decision, based on these estimations, as to which class this object belongs. The most crucial question with respect to pattern recognition is: what parameters are the most important?

As far as applications are concerned, one can think of character recognition, but also of signal classification (cardiographs etc.) and e.g. the classifications of data obtained by oil-exploration.

An interesting list of possible applications of parameter estimation given by Clymer (1958) includes the following topics:

A. Aircraft and Automatic Industries.

- determination of aerodynamic coefficients from
  a) flight tests involving complicated manoeuvres
  b) dynamic tests in wind tunnels with elastic mountings
  c) rocket sled tests
- determination of the flutter speed of an airplane or missile surface from subcritical tests of
  a) airplanes in flight
  b) wind tunnels flutter models
  c) rocket sled models

by determining the damping of the flutter mode as a function of frequency
- determination of the flexibility characteristics of any structure
- determination of servo internal characteristics and output impedance
- determination of the characteristics of the interior ballistics of guns, rocket capsules, and the like
- determination of the dynamic characteristics of tires
- determination of the nonlinear characteristics of airplane control sticks, rudder pedals, and steering wheel linkages
- determination of the dynamic characteristics of landing gears and automotive suspension systems
- determination of the dynamic characteristics of turbojet, turboprop, reciprocating, free-piston, and all other engines, compressors, etc.
- determination of the parameters governing transient air flow through a porous material such as a fabric
- determination of the course and speed of a target as an aid to tracking of fire control
- determination of the characteristics of a "satellite" and its upper atmosphere environment, knowing the history of its motion (i.e., its orbit and perturbations)
- design of aircraft landing control systems, midair refueling systems and other special purpose computers using prediction

B. Process Industries (chemicals, fuels, materials, etc.)
- determination of the chemical kinetic and other parameters which
determine the reaction rates in batch or continuous processes
- determination of: (1) the half lives of the constituents, or
  (2) the composition of a radiochemical mixture
- determination of the dynamic rheological characteristics of
  materials undergoing deformation or flow
- determination of the lumped parameters in sets of ordinary
  differential questions used to approximate distributed-
  parameter systems such as heat exchangers
- determination of the characteristics of combustion processes
- determination of neutron process parameters in nuclear
  reactors
- determination of the dynamic parameters of instruments,
  transducers, and process control system components

C. Miscellaneous
- determination of dynamic parameters in biological population
  systems
- determination of the characteristics of systems for which
  forecasts are to be made, such as dynamic economic systems
  (sales of product, growth of new industry etc.)
- determination of the dynamic characteristics of a human
  operator of various devices when he is a part of a loop
- determination of the dynamic characteristics of the regulatory
  system in human or animal physiology.

Summarizing, one may view the parameter and state estimation as an
extension of the classical measurement theory to more complicated
situations and tasks, made possible by the advent of modern computing
equipment.
1.7 A wider perspective

In a number of fields of physical and engineering science theorems have been derived with respect to the maximum result obtainable. Some examples:

- thermodynamics: the maximal efficiency of energy conversion in a given situation is determined by the Carnot-cycle;
- physics: the uncertainty relations of Heisenberg indicate the limit of accuracy that can be obtained by simultaneous measurement of certain quantities;
- communication theory: the Shannon-relations indicate the (maximal) channel capacity.

In addition in the fields of parameter and state estimation there is a need for such limiting theorems. It would be most useful if it were possible to indicate the maximum amount of "information" about a parameter or a state vector that could be derived in clearly specified situations. In that way different estimation techniques or schemes applied to a given situation, could be compared just as communication techniques can be evaluated by communication (information) theory.

In fact, the comparison of parameter and state estimation with the communication problem suggests itself quite strongly. The values to be estimated are the messages sent by an information source; the transmitter, channel and receiver represent the information process scheme. The (channel) noise stands for the disturbances in process and measurement devices. Appealing as this idea might be, a well developed theory along these lines seems not yet to be available, in
spite of some problem statements made (e.g. Zaborszky, 1966) and much material being available (Lindley, 1956; Kullback, 1959; Rényi, 1966; Barnes, 1968; Weidemann, 1969; Weidemann and Stear, 1970). Essential difficulties due to nonlinearities are met if one tries to describe the development of crucial probability functions with respect to estimation time.

The best limiting theorem available at present is the Cramér-Rao inequality, cf. chapter II, which provides knowledge of the accuracy that can be obtained under certain prespecified conditions. It does not give, however, the possibility of taking into account the available a priori knowledge on the parameters before the estimation measurements and data processing.


Lindley, D.V. (1956). On a measure of the information provided by an experiment. 


Chapter 2  BASIC APPROACHES TO THE PROBLEM;
Statistical and engineering

This chapter gives a survey of the parameter estimation problem. The purpose is mainly to indicate the relations and distinctions between the different estimation techniques. This chapter should provide the reader with a clear picture of the organization of this text. There is nothing in this chapter that will not be discussed more fully and adequately in the chapters to follow. Consequently, the reader should not be alarmed by notions and discussions that yet may be unfamiliar to him. On the other hand he is advised to return to this chapter once in a while to develop his overall picture of this field.

Some general aspects of the estimation problem, treated in section 2.1, include process input signals, the a priori knowledge about the process and the choice of estimation schemes. Attention is drawn to the existence of two classes of implementation in section 2.2.

In section 2.3, the statistical considerations that result in estimates providing explicit answers (expected value, variance or confidence interval) for the parameters to be determined are outlined.

In section 2.4, attention is devoted to model-adjustment approaches. These imply the need of deriving partial derivatives of the error criterion with respect to the parameters to be adjusted. Some ways of obtaining these derivatives are given.

A particularly interesting situation arises if the error is a linear function or functional of the parameters to be adjusted. This case leads to the generalized model approach.
2.1 General aspects of the problem

Process input signals. The estimating schemes are partly determined by the type of signals that are available. The signals may be either "continuous in time" (non sampled) or sampled. With the proper technical means the sampled and continuous forms can be converted into each other, provided that the sampling frequency \( f_s \) satisfies the Whittaker-Shannon relation:

\[
f_s > 2 f_h
\]

where \( f_h \) is the highest frequency occurring in the band-limited continuous signal. Under that condition, the amount of information available does not decrease through the sampling process. Ways of describing signals and signal characteristics are discussed in chapter 3.

A list of symbols used in this text is given in Appendix A.

In cases where the addition of a test signal is permissible, the amount of freedom in choosing such a signal is an important aspect. In those circumstances, one has to consider: the bandwidth of the signal, the maximum permitted energy and amplitude, and the properties with respect to the generation and the information processing of the test signal; cf. chapter 10.

A priori knowledge about the process. In chapter 1, the importance of initial knowledge of the process is indicated. That knowledge is represented in the structure and possible parameter values of the physical or mathematical (a priori) "model" of the process. As regards implementation, the availability of such knowledge may amount to a difference in complexity and investment of an order of magnitude. The same is often true for the distinction between nonlinear and linear aspects of the process. Some knowledge on the speed of parameter variation is a basic necessity as this may determine the feasibility of any type of implementation for the problem.

Although the "black box approach" is attractive from the theoretical point of view it often leads to solutions which are highly wasteful due to neglecting some of the knowledge which is available. In many practical situations
there may also be more measurable process quantities than merely one output. If no use is made of these signals it inevitably means a loss of available information, (unless the dynamic relation between those signals and the output is known deterministically e.g., a first order low pass filter with a time constant of $\tau$ seconds, and if no noise is added). The use of such additional "outputs" has to be constantly kept in mind. Based on the knowledge of the process and the purpose of the estimation, it is mostly clear what must be demanded of a model. The parameters of the model must preferably be such that their values are a direct measure of the parameters of interest. Now, there are many different model representations available: differential equation model, transfer function model, series expansion models in time or frequency domain, to mention only some of the most common ones. Such representations are discussed in chapter 4. The choice of the type of representation is strongly case-determined.

**Estimation scheme.** For each particular type of description the requirements posed on the implementation of the estimation problem will be to find the characteristics of the process:
- accurately, not affected by additive noise and possible unwanted non-linearities in the system
- rapidly, yet under the condition that the system always has to be stable
- economically, using an instrumentation that is feasible and acceptable as regards costs.

The estimation scheme should preferably satisfy the following requirements:
- it should be mathematically tractable
- it should be easily implemented
- it should be generally applicable
- it should yield an "optimum" estimate
- it should yield an acceptable speed of convergence.
From the above list, which of course could be expanded, it will be obvious that the choice of an estimation scheme is governed by many considerations. Some of the general statistical and convergence notions are discussed in chapter 5.

For the reader who likes to review the notions from probability theory a survey is given in Appendix B.

2.2 Classes of implementation.

A distinction can be made with respect to the type of implementation used for solving the parameter estimation problem. In principle one may distinguish two different classes (Eykhoff, 1963):

- Implementation of mathematical relations that result in numerical quantities (e.g., coefficients of the differential equations; characteristic frequencies and damping ratios of poles and zeros; a number of points of the impulse response). This implementation will be denoted as: "using explicit mathematical relations".

- Implementation in the form of a (physical or mathematical) model of which the parameters are controlled in such a way, that the model characteristics approach the characteristics of the system under study in some pre-defined sense. This will be called: "using a model-adjustment technique".

This difference is illustrated by the block diagrams of Figure 2.1a and b; one realization of these two classes is shown in Fig. 2.1c and d.

In analytical terms the difference amounts to the following. Suppose the process output \( y(t) \) is given by:

\[
y(t) = f(u,b,n)
\]  
(2.1a)

and the output \( y_M \) of an (assumed) model by:

\[
y_M(t) = f(u,\theta,0)
\]  
(2.1b)

where
\( b \) is the \( m+1 \)-dimensional vector of process parameters, i.e. a shorthand notation for the parameters \( b_0, \ldots, b_m \).

\[ u = u(t) \text{ is the process and model input} \]

\[ n = n(t) \text{ is the noise (disturbance)} \]

\( \beta \) is the \( m+1 \)-dimensional vector of parameter estimates.

\( f \) is a functional relationship between input and output quantities.

This choice implies that we assume that the process and the model can be described by the same "structure" and that the model does not suffer from an (appreciable) amount of noise.

As an example, let us assume that the correspondence between process and model is measured according to an error criterion:

\[
E \{ y, y_M; \beta \} = \int_{t-T}^{t} (y(t) - y_M(t; \beta))^2 \ dt
\]

Now one can follow two strategies:

- Put

\[
\frac{\partial E}{\partial \beta_i} = 0 \quad \text{for } i = 0, \ldots, m.
\]

This is a necessary condition for obtaining the minimal error. These \( m+1 \)-equations with the \( m+1 \)-unknown estimates \( \beta_0, \ldots, \beta_m \) can be solved for the \( \beta \)'s thus providing the explicit mathematical relations.

(The additional requirement that \( \frac{\partial^2 E}{\partial \beta_i^2} > 0 \) is tacitly assumed to be fulfilled).

- Put

\[
\frac{\partial E}{\partial \beta_i} \to 0 \quad \text{for } i = 0, \ldots, m.
\]

A convergence towards zero can be obtained if \( \frac{\partial E}{\partial \beta_i} \) can be determined by a suitable instrumentation and if these derivatives are used for the adjustment of a physical model.

Instead of these names chosen for the two classes, the following distinctions could also be used:
Class I
- using explicit mathematical relations or
- explicit methods or
- open loop methods or
- direct methods or
- accumulative or "one-shot" procedures

where the solution:
- is available after a finite number of elementary operations
- requires considerable memory
- is not available in an approximate form as an intermediate result
- open loop with respect to the estimate

Class II
- using model-adjustment techniques or
- implicit methods or
- closed loop methods or
- iterative methods

where the solution:
- is available after (in principle) an infinite number of elementary operations
- requires less memory
- is available in an approximate form as an intermediate result
- closed loop with respect to the estimate

The reasoning behind these aspects will be clarified in
Chapter 6 for sampled signals | Chapter 7 for sampled signals
Chapter 8 for non-sampled signals | Chapter 9 for non-sampled signals

These differences are of great importance in engineering sense.
The number of operations and the memory required are relevant with respect to the computational means chosen for the implementation, viz. digital, analog, or hybrid. The availability of intermediate results may be essential for a particular type of application e.g., in those cases where new measurements come available, and thus have to be used for on-line updating of the estimates (tracking); particularly for control purposes such intermediate results are important. Consequently the distinction between off line and on line methods is relevant. The model offers the possibility of using a priori knowledge about the process (structure, initial estimates of parameters).
The accuracy of the solution depends more strongly on the quality of some of the implementation components for class I than for class II, due to the
properties of the (self-correcting) closed loop that can be recognized in class II.

The diagrams in figure 2.1 are rather general ones; there is not only much freedom as regards the description of the signals, cf. chapter 3, but they apply also to a wide variety of process descriptions (models), e.g., impulse response model, transfer function model, orthogonal series expansion model, state equation model. There is a large body of literature on these representations of processes, cf. chapter 4.

2.3 Classes of statistical approaches
Among the problems studied in statistics are:
- the condensation of information, given by a large set of data, to a few characteristic numbers (data reduction).
- the sampling from a (conceptually) large population and the drawing of inferences about the properties of this large population from the properties of the sample taken. The goals of this procedure may be the following: 1) to test a statistical assumption, e.g. that a property of the population has a probability function of a certain class (hypothesis testing); 2) to find numerical values for the estimate and the variance of an unknown parameter in the distribution function of the population (population-parameters).
- the analysis of the relationship between two properties, e.g., to find numerical values for the estimate and its variance for coefficients in an assumed relationship between quantities \(u\) and \(y\) (dependence-parameters).

The process-parameter estimation problem may be approached as a study of population-parameters or of dependence-parameters. The latter point of view is obvious if \(u\) and \(y\) represent the data available from the process input and output signals. The interpretation of a population-parameter study is arrived
at by using the method of maximum likelihood estimates.

The results desired from the estimation procedure may be of a different kind, dictated by the type of application at hand and the available a priori knowledge. Consider a process represented in a state description form:

\[ X(t) = f(\bar{X}, u, b) \]
\[ y = g(\bar{X}, u, b, n) \]  \hspace{1cm} (2.3a)

with

\[ u = u(t) = \text{input signal} \]
\[ y = y(t) = \text{output signal} \]
\[ n = n(t) = \text{noise} \]

\[ X = X(t) = \text{process state vector} \]
\[ b = \text{process parameter vector} \]
\[ x(t) = [x_1(t) \ldots x_n(t)] \]
\[ b = [b_0 \ldots b_m] \]

and a model description of the same form:

\[ \dot{\bar{X}} = f(\bar{X}, u, \beta) \]
\[ y_M(t) = g(\bar{X}, u, \beta, \sigma) \]  \hspace{1cm} (2.3b)

with

\[ \bar{X} = \bar{X}(t) = \text{model state vector} \]
\[ \beta = \text{model parameter vector} \]

A simplified linear case is sketched in fig. 2.2. Now the estimation problem may be formulated as a minimization of:
Some function(al) of $e = y - y_M$, the "error" between the output of the

process and a model, can be minimized as $e$ can be made measurable by proper
implementation or by programming the model on a computer. Depending on the model and on the input signal, this error provides some measure of the correspondence of the parameter and/or state vectors of process and model. In some cases, this correspondence of input-output relations is more important than parameter correspondence, particularly if the model is simpler, e.g., of a lower order, than the process.

If we focus our attention on $b$ and/or $x$ directly then we notice that $b$ and $x$ are inaccessible for direct measurement. Consequently, one can only minimize the expectation of $E \{ b - \beta \}$ or $E \{ x - \xi \}$. This, however, is possible only if some a priori knowledge is available with respect to probability densities.

As in many other situations, we are tempted to look for the "optimal" or
"best" estimation procedure. The definition of "best", however, depends on the type of application, on the requirements as regards the properties of the estimate, on the a priori knowledge about the process and the additive noise. In the statistical literature, a number of different estimation procedures have been developed. These methods differ predominantly in the criteria used for defining optimality and in the use of available a priori knowledge. It is unfortunate that the choice between these criteria has aspects that are more or less subjective and that the mathematical approach is strongly dependent on these criteria. In chapter 5, these estimation procedures and their relations are discussed. Here a brief preview is in order. Again it should be stressed that it suffices that from a first reading of this chapter the reader retains merely a general idea about the methods and the expressions that come into the picture.

A simple approach can be represented by a block diagram according to fig. 2.3. The case of sampled time functions is considered. The quantities \( u, u, y, \bar{n} \) and \( y \) represent vectors consisting of samples, e.g.

\[
y' = \begin{bmatrix} y(\theta), y(2\theta), \ldots, y(k\theta) \end{bmatrix}^{T}
\]  

(2.4)

\( y' \) denotes the transpose of the (column) vector \( y \). In the remaining part of the text the sample-interval \( \theta \) will be constant and will be deleted.

The quantities \( b \) and \( \beta \) represent vectors consisting of the process parameters and model parameters respectively:

\[
b' = \begin{bmatrix} b_0, b_1, \ldots, b_m \end{bmatrix}
\]

\[
\beta' = \begin{bmatrix} \beta_0, \beta_1, \ldots, \beta_m \end{bmatrix}
\]

Consequently, one finds:

\[
y = y(\underline{u}, \underline{b}, \bar{n})
\]

(2.5a)

\[
y_M = \underline{U}/\beta
\]

(2.5b)
In the matrix \( U \), each column \( u_i \) consists of sample values of the outputs of the blocks \( F_i \). These blocks are chosen on the basis of the process representation needed, cf. Chapter 4.

Now one can define the problem as finding the estimate \( \hat{\beta} \) of the process parameter vector \( \beta \), based on observations of \( y \).

Based on the available a priori knowledge one can make a distinction between different kinds of estimates:

a) **least-squares estimation.** This does not require any additional knowledge.

The error criterion is defined as:

\[
E = e' I e = \sum_{j=1}^{k} e^2(j)
\]

(2.7)

with

\[
e = y - y_M
\]

(2.8)

\( I \) is the identity matrix. From the minimization of the error-criterion the so-called "normal equation" for the estimator \( \hat{\beta} \) is found to be:

\[
\hat{\beta} = [U'U]^{-1} U'y
\]

(least-squares estimator) \(2.9\)

where again \( U' \) is the transpose of \( U \).

b) **Markov or generalized least-squares estimation.** Now it is assumed that we have some knowledge about the additive noise \( n \), viz. its covariance matrix.
It can be proved that if $N \neq I$ (identity matrix) a better linear estimator can be found by minimizing
\[ E = e' N^{-1} e \] resulting in
\[ \hat{\beta} = [U' N^{-1} U]^{-1} U' N^{-1} y \] (Markov or generalized least-squares estimator).

**c) maximum likelihood estimation.** It has already been mentioned that the estimation problem can also be approached as being a population-parameter study. Now the attention is focused on the output vector $y$.

**A priori** (before the measurement) we know that the samples of $y$ are random variables, that have the joint probability function
\[ p \{ y(1), \ldots, y(k); b \} \] (2.13)
cf. Appendix B; $y$ depends on $b$.

**A posteriori** (after the measurement) we know the observed sample values of these random variables, say $y(1) = c_1, \ldots, y(k) = c_k$. From this we want to determine an estimate $\beta$. The functional relationship between $e' = \{ c_1, \ldots, c_k \}$ and $\beta$ is the same as given in (2.13). In order to accentuate the dependence on the observed sample values we write this dependence as
\[ L \{ c_1, \ldots, c_k; \beta \} \] (2.14)
This is called the likelihood function. As the sample values are given by the measurement, this $L$ is a function of $\beta$ only. For the sake of convenience, $\ln L$ is usually considered; as the logarithmic function is monotonic, the maximum of $L$ and the maximum of $\ln L$ occur at the same value of $\beta$. This value of $\beta$ can be obtained by solving:
Such a solution $\hat{\beta}$ is called a maximum likelihood estimate of $\beta$, if $\hat{\beta}$ actually depends on the observed sample values $c_1, \ldots, c_k$.

From this outline it follows that not only the covariance matrix of the noise has to be known but also the joint probability density function of its samples.

d) **Minimum risk estimation.** In this case, even more a priori information is needed. As outlined in chapter 5, the probability density functions of $u$, $n$ and $b$ have to be given.

It is a matter of theoretical and engineering interest to consider how much available a priori knowledge will be used in practice. Generally speaking, the use of more a priori knowledge may lead to a better estimation scheme; the costs of implementation, however, may easily become prohibitive.

Whatever estimation technique is being used, it has to be based on observations, which are random variables due to the additive noise. Consequently, any estimate, being a function of these observations (statistic), is a random variable too.

### 2.4 Model-adjustment techniques

"Odd-function" and "even-function" error. The model-adjustment technique has been interpreted with fig. 2.1b as the task to adjust the model $M$ in such a way that the transfer characteristics of $M$ approach those of the process $P$ in some predefined sense. The crucial part of fig. 2.1b is the block denoted by I. This block processes the information that is available. The error may be chosen as $e(t) = y(t) - y_M(t)$ and the error criterion may be the minimization of any even-functional $E$ of this error e.g.,
\[ E = \int_{t_a}^{t_b} e^2(t, b, \beta) \, dt \quad (2.16) \]

where again \( b \) is the parameter of the process under consideration and \( \beta \) is the corresponding model parameter. Such an even-functional leads to the type of relation indicated in fig. 2.4 between \( E \) and the value of the parameter that has to be adjusted. This brings us to the distinction between the so called "odd-function" and "even-function" control. (Eykhoff, 1960).

Let us look at the human eye, which shows both these different types of control; cf. fig. 2.5a. The iris controls the illumination level of the retina in an ordinary control-system way; if the level is too high, then the error signal causes the iris to contract and vice versa. The sign of the error signal is a direct indication of the direction in which the iris has to change; cf. fig. 2.5b. The error signal goes to zero.

The lens controls the definition of the image on the retina. Now there is no error signal that changes sign when the lens accommodates from too far a distance to too close or vice versa. The function of definition versus amount of accommodation, and thus the error signal, has an extremum; cf. fig. 2.5c. We face the same situation with the error defined for the

\[ \frac{\partial E}{\partial b} = 0 \]

In this situation a number of different clues are being used for the control of accommodation.

In this situation a number of different clues are being used for the control of accommodation. This situation seems to be characteristic of adaptive and optimizing feedback control systems as well. Figure 2.6 gives a comparison of the goals of parameter estimation, optimization and adaption. Figure 2.7 shows the main difference between the ordinary control and this "adaptive" type of control. The two types of feedback systems will be called, with regard to the fundamental difference, "odd-function" (error) systems and "even-function" (error) systems. From 2.8 it is evident that when the sign of the second derivative is known (maximum or minimum of the function)
for an even-function system at least two observations are required to
locate the direction of the extremum as regards to these observations;
c.f. fig. 2.8c. In order to carry out control action we have to reduce the
information of the two or more observations to one quantity. This
data reduction changes the "even-function" error into an "odd-function"
error. After this reduction has been applied, ordinary control prin-
ciples based on this odd-function error can be used.

The direction towards the optimum adjustment is shown by the first
derivative with respect to $\beta$. For continuous (non-sampled) signals,
using expression (2.16) for the integral squared error one finds:

$$\nabla_\beta E = \frac{\partial E}{\partial \beta} = 2 \int_0^t c(t, b, \beta) \frac{\partial e}{\partial \beta} \, dt \quad (2.17)$$

Neither $\partial E/\partial \beta$ nor $\partial e/\partial \beta = -\partial w/\partial \beta$ are measurable explicitly without
additional instrumentation. Some techniques for obtaining the desired
information are summarized in the following. If such partial derivatives
can be determined then by "hill climbing" techniques, c.f. chapter 5, on
the function(al) $E$ one may find the "best" value of $\beta$.

For sampled signals one finds that if the representation chosen is adequate
for describing the process behaviour then, c.f. fig. 2.3:

$$y = U b + n \quad \text{process} \quad (2.18a)$$

$$y_M = U \beta \quad \text{model}$$

and consequently the error is found to be

$$e = y - y_M = U (b - \beta) + n \quad (2.18b)$$

Consequently a quadratic form

$$E = e' e \quad (2.19)$$

will be minimized if $b = \beta$. Again $\nabla_\beta E = \partial E/\partial \beta$ has to be found.
Assuming that $\partial E / \partial \beta$ can be determined, then the task is to find an implementation that gives an automatic adjustment of the parameters $\beta$. There are several possible ways of using this information for obtaining that parameter adjustment scheme. The adjustment can be continuous or intermittent.

- **Continuous adjustment scheme.** A favoured policy is obtained by choosing

$$ \dot{\beta} = - \Gamma \frac{\partial E}{\partial \beta} $$

(2.20)

Such a choice leads to the gradient method. The coefficient $\Gamma$ represents a "gain factor", governing the speed of the adjustment. We will find that this coefficient may be constant or time dependent; it may be a scalar $\Gamma$ viz. $\Gamma(t)$ or a matrix $\Gamma$ viz. $\Gamma(t)$.

Strictly speaking, $\beta$ has to be time-invariant while the gradient is determined. This is not the case in continuous adjustment schemes. For that reason eq. (2.20) offers an approximate description of the adjustment dynamics only if that adjustment is comparatively slow.

- **Intermittent adjustment scheme.** The problem of the gradient determination just mentioned does not occur if one uses an intermittent scheme: measuring the gradient while $\beta$ is kept constant, adjusting $\beta$, measuring again etc. This problem is studied in the theory of stochastic approximation (Tsypkin, 1966), which may lead to an adjustment algorithm as for example:

$$ \beta(i+1) = \beta(i) - \Gamma(i) \nabla \beta E(i) $$

(2.21)

with

$\beta(i) =$ the parameter vector of the model after the $i$-th adjustment

$E(i) =$ the value of the error criterion after the $i$-th sequence of error samples

$\Gamma(i) =$ a "gain factor" governing the speed of convergence.
A simple example may indicate again the connection between model adjustment techniques and the use of explicit mathematical relations. Assume that the adjustment criterion is:

$$\min \quad \mathbf{e}'\mathbf{e}$$

After the i-th model adjustment we know from eq. (2.9) that the optimal parameter is given by:

$$\hat{\beta} = [\mathbf{U}'\mathbf{U}]^{-1}\mathbf{U}'(\mathbf{y}_M + \mathbf{e}) =$$

$$= [\mathbf{U}'\mathbf{U}]^{-1}\mathbf{U}'\mathbf{\beta}(i) + [\mathbf{U}'\mathbf{U}]^{-1}\mathbf{U}'\mathbf{e} =$$

$$= \mathbf{\beta}(i) + [\mathbf{U}'\mathbf{U}]^{-1}\mathbf{U}'\mathbf{e}$$

This is the best least squares estimation, given the "a priori" model adjustment and the particular series of error samples. For consecutive model adjustments one finds; cf. equation (2.21):

$$\hat{\beta}(i+1) = \hat{\beta}(i) + [\mathbf{U}'\mathbf{U}]^{-1}\mathbf{U}'\mathbf{e}(i)$$

A measure for the "uncertainty" of this estimate, $\text{cov}[\hat{\beta}(i)]$, can also be determined.

In the same way, $\min \mathbf{e}'\mathbf{N}^{-1}\mathbf{e}$ could be used as the model adjustment criterion. Convergence schemes, that is the way in which the gradients are used by the implementation, are discussed in chapter 5.

Techniques for obtaining the partial derivatives or gradients. In principle the following approaches are available:

a) The use of parameter influence coefficients or parameter sensitivity functions. In recent years this notion has gained much interest, cf. chapter 9. A simple example following Meissinger (1960) may indicate the procedure. Let us assume the model to be of the first order having one unknown parameter $\beta$ and an output time function $y_M = w$:

$$\frac{dW}{dt} + \beta W = u$$

(2.22)
and the error to be \( e = y - w \). Thus \( \partial e / \partial \beta = - \partial w / \partial \beta \). As we are interested in the behaviour for \( -\infty < t < \infty \) initial conditions do not concern us. Formally \( w = w(t, \beta) \) and

\[
w(t, \beta) = w(t, \beta_0) + \frac{\partial w}{\partial \beta} \Delta \beta + \frac{1}{2!} \frac{\partial^2 w}{\partial \beta^2} (\Delta \beta)^2 + \ldots \tag{2.23}
\]

Only the first order extrapolation will be used. The corresponding term can be obtained from the differential equation (2.22) by differentiation with respect to \( \beta \):

\[
\frac{\partial^2 w}{\partial \beta^2} + \beta \frac{\partial w}{\partial \beta} + w = 0
\]

Interchanging the order of differentiation and substituting \( v = \partial w / \partial \beta \) one finds:

\[
\frac{\partial v}{\partial t} + \beta v = - w
\]

This is called the sensitivity equation with respect to \( \beta \). For varying \( \beta \) the quantity \( v \) is an approximation to the gradient \( \partial w / \partial \beta \) which is a function of time. Equations (2.22) and (2.25) can be instrumented according to fig. 2.9 where the time function \( v \) is generated. Equations such as (2.17) and (2.20) or (2.21) can be instrumented to give a system which converges towards \( b \) under a few restrictions. This is discussed in detail in chapter 9.

b) The use of two models with parameters \( \beta \) and \( (\beta + \Delta \beta) \). The first derivative with respect to \( \beta \) can be approximated by using two models with parameters \( \beta \) and \( (\beta + \Delta \beta) \) respectively. The difference of the outputs \( w_1 \) and \( w_2 \) of these models then does give a quantity \( \Delta w / \Delta \beta = - \partial e / \partial \beta \) that can be used as an approximation of \( \partial e / \partial \beta \) in an equation such as (2.17) and (2.20) or (2.21); cf. fig. 2.10. This technique is discussed in chapter 9 as well.

c) The use of one model with a parameter perturbation. A third method is based on:

\[
\frac{dE}{dt} = \frac{\partial E}{\partial \beta} \frac{\partial \beta}{dt} + \frac{\partial E}{\partial t}
\]

(2.26)
From this expression it follows that if the model parameter is disturbed according to a known time function (test signal) and the resulting disturbance $\frac{dE}{dt}$ can be measured, then information on $\frac{\partial E}{\partial \beta}$ can be derived; cf. fig. 2.11. This technique is studied in chapter 9.

d) The use of a generalized model. It will be clear that the methods mentioned up to now for obtaining $\frac{\partial E}{\partial \beta}$, $\frac{\partial \omega}{\partial \beta}$ or an approximation of these quantities are putting quite heavy demands on the amount of implementation or computation. This is even more so if there are a number of parameters that have to be determined. Generally speaking, for each parameter to be estimated one needs another additional model or another model parameter perturbed by a suitably chosen test signal. For this reason it is worthwhile looking for an arrangement where the error is $\text{linear-in-the-parameters}$ to be adjusted, for example:

$$e = \sum_{i=0}^{n} \alpha_i \gamma_i + \sum_{j=0}^{m} \beta_j \nu_j$$

(2.27)

In that case the following holds:

$$\frac{\partial e}{\partial \alpha_i} = \gamma_i \quad \frac{\partial e}{\partial \beta_j} = \nu_j$$

(2.28)

i.e., the parameters sensitivity functions are already explicitly present in the model. Figure 2.12 gives an illustration of such a case. The input signal $u(t)$ is stationary and in most cases considered to be stochastic. The blocks $F_0$, $F_m$ and $G_0$, $G_n$ are operators (e.g., linear transfer functions, nonlinear operators). These blocks together with the "potentiometers" $\alpha_0$, $\alpha_n$ and $\beta_0$, $\beta_m$ and the adder form a "generalized model" of the process. The "potentiometers" can be adjusted to positive and negative coefficient values. The output $e(t)$ will be called a "generalized error".
The reader will appreciate the convenience of a model that leads to an error which is linear in the parameters to be estimated. Fig. 2.13 a, b, and c give examples of such situations:

\[ a) \quad e(t) = y(t) - \sum \beta_j u(t - \tau_j) \]
\[ b) \quad e(t) = y(t) - \sum \beta_j u(t) \]
\[ c) \quad e(t) = y(t) - \sum \alpha_i y(t) - \sum \beta_j u(t) \]

Note that a) is a special case of b) and b) is a special case of c).

This does not necessarily imply that the input-output relationship of the model has to be linear. In connection with figure 2.12, it has already been pointed out that the blocks \( F_i \) and \( G_j \) may be nonlinear.

Figure 2.13d gives an example where the error is nonlinear in the parameters to be estimated:

\[ d) \quad e(t) = y(t) - C[z(t), \alpha, \beta] \]

where \( C \) is the dynamic operator provided by the model and \( \alpha \) and \( \beta \) are the parameter vectors with components \( \alpha_i \) and \( \beta_j \).

2.5 Parameter and state estimation

In quite a number of applications, the notion of state as mentioned in chapter 1 plays an important part; this is particularly true for optimal control problems. As the state variables need not be measurable directly and as the quantities that are measurable will be subject to noise, the process state can be estimated only. This estimation problem is adequately solved for linear processes for which the dynamic properties are known completely (including the numerical values of the process parameters) (Kalman, Bucy, 1961). Many approximative schemes for nonlinear processes have been proposed. These estimation techniques are summarized in chapter 12.

The problem becomes more complicated if the parameter values of the process are not known beforehand but have to be estimated together with the state variables. This type of problem is discussed in chapter 13.
The remaining chapters are devoted to applications in different fields of engineering science.

Conclusion
Now the reader will be in a position to appreciate the organization of this book. He is strongly recommended to read again the preface, including the organization chart of the part on parameter estimation.

In periodical literature, all kind of estimation techniques are presented casewise.

In studying the following parts of the book, the reader will notice many relations between the different estimation techniques. By having read this chapter the reader is well-stimulated to recognize such relations.

For survey papers in this field the reader may consult: Eykhoff (1964); Eykhoff, Van der Grinten, Kwakernaak and Veltman (1966); Cuenod and Sage (1967); Eykhoff (1967); Strobel (1967, 1968); Aleksandrovskii and Deich (1968); Strobel (1968); Bekey (1969); Balakrishnan and Peterka (1969); Åström and Eykhoff (1970).

Much information on identification and estimation can be found in:
- "Identification in automatic control systems".
  Preprints IFAC symposium, Prague, june 1967
  published by Academia, Prague.
- "Identification and process parameter estimation".
  Preprints IFAC symposium, Prague, june 1970
  published by Academia, Prague.

Monographs of general interest are: Lee (1964); Raibman (1970); Sage and Melsa (1971).
References


using implementation of an explicit mathematical relation

(a) mathematical relation

(b) using a model-adjustment technique

relation between these techniques e.g. based on timedelay model (impulse response) representation:

(c) correlation

(d) adjustment of timedelay model

fig 2.1

fig. 2.2
fig. 2.3

fig. 2.4
PARAMETER-ESTIMATION
we want to know

OPTIMIZATION
we want to correct
with respect to setpoints

ADAPTATION
we want to correct
with respect to system-
dynamics

disturbances: signal, system
performance criterion

process dynamics

efficiency, yield, etc

set point

adjustable parameter

increasing intensity of incident light

desired level

iris opening

definition of the image on retina

increasing distance of object

amount of accommodation

illumination on retina

fig 25

fig 26
### I: ORDINARY FEEDBACK CONTROL

- **General scheme**: [Diagram]
- **"error gain" function**: [Graph]
- **"odd" function**: [Graph]
- **Principal property**: 1st derivative does not change sign.
- **A priori knowledge**: Null exists if sign of 1st derivative is known
- **Sufficient information for feedback action**: One observation

### II: ADAPTIVE AND SELF OPTIMIZING FEEDBACK CONTROL

- **Many different schemes**
- **Performance**: [Graph]
- **"even" function controlled variable**: [Diagram]
- **2nd derivative does not change sign.**
- **Extreme exists**: If sign of 2nd derivative is known
- **Sufficient information for feedback action**: Minimum two observations (and data reduction)

---

**Fig 27**

1. **One observation**
   - [Diagram](a)
2. **Two observations**
   - [Diagram](b)
3. **Direction towards optimum**
   - [Diagram](c)
fig. 2.9

fig. 2.10

fig. 2.11
3.1 Classes of signals.

The basis of estimation and identification lies in the mathematical description of the relation between some (set of) "input" and (set of) "output" time functions of the process. In choosing the type of description which will be used for the signals, one has to weigh both aspects of signal representation:
- the display of the information-bearing attributes of the signals and
- the giving of means for studying the transmission properties of the processes.

Emphasis on either one aspect may lead to a different choice.

The signals occurring in estimation schemes may be distinguished as "continuous" (a better term would be nonsampled) and sampled. These refer to the representation of the information as a function of time; cf. fig. 3.1. In numerous situations the sampling of signals may be useful and/or necessary. This is the case when the information processing is of an intermittent type (e.g., by digital computer) or when time-sharing is used for an information-transportation channel.

Analytically speaking by sampling one generates from a time function $x(t)$ a vector of sample values:

$$X' = \{ x(1), x(2), \ldots, x(k) \}$$

From this definition it follows that the sampling operation is a linear one, for if the sampling operation is indicated by $\mathcal{J}[\quad]$ then

$$\mathcal{J} [\alpha x(t)] = \alpha \mathcal{J} [x(t)] \quad \text{for each } \alpha$$

$$\mathcal{J} [\alpha x(t) + \beta y(t)] = \alpha \mathcal{J} [x(t)] + \beta \mathcal{J} [y(t)] \quad \text{for each } \alpha \text{ and } \beta$$

provided that the sampling of both signals occurs simultaneously.

With respect to amplitude, a distinction can be made between analog, quantized and binary. This is also illustrated by fig. 3.1. The binary signal is the
limiting case of a quantized signal. The quantizing operation is nonlinear (Watts, 1962); if the quantizing operator is indicated by \( \mathcal{Q} \) then the expressions

\[
\mathcal{Q}[\alpha x(t) + \beta y(t)] \quad \text{and} \quad \alpha \mathcal{Q}[x(t)] + \beta \mathcal{Q}[y(t)]
\]

for each \( \alpha \) and \( \beta \) are not necessarily equal. This nonlinearity is responsible for many mathematical problems when dealing with quantized signals.

### 3.2 Orthogonal functions

Functions of a rather general nature can be described by the development in a set of orthogonal functions \( \{ g_i(t) \} \). In this way, such an "arbitrary" function can (approximately) be described by a finite number of coefficients. Another advantage of orthogonal expansions will appear later on in this book with respect to convergence properties of adjustment procedures.

A discrete set of orthogonal functions \( g_1(t), g_2(t), \ldots \) over the interval \((a, b)\) satisfies the conditions:

\[
\int_a^b g_i(t) g_k(t) \, dt = \delta_{ik} \quad \left\{ \begin{array}{l}
c = 0 \quad \text{if} \quad i \neq k \\
\neq 0 \quad \text{if} \quad i = k
\end{array} \right.
\]

If \( d_{ik} = \delta_{ik} \) (Kronecker delta; \( \delta_{ik} = 1 \) for \( i = k \)) then the set is called orthonormal. Such a set is called complete if no non-zero function \( y(t) \) can be found such that:

\[
\int_a^b y(t) g_i(t) \, dt = 0 \quad \text{for all} \quad i
\]

If a set is complete then a time function \( x(t) \) can be expressed in these orthonormal functions:

\[
x(t) = \sum_{i=1}^{\infty} c_i g_i(t)
\]

where

\[
c_i = \int_a^b x(t) g_i(t) \, dt
\]

The Fourier series is an example of the use of orthogonal functions.
A continuous set of orthonormal functions is given by:

\[
\int_{a}^{b} g_{u}(t) q_{v}(t) \, dt = \delta \left( u - v \right) \quad \left\{ \begin{array}{l}
\delta (t) = 0 \quad \text{for} \quad t \neq 0 \\
\int_{-\infty}^{\infty} \delta (t) \, dt = 1 \end{array} \right.
\]  

(3.5)

A time-function \( x(t) \) can be expressed as follows:

\[
x(t) = \int_{-\infty}^{\infty} c_{v} q_{v}(t) \, dv
\]

(3.6)

where

\[
c_{v} = \int_{a}^{b} x(t) q_{v}(t) \, dt
\]

(3.7)

One member of this family is the Fourier transform.

For discrete sets of orthonormal functions, the following properties are of interest:

- The approximation of the signal by the sum of a finite number of orthonormal functions in the minimum mean-square error sense leads to the same coefficients as given by equation (3.4). Take a set of \( m \) orthonormal functions:

\[
E = \int_{a}^{b} \left\{ x(t) - \sum_{i=1}^{m} y_{i} q_{i}(t) \right\}^{2} dt
\]

(3.8)

Minimization of \( E \) with respect to \( y_{k} \) requires as a necessary condition:

\[
o = \frac{dE}{dy_{k}} = -2 \int_{a}^{b} \left\{ x(t) - \sum_{i=1}^{m} y_{i} q_{i}(t) \right\} q_{k}(t) dt
\]

Consequently

\[
\int_{a}^{b} x(t) q_{k}(t) dt = \sum_{i=1}^{m} y_{i} \int_{a}^{b} q_{i}(t) q_{k}(t) dt = y_{k}
\]

(3.9)

because of the orthonormality. A comparison of equation (3.9) with equation (3.4) shows that \( y_{i} = c_{i} \). This implies that \( \sum_{i} c_{i} g_{i}(t) \) gives a smaller mean square error than any other linear combination of these orthonormal functions.

- Note that the value of \( y_{i} \) does not depend on \( m \), the number of orthonormal functions in the set used. Consequently, an improvement by using \( n \) \((n > m)\) functions does not alter the coefficients \( c_{i}, \ldots, c_{m'} \). This is a desirable characteristic indeed.
- By using $\gamma_i = c_i$ in equation (3.8) it follows that:
  $$
  \min E = \int_a^b x(t)^2 \, dt - \sum_{i=1}^m c_i^2 \geq 0
  $$
or
  $$
  \sum_{i=1}^m c_i^2 \leq \int_a^b x(t)^2 \, dt
  \tag{3.10}
  $$

For $m \to \infty$ this is the Bessel inequality. Thus $c_i \to 0$ for $i \to \infty$ provided

- If the minimum error approaches zero for $m \to \infty$ then:
  $$
  \sum_{i=1}^m c_i^2 = \int_a^b x(t)^2 \, dt
  $$

which is called the Parseval equality.

Starting from a set of functions \{ $f_i(t)$ \} defined over the interval $(a,b)$, one can derive a set of orthonormal functions \{ $g_i(t)$ \} by applying the well known Schmidt procedure. The first function is chosen as:

$$
q_1(t) = \frac{f_1(t)}{\| f_1(t) \|}
$$

with

$$
\| f_1(t) \| = \sqrt{\int_a^b f_1(t)^2 \, dt}
$$

As a result of this choice $\| g_1(t) \| = 1$, i.e. the function is "normed".

If $f_2(t)$ is independent of $f_1(t)$, then one can choose a coefficient $\alpha$ in such a way that $f_2(t) - \alpha g_1(t)$ is orthogonal to $g_1(t)$:

$$
\int_a^b \left\{ f_2(t) - \alpha q_1(t) \right\} q_1(t) \, dt = 0
$$

from which follows

$$
\alpha = \int_a^b f_2(t) q_1(t) \, dt
$$

and

$$
q_2(t) = \frac{f_2(t) - \alpha q_1(t)}{\| f_2(t) - \alpha q_1(t) \|}
$$

Similarly $g_2(t)$ can be generated so that this signal is orthogonal to $g_1(t)$ ...... $g_{i-1}$ and also $\| g_i(t) \| = 1$.

The notions of orthogonality and orthonormality can be extended by introducing a weighting function $w(t)$; this offers the possibility of
emphasizing the contributions to the mean square error in a predetermined way. The equivalent expression for eq. (3.1) and (3.8) are:

\[ \int_a^b q_i(t) q_k(t) w(t) \, dt = \delta_{ik} \]  

(3.14)

\[ E = \int_a^b \left\{ x(t) - \sum_{i=1}^{M} c_i q_i(t) \right\}^2 w(t) \, dt \]

Depending on the choice of the interval \((a, b)\) and the weighting function \(w(t)\), one obtains the sets of orthogonal functions given in table 3.1. For further aspects of orthogonal functions the reader is referred to Courant-Hilbert (1968), Sneddon (1961), Lee (1963). Examples of practical applications are found in the literature e.g. Lampard (1955), Kitamori (1960), Barker and Hawley (1966), Roberts (1966 and 1967) and others, where variance functions are identified as coefficients in orthogonal series expansions. Recent examples are Schwartz (1969), Gorecki and Turowicz (1969), Dorati and Milanese (1970).
<table>
<thead>
<tr>
<th>discrete sets</th>
<th>interval ((a,b))</th>
<th>weighting function (w(t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourier</td>
<td>(-\pi \leq t \leq \pi)</td>
<td>1</td>
</tr>
<tr>
<td>Legendre polynomials</td>
<td>(-1 \leq t \leq 1)</td>
<td>1</td>
</tr>
<tr>
<td>Tschebychef polynomials</td>
<td>(-1 \leq t \leq 1)</td>
<td>(\frac{1}{\sqrt{1-t^2}})</td>
</tr>
<tr>
<td>Laguerre polynomials</td>
<td>(0 \leq t &lt; \infty)</td>
<td>(e^{-t})</td>
</tr>
<tr>
<td>Laguerre functions</td>
<td>(0 \leq t &lt; \infty)</td>
<td>(e^{-t^2})</td>
</tr>
<tr>
<td>Hermite polynomials</td>
<td>(-\infty &lt; t &lt; \infty)</td>
<td>(e^{-t^2})</td>
</tr>
<tr>
<td>Hermite functions</td>
<td>(-\infty &lt; t &lt; \infty)</td>
<td>(e^{-t^2})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>continuous sets</th>
<th>interval ((a,b))</th>
<th>weighting function (w(t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourier</td>
<td>(-\infty &lt; t &lt; \infty)</td>
<td>(e^{j\omega t})</td>
</tr>
<tr>
<td>Dirac</td>
<td>(-\infty &lt; t &lt; \infty)</td>
<td>(\delta(t-\nu))</td>
</tr>
</tbody>
</table>

Table 3.1: Sets of orthonormal functions.
3.3 Description of deterministic signals

The description of signals in the frequency-domain by Fourier-, Laplace- or z-transform is assumed to be well known. The definitions given in table 3.2 will be used. It is of importance to note that due to the Parseval theorem, the orthogonality in the time-domain also implies orthogonality in the frequency-domain.

One way of describing signals is by means of orthogonal components as indicated in the previous section. The coefficients of these components can be found e.g. from equation (3.4)

\[ c_i = \int_a^b x(t) q_i(t) \, dt \quad (3.15) \]

This can be instrumented in several ways:
- using a multiplication of the signal \( x(t) \) with a member of the set of orthonormal functions and integrating the multiplier output over the interval \( a \leq t \leq b \); fig. 3.2. The orthonormal function can be programmed in function generators or they can be obtained as the solutions of differential equations, programmed on a computer.
- using a set of stable filters. Figure 3.3 shows an example of filters that have orthogonal impulse responses, i.e. if the input of the filter is a \( \delta(t) \) function, then the outputs are the respective time functions of an orthogonal set. However, as a filter leads to a convolution of input signal and the filter impulse response,

\[ y_i(t) = \int_0^\infty x(t-\theta) q_i(\theta) \, d\theta \quad (3.16) \]

a straightforward instrumentation does not lead to the desired result.

At \( t = 0 \) this equation reduces to
\[ y_i(0) = \int_{-\infty}^{\infty} x(-\theta) q_i(\theta) \, d\theta \] 

Comparing this with equation (3.4) one observes that

\[ y_i(0) = c_i \quad \text{if} \quad x(-\theta) = x(\theta) \]

Consequently the signal to be developed has to be recorded and played in the reverse direction; fig. 3.3 b. For parameter estimation this is, of course, an unwanted feature.

- using a set of unstable filters. Equation (3.15) can also be written as:

\[ y_i(T) = \int_{0}^{T} x(\theta) q_i(T-\theta) \, d\theta \]

if \( x(t) = 0 \) for \( t < 0 \) and if one chooses for convolution the time interval \( (0, T) \); fig. 3.4. Comparing this again with equation (3.4) shows that

\[ y_i(T) = c_i \quad \text{if} \quad q_i(T-\theta) = q_i(t) \]

The filters to be used are unstable ones. This can be realized by analog computer programming, using suitable clamping techniques.

Another way of describing signals in the time-domain is by decomposing them in terms of the form \( a \exp(\alpha t) \) and \( b \exp(\beta t) \sin(\omega t + \phi) \), etc. These signals are not orthogonal. They are called "generalized frequencies" and derive their importance from the fact that the solution of linear homogeneous differential equations with constant coefficients consists of (only) such terms.
<table>
<thead>
<tr>
<th>Fourier Transform</th>
<th>Laplace Transform (Two-sided)</th>
<th>Laplace Transform (One-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F(j\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} , dt )</td>
<td>( F(s) = \int_{-\infty}^{\infty} f(t) e^{-st} , dt )</td>
<td>( F(s) = \int_{0}^{\infty} f(t) e^{-st} , dt )</td>
</tr>
<tr>
<td>( F(j\omega) ) is defined for ( \omega ) real and exists if</td>
<td>( s = \sigma + j\omega )</td>
<td>( s = \sigma + j\omega )</td>
</tr>
<tr>
<td>( \int_{-\infty}^{\infty}</td>
<td>f(t)</td>
<td>, dt ) exists</td>
</tr>
<tr>
<td>( \int_{-\infty}^{\infty}</td>
<td>f(t)e^{-st}</td>
<td>, dt ) exists</td>
</tr>
<tr>
<td>This integral exists in some allowable region of convergence ( \sigma_1 &lt; \sigma &lt; \sigma_2 )</td>
<td></td>
<td>This integral exists in some allowable region of convergence ( \sigma &gt; \sigma_1 )</td>
</tr>
</tbody>
</table>

\[ f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(j\omega)e^{j\omega t} \, d\omega \]

which may be evaluated directly or from tables

or

\[ f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(j\omega)e^{j\omega t} \, d(j\omega) \]

\[ = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(s)e^{st} \, ds \mid s = j\omega \]

For \( t < 0 \)

\[ f(t) = \frac{1}{2\pi} \oint_{C_1} F(s)e^{st} \, ds \]

\[ = - \text{(Sum of residues of poles with real parts to right of } \sigma_0 \) \]

For \( t > 0 \)

\[ f(t) = \frac{1}{2\pi} \oint_{C_2} F(s)e^{st} \, ds \]

\[ = \text{Sum of residues of poles with real parts to left of } \sigma_0 \]

Note: As a consequence of the definition of the one-sided transform an acceptable \( \sigma_0 \) is automatically to the right of the poles of \( F(s) \).

Note: Residues are defined for counterclockwise contours. Contour \( C_1 \) is clockwise. The integral around \( C_1 \) is, therefore, minus the sum of the residues of the enclosed poles.

\[ \text{Region of Convergence} \]

\[ \text{Region of Convergence} \]

Table 3.2 Definition of different transforms with their inverses.
### Z-Transform (Two-sided)

\[ f^*(t) = \sum_{n=-\infty}^{\infty} f(n) \delta(t - nT) \]

\[ F(z) = \sum_{n=-\infty}^{\infty} f(n) z^{-n} \]

or

\[ F(z) = F^*(z) \bigg|_{z = e^{i\pi}} \]

This exists in an allowable annulus of convergence outside a circle centered at the origin.

\[ f(n) = \frac{1}{2\pi i} \oint_C F(z) z^{n-1} \, dz \]

where \( C \) is a circle of radius \( r_1 \) such that \( r_1 < r_0 < r_2 \).

For \( n \geq 0 \)

\[ f(n) = \text{sum of residues of } F(z) z^{n-1} \text{ inside } C. \]

For \( n < 0 \)

\[ f(n) = \text{sum of residues of } F(z) z^{n-1} \text{ outside } C. \]

Note: the formula for \( f(n) \) is the formula for the coefficients of the Laurent series for \( F(z) \) (see Appendix).

### Z-Transform (One-sided)

\[ f^*(t) = \sum_{n=0}^{\infty} f(n) \delta(t - nT) \]

\[ F(z) = \sum_{n=0}^{\infty} f(n) z^{-n} \]

or

\[ F(z) = F^*(z) \bigg|_{z = e^{i\pi}} \]

This exists in an allowable region of convergence outside a circle centered at the origin.

\[ f(n) = \frac{1}{2\pi i} \oint_C F(z) z^{n-1} \, dz \]

where \( C \) encloses all the poles of \( F(z) z^{n-1} \).

\[ f(n) = \text{sum of the residues of } F(z) z^{n-1} \]

Note: the formula for \( f(n) \) is the formula for the coefficients of the Laurent series for \( F(z) \) (see Appendix).

---

**Appendix. Complex-Variable Formulas**

1. \( \oint_C f(z) \, dz = 0 \) if \( f(z) \) is analytic on and inside the closed contour \( C \).

2. \( \oint_C g(z) \, dz = 2\pi i \sum \text{[residues of poles of } g(z) \text{ inside } C] \).

If \( g(z) \) has an \( n \)-th order pole at \( z_0 \) then

\[ \text{residue of } z_0 = \frac{1}{(n-1)!} \left[ \frac{d^{n-1}}{dz^{n-1}} \left( (z-z_0)^n g(z) \right) \right]_{z=z_0} \]

3. If \( f(z) \) is analytic on \( C_1 \) and \( C_2 \) and in the region between them, then \( f(z) \) can be represented in that region by a Laurent Series.

\[ f(z) = \sum_{n=-\infty}^{\infty} A_n (z - z_0)^n \]

where \( A_n = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-z_0)^{n+1}} \, dz \).

\( C \) is any curve between \( C_1 \) and \( C_2 \) and enclosing \( z_0 \).
3.4 Description of stochastic signals.

For stochastic time functions, fig. 3.5 shows a few relations together with the symbols used for the quantities under consideration. These notions are summarized in appendix B.

The most important ones are:

the correlation function, general definition:

$$\psi_{xy}(t_1, t_2) = \int_{-\infty}^{\infty} x(t_1) y(t_2) p(x, t_1; y, t_2) \, dx \, dy$$

for stationary signals:

$$\psi_{xy}(\tau) = \int_{-\infty}^{\infty} x(t) y(t-\tau) p(x, y; \tau) \, dx \, dy$$

for ergodic signals this correlation function can also be found as a time average:

$$\psi_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T}^{T} x(t) y(t-\tau) \, dt$$

for finite-interval observation:

$$\psi_{xy}(\tau; T) = \frac{1}{T} \int_{-T}^{T} x(t) y(t-\tau) \, dt$$

which is, of course, a random variable.

the power density functions

for stationary signals:

$$\bar{P}_{xy}(s) = L_2 \left[ \psi_{xy}(\tau) \right]$$

or

$$\bar{P}_{xy}(j\omega) = F \left[ \psi_{xy}(\tau) \right]$$

Note the use of the j-operator in the argument of the F transformed variable

For studying random signals in systems the assumption of a Gaussian (normal) distribution of the random processes is used in the majority of cases because:

- the origin of the random signal, via the central-limit theorem, warrants this assumption;
- passing a Gaussian random signal through a linear element results in a random signal with a Gaussian distribution;
- other probability-density distributions can be thought of as being derived from a Gaussian random signal by passing it through a nonlinear element;
- a one dimensional Gaussian distribution is completely determined if only the statistical mean \( \mu \) and the variance \( \sigma^2 \) are given:
\[
\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}
\]
For a stationary Gaussian random process the second order (joint) probability density function for \( x(t) \) and \( y(t) = x(t - \tau) \) is also Gaussian and can be expressed in terms of \( \psi_{xx}(0) \) and \( \psi_{xx}(\tau) \).
Similar statements hold for higher order probability density functions.

For stationary random signals, the effect upon the signal means and variances because of dynamic operations (transfer functions), adders and multipliers are given in table 3.3. All these relations are assumed to be well-known, except perhaps that of the variance of the multiplier output; this relationship is derived in Appendix E.
<table>
<thead>
<tr>
<th>Time Function</th>
<th>( \mathcal{F} )-Transform</th>
<th>( \mathcal{L} )-Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x(t) )</td>
<td>( X(j\omega) )</td>
<td>( X(s) )</td>
</tr>
<tr>
<td>( y(t) )</td>
<td>( Y(j\omega) )</td>
<td>( Y(s) )</td>
</tr>
<tr>
<td>( g(t) )</td>
<td>( G(j\omega) )</td>
<td>( G(s) )</td>
</tr>
</tbody>
</table>

**Dynamic Operator**

\[
y(t) = g(t) \ast x(t) = \int g(t-\tau) x(\tau) d\tau = \int [g(t) x(t-\tau)] d\tau
\]

\[
Y(j\omega) = G(j\omega) X(j\omega)
\]

\[
Y(s) = G(s) X(s)
\]

**Nonlinear Operator**

\[
y(t) = \varphi[x(t)]
\]

**Adder**

\[
z(t) = x(t) + y(t)
\]

\[
z(j\omega) = X(j\omega) + Y(j\omega)
\]

\[
z(s) = X(s) + Y(s)
\]

**Multiplier**

\[
z(t) = x(t) y(t)
\]

\[
z(j\omega) = X(j\omega) Y(j\omega)
\]

\[
z(s) = X(s) Y(s)
\]

---

**Correlation**

\[
\psi(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T}^{T} x(t) y(t+\tau) dt
\]

(time average; if ergodic)

---

**Table 3.3**
<table>
<thead>
<tr>
<th>Correlation function</th>
<th>Power spectral density function</th>
<th>Probability density function</th>
<th>Characteristic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_{xx}(\tau) )</td>
<td>( \tilde{P}_{xx}(s) )</td>
<td>( P_x(x) )</td>
<td>( \Phi_x(u) )</td>
</tr>
<tr>
<td>( \psi_{xy}(\tau) )</td>
<td>( \tilde{P}_{xy}(s) )</td>
<td>( P_y(y) )</td>
<td>( \Phi_y(v) )</td>
</tr>
<tr>
<td>( g(\tau) )</td>
<td>( G(s) )</td>
<td>( \rho_{xy}(x,y) )</td>
<td>( \rho_{xy}(u,v) )</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\psi_{x}(r) &= g(r) \ast \psi_{x}(t) \\
\psi_{y}(r) &= g(r) \ast \psi_{y}(t)
\end{align*}
\]

\[
\begin{align*}
\psi_{yx}(\tau) &= \psi_{xy}(\tau) + \psi_{yy}(\tau) \\
\tilde{P}_{yx}(s) &= \tilde{G}(s) \tilde{P}_{xy}(s) \\
\tilde{P}_{yy}(s) &= \tilde{G}(s) \tilde{P}_{xy}(s) \\
\tilde{P}_{yy}(s) &= \tilde{G}(s) \tilde{P}_{xy}(s)
\end{align*}
\]

Gaussian:
\[
\begin{align*}
\psi_{xy}(\tau) &= \psi_{xx}(\tau) \xi_{xy}(0) \\
\psi_{xx}(\tau) &= \psi_{xx}(\tau) + \xi_{xy}(s) + \frac{1}{3} \psi_{xx}(s) + \ldots
\end{align*}
\]

\[
\begin{align*}
\psi_{zz}(\tau) &= \psi_{xx}(\tau) + \psi_{yy}(\tau) \\
\tilde{P}_{zz}(s) &= \tilde{P}_{xx}(s) + \tilde{P}_{yy}(s) \\
\tilde{P}_{yy}(s) &= \tilde{G}(s) \tilde{P}_{xy}(s)
\end{align*}
\]

Gaussian; \( \psi_{xy}(0) = 0 \):
\[
\begin{align*}
\psi_{xx}(\tau) &= \psi_{xx}(\tau) \\
\tilde{P}_{xx}(s) &= \tilde{G}(s) \tilde{P}_{xy}(s) + \tilde{P}_{yy}(s)
\end{align*}
\]

\[
\begin{align*}
\psi_{xy}(\tau) &= \frac{1}{i2\pi} \int_{-\infty}^{\infty} e^{i\omega \tau} \xi_{xy}(\omega) d\omega \\
\psi_{xy}(\tau) &= \frac{1}{i2\pi} \int_{-\infty}^{\infty} \tilde{P}_{xy}(\omega) d\omega \\
\psi_{xy}(\tau) &= \frac{1}{i2\pi} \int_{-\infty}^{\infty} e^{i\omega \tau} \xi_{xy}(\omega) d\omega \\
\psi_{xy}(\tau) &= \frac{1}{i2\pi} \int_{-\infty}^{\infty} \tilde{P}_{xy}(\omega) d\omega
\end{align*}
\]

\[
\begin{align*}
\Phi_{uv} &= \mathbb{E}[e^{i(uX + vy)}] \\
\Phi_{uv} &= \mathbb{E}[e^{i(uX + vy)}] \\
\Phi_{uv} &= \mathbb{E}[e^{i(uX + vy)}] \\
\Phi_{uv} &= \mathbb{E}[e^{i(uX + vy)}]
\end{align*}
\]

\[
\begin{align*}
\beta_{uv} &= \mathbb{E}[x(t) y(t)] \\
\beta_{uv} &= \mathbb{E}[x(t) y(t)] \\
\beta_{uv} &= \mathbb{E}[x(t) y(t)] \\
\beta_{uv} &= \mathbb{E}[x(t) y(t)]
\end{align*}
\]
<table>
<thead>
<tr>
<th>General moments</th>
<th>Central moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{10} )  ( \alpha_{20} )  ( \alpha_{30} )</td>
<td>( \beta_{10} )  ( \beta_{20} )  ( \beta_{30} )</td>
</tr>
<tr>
<td>( \alpha_{11} )  ( \alpha_{21} )  ( \alpha_{31} )</td>
<td>( \beta_{11} )  ( \beta_{21} )  ( \beta_{31} )</td>
</tr>
<tr>
<td>( \alpha_{01} )  ( \alpha_{02} )  ( \alpha_{03} )</td>
<td>( \beta_{01} )  ( \beta_{02} )  ( \beta_{03} )</td>
</tr>
</tbody>
</table>

\[ \alpha_{10} = \mu_x \]

\[ \alpha_{01} = \mu_y \]

\[ \mu_y = G(0) \mu_x \]

\[ \beta_{10} = \beta_{01} = 0 \]

\[ \beta_{20} = \sigma_x^2 \]

\[ \beta_{02} = \sigma_y^2 \]

\[ \beta_{12} = \sigma_{xy}^2 \]

\[ \sigma_y^2 = \psi_{yy}(0) \]

\[ \mu_z = \mu_x \pm \mu_y \]

\[ \sigma_z^2 = \sigma_x^2 + \sigma_y^2 \pm 2 \sigma_{xy}^2 \]

\[ \mu_z = \psi_{xy}(0) \]

**Gaussian**

\[ \sigma_z^2 = \sigma_x^2 \sigma_y^2 + \sigma_{xy}^4 \]

\[ \beta_{n0} = \sum_{k=0}^{n} (-1)^k \beta_{n-k} \]

\[ \psi^{(\eta)}(u) \mid u=0 \]

\[ \beta_{n0} = \sum_{k=0}^{n} (-1)^k \psi^{(\eta)}(u) \]

\[ \alpha_{10} \alpha_{(n-1)0} \]
References


Additional literature


Fig. 3.1

Fig. 3.2

Fig. 3.3

Fig. 3.4
A stochastic time function is given by $x(t)$. The ensemble element is an ensemble of stochastic processes.

The correlation function $\psi_{xy}(t_1, t_2)$ is used to analyze the relationship between two processes. The expectation $\mu_x(t)$ is calculated for each process.

Stationarity is achieved when the probability density function remains constant over time.

The centered signal is given by $x' = x - \mu_x$ and $y' = y - \mu_y$.

Fig. 3.5
In chapter one, the importance of "modelbuilding" was stressed. In this chapter the field of possible model choices is scanned.

First the classes of process models are distinguished according to the well known aspects linear/varylinear/nonlinear and nonparametric/parametric. These types of description are summarized in the sections 4.2, 4.3 and 4.4.

For identification techniques the concepts of controllability, observability and identifiability, as summarized in section 4.5, are relevant.

In section 4.6 attention is focussed on the concept "linearity-in-the-parameters" which concept is of importance in estimation procedures, and on canonical forms having the least number of parameters.
4.1 **Classes of process models.**

Stimulated by the theory of classical mechanics, an important method of interpreting the behaviour of processes has been by means of differential equations, expressed in terms of input and output quantities $u = u(t)$ and $y = y(t)$, respectively:

$$\frac{d^n y}{dt^n} + \cdots + a_n \frac{dy}{dt} + b_n y = \frac{d^n u}{dt^n} + \cdots + b_n u + c$$

(4.1)

together with the initial conditions $\frac{dy}{dt}$ at $t=0$ for $i = 1, \ldots, n$. For processes that include a pure time delay $u = u(t-\tau)$. One may arbitrarily choose one of the coefficients, e.g., $a_0 = 1$.

In a **linear** case the coefficients $a_i$, $b_j$ do not depend on $u$ and $y$ and their derivatives. If, in addition, they do not depend on time either, we have a case of **constant coefficients**. This is the most tractable case. These coefficients may, however, depend on time; in that case the equation is called **time varying** or **vary-linear**.

If any $a_i$ or $b_j$ does depend on $u$, $y$, or their derivatives, the process is **nonlinear**.

The dominating distinction between **linear** and **nonlinear** processes is the principle of superposition which holds only for the former. This principle implies that if

- $y_1$ is the output due to input $u_1$ and
- $y_2$ is the output due to input $u_2$ then
- $\alpha y_1 + \beta y_2$ is the output due to input $\alpha u_1 + \beta u_2$.

A simple example may serve as an illustration:

<table>
<thead>
<tr>
<th>Linear diff. eq.</th>
<th>Nonlinear diff. eq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y' + ay = u$</td>
<td>$y' + ay^3 = u$</td>
</tr>
</tbody>
</table>

consequently

consequently
\[
\begin{align*}
\dot{y}_1 + ay_1 &= u_1 \\
\dot{y}_2 + ay_2 &= u_2 \\
(\dot{y}_1 + \dot{y}_2) + a(y_1 + y_2) &= u_1 + u_2
\end{align*}
\]

if \( u_1 + u_2 = u_3 \)
\( y_1 + y_2 = y_3 \) and from this
\( \dot{y}_1 + \dot{y}_2 = \dot{y}_3 \)
then
\[
\dot{y}_3 + ay_3 = u_2
\]

\[
\begin{align*}
\dot{y}_1 + a\dot{y}_1^3 &= u_1 \\
\dot{y}_2 + a\dot{y}_2^3 &= u_2 \\
(\dot{y}_1 + \dot{y}_2) + a(\dot{y}_1^3 + \dot{y}_2^3) &= u_1 + u_2
\end{align*}
\]

if \( u_1 + u_2 = u_3 \)
\( y_1 + y_2 = y_3 \) and from this
\( \dot{y}_1 + \dot{y}_2 = \dot{y}_3 \)
then
\[
(y_1^3 + y_2^3) \neq (\dot{y}_1 + \dot{y}_2)^3 = \dot{y}_3^3
\]

and
\[
\dot{y}_3 + a\dot{y}_3^3 \neq u_2
\]

In the linear case, the signal sums obey the same differential equation as the original signals; in the nonlinear case this does not hold.

The principle of superposition holds as well for the output signal due to initial conditions (general solution) as for the output signal due to an input (particular solution).

Note that in the linear case the superposition (linearity) is preserved even if the coefficient \( a \) of the differential equation is a function of time. In the engineering sense, however, such time-varying coefficients (parameters) make an important difference.

For that reason this case will be discussed separately under the heading vary-linear processes, sect. 4.3.

The difficulties of solving nonlinear differential equations are well known. So is the incompleteness of the collection of methods for solving them. The "classical" approaches for studying such systems, viz. by phase plane analysis and by describing function approximation, are of little use for our problem of process representation. As one of the possible approaches for a class of nonlinear systems, the Volterra series is discussed in section 4.4.
For processes, important types of relations are: (1) those between input and output signals and (2) those expressed in state variables. Let us consider the first category through the expression:

\[ F \{ u, \dot{u}, \ddot{u}, \ldots, u^{(n)}, y, \dot{y}, \ddot{y}, \ldots, y^{(n)} \} = 0 \]  \hspace{1cm} (4.2)

where \( F \) is an analytic function in all variables \( u, \dot{u}, \ldots, y, \dot{y}, \ldots \). As both the input and the output signals are found only implicitly in the differential equation the analysis will in general be devoted to obtaining an explicit relation between \( u(t) \) and \( y(t) \). Such a relation will be denoted by

\[ y(t) = \mathcal{O}[u(t)]_t \]  \hspace{1cm} (4.3)

In mathematical terms the operator \( \mathcal{O} \) can be a functional, i.e. a rule which makes a value correspond to each function over a given interval. As indicated by (4.3), the interval of interest for \( u(\tau) \) may be \( -\infty < \tau \leq t \). In the following we will consider the operator \( \mathcal{O} \) for a number of simple cases. (Van Harten, 1960).

1) \[ F = y - b u = 0 \]
In this trivial case there is a linear relation between \( u(t) \) and \( y(t) \); the process does not include dynamical elements that accumulate energy. For this reason such a system is known as linear-without-memory. In this case the functional \( \mathcal{O}[u(t)] \) reduces to a linear function of \( u \):

\[ y(t) = \mathcal{O}[u(t)]_t = b u(t) \]

2) \[ F = \ddot{y} + a_1 \dot{y} + a_2 y - \ddot{u} - b_1 u = 0 \]
\[ a_1, a_2, b_1 \text{ constant} \]
From the theory of linear differential equations it follows that \( y(t) \).
can be written explicitly by the functional
\[
y(t) = \mathcal{O}
\begin{bmatrix} u(t) \end{bmatrix}^T = \int_{-\infty}^{t} h(\tau) u(t-\tau) \, d\tau
\]
The process includes dynamical elements that provide some kind of memory; consequently, there is a dependence on previous values of \( u(t) \) (linear-with-memory). One of the ways of determining the weighting function or impulse response \( h(t) \) from the differential equation is by use of the Laplace transform theory;
\[
h(t) = \mathcal{L}^{-1}\left[\frac{s + b_2}{s^2 + a_3 + a_2}\right]
\]
Another approach for obtaining \( h(t) \) is indicated in section 4.2.

3) \( F = y + ay^3 - u = 0 \) \( a \) is constant
This represents a nonlinear relation between process input and output (e.g., a saturating amplifier; we restrict the discussion to a monotonic function). There are no dynamical elements that provide a memory. (non-linear-without-memory). Consequently the functional again reduces to a function:
\[
y(t) = \mathcal{O}
\begin{bmatrix} u(t) \end{bmatrix}^T = g\{u(t)\}
\]
This function can be obtained by substitution or iteration procedures or by a simple Taylor development. For this example it follows that
\[
y = u - au^3 + 3a^2u^5 - 12a^3u^7 + \ldots
\]
This equation holds only for \( u \) inside the region of convergence. Although the implicit relation is very simple the explicit one is rather awkward. In some cases a graphical determination has advantages (McFee, 1961).
4) \( F = \dddot{y} + a_1 \ddot{y} + a_2 y + f(y) - u = 0 \quad a_1, a_2 \text{ constant} \)

This equation describes the behaviour of the process indicated in fig. 4.1 which is nonlinear and has dynamic elements (nonlinear-with-memory).

The nonlinearity itself is assumed to have no memory; this restriction excludes hysteresis-type of nonlinearities. In section 4.4 it will be shown that for this case an explicit relation can be written in the following form (Volterra series)

\[
y(t) = \mathcal{O}[u(t)] = \\
= \int_{-\infty}^{t} h_1(\tau) u(t-\tau) \, d\tau + \int_{-\infty}^{t} \int_{-\infty}^{t} h_2(\tau_1, \tau_2) u(t-\tau_1) u(t-\tau_2) \, d\tau_1 \, d\tau_2 + \\
+ \int_{-\infty}^{t} \int_{-\infty}^{t} \int_{-\infty}^{t} h_3(\tau_1, \tau_2, \tau_3) u(t-\tau_1) u(t-\tau_2) u(t-\tau_3) \, d\tau_1 \, d\tau_2 \, d\tau_3 + \ldots
\]

These examples may also serve as illustrations of a discussion pertinent to the use of models as given by
McGhee (1963). He coins the terms "function space description" and "parameter space description". The first uses the idea of transformation defined over a function space. This function space provides a representation of the process input signal. The representation may be such that the signal over a fixed time interval is given by one point in space, e.g., a Hilbert space. A representation can also be found for the signal using an interval of time that "slides" along the t-axis; in that case one finds a curve in the function space. Examples of such representations can be found in e.g., the Fourier series expansion, the signal space expansion (using time samples), the Laguerre function expansion.

The transformation that is defined over this space follows from the dynamics of the process. The process output signal may be represented on a similar space or on the one-dimensional amplitude space. In these terms the estimation problem is to find what transformation from input function space to output function space characterizes the process.

As no information about the physical structure of the process or its assumed mathematical equations is used, this approach is of the black box identification type. The parameter space description starts from an assumed mathematical description of the process dynamics. This description is a parametric model of finite dimensionality. The coordinates of the parameter space are the numerical values of the quantities that determine the "output" of the model. If, for example, the assumed description is an ordinary differential equation, then the coordinates may be the values of the coefficients and the values of the initial conditions. If there is no forcing function, then from this one point in the parameter space one can predict the process output. If there is a forcing function then the unknown parameters of the signal increase the dimensionality of the parameter space. The dimensionality remains finite, while in principle an infinite number of parameters had to be determined in the function space description.
Consequently, one may distinguish:

- **non parametric models**, for example:
  - impulse responses;
  - transfer functions if the numbers of coefficients are not given beforehand;
  - covariance functions;
  - spectral densities;
  - Volterra series

- **parametric models**; for example:
  - differential equations of pre-determined form and order;
  - state models.

The parametric models can give results with large errors if the order of the model does not agree with the order of the process. The nonparametric representations have the advantage that it is not necessary to specify the order of the process explicitly. These representations are, however, intrinsically infinite dimensional which means that it is frequently possible to obtain a model such that its output agrees exactly with the process output.

Interesting aspects of parametric versus nonparametric models are found in the literature on time series analysis. See for example Mann and Wald (1943), Whittle (1963), Grenander and Rosenblatt (1957), Jenkins and Watts (1963). Needless to say, the models must of course finally be judged with respect to the ultimate aim.
4.2 Linear models

For systems operating on continuous-time functions, fig. 4.2 gives a number of relations between different types of descriptions. It also shows several testing procedures and the type of information that is obtained from these tests. In the following sections some of these relations are reviewed and expanded.

Deterministic signals; non-sampled.

Input-output description. As a result of the superposition principle, the dynamic behaviour of a linear process can be described by means of characteristic time-functions: impulse-response function and/or step-response function. These are the solutions of the differential equation (4.1) when all initial conditions are zero and when $u$ is defined respectively as:

$$ u(t) = \delta(t) $$

$$ \int_{-\infty}^{t} \delta(t') \, dt' = 0 \quad \text{for } t \neq 0 \quad \text{(unit impulse function)} \tag{4.4} $$

and

$$ u(t) = 0 \quad \text{for } t < 0 $$

$$ u(t) = 1 \quad \text{for } t \geq 0 \quad \text{(unit step function)} \tag{4.5} $$
Physically each time function u can formally be considered to consist of such impulse or step functions. This leads to the convolution integral which, for constant coefficients and zero initial conditions, can be written as:

\[ y(t) = \int_{-\infty}^{\infty} h(\tau) u(t-\tau) \, d\tau = \int_{-\infty}^{t} h(t-\tau) \, u(\tau) \, d\tau \quad (4.6) \]

If \( u(t) = \delta(t) \), then

\[ y(t) = \int_{-\infty}^{\infty} h(\tau) \delta(t-\tau) \, d\tau = h(t) \quad (4.6a) \]

which explains the name impulse response (weighting function) for \( h(t) \).

Instead of solving the differential equation for \( y \) if \( u = \delta(t) \) in order to obtain \( h(t) \) one can also use the following procedure.

Start with equation (4.6) as an assumption. Then differentiate this equation to form all derivatives that are needed in equation (4.1). Substitution of these expressions in (4.1) and the equating of the coefficients gives an expression for \( h(t) \). As an example consider the simple first order differential equation:

\[ \dot{y} + a y = u \quad y(0) = 0 \quad u(t) = 0 \quad \forall t \leq 0 \]

with

\[ y(t) = \int_{0}^{t} h(t-\tau) \, u(\tau) \, d\tau \]

Consequently, using Leibniz's rule:

\[ \text{If} \]

\[ y(t) = \int_{a(t)}^{b(t)} f(\tau, t) \, d\tau \]

where \( a \) and \( b \) are differentiable functions of \( t \) and where \( f(\tau, t) \) and \( \partial f(\tau, t)/\partial t \) are continuous in both \( t \) and \( \tau \), then is

\[ \frac{dy}{dt} = f(\dot{b}(t), t) \frac{\partial f}{\partial t} \dot{a} + f(\dot{a}(t), t) \frac{\partial f}{\partial \tau} + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t} \, d\tau \]

\[ \dot{y}(t) = h(t) \, u(t) + \int_{0}^{t} \frac{d}{dt} h(t-\tau) \, u(\tau) \, d\tau \]

These expressions for \( y \) and \( \dot{y} \) are substituted in the differential equation. As the resulting equation has to hold for all \( t \) and as this must be the case for arbitrary \( u(t) \) it follows that the corresponding parts of the left and
right hand sides have to be equated:

\[ h(0) u(t) = u(t) \quad \Rightarrow \quad h(0) = 1 \quad (4.7) \]

\[ \int_0^t \left\{ \frac{\partial h(t-r)}{\partial t} + a h(t-r) \right\} u(t) \, dr = 0 \]

Again this equation has to be true for all time functions \( u(t) \). It follows that:

\[ \frac{d h(t)}{dt} + a h(t) = 0 \quad \Rightarrow \quad h(t) = c e^{-at} \quad (4.8) \]

Combining (4.7) and (4.8) leads us to

\[ h(t) = e^{-at} \]

In this way we have derived the well known result without using the properties of the symbolic impulse "function" \( \delta(t) \) as the independent variable. The same type of reasoning holds for \( y(0) = b \neq 0 \).

**State space description.** In modern system and control theory the concept of state plays an important role. (Zadeh and Desoer, 1963; Tou, 1964; Elgerd, 1967). In table 4.1 some of the well known relationships between time and frequency domain state-description are summarized and compared to the input-output relationships. Disturbances (noise signals) are not included in these descriptions.

The state and the output equation are schematically represented in fig. 4.3, where possible disturbances are represented by \( y(t) \) and \( n(t) \).

The matrices are named:

- \( A = \) system matrix
- \( B = \) distribution matrix
- \( C = \) output or measurement matrix
- \( D = \) input-output matrix

The direct solution of the state equation in the time domain, neglecting \( y(t) \) and \( n(t) \), goes along the following lines:
### Time Domain

**Input-output relationships (2nd order example)**

\[ \ddot{y} + a_1 \dot{y} + a_2 y = u \]

\[ \begin{align*}
\dot{y}(0) & = y(0) \\
y(0) & = y(0)
\end{align*} \]

**Direct solution**

\[ y(t) = h_0 y(0) + h_1 \dot{y}(0) + \int_0^t h_2(t-\tau) u(\tau) d\tau \]

**General solution**

**Particular solution**

### Frequency Domain

**Transfer function**

\[ Y(s) = \frac{s^2 + a_1 s + a_2}{s^2 + q s + q^2} Y(0) + \frac{1}{s^2 + q s + q^2} \dot{y}(0) + \frac{1}{s^2 + q s + q^2} U(s) \]

### State Relationships (General Example)

**State equation**

\[ \dot{x} = Ax + Bu \]

**Output equation**

\[ y = Cx + Du \]

\[ x(0) = x(0) \]

**Initial condition**

**Direct solution**

\[ x(t) = \phi(t) x(0) + \int_0^t \phi(t,v)Bu(v) dv \]

**General solution**

**Particular solution**

**State transition matrix**

\[ \phi(t) = e^{At} \]

**Principal matrix**

\[ [sI - A]^{-1} \]

### Table 4.1
\[
\dot{\mathbf{x}} - A \mathbf{x} = B u
\]
\[
\frac{d}{dt} \left( e^{-At} \mathbf{x} \right) = e^{-At} \left[ \dot{\mathbf{x}} - A \mathbf{x} \right] = e^{-At} \mathbf{B} u
\]
For the definition of the matrix exponential c.f. Appendix C. Consequently, integration over \((0,t)\) gives:
\[
\mathbf{e}^{-At} \mathbf{x} - \mathbf{x}(0) = \int_0^t \mathbf{e}^{-\Delta t} \mathbf{B} u(\tau) d\tau
\]
or
\[
\mathbf{x} = \mathbf{e}^{At} \mathbf{x}(0) + \int_0^t \mathbf{e}^{A(t-\tau)} \mathbf{B} u(\tau) d\tau
\]
(4.9)

The state concept is closely related to the mode concept. For an illustration of this point consider only the set of homogeneous equations
\[
\dot{\mathbf{x}} = A \mathbf{x}
\]
(4.10)
with the solution
\[
\mathbf{x}(t) = \mathbf{e}^{At} \mathbf{x}(0)
\]
Instead of this formal solution we can look for special initial conditions under which one and only one mode is excited, i.e. where
\[
\mathbf{x}(t) = \mathbf{r} \mathbf{e}^{\lambda t}
\]
This substituted in equation (4.10) gives:
\[
\lambda r \mathbf{e}^{\lambda t} = A \mathbf{r} \mathbf{e}^{\lambda t}
\]
or
\[
\left[ A - \lambda I \right] \mathbf{r} = 0
\]
(4.11)
A non zero solution to this set of algebraic equations will exist if the determinant equals zero, that is if
\[
| A - \lambda I | = 0
\]
(4.12)
The roots of this algebraic equation are called the characteristic or eigen values \( \lambda_1, \ldots, \lambda_n \). Let us assume that all \( \lambda \)'s are distinct, then to each eigen value \( \lambda_i \) there corresponds a characteristic or eigen vector \( r_i \) which is found by solving equation (4.11). A simple example is given in Table 4.2.

If the eigen values are distinct, then it is possible to choose a set of state space coordinates such that the differential equations in these new coordinates become independent. For the theory on multiple eigen values and the generalized eigen vectors associated with these eigen values cf. Friedman (1956).

We still have to show the equivalence of the state-transition matrix \( \Phi(t) \) as given in Table 4.1, and the formal solution \( e^{At} \). This is done for the example in Table 4.2. Making use of the Cayley-Hamilton theorem (Friedman, 1956; p. 121) one can write for this 2nd order process, provided the eigen values are distinct:

\[
\begin{align*}
e^{At} &= \alpha I + \beta A \\
\alpha \text{ and } \beta &\text{ have to be determined. As a result of that theorem, this can be done using the eigen values instead of } A.
\end{align*}
\]

\[
\begin{align*}
\varepsilon^{-t} &= \alpha - \beta \\
\varepsilon^{-2t} &= \alpha - 3\beta
\end{align*}
\]

and it follows:

\[
\begin{align*}
\alpha &= \frac{1}{2} \left\{ \varepsilon^{-t} - \varepsilon^{-2t} \right\} \\
\beta &= \frac{1}{2} \left\{ \varepsilon^{-t} - \varepsilon^{-3t} \right\}
\end{align*}
\]
\[
\Phi(t) = \left\{ 
\begin{bmatrix}
    e^{At}
\end{bmatrix}
\right.
\]

\[
\begin{bmatrix}
    h_0(t) & h_1(t) \\
    h_0'(t) & h_1'(t)
\end{bmatrix}
\]

(second order case)

Example:
\[
A = \begin{bmatrix}
    0 & 1 \\
    -3 & -4
\end{bmatrix}
\]

\[
\Phi(t) = \frac{1}{2} \begin{bmatrix}
    3e^{-t} - e^{-3t} & e^{-t} - e^{-3t} \\
    -3e^{-t} + 3e^{-3t} & -e^{-t} + 3e^{-3t}
\end{bmatrix}
\]

\[
sI - A = \begin{bmatrix}
    5 & -1 \\
    3 & 5 + 4
\end{bmatrix}
\]

\[
\begin{bmatrix}
    sI - A
\end{bmatrix}^{-1} = \begin{bmatrix}
    \frac{s + 4}{s^2 + 4s + 3} & \frac{1}{s^2 + 4s + 3} \\
    \frac{-3}{s^2 + 4s + 3} & \frac{5}{s^2 + 4s + 3}
\end{bmatrix}
\]

eigenvalues and eigenvectors:

\[
|A - \lambda I| = \lambda^2 + 4\lambda + 3 = 0
\]

\[
\lambda_1 = -1 \Rightarrow \begin{bmatrix} 1 & 1 \\ -3 & -3 \end{bmatrix} \xi_1 = 0 \Rightarrow \xi_1 = \alpha_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\]

\[
\lambda_2 = -3 \Rightarrow \begin{bmatrix} 3 & 1 \\ -3 & -3 \end{bmatrix} \xi_2 = 0 \Rightarrow \xi_2 = \alpha_2 \begin{bmatrix} 1 \\ -3 \end{bmatrix}
\]

no input, arbitrary \(x(0)\):

\[
x(t) = \xi_1 \xi_1 e^{-t} + \xi_2 \xi_2 e^{-3t}
\]

if \(x(0) = \xi_1 \xi_1 + \xi_2 \xi_2\)

Table 4.2
Consequently:

\[
\mathcal{E} \left[ \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} \right] t = \frac{1}{2} \begin{bmatrix} 3 e^{-t} - e^{-3t} \\ 0 \end{bmatrix} \left[ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} e^{-t} - e^{-3t} \\ 1 \end{bmatrix} \right] = \frac{1}{2} \begin{bmatrix} 3 e^{-t} - e^{-3t} & e^{-t} - e^{-3t} \\ -3 e^{-t} + 3 e^{-3t} & -e^{-t} + 3 e^{-3t} \end{bmatrix}
\]

which is the same as \( f(t) \) given in Table 4.2.

### Deterministic signals; sampled

For the description of the dynamic behaviour of processes in which there are only sample values available of input and output signals one can use difference equations instead of differential equations. Such situations arise naturally when:

- certain types of measuring instruments are used (e.g. those performing a relatively slow chemical analysis);
- using measuring instruments and/or communication channels in a time sharing fashion;
- using digital data reduction or communication equipment;
- the data are of a discrete nature (as e.g., in economics).

Another application may be found in continuous processes and signals, where a sampled data approach is used for mathematical simplicity. For sampled signals a relation scheme analogously to fig. 4.2 can be given. In this case difference equations, unit-sample response functions and Z-transforms are used. Note that we restrict ourselves to that situation where both process input and output are given in sampled form (cf. Elgerd, 1967, p. 393).

**Input-output description.** Let the sample values of input and output signals be denoted by \( u(k) = u(k\theta) \) and \( y(k) = y(k\theta) \) respectively. Then the difference equation may be of the form:

\[
a_0 y(k) + a_1 y(k+1) + \ldots + a_n y(k+n) = b_0 u(k) + \ldots + b_n u(k+n)
\]  (4.13)
plus initial conditions, where again $a_0 = 1$ may be chosen arbitrarily.

(Note that the difference equation also can be expressed in terms of $a_i y(k-i)$, etc.).

Again the linear, constant-coefficient case will be reviewed. Superposition holds because of the linearity, and the dynamic behaviour of the process can be described by a sampled time function $h(k)$, the unit-sample-response function. The convolution integral goes over into a summation:

$$y(k) = \sum_{i=-\infty}^{k} h(k-i) u(i)$$  \hfill (4.14)

Again $h(k)$ can be derived by a method analogous to that presented for the non-sampled case. As an example the following simple difference equation is taken:

$$y(k+1) + a y(k) = u(k)$$ \hfill $y(0) = 0$

For equation (4.14):

$$y(k+1) = h(0) u(k+1) + h(1) u(k) + h(2) u(k-1) + ...$$

$$a y(k) = a h(0) u(k) + a h(1) u(k-1) + ...$$

$$u(k) = y(k+1) + a y(k) = h(0) u(k+1) + \{h(1) + a h(0)\} u(k) + \{h(2) + a h(1)\} u(k-1) + ...$$

This equation has to hold for all sample sequences $u(i)$. Consequently

$$h(0) = 0$$

$$h(1) = 1$$

$$h(i+1) + a h(i) = 0$$ \quad \text{for } i = 1, 2, 3, ...

The solution to this set of equations is:

$$h(1) = 1$$

$$h(2) = -a h(1) = -a$$

$$h(3) = -a h(2) = a^2$$

$$h(i) = (-a)^{i-1}$$ \quad \text{for } i = 1, 2, 3, ...

Please note that this $h(k)$ is not necessarily identical with $h(t)$ at $t=k$ for the non-sampled case, if this one is being approximated by a sampled version. The same symbol is used to accentuate the similarities between both cases.
For stable processes, only a limited number of points of the unit-sample response function have to be given if \( h(i) \neq 0 \) for \( i \geq L \). Consequently

\[
y(k) \approx \sum_{i=0}^{L} h(i) u(k-i) \quad k \geq L
\]

(4.14a)

The system behaviour is now characterized by the set of values \( \{h(0), \ldots, h(L)\} \).

Introducing the symbol \( z \) as a shift operator, i.e.

\[
z y(k) = y(k+1)
\]

and

\[
Y(z) = \sum_{k=0}^{\infty} y(k) z^{-k}
\]

to the difference equation (4.13) leads to:

\[
\left\{ \begin{array}{c}
a_0 z^k + a_1 z^{k+1} + \ldots + a_n z^{k+n} \\
\end{array} \right\} Y(z) = \left\{ \begin{array}{c}
b_0 z^k + b_1 z^{k+1} + \ldots + b_n z^{k+n} \\
\end{array} \right\} U(z)
\]

(4.15)

where \( z^k \) can be divided out. The first \( n \) values of \( y \) are assumed to be zero.

From eq. (4.15) follows:

\[
H(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z + \ldots + b_n z^n}{a_0 + a_1 z + \ldots + a_n z^n}
\]

(4.16)

which is called the sample transfer function. By long division one obtains:

\[
H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \ldots
\]

(4.17)

where the coefficients \( h_i = h(i) \) as in eq. (4.14).

State space description. Instead of the difference equation (4.13) one may use again a vector difference equation. In table 4.3, some relationships between the time and \( z \)-domain description are summarized and compared to the input-output relationship.

The direct solution of the state equation in the time domain is straightforward:
**Time domain**

**Input-output relationships (2nd order example)**

\[
\begin{align*}
y(k+2) + a_1 y(k+1) + a_0 y(k) &= u(k) \\
y'(k) &= y(k)
\end{align*}
\]

- **General solution**
  \[
y(k) = h_o(k) y(k) + h_1(k) y'(k) + \sum_{i=0}^{k} h_i(k-i) u(i)
\]
- **Particular solution**

**Z-domain**

\[
y(z) = \frac{z^2 + a_2 z + a_0}{z^2 + a_2 z + a_0} y(0) + \frac{z}{z^2 + a_2 z + a_0} y'(0) + \frac{1}{z^2 + a_2 z + a_0} u(z)
\]

**State relationships (general example)**

\[
\begin{align*}
x(k+2) &= A x(k) + B u(k) \\
x'(k) &= C x(k) + D u(k)
\end{align*}
\]

- **Initial condition**
  \[
x(0)
\]

**Direct solution**

\[
x(k) = \Phi(k) x(0) + \sum_{i=0}^{k-1} \Phi(k-1-i) B u(i)
\]

- **General solution**
- **Particular solution**

\[
\Phi(k) = A^k
\]

**Table 4.3**
\[ x(k+1) = A x(k) + B u(k) \]

leads to:

\[
\begin{align*}
    x(1) &= A x(0) + B u(0) \\
    x(2) &= A x(1) + B u(1) = A^2 x(0) + AB u(0) + B u(1) \\
    &\ddots \\
    x(k) &= A^k x(0) + \sum_{i=0}^{k-1} A^{k-1-i} B u(i)
\end{align*}
\]

For a simple example cf. Table 4.4.
\[
\phi(k) = \begin{cases} 
A^k \\
h_0(k) \\ h_1(k) \\
h_0(k) \\ h_1(k)
\end{cases} \text{ (second order case)}
\]

**Example**
\[
A = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}
\]
\[
\phi(k) = \frac{1}{2} \begin{bmatrix} 3(-1)^k(-3)^k & (-1)^k(-3)^k \\ -3(-1)^k+3(-3)^k & -(-1)^k+3(-3)^k \end{bmatrix} = \begin{bmatrix} 0, -3, \ldots \} \{ 0, 1, -4, \ldots \} \\ 0, -3, 12, \ldots \} \{ 1, -4, 13, \ldots \}
\]
unstable

\[
zI - A = \begin{bmatrix} z & -1 \\ 3 & z+4 \end{bmatrix}
\]
\[
z[zI - A]^{-1} = \begin{bmatrix} \frac{z^2+4z}{z^2+4z+3} & \frac{z}{z^2+4z+3} \\ \frac{-3z}{z^2+4z+3} & \frac{z^2}{z^2+4z+3} \end{bmatrix}
\]

Just as in the continuous case, a development using eigenvalues and eigenvectors can be given, together with the corresponding modes.

**Table 4.4**
Stochastic signals.

In principle, the description given in the previous sections also holds for stochastic signals. With such signals there is little sense in asking for the instantaneous value of the amplitude. Consequently, the most useful questions are those with respect to the statistical properties and parameters, e.g., spectral density, expected value, variance, correlation function, etc. (cf. Appendix B and section 3.4).

The input signal of a linear system is \( u(t) \), its output signal \( y(t) \).

These signals are related through the convolution integral:

\[
y(t) = \int_{-\infty}^{\infty} h(\theta) u(t-\theta) \, d\theta
\]

with

\[
h(t) = 0 \quad \text{for} \quad t < 0
\]

We find the first moment by averaging eq. (4.18) over the ensemble (taking the expectation):

\[
\mathbb{E}[y(t)] = \int_{-\infty}^{\infty} h(\theta) \mathbb{E}[u(t-\theta)] \, d\theta
\]

i.e.

\[
\int_{-\infty}^{\infty} y(\tau) h(\tau, y) \, d\tau \, dy = \int_{-\infty}^{\infty} h(\theta) \int_{-\infty}^{\infty} u(t-\theta) p(y) \, du \, dy \, d\theta
\]

or

\[
\mu_y(t) = \int_{-\infty}^{\infty} h(\theta) \mu_u(t-\theta) \, d\theta
\]

for a general (non stationary) stochastic process;

\[
\mu_y = \mu_u \int_{-\infty}^{\infty} h(\theta) \, d\theta
\]

if \( u(t) \) is a stationary process;

\[
\mu_y = 0
\]

if \( u(t) \) is a centered process (\( \mu_u = 0 \)).

For the second moment we can multiply eq. (4.18) with a suitably chosen time function and take the expectation:

\[
\mathbb{E}[u(t-\tau) y(t)] = \int_{-\infty}^{\infty} h(\theta) \mathbb{E}[u(t-\tau) u(t-\theta)] \, d\theta
\]
or

\[ \psi_{uy}(t, t) = \int_{-\infty}^{+\infty} h(\theta) \psi_{uu}(t, t - \theta) d\theta \]

for a general (non stationary) stochastic process;

\[ \psi_{uy}(t) = \int_{-\infty}^{+\infty} h(\theta) \psi_{uu}(t - \theta) d\theta \overset{df.}{=} h(t) * \psi_{uu}(t) \]

for a stationary process.

Along these lines one can find the following relations:

\[ \begin{align*}
\psi_{uy}(t) &= h(t) * \psi_{uu}(t) \\
\psi_{yu}(t) &= h(-t) * \psi_{uu}(t) \\
\psi_{yy}(t) &= h(t) * \psi_{yy}(t) \\
\psi_{yy}(t) &= h(-t) * \psi_{yy}(t)
\end{align*} \tag{4.20} \]

which can be represented in the above diagram. Note that for \( \mu = 0 \):

\[ \psi_{uu}(0) = \sigma_u^2 \quad \text{and} \quad \psi_{yy}(0) = \sigma_y^2 \]

By \( L_2 \) transformation, one obtains the corresponding relations in the frequency domain:

\[ \begin{align*}
\tilde{\Phi}_{uy}(s) &= H(s) \tilde{\Phi}_{uu}(s) \\
\tilde{\Phi}_{yu}(s) &= H(-s) \tilde{\Phi}_{uu}(s) \\
\tilde{\Phi}_{yy}(s) &= H(s) \tilde{\Phi}_{yy}(s) \\
\tilde{\Phi}_{yy}(s) &= H(-s) \tilde{\Phi}_{yy}(s)
\end{align*} \tag{4.21} \]

which is represented in the above diagram.
From the auto- and cross correlation function the process impulse response can be determined by "deconvolution". Generally speaking it is not a trivial problem to determine $h(t)$ from eq. (4.20) if $\psi_{uu}(t)$ and $\psi_{uy}(t)$ are given by correlation measurements. Special purpose analog equipment has been built for this problem (Wallman, 1950; Goodman, 1956). In addition routines for the digital computer are available. If $u(t)$ is white noise then

$$\psi_{uy}(t) = \int_{-\infty}^{+\infty} c \delta(t-\theta) h(\theta) d\theta = c h(t)$$

with $c = \overline{\psi_{uu}(0)}$, and deconvolution is not needed.

Instead of using the impulse response one may also characterize a process by the time moments of that impulse response, defined as

$$m_i = \int_{-\infty}^{+\infty} t^i h(t) dt$$

$m_0$ is the area under the impulse response curve and is equal to the static gain of the process; $m_1$, the first moment of impulse response area with respect to the $t=0$ axis, can be interpreted as a mean delay time $t$

$$t_\mu = \frac{m_1}{m_0}$$

$m_2$, the moment of inertia of the impulse response area with respect to the $t=0$ axis, can be interpreted as a dispersion time $t_\sigma$:

$$t_\sigma = \sqrt{m_2 - \left(\frac{m_1}{m_0}\right)^2}$$

The transfer function may be expanded in a series:

$$H(s) = \int_{-\infty}^{+\infty} h(t) e^{-st} dt =
\int_{0}^{\infty} h(t) \left\{ 1 - st + \frac{s^2 t^2}{2} - \ldots \right\} dt =
= m_0 - m_1 s + \frac{m_2 s^2}{2} - \ldots$$

If this series converges then the transfer function is characterized by these moments, cf. Goodman and Hillsley (1958).
From the power spectra density functions the transfer function follows as:

\[ H(j\omega) = \frac{\Phi_{yy}(j\omega)}{\Phi_{ww}(j\omega)} \]  

(4.22)

4.3 Varylinear models

As mentioned in section 4.1, the time dependence of coefficients of a (linear) differential or difference equation does not destroy the linearity. Consequently, the types of description given in the previous section remain valid. The simple Laplace transform, however, does not offer much help. The useful relations can be summarized as follows:

**Input-output description**

\[ y(t) = \int_{-\infty}^{t} h(t, \tau) u(\tau) \, d\tau \]  

(4.23)

where \( h(t, \tau) \) can be found through substitution of \( y, \dot{y}, \) etc. in the differential equation and equating coefficients. For a survey of techniques used in describing time-varying networks cf. Zadeh (1961), where 138 references to the literature are given.

**State space description**

\[
\begin{align*}
\dot{x}(t) &= A(t)x(t) + B(t)u(t) \\
y(t) &= C(t)x(t) + D(t)u(t)
\end{align*}
\]  

(4.24)

from which:

\[
\begin{align*}
x(t) &= \phi(t, t_0)x(t_0) + \int_{t_0}^{t} \phi(t, \tau)B(\tau)u(\tau) \, d\tau \\
\phi(t, t_0) &= I - \int_{t_0}^{t} \phi(t, \tau)A(\tau)\phi(\tau, \tau) \, d\tau
\end{align*}
\]  

(4.25)

with

\[
\begin{align*}
\frac{d}{dt} \phi(t, \tau) &= A(t)\phi(t, \tau) \\
\frac{d}{d\tau} \phi(\tau, t) &= -A'(\tau)\phi'(\tau, t) \quad \text{adjoint equation}
\end{align*}
\]  

(4.26, 4.26a)

\[
\phi(t, \tau)\phi(\tau, t) = I
\]  

(4.27)

For sampled signals the state space representation may be given as:

\[
\begin{align*}
x(k+1) &= A(k)x(k) + B(k)u(k) \\
y(k) &= C(k)x(k) + D(k)u(k)
\end{align*}
\]  

(4.28)

and corresponding expressions for the transition matrix \( \phi(k,i) \).
4.4 Nonlinear models

In section 4.1, the importance of linearity and the associated principle of superposition has been stressed. In such a case the knowledge of one characteristic process time function (e.g., impulse response) is sufficient to determine the process output for arbitrary input signals. Such a procedure has advantages which we would also like to have for the description of nonlinear processes (George, 1959) viz:

- it gives an explicit input-output relationship
- it facilitates the discussion of combinations of systems
- it allows the consideration of random inputs.

For some classes of nonlinear systems these requirements are fulfilled by the Volterra series.

**Volterra representation** (Volterra, 1959)

Much of the material in this section is taken from Van Harten (1966).

Through the use of a Volterra series, whose kernels in addition to the ordinary linear impulse response are higher order impulse responses, it is possible to characterize a nonlinear system in a manner which yields a good deal of physical insight into the problem. Moreover, the method of solution has the great conceptual advantage of presenting a nonlinear system as a rather straightforward generalization of the linear case, although the actual mechanics may be much more involved. Said in a converse manner, the Volterra series method treats the linear case as a subclass of the nonlinear case, a very desirable property indeed (Zadeh, 1953).

This advantage can hardly be overstressed because of the great deal of effort and time engineers have spent on solving linear systems in both the time and frequency domains; the Volterra series makes it
possible to utilize these concepts and way of thinking for nonlinear systems.

In the analysis and synthesis of nonlinear systems the Volterra functional representation is the most satisfactory and the most practical of the existing representations of nonlinear systems. The amount of literature available on the Volterra series representation is extensive; many authors have contributed to the analysis and synthesis.

Let us consider a general nonlinear system. We will assume that the input $u(t) = 0$ for $t<0$. Moreover we will assume that the memory of the system is finite, that is to say, it is granted that the value $u(t_0 - \tau)$ with $\tau$ sufficiently large no longer adds to the response $y(t_0)$. These assumptions follow naturally as a result of stability as well as physical realizability of the system. (An ideal integrator has an infinite memory). The input quantity can be approximated through rectangular pulses with finite dimensions. The samples from $u(t)$ for $t \leq t_0$, taken periodically with time intervals $\Delta t$, will be called $u_1, u_2, \ldots, u_N$ respectively; cf. fig. 4.4. $N$ will be taken so large that $u_n$ for $n > N$ will no longer add to the value of $y(t)$. The response $\tilde{y}(t_0) = f(u_1, u_2, \ldots, u_N)$ is thus approximated as a function of $N$ variables.

A multidimensional Taylor development gives us:

$$
\tilde{y}(t) = (a_1 u_1 + a_2 u_2 + \ldots + a_N u_N) + \\
+ (a_{12} u_1^2 + a_{13} u_1 u_2 + \ldots + a_{NN} u_N^2) + \\
+ (a_{111} u_1^3 + a_{112} u_1^2 u_2 + \ldots + a_{NNN} u_N^3) + \ldots = \\
= \sum_{i=1}^{N} a_i u_i + \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} u_i u_j + \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} a_{ijk} u_i u_j u_k + \ldots = \\
= \tilde{y}_{lin.} + \tilde{y}_{quadr.} + \tilde{y}_{cub.} + \ldots
$$

(4.29)

Here it is assumed that the initial conditions are zero.

Under the previously named conditions and if $\Delta t$ is small in comparison to the "time constant" of the system $\tilde{y}(t)$ will be a close approximation of $y(t)$. 


Should all \( a_{ij} \) and \( a_{ijk} \) be nought then only \( y_{lin} = \sum_{i=1}^{N} a_i u_i \) will remain and this is an approximation of the convolution integral for linear systems.

The latter is shown as follows:

We introduce \( a_i = h_i \Delta t \) where \( h_i \) is the height and \( \Delta t \) is the breadth of the pulse. Therefore:

\[
\tilde{y}_{lin} = h_i u_i \Delta t + h_j u_j \Delta t + \cdots = \sum_{i=1}^{N} h_i u_i \Delta t
\]

(4.30)

Taking into account the direction of the numbering of the \( u_i \), cf. fig. 4.4., this sum for \( \Delta t \rightarrow 0 \) and \( N \rightarrow \infty \) and \( N \Delta t = t_0 \) is converted into:

\[
y_{lin} = \int_{t_0}^{t} h_i(t) u(t-\tau) d\tau
\]

(4.31)

Should two pulses be applied to the system then the contribution to the linear sum follows the superposition principle: \( h_i u_i \Delta t + h_j u_j \Delta t \)

However, in sums that are quadratic, cubic etc., the interaction between the pulses is also evident.

Suppose \( a_{ij} = h_{ij} \Delta t \Delta t, \) then there is a contribution to \( \tilde{y} \) _quadr._

\[
h_{ij} u_i \Delta t \Delta t + 2h_{ij} u_i u_j \Delta t \Delta t + h_{ijj} u_j \Delta t \Delta t
\]

(4.32)

In the limit \( \tilde{y} \) _quadr._ becomes:

\[
y_{quadr} = \int_{0}^{t} \int_{0}^{t} h_{ij}(\tau_1, \tau_2) u(t-\tau_1) u(t-\tau_2) d\tau_1 d\tau_2
\]

(4.33)

In this way

\[
\tilde{y}(t) = \tilde{y}_{lin} + \tilde{y}_{quadr} + \tilde{y}_{cub}
\]

is converted into the Volterra series

\[
y(t) = \int_{0}^{t} h_i(\tau) u(t-\tau) d\tau + \int_{0}^{t} \int_{0}^{t} h_{ij}(\tau_1, \tau_2) u(t-\tau_1) u(t-\tau_2) d\tau_1 d\tau_2 + \cdots
\]

(4.34)

The description using the Volterra series is therefore a direct generalization of the description of the linear systems using a convolution integral. The impulse response \( h(\tau) \) of the linear system gives place to the impulse responses (kernels)

\[
h_1(\tau), h_2(\tau_1, \tau_2), h_3(\tau_1, \tau_2, \tau_3), \text{ etc.}
\]

The term \( h_1 \) is called an impulse response of the i-th degree.
The integral
\[
\int_{\tau_1}^{t} \cdots \int_{\tau_i}^{t} \cdots \int_{\tau_n}^{t} h_i(\tau_1, \tau_2, \ldots, \tau_i) u(t - \tau_i) \cdots u(t - \tau_1) \, d\tau_i \cdots d\tau_1
\]
is called the functional of the \(i\)-th degree.

The Volterra series \((4.34)\) is known as a functional power series or a functional development.

The correspondence with the power series is apparent. If a system has no dynamic elements or if, in the case of an harmonic input signal, the frequency approaches zero then the series \((4.34)\) gives place to a power series in \(u\). Let us suppose for example that there are no dynamics so that
\[
h_i(t) = c_i \delta(t)
\]
then the functional series takes the following form:
\[
y(t) = c_1 u(t) + c_2 u(t) u(t) + \cdots
\]
We shall see later that the behaviour of the power series gives us an insight into the kind of nonlinearity. In certain cases it is possible to investigate the convergence of the Volterra series using a power series.

If the Volterra series continues infinitely then according to Fréchet's theorem the functional power series can represent the input-output relation of all continuous nonlinear systems. The functional series is unique and converges for \(|u(t)| < r\); \(r\) being the radius of convergence. As it seldom happens that the series can be successfully brought into a closed form, in most cases only a series can be used which breaks after several terms, or which gives a satisfactory estimate after a few terms.

The kernels of the Volterra series for physical realizable systems have the following properties

a) \(h_n(\tau_1, \tau_2, \ldots, \tau_n) = 0\) for \(\tau_i < 0\) \(i = 1, 2, 3, \ldots, n\)

b) \(\lim_{\tau_i \to -\infty} h_n(\tau_1, \ldots, \tau_n) = 0\) \(i = 1, 2, 3, \ldots, n\)

c) \(h_n(\tau_1, \ldots, \tau_n)\) are or can be made symmetrical.
State space description. Along the same lines as in the previous sections, the state space representation for nonlinear processes can be chosen as:

\[
\dot{x}(t) = \left\{ \begin{bmatrix} x(t), u(t), a(t), v(t), t \end{bmatrix} \right\} \\
y(t) = \left\{ \begin{bmatrix} x(t), u(t), n(t), t \end{bmatrix} \right\}
\]

(4.64)

with

- \( x(t) \) = process state vector
- \( u(t) \) = measurable input vector
- \( a(t) \) = process parameter vector
- \( v(t) \) = unmeasurable input vector (process noise)
- \( n(t) \) = (measurement) noise vector
- \( y(t) \) = observable process vector

4.5 Controllability, observability, identifiability

Since the introduction of these concepts by Kalman (1960) they have been discussed by many authors, e.g. Elgerd (1967), Lee (1964). The reader is referred to the publications cited at the end of this chapter for a thorough discussion of this subject. Here we only summarize the results for processes described by (stationary) linear difference equations. For

The process under discussion is described by:

\[
\begin{align*}
    x(k) &= A x(k) + B u(k) \\
    y(k) &= C x(k)
\end{align*}
\]  

(4.65)

where \( x \) is an \( n \)-dimensional vector.

A process is called **controllable** if it is possible to find an (unconstrained) control vector which brings the system from any initial state to any specified final state in finite time.

Consequently, we ask for the condition under which we can determine the control necessary to drive the system from \( x(o) \) to a **preassigned** state \( x(n) \).

\[
\begin{align*}
    x(1) &= A x(o) + B u(o) \\
    x(2) &= A x(1) + B u(1) = A^2 x(o) + A B u(o) + B u(1) \\
    x(n) &= A^n x(o) + A^{n-1} B u(o) + \cdots + B u(n-1)
\end{align*}
\]

or

\[
\begin{align*}
    x(n) - A^n x(o) &= \left[ A^{n-1} B \mid A^{n-2} B \mid \cdots \mid B \right] \\
                   &\quad \left[ \begin{array}{c} u(o) \\ u(1) \\ \cdots \\ u(n-1) \end{array} \right]
\end{align*}
\]  

(4.66)

As \( A, x(o) \) and \( x(n) \) are known the left hand side of eq. (4.66) is determined.

A unique solution of \( u \) vector can be found if the matrix

\[
\left[ A^{n-1} B \mid A^{n-2} B \mid \cdots \mid B \right]
\]

has rank \( n \). In that case \( A, B \) is called a controllable pair.

A process is called **observable** if from the measurements of the output it is possible to determine the state of that process.

Consequently we ask for the condition under which we can determine \( x(o) \) by measuring \( y \).

\[
\begin{align*}
    y(o) &= C x(o) \\
    y(1) &= C x(1) = C A x(o) \\
    y(n-1) &= C A^{n-1} x(o)
\end{align*}
\]

or taking the transpose

\[
\begin{align*}
    \begin{bmatrix} y(o) \\ y(1) \\ \cdots \\ y(n-1) \end{bmatrix} &= x(o) \begin{bmatrix} C \mid A^t C \mid \cdots \mid A^{n-1} C \end{bmatrix}
\end{align*}
\]  

(4.67)
As the \( \mathbf{y} \) vectors are known, a unique solution for \( \mathbf{x}(0) \) can only be found if the matrix
\[
\begin{bmatrix}
\mathbf{C}' & \mathbf{A}' \mathbf{C}' & \cdots & \mathbf{A}'^{n-1} \mathbf{C}'
\end{bmatrix}
\]
has rank \( n \). In that case \( \mathbf{A}, \mathbf{C} \) is called an observable pair.

It has been shown by Brockett (1965) that the introduction of state variable feedback to a process may change the observability of that process.

A process is called identifiable if from the measurements of the state variables, it is possible to determine the system matrix \( \mathbf{A} \).

\[
\begin{align*}
\mathbf{x}(1) &= \mathbf{A} \mathbf{x}(0) \\
\mathbf{x}(2) &= \mathbf{A}^2 \mathbf{x}(0) \\
&\vdots \\
\mathbf{x}(n) &= \mathbf{A}^n \mathbf{x}(0)
\end{align*}
\]
or
\[
\begin{bmatrix}
\mathbf{x}(1) \\
\mathbf{x}(2) \\
&\vdots \\
\mathbf{x}(n)
\end{bmatrix} = \mathbf{A} \begin{bmatrix}
\mathbf{x}(0) \\
\mathbf{A} \mathbf{x}(0) \\
&\vdots \\
\mathbf{A}^{n-1} \mathbf{x}(0)
\end{bmatrix}
\]

(4.68)

As the \( \mathbf{x} \) vectors are known a unique solution for \( \mathbf{A} \) can be found only if the following matrix has rank \( n \):
\[
\begin{bmatrix}
\mathbf{x}(0) \\
\mathbf{A} \mathbf{x}(0) \\
&\vdots \\
\mathbf{A}^{n-1} \mathbf{x}(0)
\end{bmatrix}
\]

The interested reader will find in the literature more elaborate studies of these concepts and finer distinctions as e.g., for controllability, which may be either: state- or output-, completely- or totally-, strongly- or weakly-. The observability may be either completely- or totally- (Kreindler and Sarachik, 1964).
Fischer and Vamos (1967) have extended the controllability concept to what they call $\mathcal{J}$-controllability, in which case an initial state $x(0)$ can be driven to the $\mathcal{J}$-environment of a desired given state.

Aoki (1967) has extended the definitions of controllability, observability, and identifiability to stochastic situations. The question of identifiability has also been discussed by Balakrishnan (1969) and by Staley and Yue (1970). Closely related to identifiability is the concept of determinable classes (Root, 1965; Prosser and Root, 1966; Root, 1970).

For illustrating these concepts a simple second order system will be chosen:

\[
\begin{align*}
    x(k+1) &= A x(k) + B u(k) \\
    y(k) &= C x(k)
\end{align*}
\]

Controllability:

\[
\begin{bmatrix}
    A \\
    B
\end{bmatrix}
\]

Controllable if rank = 2; uncontrollable if $\alpha_{11} = 0$. From fig. 4.13 it is clear that when $a_{21} = 0$, we do not have control over the state $x_2$. Note that all other parameters $a_{ij}$ may be zero for the system to remain controllable.

Observability:

\[
\begin{bmatrix}
    C' \\
    A' C'
\end{bmatrix}
\]

Observable if rank = 2; unobservable if $\alpha_{21} = 0$. From fig. 4.14 it is clear that when $a_{12} = 0$, the output variable $y$, does not provide any knowledge about $x_2$. Note again that all other parameters may be zero for the system to remain observable.
Identifiability:

\[
\begin{bmatrix}
X(0) \\ A X(0)
\end{bmatrix} =
\begin{bmatrix}
x_1(0) \\ x_2(0)
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} \\ a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
x_1(0) \\ x_2(0)
\end{bmatrix}
\]

Identifiable if rank = 2; unidentifiable, if the determinant of this matrix equals zero. This is the case if both columns of the matrix are linearly dependent. We can distinguish:

- a trivial case: \(x_1(0) = x_2(0) = 0\), that is, the process at rest can not be identified.
- a non trivial case:

\[
\lambda \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = 0
\]

This is the approach for finding the eigen values \(\lambda_1\) and \(\lambda_2\) and the corresponding eigen vectors \(r_1\) and \(r_2\). If \(x(0) = \alpha r_1\) then only the mode \(\exp[\lambda_1 t]\) of the process is excited by \(x(0)\) and the mode \(\exp[\lambda_2 t]\) is not identifiable. If \(x(0) = \beta r_2\) then only the mode \(\exp[\lambda_2 t]\) can be identified. Consequently, the process is identifiable only if all modes of the process are excited by \(x(0)\).

In Chapter II other notions of identifiability are given.

4.6 Models based on these descriptions

In the previous sections, several types of description of process characteristics have been mentioned. Now the models or simulations based on these descriptions will be considered briefly. Only general ideas are presented, as the choice of a model is strongly object and goal oriented.

Linearity-of-the-dynamics versus linearity-in-the-parameters.

In section 4.1, a distinction between linear and nonlinear has been made on the criterion of dynamic behaviour, i.e., the relation between the dependent and the independent time-variables. For parameter estimation another distinction between linearity and nonlinearity is of as much importance, viz. with respect to the relation between the dependent variables.
and the parameters. A model is said to be linear-in-the-parameters if the (generalized) error is linear in the parameters.

Apparently, these two notions of linearity have no immediate relation as can be seen from the following examples. We assume a process with input signal $u$ and output signal $y$. Then the "model" may be chosen to form an "error" $e$ between process and model output in the following way:

<table>
<thead>
<tr>
<th>model; in-the-parameters:</th>
<th>process:</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear</td>
<td>$\dot{y} + ay = u$</td>
</tr>
<tr>
<td>$e = y_0 \dot{y} + ay - u$</td>
<td>$\dot{w} + \alpha w = u$</td>
</tr>
<tr>
<td>$e = y_0 \dot{y}^2 + ay - u$</td>
<td>$\dot{w} + \alpha w^2 = u$</td>
</tr>
</tbody>
</table>

The two different uses of the terms linear and nonlinear may cause some confusion. This is due to the mixing of concepts from the fields of system theory and regression analysis. Henceforth we will use the term "linear" for the dynamic behaviour and use "linear-in-the-parameters" for the other type.

In connection with estimation schemes the great importance of linearity-in-the-parameters will become clear. A typical example is the identification of a discrete-time linear system when the output is disturbed with white measurement noise. The representation of the system by the coefficients of the pulse transfer function leads to a nonlinear regression problem while the representation of the model by coefficients of a generalized model or by the ordinates of the weighting function leads to an
estimation problem which is "linear-in-the-parameters". It also pays to try to find transformations of the variables to obtain such a linearity if possible. Some simple examples may illustrate this.

\[ Z = \frac{\alpha_2 x_2 + \alpha_1 x_1}{\alpha x_1 x_2} \Rightarrow \begin{cases} z = y \\ \frac{1}{\alpha y} = \beta_1 \\ \frac{1}{\alpha z} = \beta_2 \end{cases} \Rightarrow y = \beta_1 u_1 + \beta_2 u_2 \]

reciprocal transformation

\[ Z = \exp(x_1, x_2) \Rightarrow \begin{cases} \log z = y \\ \log x_1 = u_1 \\ \log x_2 = u_2 \end{cases} \Rightarrow y = \beta_0 + \beta_1 u_1 + \beta_2 u_2 \]

logarithmic transformation

Such nonlinear expressions that can be made linear-in-the-parameters through transformation are called intrinsically linear. If such a linearization is not possible then the term intrinsically nonlinear is used. It may pay to make transformations even if the system is intrinsically non-linear; see e.g., Diskind (1969).

Linear input-output models

Linear models naturally represent the most extensively developed area in the field of identification. Now we will consider linear processes as well as "linear environments", that is, environments that can be characterized by linear stochastic models. In most control problems the properties of the environment will be just as important as the system dynamics, because it is the presence of disturbances that creates a control problem in the first place.

If classical design techniques are to be used, the model can be characterized by a transfer function or by an impulse response. Many recently developed design methods will however require a state model, i.e., a parametric model.

Several problems naturally arise:

- Suppose the impulse response is desired. Should this be identified directly or is it "better" to identify a parametric model and then compute the impulse response?
Assume that a parametric model is desired. Should this be fitted directly or is it "better" to first determine the impulse response and then fit a parametric model to that?

Since a parametric model contains the order of the system explicitly, what happens if the wrong order is assumed in the problem formulation?

There are not yet any general answers to these problems. Special cases have been investigated by Gustavsson (1969) in connection with identification of nuclear reactor and distillation tower dynamics as well as on simulated data.

For linear processes the well known analog computer techniques provide a straightforward way of modelling ordinary differential equations such as

\[
\frac{d^n y}{dt^n} + \ldots + \alpha_{n-1} \frac{dy}{dt} + \alpha_n y = b_0 \frac{d^{n-1} u}{dt^{n-1}} + \ldots + b_{n-1} \frac{du}{dt} + b_n u
\]

This leads to a transfer function

\[
H(s) = \frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \ldots + b_{n-1} s + b_n}{s^n + \alpha_1 s^{n-1} + \ldots + \alpha_{n-1} s + \alpha_n}
\]  

(4.69)

and a computer diagram as given in fig. 4.15. If the state variables are chosen as \(x_{n-1}, \ldots, x_0\) then those quantities are directly measurable in this model. If \(y\) and its derivatives are chosen as state variables then they can be formed as linear combinations of these \(x\) variables. In the case \(\alpha_i = 0\) for \(i = 1, \ldots, m\) then \(x_i = \frac{y}{s}\). Note that \(H(s)\) is non-linear-in-the-parameters \(\alpha_i\).

The transfer function of a linear (ized) process, e.g.

\[
H(s) = \frac{(s + a_1)(s + a_2)}{(s + e_1)(s + e_2)(s + e_3)}
\]

(4.70)

can be decomposed in several ways, cf. fig. 4.16:

- as a series combination of simple transfer functions having only one or two poles and perhaps a zero of the overall transfer function.

\[
H(s) = \frac{s + a_1}{s + e_1} \cdot \frac{s + e_2}{s + e_2} \cdot \frac{1}{s + e_3}
\]

(4.71)
as a parallel combination of simple transfer functions, each weighted by a particular coefficient $\beta_i$: $$H(s) = \frac{\beta_i}{s + e_1} + \frac{\beta_2}{s + e_2} + \frac{\beta_3}{s + e_3} \quad (4.72)$$ as a (parallel) combination of orthogonal transfer functions: $$H(s) = \sum_{i=1}^{n} \beta_i F_i(s) \quad (4.73)$$ as an example may serve the Laguerre representation $$F_i(s) = \frac{1}{s + \alpha} \left(\frac{s - \alpha}{s + \alpha}\right)^i$$ which leads to the block diagram of fig. 4.17.

The modelling techniques mentioned before can also be applied to systems in which information of the signals is available only as samples at discrete intervals of time. Figure 4.18 gives one example of a simple first order system with the transfer function: $$H(z) = \frac{Z}{z + \alpha} \quad \text{with} \quad Z = e^{s \theta}$$ A higher order transfer function $$H(z) = \frac{b_0 + b_1 z + \cdots + b_n z^n}{1 + a_1 z + \cdots + a_n z^n} \quad (4.74)$$ can be instrumented using delay units analogously to fig. 4.15, and in addition it can be decomposed analogously to fig. 4.16. A pure time delay can easily be incorporated: $$q(z) = z^{-k} H(z)$$ This type of process description can well be used for control purposes. If the process is contaminated with non-white additive noise, then the noise properties can also be described (and estimated) in terms of such a transfer function with a white-noise input.

The impulse response $h(t)$ provides a general, non-parametric representation for a linear process, because it requires not more a priori information than that the process is linear.
A pure time delay can easily be accommodated. It can be instrumented in several ways. A condition, however, is that the model enables the convolution of \( h(t) \) with \( u(t) \). This leads to the following possibilities:

- \( h(t) \) given as a photographic mask and \( u(t) \) as a film track
- \( h(t) \) given as a series of coefficients of orthonormal terms; cf.

Chapter 3:

\[
\begin{align*}
    h(t) &= \sum \beta_i g_i(t) \\
    \int_a^b g_i(t) g_j(t) \, dt &= \delta_{ij}
\end{align*}
\]

where \( \delta_{ij} = 1 \) if \( i = j \) and \( \delta_{ij} = 0 \) if \( i \neq j \).

The usefulness of the representation as given by eq. (4.75) depends to a large extent on the "match" between the impulse response and the type of (orthonormal) functions chosen. In a favourable case, only a few parameters \( \beta_i \) unequal to zero may give a satisfactory approximation; otherwise it may take a rather large number. It has been proposed (Wolff and Dietz, 1962; Izawa and Furuta, 1967) that the functions \( g_i(t) \) be made dependent on other parameters, e.g., \( g_i(t; \alpha) \). In that case, one needs to determine the best values of the \( \beta \) and \( \alpha \) parameters, complicating the estimation procedure. Attempts have been made to extend the orthogonal models to multi-variable processes (Barker, 1965).

- \( h(t) \) given as a number of points \( \beta_0 = h(0), \beta_1 = h(\theta), \beta_2 = h(2\theta), \ldots \)

This leads to the delay line synthesizer (Goodman and Reswick, 1956).

Instrumented according to fig. 4.19, the convolution is approximated by

\[
    \psi(t) = \sum_{i=0}^{h} \beta_i u(t - i\theta)
\]

A disadvantage is the rather large number of points (parameters) needed,
The approximation of the impulse response may take several forms, cf. fig. 4.20:

\[ h_i(t) \text{ as a (continuous) function} \]
\[ h_i(t) = \sum h_k \delta(t-k) \]
\[ h_2(t) = h_1(t) \]
\[ h_3(t) = h_1(t) + \frac{\delta - \theta}{\tau} [h_1(t+\theta) - h_1(t)] \text{ for } 0 \leq t-\theta < \theta \]

Linearity-in-the-parameters enables us to represent the sampled-date case in the form

\[ y = U b \]

with

\[ U = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix} \]

Note that

\[ U = \begin{bmatrix} u_0 & \cdots & u_m \end{bmatrix} \]

and that no assumption has yet been made with respect to the relations between \( u_j \) and \( u_j \). They may represent the same sample values shifted in time (delayline outputs), but they may as well stand for output samples of nonlinear filters. If the "delay line" representation is chosen then

\[ U = \begin{bmatrix} u_0 \end{bmatrix} \]

where \( u(i) = 0 \) for \( i \leq 0 \) is assumed.

In addition, the following can be used:

\[ y = \mathcal{B} u \]

with

\[ y = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix} \]

and

\[ U = \begin{bmatrix} b_0 & 0 & \cdots & 0 \\ b_1 & b_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ b_m & b_{m-1} & \cdots & b_0 & 0 \end{bmatrix} \]

where \( u(i) = 0 \) for \( i \leq 0 \) is assumed.
This permits a direct notation of the covariance of \( y \), viz.

\[
\text{cov}[y] = \mathcal{B} \text{cov}[y] \mathcal{B}^\prime 
\]  

(4.80)

**Linear state models; canonical forms.**

Finally, our attention will be focussed on the important type of state model \( S(A, B, C, D) \) defined by:

\[
\begin{align*}
\dot{x} &= A x + B u \\
y &= C x + D u
\end{align*}
\]  

(4.81)

where \( x \) is an \( n \)-vector, the input \( u \) is a \( p \)-vector and the output \( y \) is an \( r \)-vector. It is well known that the systems \( S(A, B, C, D) \) and \( S(TAT^{-1}, TB, CT^{-1}, D) \) where \( T \) is a nonsingular matrix are equivalent in the sense that they have the same input-output relation. It is also easy to verify that the systems \( S(A, B, C, D) \) and \( S(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) \) are equivalent in the sense that they have the same input-output relation if

\[
\begin{align*}
\tilde{D} &= \tilde{D} \\
\tilde{C} A^k \tilde{B} &= \tilde{C} \tilde{A}^k \tilde{B} \\
k = 0, 1, \ldots, n
\end{align*}
\]  

(4.82)

The relations between the different representations were clarified by Kalman's work; see e.g., Kalman (1963). The impulse response and the transfer function represent only the part of the system \( S \) which is completely controllable. It is thus clear that only the completely controllable and completely observable part of a state model \( S(A, B, C, D) \) can be determined from input-output measurements.

The problem of determining a state model from the impulse response is more subtle, even if we disregard the fact that only the controllable and observable subsystem can be determined from the impulse response. The problem of assigning a state model of the lowest possible order which has a given impulse response has been solved by Ho and Kalman (1966). See also Kalman, Falb and Arbib (1969). Again the solution is not unique. The model \( S(A, B, C, D) \) contains
parameters. The fact that the input-output relation is invariant under a linear transformation of the state variables implies that all $N_1$ parameters cannot be determined from input-output measurements. To obtain unique solutions as well as to be able to construct efficient algorithms, it is therefore of great interest to find representations of the system which contain the smallest number of parameter i.e., canonical representations.

Canonical forms for linear deterministic models. Canonical forms for linear systems are discussed e.g., by Kalman (1963). When the matrix $A$ has distinct eigenvalues, canonical forms can be obtained as follows. By a suitable choice of coordinates the matrix $A$ can be brought to diagonal form.

\[
\begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
+ \begin{bmatrix}
b_{11} & b_{12} & \cdots & b_{1p} \\
b_{21} & b_{22} & \cdots & b_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
b_{n1} & b_{n2} & \cdots & b_{np}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
\vdots \\
u_n
\end{bmatrix}
\tag{4.84}
\]

This representation contains $n + np + nr + pr$ parameters of which $n$ are redundant since all state variables can be scaled without affecting the input-output relations. The input-output relation can thus be characterized by
\[ N_z = n(p+r) + pr \] (4.85)

parameters. Since the system is completely controllable and observable there is at least one non zero element in each row of the \( B \) matrix and of each column of the \( C \) matrix. The redundancy in (4.84) can thus be reduced by imposing conditions such as

\[ \max_j b_{ij} = 1 \quad i = 1, 2, \ldots, n \] (4.86)

\[ \sum_j |b_{ij}| = 1 \quad i = 1, 2, \ldots, n \] (4.87)

or similar conditions on the \( C \)-matrix. When the matrix \( A \) has multiple eigenvalues the problem of finding a minimal parameter representation is much more complex. If \( A \) not only has multiple eigenvalues but is cyclic as well (i.e. there exist a vector \( x \) such that the vectors \( x, Ax, A^2 x, \ldots, A^{n-1} x \) span the \( n \)-dimensional space), the matrix can be transformed to companion form and a minimal parameter representation is then given by

\[
\begin{align*}
\dot{x} &= \begin{bmatrix}
-a_1 & 1 & 0 & \cdots & 0 \\
-a_2 & 0 & 1 & \cdots & 0 \\
& \ddots & \ddots & \ddots & \vdots \\
-a_{n-1} & 0 & 0 & \cdots & 1 \\
-a_n & 0 & 0 & \cdots & 0 
\end{bmatrix} x + \begin{bmatrix}
b_{i1} \\
b_{i2} \\
\vdots \\
b_{in}
\end{bmatrix} u \\
y &= \begin{bmatrix}
c_{11} & \cdots & c_{1n} \\
& \ddots & \ddots & \vdots \\
c_{r1} & \cdots & c_{rn}
\end{bmatrix} x + \begin{bmatrix}
d_{11} \\
\vdots \\
d_{r1}
\end{bmatrix} u
\end{align*}
\] (4.88)
where \( n \) additional conditions, e.g. of the form (4.86) or (4.87) are imposed on the elements of the matrices \( \mathbf{B} \) and \( \mathbf{C} \).

In the case of processes with one output, the additional conditions are conveniently introduced by specifying all elements of the vector \( \mathbf{C} \) e.g. 
\[
\mathbf{C}' = [1 \ 0 \ \ldots \ \ 0].
\]

The canonical form then becomes
\[
\begin{align*}
\mathbf{Y}(s) &= \left[ \begin{array}{c}
\mathbf{c}'_1 + \frac{\mathbf{b}_1 s^{n-1} + \mathbf{b}_2 s^{n-2} + \ldots + \mathbf{b}_n}{s^n + a_1 s^{n-1} + \ldots + a_n}
\end{array} \right] \mathbf{U}_i(s) + \\
&+ \\
&+ \left[ \begin{array}{c}
\mathbf{c}'_p + \frac{\mathbf{b}_1 s^{n-1} + \mathbf{b}_2 s^{n-2} + \ldots + \mathbf{b}_n}{s^n + a_1 s^{n-1} + \ldots + a_n}
\end{array} \right] \mathbf{U}_p(s)
\end{align*}
\]

where \( \mathbf{Y} \) and \( \mathbf{U}_i \) denote the Laplace transforms of \( y \) and \( u_i \). A canonical representation of a process of the \( n \)th order with \( p \) inputs and one output can thus be written as

\[
\begin{align*}
\frac{\partial^n y}{\partial t^n} + a_1 \frac{\partial^{n-1} y}{\partial t^{n-1}} + \ldots + a_n y &= \\
= \left[ b'_0 \frac{\partial^n u_i}{\partial t^n} + \ldots + b'_n u_i \right] + \\
&+ \\
&+ \left[ b'_{0,p} \frac{\partial^n u_p}{\partial t^n} + \ldots + b'_{n,p} u_p \right]
\end{align*}
\]

An analogous form for systems with several outputs is
\[ \frac{d^n y}{dt^n} + A_1 \frac{d^{n-1} y}{dt^{n-1}} + \cdots + A_n y = \]
\[ = \left[ B_{0t}, \frac{d^n u_t}{dt^n} + \cdots + B_n, u_t \right] + \]
\[ + \cdots \cdots \cdots + \]
\[ + \left[ B_{0q}, \frac{d^n u_q}{dt^n} + \cdots + B_{nq}, u_q \right] \]

(4.91)

This form has been used among others by Wong et al. (1968) and Rowe (1968). The determination of the order of the process (4.91), which in general is different from \( n \), as well as the reduction of (4.91) for state form has been done by Tuel (1966). Canonical forms for linear multivariable systems have also been studied by Luenberger (1967) and Bucy (1968). Analogous results hold for discrete time systems.

When the matrix \( A \) has multiple eigenvalues and is not cyclic it is not clear what a "minimal parameter representation" means. The matrix \( A \) can of course always be transformed to Jordan canonical form (Friedman, 1956). Since the eigenvalues of \( A \) are not distinct the matrix \( A \), strictly speaking, be characterized by fewer than \( n \) parameters. The one's in the superdiagonal of the Jordan form can, however, be arranged in many different ways depending on the internal couplings which leads to many different structures.

Canonical forms for linear stochastic models. We will now discuss canonical forms for stochastic systems. To avoid the technical difficulties associated with continuous-time white noise we will present the results for discrete time systems. The analogous results are, however, true also for continuous time systems. Consider the system
\[
\begin{align*}
\mathbf{x}(k+1) &= A \mathbf{x}(k) + B \mathbf{u}(k) + \mathbf{v}(k) \\
\mathbf{y}(k) &= C \mathbf{x}(k) + D \mathbf{u}(k) + \mathbf{n}(k)
\end{align*}
\] (4.92)

where \( k \) takes integer values. The state vector \( \mathbf{x} \), the input \( \mathbf{u} \) and the output \( \mathbf{y} \)
have dimensions \( n, p \) and \( r \); \( \{v(k)\} \) and \( \{n(k)\} \) are sequences of independent
random vectors with zero mean values and covariance \( \mathbf{V} \) and \( \mathbf{N} \).

Since the covariance matrices are symmetric the model (4.92) contains

\[
N_3 = n^2 + np + nr + pr + \frac{1}{2} n(n+1) + \frac{1}{2} r(r+1) = \\
= n \left( \frac{3}{2} n + \frac{1}{2} + p + r \right) + r \left( p + n + \frac{1}{2} \right)
\] (4.93)

parameters. Two models of the type (4.92) are said to be equivalent if: (i)
their input-output relations are the same when \( n = 0 \) and \( v = 0 \) and (ii) the
stochastic properties of the outputs are the same when \( u = 0 \). The parameters
of \( A, B \) and \( C \) can be reduced by the techniques applied previously. It still
remains to reduce the parameters representing the disturbances. This is accom-
plished e.g., by the Kalman filtering theorem. It follows from this that the
output process can be represented as

\[
\begin{align*}
\hat{\mathbf{x}}(k+1) &= A \hat{\mathbf{x}}(k) + B \mathbf{u}(k) + K \mathbf{r}(k) \\
\mathbf{y}(k) &= C \hat{\mathbf{x}}(k) + D \mathbf{u}(k) + \mathbf{r}(k)
\end{align*}
\] (4.94)

where \( \hat{\mathbf{x}}(k) \) denotes the conditional mean of \( \mathbf{x}(k) \) given \( \mathbf{y}(k-1), \mathbf{y}(k-2), \ldots \), and
\( \{\mathbf{r}(k)\} \) is a sequence of independent equally distributed random variables
with zero mean values and covariance \( \mathbf{R} \).
The single output version of the model (4.94) was used in Åström (1964). Kailath (1968) calls (4.94) an innovations representation of the process. A detailed discussion is given in Åström (1970). The model (4.94) is also used by Mehra (1969).

Notice that if the model (4.94) is known, the steady state filtering and estimation problems are very easy to solve. Since $K$ is the filter gain it is not necessary to solve any Riccati equation. Also notice that the state of the model (4.94) has a physical interpretation as the conditional mean of the state of (4.92). If $A$ is chosen to be in diagonal form and if conditions such as (4.86) are introduced on $B$ and $C$ the model (4.94) is a canonical representation which contains

$$
N_{y} = n(p+2r) + r(p + \frac{c}{z} + \frac{1}{z})
$$

(4.95)

parameters.

For systems with one output, where the additional conditions are $\zeta = [1 \ 0 \ \ldots \ 0]$, the equation (4.94) then reduces to

$$
y(k) + a_{1}y(k-1) + \ldots + a_{n}y(k-n) = \\
= [b_{1}^{*}u_{1}(k) + \ldots + b_{n}^{*}u_{n}(k-n)] + \\
+ \ldots + \\
+ [b_{1}^{*}u_{1}(k) + \ldots + b_{n}^{*}u_{n}(k-n)] + \\
r(k) + c_{1}r(k-1) + \ldots + c_{n}r(k-n)
$$

(4.96)

By introducing the shift operator $z$ defined by

$$
z \cdot y(k) = y(k+1)
$$

(4.97)

the polynomials
\[ A(z) = z^n + a_n z^{-n} + \ldots + a_1 z^{-1} + a_0 \]

\[ B_i(z) = b_{ni} z^n + b_{ni} z^{-n} + \ldots + b_{ni} \]

\[ C(z) = z^n + c_n z^{-n} + \ldots + c_1 z^{-1} + c_0 \]  \hspace{1cm} (4.98)

and the corresponding reciprocal polynomials

\[ A^*(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n} = z^{-n} A(z) \]

\[ B_i^*(z^{-1}) = b_{ni} + b_{ni} z^{-n} + \ldots + b_{ni} z^{-n} = z^{-n} B_i(z) \]

\[ C^*(z^{-1}) = 1 + c_1 z^{-1} + \ldots + c_n z^{-n} = z^{-n} C(z) \]  \hspace{1cm} (4.99)

the equation (4.96) can be written as

\[ A^*(z^{-1}) y(k) = \sum_{i=1}^{k} B_i^*(z^{-1}) u_i(k) + C^*(z^{-1}) r(k) \]  \hspace{1cm} (4.96')

or

\[ A(z) y(k) = \sum_{i=1}^{k} B_i(z) u_i(k) + C(z) r(k) \]  \hspace{1cm} (4.96'')

This canonical form of an n-th order system was introduced in Åström, Bohlin and Wensmark (1965) and has since then been used extensively. The corresponding form for multivariable systems is obtained by interpreting \( y \) and \( u_i \) as vectors and \( A, B_i, \) and \( C \) as polynomials whose coefficients are matrices.

Such models have been discussed by Eaton (1967), Kashyap (1970), Rowe (1968), Vališ (1970).

The following canonical form

\[ y(k) = \frac{B_1(z)}{A_1(z)} u_1(k) + \frac{B_2(z)}{A_2(z)} u_2(k) + \ldots + \]

\[ + \frac{B_p(z)}{A_p(z)} u_p(k) + \frac{C(z)}{A(z)} r(k) \]  \hspace{1cm} (4.100)
has been used by Bohlin (1968) and Steiglitz and McBride (1965) as an alternative to (4.96).

The choice of model structure can greatly influence the amount of work required to solve a particular problem. We illustrate this by:

**A filtering example.** Assume that the final goal of the identification is to design a predictor using Kalman filtering. If the process is modeled by

\[
\begin{align*}
x(k+1) &= A x(k) + v(k) \\
y(k) &= C x(k) + w(k)
\end{align*}
\]

(4.101)

where \( \{n(k)\} \) and \( \{v(k)\} \) are discrete-time white noise with covariances \( N \) and \( V \), the likelihood function for the estimation problem can be written as

\[
-\log L = \frac{1}{2} \sum_{k=1}^{n} \left[ \bar{y}'(k) V^{-1} \bar{y}(k) + n(k) N^{-1} n(k) \right] + \\
+ \frac{n}{2} \log (\det V)(\det N) + \text{const.}
\]

(4.102)

where the system equations are considered as constraints. The evaluation of gradients of the loss function leads to two point boundary value problems. In addition, when the identification is done, the solution of the Kalman filtering problem requires the solution of a Riccati equation. Assume instead that the process is identified using the structure

\[
\begin{align*}
v(k+1) &= A v(k) + K e(k) \\
y(k) &= C v(k) + r(k)
\end{align*}
\]

(4.103)
the likelihood function then becomes

\[ - \log L = \frac{1}{2} \sum_{k=1}^{n} \sigma'(k) R^{-1} \sigma(k) + \frac{n}{2} \log \det R \]

(4.104)

The evaluation of gradients of the loss function in this case is done simply as an initial value problem. When the identification is done the steady state Kalman filter is simply given by

\[
\hat{x}(k+1) = A \hat{x}(k) + K \left[ y(k) - A \hat{x}(k) \right]
\]

(4.105)

Hence if the model with the structure (4.103) is known there is no need to solve a Riccati equation in order to obtain the steady state Kalman filter.
Nonlinear models

For a wide class of nonlinear processes a rather general, non-parametric representation is given by the Volterra series, using impulse responses of increasing dimensionality.

Approximation of these functions by a finite number of points leads to a model that is linear-in-the-parameters; c.f. Eykhoff (1963), Alper (1965), Roy and Sherman (1967). For most practical cases, the number of parameters needed for this description is too large. Assume that for a process the impulse response of the first degree, \( h_1(\tau) \), can be approximated by say 10 sample values (parameters). Then taking into account the symmetry \( h_2(\tau_1,\tau_2) \) would need \( \frac{1}{2} \cdot 10 \cdot (10+1) = 55 \) parameters, \( h_3(\tau_1,\tau_2,\tau_3) \) would need 220 parameters, etc. if these impulse responses of higher degree have to have the same "resolution". Apparently the dimensionality of the parameter estimation problem becomes excessive. Moreover the physical interpretation of these parameters probably will hardly be possible.

Consequently, more than for linear processes, for which non-parametric models are feasible, one needs to choose a parametric model when studying nonlinear processes. The use of a priori knowledge due to physical insight into the process is of high importance.

A scheme for deriving a parametric model from observations has been proposed by Ivakhnenko (1970), Ivakhnenko and Koppa (1970), under the name Heuristic Self-Organization or Group Method of Data Handling. The method consists of using a hierarchy of "partial" models instead of a "complete" model that consists for example, of a nonlinear polynomial of higher order.
The choice of the particular type of (canonical) form to be used is dictated by a number of considerations, e.g., the interest in: the eigenvalues, the coefficients of difference equation, certain critical parameters, the minimum number of parameters, etc.

Some examples of models for industrial processes are given by Van der Grinten in the IFAC survey paper (Eykhoff et al., 1966), cf. the reference at the end of chapter 1. From the same reference is the following quotation from Kwakernaak:
"To show that the notion of state can be advantageously employed in models of varying complexity we quote some examples which are picked from different areas of application in a rather random fashion.

- Models of electrical networks have as their state variables capacitance voltages and inductance currents.
- Mechanical systems can be characterized by models having positions and velocities as their state variables.
- Thermal systems with material flow have enthalpies as their state variables.
- A simple model for the design of a longitudinal guidance system for aircraft landing employs as state variables pitch, pitch rate, altitude and rate of ascent of the aircraft.
- The model of a multi-machine power system has as its state variables magnetic fluxes, rotor angles and machine speeds.
- In process engineering also examples are known in which the concept of state is applied. A lumped-parameter model of a superheater includes as its state variables the output steam enthalpy and the tube wall temperature.
- A model of a distillation column has as its state variables the liquid composition and the enthalpy of the saturated liquid at each tray. Numerical methods can be employed to obtain verification.

Also in other fields the notion of state finds application in dynamic model building:

- Simulation of traffic on a 'microscopic' scale can be done with a model in which the positions and speeds of individual automobiles are recorded.
- Dynamic models in industrial engineering also employ the notion of state. A model of a production-distribution system developed by Forrester has as its state variables (denoted as 'levels'): "inventories, number of employees, unfilled orders, stocks of capital equipment, bank balances, orders in
transit in communication channels, goods in shipment and unfilled orders for new employees”.

It appears that the notion of state can be advantageously employed for the description of systems of a widely varying nature. It is of great help in understanding the operation of the system and in conceiving the model. Digital computers are very suitable to carry out the numerical manipulations necessary for simulation".


Ho, B.L. and R.E. Kalman (1966). Effective construction of linear state-variable model from input/output functions. Regelungstechnik 14, 545-548.


Additional literature

4.2. Linear models.


State Space Analysis: an Introduction, 


4.3. Varylinear models.


4.4. Nonlinear models.


4.5. Controllability and observability, identifiability.


.6. Models based on these descriptions

Ilper, P. (1965). A consideration of the
control, AC-10, 322-327.

Davison, E.J. (1968). A new method for sim-
plifying large linear dynamic systems. (cor-
respondence). IEEE Trans. autom. control,

complex domain analysis. Preprint JACC,
'70-781.

autom. control, AC-8, 347-357.

identification of dynamics of the Ågesta
reactor and comparison with results of
spectral analysis. Report 6903. Division of
Automatic Control, Lund Institute of Techn.,
Lund, Sweden.

ation in problems of engineering cy-
bernetics. Automatica, 6, 207-219.

The group method of data handling for the
solution of the various interpolation pro-
blems of cybernetics. 2nd IFAC symp.
"Identification and process parameter esti-
mation", Prague, paper 2.1.

Kashyap, R.L. (1970). A new method of re-
cursive estimation in discrete linear
systems. IEEE Trans. autom. control, AC-15,
to appear.

Koepcke, R.W. (1963). Solution to sampled,
minimum-time problem. Paper JACC, 94-100;
also in Trans. ASME (J. Basic Engng),

Roy, R.J. and J. Sherman (1967). A learning
 technique for Volterra series representation.
(short paper). IEEE Trans. autom. control,
AC-12, 761-764.
Fig. 4.1

Fig. 4.3
Fig. 4.10

Fig. 4.11

\[
\frac{u(t)}{U(s)} = \frac{1}{s+1} \frac{s}{s+1} \frac{s}{s+1} \ldots \frac{s}{s+1} \frac{-1}{s+1} \frac{-s}{s+1}
\]
fig. 4.15

\[
\begin{align*}
&\frac{s + d_1}{s + e_1} \quad \frac{s + d_2}{s + e_2} \quad \frac{1}{s + e_3} \\
\text{(a)}
\end{align*}
\]

fig. 4.16

\[
\begin{align*}
&\frac{1}{s + e_1} \\
&\frac{1}{s + e_2} \\
&\frac{1}{s + e_3} \\
\text{(b)}
\end{align*}
\]

fig. 4.17
Chapter 5  SOME PROBABILISTIC NOTIONS;
Estimation theory, convergence schemes, stochastic approximation.

In this chapter, some ideas are discussed that are of great importance for parameter estimation. First the estimators are considered from the statistical point of view. Evidently there is a close relation between the amount of knowledge of the process that is initially available and the type of "optimum" procedure that can be used to estimate the parameters. The characteristics of different kinds of estimators are considered:

- the least squares estimator
- the Markov estimator
- the maximum likelihood estimator
- the Bayes' estimator

The amount of assumed initial knowledge available on the process increases in this order. For the least squares estimator, the only assumption is that the dynamics of the process can be approximated sufficiently by the "model" chosen. The Markov estimator requires as well the knowledge of the co-variance matrix of the noise.

For the maximum likelihood estimator, one needs to know the probability density function of the stochastic process from which the sample values are taken.

The Bayes', minimum cost- or minimum risk procedure requires a knowledge of the a priori probability density functions of the unknown parameters, and of the costs of making errors. Properties such as efficiency, asymptotic efficiency, consistency, bias etc. are indicated.

Starting from the Bayes' estimator one can derive the other estimators as special cases when there is less a priori knowledge available. In this way, relations between the different estimators are derived in this chapter.
In many cases, for example, in the implicit or model-adjustment methods, the estimation techniques really are optimization methods. Such an optimization may be the search for a minimum of a cost function, a maximum probability, a minimum squared error, etc.

For that reason optimum seeking or convergence schemes are discussed in this chapter as well. These schemes can be categorized as deterministic or stochastic, depending on the absence or presence of stochastic disturbances.
5.1 Estimation theory

The cases to be considered can be illustrated by fig. 5.1. In this section only the central ideas are summarized; a more detailed discussion is given in chapter II. Here we restrict the attention to parameter estimation.

We want to derive an estimator, i.e. a relationship

$$\beta = \beta \{ u(t), \ldots, u(k); y(t), \ldots, y(k) \} = \beta (u, y)$$  \hspace{1cm} (5.1)

such that a numerical value (estimate) can be assigned for the process parameter $b$ under consideration. This estimate depends on the sequences of observations. When a number of estimates $\{ \beta_0, \ldots, \beta_n \} = \beta'$ have to be found for the process parameters $b$, then the relationship is indicated by the vector expression:

$$\beta = \beta (u, y)$$

**Characteristics of estimators.** While estimating we are interested in the development of our knowledge about the parameters of interest. That knowledge can be expressed in terms of the probability density function $p(\beta; T)$ or $p(\beta; k)$. This function depends on the length of the observation interval $T$ or, equivalently, on the number $k$ of signal samples processed. This is indicated in fig. 5.2 for a one-dimensional case; $T_1 < T_2 < T_3$. This probability density function is the most complete type of knowledge that can be derived through application of statistical techniques. Working with this particular type of knowledge, however, is quite complicated and impractical, particularly if the dimensionality of $\beta$ is higher than two and a multi-dimensional presentation would be needed to give the results to the experimenter. For that reason, in most cases the interest is reduced from the complete probability density function to its most-significant characteristics:

* If necessary the reader will find an interpretation of probabilistic notions in appendix B.
expected value \( \mathbb{E}[\beta] \)

bias \( \mathbb{E}[\beta] - \beta \)

covariance \( \text{cov}[\beta] = \mathbb{E}[\{\beta - \mathbb{E}[\beta]\}\{\beta - \mathbb{E}[\beta]\}'] \)

The reader will appreciate this simplification from functions to numerical constants by application of this reduction to the examples of figures 5.2 a and b. In this way he will also get an idea about the "information" that is preserved by this simplification. Note that if the probability density function were (multivariate) Gaussian then no "information" is lost by restricting to \( \mathbb{E}[\beta] \) and \( \text{cov}[\beta] \), as Gaussian functions are completely characterized by their first and second moments.

In statistical literature, e.g., Deutsch (1965), Nahi (1969), some desirable properties are defined for the estimators \( \hat{\beta} \) of the parameters \( \beta \):

- unbiased estimator if for each \( k \):
  \[\mathbb{E}[\hat{\beta}] = \beta\]

- consistent estimator if for increasing \( k \):
  \[\lim_{k \to \infty} P\{|\beta - \hat{\beta}| > \varepsilon\} = 0\]

with \( \varepsilon \) arbitrary small; i.e., \( \hat{\beta} \) converges in probability (or stochastically) to the true value \( \beta \).
efficient estimator if for all unbiased estimators \( \hat{y} \)

\[
\text{cov}[\hat{\beta}] = \mathbb{E}[\{\hat{\beta} - \beta\}\{\hat{\beta} - \beta\}'] \leq \mathbb{E}[\{y - \beta\}\{y - \beta\}'] = \text{cov}[y]
\]
or

\[
\det[\text{cov}[y] - \text{cov}[\hat{\beta}]] \geq 0
\]

- sufficient estimator if for all other estimators \( \hat{y} \) it is true that \( p(y|\hat{\beta}) \)

is not dependent on \( \beta \).

If the first and third property hold only for \( k \to \infty \) then they are called asymptotic unbiasedness and efficiency. Referring again to fig. 5.2 we note that a) suggests an unbiased case and b) a biased, but possibly asymptotically unbiased, case.

Note that another important characteristic of these ways of description is, whether or not they provide us with a possibility of introducing a priori knowledge which might be available from physical insight or from prior measurements, including our "confidence" in this knowledge.

Bayes' estimators. As a starting point for deriving estimators, we will choose a situation where much a priori knowledge is available (Maslov, 1963), viz.:

a) the probability density function of the noise \( n \). From this function the probability density of the measurements \( y \) follows; this function is dependent on the process parameters \( \beta \), and is denoted as \( p(y|\beta) \). It is understood, of course, that \( p(y) \) also depends on \( u \).

b) the probability density functions of the parameter values \( \beta \). This function is written as \( q(\beta) \).

c) the cost of choosing the value \( \hat{\beta} \) for the estimate if the true value of the process parameters is \( \beta \). This cost or loss function \( C(\beta, \hat{\beta}) \) has a minimum for \( \hat{\beta} = \beta \).

After a discussion of this most elaborate case we will consider successively the effect of dropping the assumptions c), b) and a).
Bayes' estimator. The essence of the method is centered around Bayes' rule:

\[
p(y | b) q(b) = p(y, b) = p(b | y) p(y)
\]

For our purpose we consider

\[
p(b | y) = \frac{p(y, b)}{p(y)} = \frac{p(y | b) q(b)}{p(y)} \tag{5.2}
\]

where

\[
p^2(y) = \int_{\mathbb{R}^{m+1}} p(y, b) \, db
\]

indicates an \((m+1)\) fold integral

\[
d^{m+1}b
\]

indicates \(db_0 \, db_1 \ldots \, db_m\)

Consequently, all terms in the right hand part are known from our a priori knowledge.

This conditional probability function \(p(b | y)\) can be interpreted as:

the (a posteriori) probability density function of the parameters \(b\), given the results of the measurements on \(y\), say \(y = c\).

It may be helpful to illustrate these ideas with an example of utmost simplicity. Consider the situation of fig. 5.3 where an estimate \(\beta\) has to be found for parameter \(b\). This estimate has to be derived from a number of signal samples \(u(i)\) and \(y(i)\) where

\[
y(i) = b \, u(i) + n(i)
\]

and where the average (expected) value of \(n(i)\) is zero.

Let the probability density functions of \(n\) and \(b\) be given by \(p_n\) and \(q\).

If \(n\) and \(b\) are statistically independent then

\[
p(y, b) = p(y - ub, b) \left| \frac{\partial(y - ub)}{\partial y} \right| = p_n(y - ub) \, q(b)
\]
This is indicated in fig. 5.4. Note that this is the *a priori* knowledge, available before the measurements of \( u(i) \) and \( y(i) \) are made. After these measurements are made they provide us with a "cut" through the probability-density function, from which the *a posteriori* probability function for \( b \) follows. This new probability function now may be used as the a priori knowledge for the following measurement. In this way the development of \( q(b) \) with increasing number of observations can be followed; cf. fig. 5.5. Note that \( p_n \) does not change as the additive noise is assumed stationary.

The reader is invited to consider special cases such as \( u(k) = 0 \) and \( u(k) \to \infty \). Again the method can be generalized to more parameters; its visualization has severe limitations, however. Note that the knowledge on \( b \) is given in terms a function.

Based on the a posteriori probability density function \( p(b|y) \) one has to decide on what value will be assigned as the estimate. Application of a cost-function \( C(\beta, b) \) is illustrated in fig. 5.6 for a one dimensional case and a quadratic function \( C(\beta, b) = (\beta - b)^2 \). If \( \beta_1 \) should be chosen as the estimate, then the expected cost or loss will be

\[
\int_{-\infty}^{+\infty} (\beta_1 - b)^2 \ p(b|y) \ db \quad \text{with } y = \mathcal{C}
\]

(5.3)

This is a scalar, and can be plotted as a function of \( \beta \). This is shown in fig. 5.6. Consequently, then the best estimate is that value \( \hat{\beta} \) for which this function has a minimum value.

A similar line of thought proceeds in the following manner. The *conditional risk*
of choosing $\beta(y)$ if the true process parameter value is $b$ can be written as the expectation $\Pi$ of the cost function, with respect to the observations $y$:

$$
\Pi \text{ Note that:}
$$

\begin{align*}
\mathcal{E}_y \left[ \ldots \right] &= \int \ldots p(y) \, dk_y \\
\mathcal{E}_y | b \left[ \ldots \right] &= \int \ldots p(y | b) \, dk_y \\
\text{and} \quad \int p(y | b) \, dk_y &= 1
\end{align*}

\begin{align*}
\mathcal{E}_y | b \left[ C(\beta, b) \right] &= \int C(\beta, b) \, p(y | b) \, dk_y \\
\text{(5.4)}
\end{align*}

The average risk for this estimating situation is the expectation with respect to the probability of the value of the process parameter $b$

$$
R(\beta) = \mathcal{E}_b \left[ \mathcal{E}_y | b \left[ C(\beta, b) \right] \right] = \int \int C(\beta, b) \, p(y | b) \, q(b) \, dk_y \, d^m b \\
\text{(5.5)}
$$

The estimate that minimizes this expression is called the minimum risk- or minimum cost estimate. Because of the Bayes' relationships, this average risk can be written as:

$$
R(\beta) = \int \left. \frac{\partial}{\partial \beta} \int C(\beta, b) \, p(y | b) \, d^m b \right|_{\beta = \hat{\beta}} \\
\text{(5.6)}
$$

As $p(y) \neq 0$ the average risk $R(\beta)$ can be minimized by making the second integral as small as possible for the observed $y = c$:

$$
\min_{\beta} \int C(\beta, b) \, p(b | c) \, d^m b \\
\text{(5.7)}
$$

A necessary condition for such a minimum is simply:

$$
\frac{\partial}{\partial \beta} \int C(\beta, b) \, p(b | c) \, d^m b \bigg|_{\beta = \hat{\beta}} = 0
$$

Due to the use made of Bayes' theorem, the resulting estimate is also called Bayes' estimate.

As a simple illustration consider again the case where we want to estimate one parameter $\beta$ using a quadratic cost function $C(\beta, b) = (\beta - b)^2$. Then eq. (5.7) results in

$$
\hat{\beta} = \frac{\int b \, p(b | c) \, db}{\int p(b | c) \, db} = \int b \, p(b | c) \, db = \mathcal{E}_{b | c}[b] \\
\text{(5.8)}
$$
i.e., the optimum estimate is the (conditional) expected value of \( \beta \).

As a suitable cost function \( C(\beta, b) \) has a minimum for \( \beta = b \), as it presumably has small values for values of \( \beta \) in the vicinity of \( b \) and as

\[
\int_{-\infty}^{\infty} p(b|y) d\beta = 1
\]

it will be clear that eq. (5.7) is satisfied if \( \beta \) is chosen in the neighbourhood of that \( b \) where the conditional probability \( p(b|y) \) is maximum; cf. fig. 5.6.

Now we will drop the assumption c), i.e. the knowledge about an adequate cost or loss function \( C(\beta, b) \). In that case it is reasonable to choose \( \beta \) at that value of \( b \) for which \( p(b|y) \) is a maximum. As according to Bayes' rule:

\[
p(b|y) = \frac{p(y|b) q(b)}{p(y)}
\]

with

\[
p(y) = \int_{-\infty}^{\infty} p(y|b) q(b) d\beta
\]

both assumptions a) and b) are still being used.

**Maximum likelihood estimators.** Next we will consider the consequences of dropping both assumptions b) and c); the a priori probability \( q(b) \) of the process parameter \( b \) is also unknown. This ignorance can be expressed by assuming a uniform distribution \( q(b) = \text{const.} \) over the interval of consideration. In that case for any \( y \):

\[
\max_{b} p(b|y) = \frac{\text{const.}}{p(y)} \max_{b} p(y; b)
\]

Now \( b \) is no longer a random variable but an unknown constant parameter; this is expressed by a semicolon instead of a bar in the argument of \( p \).

The probability density function (a priori) is given by

\[
p(\{y(1), \ldots, y(k); b \})
\]
A posteriori, after the measurement with the observed sample values
\( y(1) = c_1, \ldots, y(k) = c_k \) or \( y = \mathbf{c} \) this function is noted as
\[
\mathcal{L}\{ c_1, \ldots, c_k ; \beta \}
\]
the likelihood function; cf. section 2.3.

Again the simple example of fig. 5.3 can be used as an illustration. We have
to know \( p_n \), the probability density function of \( n(i) \). In that case the measure-
ment of \( u(i) \) provides us with the knowledge sketched in fig. 5.7, i.e., the
a priori probability density of \( y(i) \) with \( b \) as parameter. Now the measurement
\( y(i) = c_i \) brings us to the situation indicated in fig. 5.8, the a posteriori
knowledge. From the likelihood function we have to assign an estimate \( \hat{\beta} \). A
reasonable and popular choice is to take that value \( \hat{\beta} \), for which \( L \) has its
maximal value.

For the more-dimensional case
the following necessary-condition for finding the maximum can be given:
\[
\frac{\partial}{\partial \beta} \mathcal{L}( \mathbf{c} ; \beta ) \bigg|_{\beta = \hat{\beta}} = 0
\]
(5.10a)
or as the logarithmic function is monotonic
\[
\frac{\partial}{\partial \beta} \ln \mathcal{L}( \mathbf{c} ; \beta ) \bigg|_{\beta = \hat{\beta}} = 0
\]
(5.10b)
This expression is called the likelihood equation. Picking the root of
this set of equations which yields the largest value for \( \mathcal{L}( \mathbf{c} ; \beta ) \) or for
\( \ln \mathcal{L}( \mathbf{c} ; \beta ) \) we are given the celebrated maximum likelihood estimate (M.L.E.)
The M.L.E. has been extensively discussed in literature. Some of its interesting properties are the following:

- **asymptotic normality**: i.e., \( p(\hat{\beta};b) \) approaches a normal distribution for \( k \to \infty \); for the meaning of \( k \) cf. eq. (5.1)

- **asymptotic unbiasedness**: i.e., \( \mathbb{E}[\hat{\beta}] = b \) for \( k \to \infty \)

- **asymptotic efficiency**: i.e., approaching the best accuracy or minimum variance as given by the Cramer-Rao (in)equality cf. chapter II.

- **consistency**: as defined before

- **invariance**: i.e., if \( \hat{\beta} \) is a M.L.E. of \( b \) then \( g(\hat{\beta}) \) is a M.L.E. of \( g(b) \)

Relations between the different estimators. It remains to be shown how the Markov and least squares estimators, mentioned in section 2.2 and discussed in chapter 6, are related to those derived in this section. Let us refer to fig. 5.9 and assume that the components of \( \mathbf{n} \) have a k-variate
Gaussian (normal) distribution i.e.,
\[ p(n) = \frac{1}{(2\pi)^{k/2} |\mathbf{N}|^{1/2}} \exp \left[ -\frac{1}{2} (\mathbf{n}^T \mathbf{N}^{-1} \mathbf{n}) \right] \]  
(5.11)
with
\[ \mathbb{E}[n] = 0 \quad \text{and} \quad \mathbb{E}[nn^T] = \mathbf{N} \]
Note that \( n \) is the only stochastic part of \( e = y - \hat{w} = y - \mathbf{U}\beta \).
Consequently we can write for the logarithm of the probability density function of \( e \):
\[ \ln p(y - \mathbf{U}\beta) = -\frac{1}{2} \ln \left\{ (2\pi)^k |\mathbf{N}| \right\} - \frac{1}{2} \left\{ (y - \mathbf{U}\beta)^T \mathbf{N}^{-1} (y - \mathbf{U}\beta) \right\} \]
Maximizing this function leads to:
\[ \frac{\partial}{\partial \beta} (y - \mathbf{U}\beta)^T \mathbf{N}^{-1} (y - \mathbf{U}\beta) \bigg|_{\beta = \hat{\beta}} = 0 \]  
(5.12)
or
\[ \mathbf{U}^T \mathbf{N}^{-1} \mathbf{U} \hat{\beta} - \mathbf{U}^T \mathbf{N}^{-1} y = 0 \]  
(5.13)
or if \( \mathbf{U}^T \mathbf{N}^{-1} \mathbf{U} \) has an inverse then eq. (5.13) can be written
\[ \hat{\beta} = \left( \mathbf{U}^T \mathbf{N}^{-1} \mathbf{U} \right)^{-1} \mathbf{U}^T \mathbf{N}^{-1} y \]  
(5.14)
This is the expression for the Markov estimate. It has the following properties:
- linearity : i.e. \( \hat{\beta} = Qy \)
- unbiasedness : i.e. \( \mathbb{E}[\hat{\beta}] = \mathbf{b} \)
- minimum variance : of all linear unbiased estimates. This variance follows from:
\[ \text{Cov} [\hat{\beta}] = \mathbb{E}[(\hat{\beta} - \mathbf{b})(\hat{\beta} - \mathbf{b})^T] \]
If knowledge of the covariance matrix of the noise is also lacking it is appropriate to choose \( \mathbf{N}^{-1} = \sigma^2 \mathbf{I} \), where \( \sigma^2 \) is the variance of the noise. This implies the assumption that the noise is white.
Consequently:
\[ \hat{\beta} = \left( \mathbf{U}^T \mathbf{U} \right)^{-1} \mathbf{U}^T y \]  
(5.15)
This is the expression for the **least squares estimator**.

The Markov and least squares estimators have been "derived" from the maximum likelihood estimate under the assumption of Gaussian noise. (This assumption of a Gaussian distribution is not necessary. In chapter 6 this estimator is derived irrespective of the Gaussian assumption). Such a derivation can be illustrated by using again the simple example of fig. 5.3.

Employing the least squares method, the estimate is chosen in such a way that the loss function, defined as

$$E(\beta) = \sum_{i=1}^{k} \{ y(i) - u(i)\beta \}^2 = [Y - u\beta]'[Y - u\beta]$$

is a minimum.

In fig. 5.10 the differences between the observations $y$ and the "model outputs" $u\beta$ are indicated. The minimization can be pursued analytically; a necessary condition for the minimum is

$$\frac{d}{d\beta} \sum_{i=1}^{k} \{ y(i) - u(i)\beta \}^2 \bigg|_{\beta=\hat{\beta}} = 0$$

or

$$\sum_{i=1}^{k} u(i) \{ y(i) - u(i)\hat{\beta} \} = 0$$

(5.16)

or

$$\hat{\beta} = \sum_{i=1}^{k} \frac{u(i) y(i)}{u(i)^2}$$

(5.17)

$\hat{\beta}$ is the optimal estimate under the conditions given. Note from (5.16) that the terms $y(i) - u(i)\beta$ are weighted with respect to $u(i)$; quite naturally the larger the input signal, the more importance is assigned to the deviation between observation $y(i)$ and "prediction" $u(i)\beta$. Equation (5.17) refers to the correlation methods.

Some other aspects. The simple example of fig. 5.3 also serves very well to illustrate the so-called "problem of input noise". Consider the system illustrated by the block diagram in fig. 5.11 where neither the input nor the output can be observed exactly.
It is well-known in statistics (see e.g. Lindley, 1965) that, unless specific assumptions are made concerning the variations in \( u, n \) and \( v \) it is not possible to estimate the dynamics of the process.

Consider, for example,
\[
y(i) = b u(i) + n(i)
\]
\[
u(i) = u(i) + v(i)
\]
Assume that \( v(i) \) and \( n(i) \) are independent stochastic variables with zero mean values. If \( v(i)=0 \) and \( n(k)=0 \) we find that the estimate of \( b \) as given by (5.17) is
\[
\hat{\beta} = \frac{\sum_{i=1}^{k} u(i) y(i)}{\sum_{i=1}^{k} u(i) \tilde{u}(i)}
\]
However, if \( n(k)=0 \) and \( v(k) \neq 0 \) we find by the same argument that the estimate of \( b \) is
\[
\hat{\beta} = \frac{\sum_{i=1}^{k} y(i) y(i)}{\sum_{i=1}^{k} y(i) \tilde{u}(i)}
\]
This corresponds to choosing \( \beta \) such that the difference between the observations \( \tilde{u}(i) \) and the predictions \( \frac{1}{\beta} y(i) \) are as small as possible in the least squares sense. See fig. 5.12. Without additional information it is of course impossible to tell which estimate to choose.

From the expression (5.17) for the least squares estimator, which provides an explicit mathematical or accumulative solution, it is quite simple to derive a model-adjustment or iteration solution
\[
\left\{ \sum_{i=1}^{k+1} u'(i) \right\} \hat{\beta}(k+1) = \sum_{i=1}^{k+1} u(i) y(i)
\]
\[
\left\{ \sum_{i=1}^{k} u'(i) \right\} \hat{\beta}(k) = \sum_{i=1}^{k} u(i) y(i)
\]
These expression for the estimates \( \hat{\beta}(k) \) and \( \hat{\beta}(k+1) \) using \( k \), viz. \( k+1 \) (pairs of) observations can be written as:

\[
\left\{ \sum_{i=1}^{k+1} u^2(i) \right\} \hat{\beta}(k+1) = \sum_{i=1}^{k} u(i) y(i) + u(k+1) y(k+1)
\]

\[
\left\{ \sum_{i=1}^{k+1} u^2(i) \right\} \hat{\beta}(k) = \sum_{i=1}^{k} u(i) y(i) + u^2(k+1) \hat{\beta}(k)
\]

\[
\left\{ \sum_{i=1}^{k+1} u^2(i) \right\} \left\{ \hat{\beta}(k+1) - \hat{\beta}(k) \right\} = u(k+1) \left\{ y(k+1) - u(k+1) \hat{\beta}(k) \right\}
\]

or

\[
\hat{\beta}(k+1) = \hat{\beta}(k) + \left\{ \sum_{i=1}^{k+1} u^2(i) \right\}^{-1} u(k+1) \left\{ y(k+1) - u(k+1) \hat{\beta}(k) \right\}
\]

The last term, \( e(k+1) \), is the error between the observation \( y(k+1) \) and the model-output \( u(k+1) \hat{\beta}(k) \). The new estimate consists of the old estimate, corrected with a proportion of the error \( e(k+1) \). What "weight" is given to the error depends on the magnitude of \( u(k+1) \) and on the sum of squares of all previous input samples ("energy" measure of the input signal).

Some of the relations between the different types of estimates as derived in this section are summarized in table 5.1. This discussion has been restricted to sampled signals. By increasing the number of samples and decreasing the sampling interval the corresponding expressions can be derived for continuous signals.
<table>
<thead>
<tr>
<th>Bayes estimation</th>
<th>unconditional maximum likelihood estimation</th>
<th>conditional maximum likelihood estimation</th>
<th>Markov estimation</th>
<th>least squares estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a priori knowledge</td>
<td>p(y</td>
<td>b)</td>
<td>p(y</td>
<td>b)</td>
</tr>
<tr>
<td>q(b)</td>
<td>q(b)</td>
<td>q(b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(\theta, b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Conditions | \min \int C(\theta, b)p(b|y)db | \max p(b|y) = \frac{\max p(y|b)q(b)}{\max \ln p(y|b)} | \max p(\theta|b) or | \hat{\theta} = (U'N^{-1}U)^{-1}U'N^{-1}y | \hat{\theta} = (U'U)^{-1}U'y |
|------------------|------------------------------------------|------------------------------------------|------------------|-------------------------|
| observations y = c | same | same | same | same | observations y = c |

<table>
<thead>
<tr>
<th>Properties</th>
<th>min. risk or cost</th>
<th>asympt. normality</th>
<th>linearity</th>
<th>linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td>unbiasedness</td>
<td>unbiasedness</td>
<td>efficiency</td>
<td>(min. variance)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relations</th>
<th>cf. text</th>
<th>if gaussian noise</th>
<th>if white noise</th>
</tr>
</thead>
</table>

Table 5.1
5.2 Deterministic convergence schemes.

As pointed out in chapter 2, the estimation problem can be formulated as a minimization of an error defined according to some criterion. This error criterion necessarily has to be such that both positive and negative errors contribute in positive sense to the criterion value. Loosely speaking this was called an "even-function error criterion". For example:

Let the criterion be given as a minimization of the loss function

\[
E \{ \beta \} = \frac{1}{2} \int_0^T (\mathbf{y} - \mathbf{w})' \mathbf{S} (\mathbf{y} - \mathbf{w}) \, dt
\]  

(5.18)

where

\( \mathbf{y} = \mathbf{y}(\mathbf{b},t) \) is the process output and
\( \mathbf{w} = \mathbf{w}(\mathbf{b},t) \) is the corresponding model output.

\( \mathbf{S} \) is a positive definite matrix, weighing the relative reliabilities of \( y_i \)'s.

In model-adjustment approaches for parameter estimation, just as in cases of optimal and self-optimizing control, the extremum of a function or functional has to be found.

Assume the error criterion to be a minimization of:

\[
E = E (\beta_0, \beta_1, \ldots, \beta_m) = E (\beta)
\]  

(5.19)

where \( \beta_j \) are the model-parameters to be adjusted. In chapter 2 some methods have been indicated to determine the partial derivatives with respect to these parameters. These derivatives can be used in a variety of ways. A rather general representation follows from the following reasoning.
The error function given by eq. (5.19) can be considered as a hyper-surface defined by the equation:

$$ E(\beta) - E_0 = 0 $$  \hspace{1cm} (5.20)

where $E_0$ is the minimum of the error function, that we want to find:

$$ E_0 = E|_{\beta = b} $$

The error function can be expressed as a Taylor expansion around the minimum:

$$ E(\beta) = E_0 + \frac{\partial E}{\partial \beta} \bigg|_{\beta = b} (\beta - b) + \frac{1}{2} (\beta - b)' \frac{\partial^2 E}{\partial \beta^2} \bigg|_{\beta = b} (\beta - b) + $$

$$ + \text{higher order terms} \quad (5.21a)$$

Because $E$ has a minimum in $\beta = b$, the following holds:

$$ \frac{\partial E}{\partial \beta} \bigg|_{\beta = b} = 0 \quad \text{and} \quad \frac{\partial^2 E}{\partial \beta^2} \bigg|_{\beta = b} \quad \text{is positive definite.} $$

Moreover in most cases the higher order terms are neglectable (if not, cf. method h mentioned below). Consequently,

$$ E(\beta) \approx E_0 + \frac{1}{2} (\beta - b)' \frac{\partial^2 E}{\partial \beta^2} \bigg|_{\beta = b} (\beta - b) \quad (5.21b)$$

After substituting (5.21b), equation (5.20) now represents a hyperparaboloid; cf. fig. 5.13.

For the convergence study of the adjustment schemes that we will discuss, some knowledge about the iso-criterion curves $E = \text{constant}$ is important. In the above quadratic case these are all hyperellipsoids. Before discussing details of the adjustment schemes, an example of a quadratic error function is given.

For a simple case a rough approximation leads to the following description:

Assume a process describable by the differential equation
\[(a_n D^n + \ldots + a_1 D + 1)y(t) = (b_m D^m + \ldots + b_1 D + b_0)u(t) \quad (5.22)\]

with \(D = \frac{d}{dt}\). A model is given by the equation:

\[(\alpha_n D^n + \ldots + \alpha_1 D + 1)w(t) = (\beta_m D^m + \ldots + \beta_1 D + \beta_0)u(t) \quad (5.23)\]

These equations can be written explicitly as \(y(t) = \ldots\) and \(w(t) = \ldots\).

In the vicinity of the optimum \(e(t) = y(t) - w(t)\) is small, if no noise is present. If \(e(t)\) is sufficiently small then \(w(t)\) can be approximated by:

\[\tilde{w}(t) = -(\alpha_n D^n + \ldots + \alpha_1 D) y(t) + (\beta_m D^m + \ldots + \beta_1 D + \beta_0) u(t) \quad (5.24)\]

Consequently,

\[e(t) \approx y(t) - \tilde{w}(t) = \left\{ (\alpha_n - \alpha_1) D^n + \ldots + (\alpha_1 - \alpha_1) D \right\} y(t) + \]

\[- \left\{ (\beta_m - b_m) D^m + \ldots + (\beta_1 - b_1) D + (\beta_0 - b_0) \right\} u(t) \quad (5.25)\]

In the neighbourhood of the optimum \(e(t)\) is approximately a linear function of \(\alpha\) and \(\beta\). By renaming \(\tau\) in practice differentials are difficult to obtain. However, an analogous scheme can be derived for integrals or semi-integrals (using \(\frac{1}{s + \sigma}\) instead of \(\frac{1}{s}\), where \(s\) is Laplace operator); cf. chapter 9.

\[\alpha_1 = a_1 = \gamma_1 \]
\[\vdots\]
\[\alpha_n = a_n = \gamma_n \]
\[\beta_0 = b_0 = \gamma_{n+1} \]
\[\vdots\]
\[\beta_m = b_m = \gamma_{n+m+1} \]

\[
\begin{align*}
D y(t) &= v_1(t) \\
D^2 y(t) &= v_2(t) \\
\vdots \\
D^n y(t) &= v_n(t) \\
D u(t) &= v_{n+1}(t) \\
D^2 u(t) &= v_{n+2}(t) \\
\vdots \\
D^m u(t) &= v_{n+m+1}(t)
\end{align*}
\]
we find the simple expressions:

\[ e(t) \propto \sum_{i=1}^{N} j_i v_i \]

\[ e^2(t) \propto \sum_{i}^{N} \sum_{j}^{N} j_i j_i' v_i v_j \]

\[ E = \int_{t-T}^{t} e^2(t) dt = \sum_{i}^{N} \sum_{j}^{N} j_i j_i' r_{ij} \]  \hspace{1cm} (5.26)

with

\[ r_{ij} = \int_{t-T}^{t} v_i(t) v_j(t) dt = r_{ji} \]

This also can be written as:

\[ E = \mathbf{y}' \mathbf{R} \mathbf{y} \]  \hspace{1cm} (5.27)

with

\[ \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \]

\[ \mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{1N} \\ \vdots & \ddots & \vdots \\ r_{N1} & \cdots & r_{NN} \end{pmatrix} \]

Comparison of equation (5.27) with (5.21b) shows, that they are equivalent for \( E_0 = 0 \) and

\[ \mathbf{R} = \frac{1}{2} \frac{\partial^2 E}{\partial \beta \partial \beta'} \bigg|_{\beta = b} \]

The adjustment schemes can be distinguished as direct search methods and gradient methods. The first class does not require an explicit evaluation of the gradient of the error- or lossfunction, whereas the second class is based on the assumption that the gradient can be determined. For surveys of optimization techniques cf. Box, Davies and Swann (1969), Wilde (1964), Bekey and McGee (1964), Kowalik and Osborne (1968).

The direct search methods can be divided into two subclasses:
α. Tabulating methods. Evaluating the error function in several points spread out in the parameter space e.g., at the junctions of a regular grid or randomly distributed.

β. Exploratory methods. Those algorithms follow implicitly the gradient path by stepwise looking for a descending direction. An example is the simplex method starting with n points at equal distances in the n-dimensional parameter space. Each iteration is made by "mirroring" the point with the largest function value on to that new point, which has also the same distances to the other ones. After some iterations the distances have to be diminished in order to prevent finding the same points again.

Another approach is the creeping random search method, where, starting from the last iteration point, a random direction is examined for smaller function values. In these methods there is a great variety of combinations between earlier iterations in the parameters. A survey can be found in White (1970). The creeping random search methods are especially important when the number of parameters is greater then about twenty.

The gradient methods can be either

- continuous adjustment schemes or
- intermittent adjustment schemes

where intermittent refers to the alternation of the measurement- and the model adjustment periods.

As a continuous adjustment technique the method of steepest descent is applied most often. In this case the model will be taken along a trajectory which, for a certain speed of adjustment, gives the most rapid reduction of the error. It turns out that this trajectory is at any point normal to the curves for constant E. For example, in the two dimensional case E = constant implies:
The equation for the tangent to the curve at the point \( P(\beta_1^p, \beta_2^p) \) is given by:

\[
\frac{dE}{\beta_1} \frac{d\beta_1}{\beta_1} + \frac{dE}{\beta_2} \frac{d\beta_2}{\beta_2} = 0
\]  

(5.28)

The vector

\[
\left( \frac{\partial E}{\partial \beta_1}, \frac{\partial E}{\partial \beta_2} \right)
\]

is perpendicular to the tangent and consequently also to the curve. For the steepest descent method the parameters are adjusted according to the equation

\[
\left( \begin{array}{c} \dot{\beta}_1 \\ \dot{\beta}_2 \end{array} \right) = -\Gamma \left( \begin{array}{c} \frac{\partial E}{\partial \beta_1} \\ \frac{\partial E}{\partial \beta_2} \end{array} \right) = -\Gamma \frac{\partial E}{\partial \beta}
\]  

(5.30)

Thus, \( P \) will follow a trajectory which is perpendicular everywhere to the curves \( E = \text{constant} \) cf. fig. 5.14.

Here \( \Gamma \) is a constant which, together with the partial derivatives, determines the rate of variation of the parameters. The "measurement" of the partial derivatives is not instantaneous and, therefore, a limitation on the rate of variation must be maintained. If too high a gain \( \Gamma \) is used, the measurement of the derivatives will degenerate and, consequently, a steepest descent trajectory will not be followed. For a still higher gain,
the system even may turn unstable. For arbitrary systems there is no generally applicable technique yet for determining stability, because our system is both time varying and nonlinear. For simple systems, however, the technique developed by Liapunov can be applied successfully. For a low gain, the approach to the optimum will be slow.

The intermittent adjustment schemes can be such that:
- the parameters are adjusted one by one
- all parameters are adjusted simultaneously.

a. Cyclic adjustments of the parameters one by one. Say that parameter $\beta_1$ is first adjusted. One possibility is then to let the adjustment system give $\beta_1$ such a value that $\frac{\partial E}{\partial \beta_1}$ is equal to zero. The same procedure will then be repeated for all parameters of the model. If necessary more cycles have to be made to obtain a satisfactory model match. In the two dimensional case with parameters $\beta_1$ and $\beta_2$ the $(\beta_1, \beta_2)$ - plane can be mapped by curves for constant $E$. As pointed out, near the optimum these will in general be concentric ellipses, the principal axes of which may be of arbitrary orientation; cf. fig. 5.14. Therefore, we must expect interaction between the model parameters and generally more cycles are needed to bring the model to the neighbourhood of the optimum. If the system has been made orthogonal through a transformation then one adjustment for each parameter may be sufficient. The problem, however, is that we do not know exactly which linear combination to use for orthogonalization. A feature of the ellipses is the angle between the lines, which represents the geometrical position of all points satisfying either

$$\frac{\partial E}{\partial \beta_1} = 0 \quad \text{or} \quad \frac{\partial E}{\partial \beta_2} = 0$$

(5.31)
The smaller this angle the more cycles are necessary to obtain a given accuracy. In fig. 5.14 a possible adjustment path is sketched (Q). If the iso-criterion curves are not such simple ellipses but a "ridge", then this method has the disadvantage that the adjustment may stop at a value \( E > E_o \). All other adjustment schemes considered will be with simultaneous parameter adjustments. For a more elaborate discussion of some of these schemes cf. Rosenbrock and Storey (1966).

b. **Steepest descent.** The adjustments are made according to the algorithm

\[
\beta(i+1) = \beta(i) - \Gamma \left. \frac{\partial E}{\partial \beta} \right|_{\beta=\beta(i)} \quad \Gamma > 0
\]

where \( \Gamma \) is a constant. Depending on the value chosen for \( \Gamma \) we obtain a slow and smooth approach to the optimum (\( \Gamma \) small, cf. path \( P_1 \) in fig. 5.15) or an oscillatory one (path \( P_2 \)). If the gradient cannot be calculated then it has to be determined by "experiments", i.e. by (small) excursions around \( \beta(i) \).

c. **Steepest descent with minimization along a line.** This is governed by the algorithm

\[
\beta(i+1) = \beta(i) - \Gamma(i) \left. \frac{\partial E}{\partial \beta} \right|_{\beta=\beta(i)} \quad \Gamma(i) > 0
\]

Now \( \Gamma(i) \) is chosen so that the system moves along the gradient direction until in that direction the minimum is reached. After that a new gradient-determination is performed, etc. Path Q in fig. 5.15 shows an example of such an adjustment. Note that necessarily the successive adjustment paths are orthogonal. Near the minimum the convergence can be slow.

It makes sense to ask under what condition one would obtain a one-step procedure to the optimum, assuming elliptic curves. For that case, it
follows from eq. (5.26) in the two-dimensional case

\[ E = r_{11} \gamma_1^2 + 2 r_{12} \gamma_1 \gamma_2 + r_{22} \gamma_2^2 \] (5.32)

\[ \frac{\partial E}{\partial \beta_i} = \frac{\partial E}{\partial \gamma_i} \quad \text{as} \quad \gamma_i = \beta_i - b_i \]

that

\[
\begin{cases}
\frac{\partial E}{\partial \gamma_1} = 2 r_{11} \gamma_1 + 2 r_{12} \gamma_2 \approx 2 r_{11} \gamma_1 & \text{if} \quad r_2 \ll r_{11} \\
\frac{\partial E}{\partial \gamma_2} = 2 r_{12} \gamma_1 + 2 r_{22} \gamma_2 \approx 2 r_{22} \gamma_2 & \text{if} \quad r_2 \ll r_{22}
\end{cases}
\]

For a one-step approach to hold it is necessary that

\[
\begin{align*}
-\gamma_1 &= \Delta \gamma_1 = - r_{11} \frac{\partial E}{\partial \gamma_1} \approx - 2 r_{11} \gamma_1 & \text{or} & \\
-\gamma_2 &= \Delta \gamma_2 = - r_{22} \frac{\partial E}{\partial \gamma_2} \approx - 2 r_{22} \gamma_2 & \text{or}
\end{align*}
\]

\[ r_1 = r_2 = r \text{ is acceptable only if } r_{11} \approx r_{22}, \text{ otherwise the "gain coefficients" have to be adapted to the } r \text{ elements.} \]

Returning to the hypersurface interpretation it is clear, that the assumptions \( r_{12} = r_{21} = 0 \) and \( r_{11} = r_{22} \) involve, that the iso-criterion lines are circles and in general the steepest descent method will be optimal, when the iso-criterion surfaces are hyperspheres. According to formula (5.21b) the iso-criterion surfaces near the minimum are approximately hyperellipsoids, so it is obvious to try to transform these ellipsoids into spheres in another space and to use the steepest descent method in that other space.

Because

\[
\frac{\partial^2 E}{\partial \beta \partial \beta'} \bigg|_{\beta = b}
\]

is symmetric and positive definite, the following holds:

\[
\frac{\partial^2 E}{\partial \beta \partial \beta'} \bigg|_{\beta = b} = S \Lambda S'
\] (5.34)
where $\Lambda$ is a diagonal matrix containing the eigenvalues and the columns of matrix $S$ are the orthonormal eigenvectors. Substitution of (5.34) in (5.21b) leads to

$$E(\z^*) = E_o + \frac{1}{2} \z^* \Lambda \z^*$$

(5.35)

where

$$\z^* = \sum' \left( \beta - b \right)$$

Define

$$\z = \Lambda^{-1/2} \z^*$$

and

$$\z^* = \Lambda^{-1} \z$$

which results in

$$E(\z) = E_o + \frac{1}{2} \z^* \z$$

(5.36)

In the $\z$-space the iso-criterion surfaces are spheres (cf. figure 5.16).

Performance of a steepest descent iteration in the $\z$ space with $\Gamma = 1$:

$$\z(i+1) = \z(i) - \frac{\delta E}{\delta \z} \bigg|_{\z = \z(i)} = \z(i) - \z(i) = 0$$

and because $\beta - b = S \Lambda^{-1} \z$ with $S$ and $\Lambda$ positive definite, $\beta - b = 0$ also holds.

Transformation back to the $\beta$ space yields:

$$\beta(i+1) = \beta(i) - \left[ \frac{\delta^2 E}{\delta \beta \delta \beta} \bigg|_{\beta - b} \right]^{-1} \frac{\delta E}{\delta \beta} \bigg|_{\beta = \beta(i)}$$

(5.37)

Only in purely quadratic error functions the second derivative will be independent of $\beta$; in many other cases, however, it can be assumed to be almost independent of $\beta$, which results in:
d. **Newton-Raphson method.** The algorithm is given by:

\[
\beta(i+1) = \beta(i) - \left[ \frac{\partial^2 E}{\partial \beta \partial \beta'} \right]_{\beta=\beta(i)}^{-1} \frac{\partial E}{\partial \beta} \bigg|_{\beta=\beta(i)} 
\]

(5.38)

This means that a Newton-Raphson iteration is directed to the center of a hyperellipsoid, as it uses the second derivative at \(\beta(i)\). This can also be interpreted as going to the minimum of a hyperparaboloid, which is the second order approximation of the hypersurface given by (5.20).

(cf. figure 5.17).

This can be shown by means of a Taylor expansion around \(\beta(i)\):

\[
E(\beta) = E(\beta(i)) + \frac{\partial E}{\partial \beta} (\beta - \beta(i)) + \frac{1}{2} (\beta - \beta(i))' \frac{\partial^2 E}{\partial \beta \partial \beta'} (\beta - \beta(i)) + \cdots
\]

(5.39)

The minimum \(\beta = \beta(i+1)\) is found by setting \(\frac{\partial E}{\partial \beta} = 0\). Consequently

\[
\left. \frac{\partial E}{\partial \beta} \right|_{\beta=\beta(i+1)} = 0
\]

which results in eq. (5.38).

Evaluation of the second derivative is quite a troublesome job, but some simplification is possible, if \(E\) is a least squares criterion like (5.18).

In that case

\[
\frac{\partial E}{\partial \beta} = - \int_{0}^{T} \frac{\partial w}{\partial \beta} G (y - w) \, dt
\]

\[
\frac{\partial^2 E}{\partial \beta \partial \beta'} = \int_{0}^{T} \frac{\partial^2 w}{\partial \beta^2} G \frac{\partial w}{\partial \beta} \, dt - \int_{0}^{T} \frac{\partial w}{\partial \beta} \frac{\partial^2 w}{\partial \beta \partial \beta'} G (y - w) \, dt
\]

\[
\text{The notation is not correct, as } \frac{\partial^2 w}{\partial \beta \partial \beta'} \text{ is a tensor.}
\]

Close to the minimum \(y \approx w\), so that the second term is neglectible.
Substitution in (5.38) gives us the e. Gauss-Newton method. The algorithm is:

\[ \beta(t+1) = \beta(t) + \left[ \int_0^T \frac{\partial w'}{\partial \beta} \frac{C}{\partial \beta} \frac{d \beta}{\beta = \beta(t)} \right]^{-1} \left[ \int_0^T \frac{\partial w'}{\partial \beta} \frac{C}{\partial \beta} (y - \overline{w}) \frac{d t}{\beta = \beta(t)} \right] \]

(5.40)

It is easy to verify, that this is a simple least squares estimate, if \( w \) is linear in \( \beta \).

The Gauss-Newton method is preferable to the steepest descent on account of the quadratic convergence, although, contrary to the steepest descent method, the convergence is not guaranteed. So a combination of the two methods incorporating both properties would be advisable. The Marquardt method satisfies this condition.

f. Marquardt method. Let the center of a hypersphere in the parameter space be given by \( \beta(t) \). We will search for the minimum on that hypersphere according to Lagrange-minimizing \( E(\beta) \) under the restriction:

\[ \Delta \beta = \beta(t+1) - \beta(t) \]

\[ \Delta \beta \Delta \beta = r^2 = \text{constant} \]

is similar to:

\[ \frac{\partial E}{\partial \Delta \beta} + \mu \Delta \beta = 0 \]

(5.41)

Handling now a Taylor expansion of \( w \) implies that the assumed hypersphere need not be a pure hypersphere:

\[ w(t+1) = w(t) + \frac{\partial w}{\partial \beta} \bigg|_{\beta = \beta(t)} \Delta \beta \]

or

\[ \frac{\partial w}{\partial \beta} \bigg|_{\beta = \beta(t)} \approx \frac{\partial w}{\partial \beta} \bigg|_{\beta = \beta(t)} \]

Consequently, eq. (5.41) will be:

\[ - \int_0^T \frac{\partial w'}{\partial \beta} \bigg|_{\beta = \beta(t)} \frac{C}{\partial \beta} (y - w(t) - \frac{\partial w}{\partial \beta} \bigg|_{\beta = \beta(t)} \Delta \beta) \frac{d t}{\beta = \beta(t)} + \mu \Delta \beta = 0 \]

which implies:

\[ \beta(t+1) = \int_0^T w(t) + \left[ \int_0^T \frac{\partial w'}{\partial \beta} \bigg|_{\beta = \beta(t)} \frac{C}{\partial \beta} \bigg|_{\beta = \beta(t)} \right]^{-1} \left[ \int_0^T \frac{\partial w'}{\partial \beta} \bigg|_{\beta = \beta(t)} (y - w(t)) \frac{d t}{\beta = \beta(t)} \right] \]

(5.42)
Further consideration of this formula shows:

\[ \mu = 0 \Rightarrow \text{pure Gauss-Newton} \]

\[ \mu = \infty \Rightarrow \text{pure steepest descent, step size zero cf. figure 5.18.} \]

By changing \( \mu \) for each iteration one can control convergence properties. A disadvantage is still the evaluation of the (approximated) second derivative of \( E \) and the inversion of this matrix, especially, when the number of parameters is large. It would be advisable to compose this inverted matrix from information obtained during iterations of simpler methods. This leads us to:

g. **Conjugated direction methods.** For the sake of convenience we first observe, that (5.21b) can be written as:

\[ E = E^* + a' \beta + \frac{1}{2} \beta' \mathcal{H} \beta \]

where

\[ \mathcal{H} = \frac{\partial^2 E}{\partial \beta \partial \beta'} \bigg|_{\beta = b} \quad a = - \mathcal{H} b \quad E^* = E_c + \frac{1}{2} b' \mathcal{H} b \]

so

\[ \frac{\partial E}{\partial \beta} \bigg|_{\beta = \beta(i)} = \frac{\partial}{\partial \beta} \beta(i) = a + \mathcal{H} \beta(i) \]

Two directions \( u \) and \( v \) are defined to be conjugate with respect to the positive definite matrix \( \mathcal{A} \) if

\[ u' \mathcal{A} v = 0 \]

(note that orthogonal vectors are conjugate with respect to the identity matrix).

An \( n \times n \) matrix \( \mathcal{A} \) (pos.def.) has at least one set of \( n \) independent vectors that are mutually conjugate, as the eigenvectors of \( \mathcal{A} \) form such a set. So in the parameter space the vector \( v \) from the initial point to the minimum can be built up by \( n \) independent conjugate directions \( d_i \):
Modification of the steepest descent method by following a conjugate direction in stead of the gradient will, as we shall see, guarantee us to reach the minimum in n iterations for a quadratic error function. The iteration formula becomes:

$$\beta(i+1) = \beta(i) + \lambda_i \ d(i)$$  \hspace{1cm} (5.44)

where \(\lambda_i\) is found by line minimization along \(d(i)\), which implies:

$$\frac{d}{dx}(i+1) \ d(i) = 0$$ \hspace{1cm} (5.45a)

As

$$\beta(i+1) = \beta(s+1) + \sum_{j=s+1}^{i} \lambda_j \ d(j)$$

expression

$$\frac{d}{dx}(i+1) \ d(s) = 0 \hspace{1cm} s = 1, 2, \ldots, i$$ \hspace{1cm} (5.45b)

holds too, because:

$$\frac{d}{dx}(s+1) + \sum_{j=s+1}^{i} \lambda_j \ d(j) = \frac{d}{dx}(s+1) + \sum_{j=s+1}^{i} \lambda_j \ d(j) = 0$$

and

$$\frac{d}{dx}(s) \ \frac{d}{dx}(i+1) = \frac{d}{dx}(s) \ \frac{d}{dx}(s+1) + \sum_{j=s+1}^{i} \lambda_j \ d(s) \ d(j) = 0$$

Consequently the gradient in the last iteration is perpendicular to all already used conjugate directions. When the next conjugate direction \(d(i+1)\) incorporates gradient \(g(i+1)\), we are sure to explore in a new dimension, which explains the hypothesis:

$$d(i+1) = -g(i+1) + y_{it} \ d(i)$$ \hspace{1cm} (5.46)
By substitution this hypothesis can easily be verified, which results in:

$$\gamma_{t+1} = \frac{\dot{g}(i+1) \cdot g(i)}{\dot{g}(i) \cdot g(i)}$$

(5.47)

Starting with the gradient as the first direction after \(n\) iterations one possesses \(n\) independent conjugate directions. In this last iteration point the gradient has to be perpendicular to all these independent directions according to (5.45b), which is impossible in the \(n\) dimensional space. Consequently, the gradient is zero and the minimum is reached.

The above algorithm has been developed by Fletcher and Reeves (1963).

Another approach, more frequently applied, is the algorithm of Davidon, Fletcher and Powell (1963). As mentioned already, the concept of this method is an updating of the inverse second derivative \(A\) in each iteration step. An explicit form for the inverse matrix can be given in terms of the conjugate directions:

$$H = \sum_{i=1}^{n} \frac{d(i) \cdot d(i)}{d'(i) \cdot H \cdot d(i)} \Rightarrow H \cdot H \cdot d(s) = d(s) \text{ for } s = 1, 2, \ldots, n$$

or

$$H \cdot H = I \Rightarrow H = H^{-1}$$

It can be proved then, that if

$$d(i) = -H(i-i) \cdot g(i)$$

(5.48)

and

$$H(i) = H(i-i) + \lambda \cdot \frac{d(i) \cdot d(i)}{g''(i) \cdot H(i-i) \cdot g(i)} - \frac{H(i-i) \cdot \gamma(i) \cdot \gamma(i) \cdot H(i-i)}{\gamma(i) \cdot H(i-i) \cdot \gamma(i)}$$

with

$$\gamma(i) = g(i+1) - g(i)$$

in \(n\) iterations \(H = A^{-1}\), which means that the minimum is reached for a quadratic criterion. A disadvantage of this method is the line mini-
mization that we need to perform in each iteration. Davidon (1968), however, introduced a similar method without line minimization. This algorithm normally reaches the minimum in the same number of iterations, but uses only one function and one gradient per iteration.

All methods mentioned above behave theoretically very well on almost quadratic surfaces, but there are many other "smooth" surfaces, where we also can expect a successful application of gradient methods. Such a category is the set of homogeneous functions:

\[ E(\beta) = \frac{1}{\delta} (\beta - b) g^t(\beta) + E_0^* \]  

(5.49)

where

\[ g(\beta) = \frac{\partial E}{\partial \beta} \]

A promising algorithm for minimizing these homogeneous functions is given by the following method.

h. Jacobson and Oksman method (1970). The idea itself is astonishingly simple: Suppose we know the function- and gradient values in n points \( \beta(i) \), where (5.49) can be written as:

\[ g^t(\beta(i)) b + E(\beta(i)) y - \gamma E_0^* = g^t(\beta(i)) \beta(i) \quad i = 1, 2, \ldots, n+2 \]

which are n+2 (or more) linear equations in n+2 parameters \( b, y \) and \( \gamma E_0^* \). When the coefficient matrix can be made regular by well-choosen \( \beta(i) \)'s, then the solution is obtained by matrix inversion (or pseudo inversion). Of course this idea can be transformed into an updating scheme like Jacobson and Oksman did. For many testfunctions they claim better results then the Powell, Fletcher, Davidon method.

Such a more rigorous test for convergence algorithms is Rosenbrock's valley, given by the expression

\[ E = 100 (\beta_2 - \beta_1)^2 + (1 - \beta_1)^2 \]
and sketched in fig. 5.19; cf. Rosenbrock and Storey (1966). In this reference a comparison can be found of several algorithms applied to a number of different "valleys".
5.3 Stochastic approximation.

The discussion of convergence schemes in section 5.2 did not take into account the disturbances introduced by additive noise in the determination of the gradient.

Stochastic approximation, as discussed in this section, can be used for several types of problems connected with optimizing, adaption, model adjustment and learning under the influence of noise. (Tsypkin, 1966).

Robbins and Monro (1951) introduced an approach to the problem of finding a point of a regression function. They start with the conditional probability function \( p(y|x) \) or distribution function \( F(y|x) \) and assume that:
- the regression function \( f(x) \) exists
- the equation \( f(x) = \lambda \) has a unique solution \( x = \theta \).

The latter condition implies that \( f(x) \) is a monotonic function.

Now the following iterative algorithm is proposed for determining the root \( x = \theta \) of the equation \( f(x) = \lambda \):

\[
X^{(i+1)} = X^{(i)} - \Gamma^{(i)} f_i(X^{(i)}) - \lambda \quad (5.50)
\]

where \( x(i) \) is the \( i \)-th estimate and the observation \( f(x(i)) \) are contaminated with an additive random disturbance; cf. fig. 5.20; \( \Gamma^{(i)} \) with \( i = 1, 2, \ldots \) is a sequence of numbers with properties that still have to be determined.

This sequence has to be chosen in such a way that \( x(i) \) converges in probability to \( \theta \), i.e. for all \( \epsilon > 0 \):

\[
\lim_{i \to \infty} P \left\{ \left| x(i) - \theta \right| \geq \epsilon \right\} = 0 
\quad (5.51)
\]

The following reasoning indicates the restrictions that result from this condition. Consider \( \{x(i+1) - \theta\}^2 \) together with equation (5.50) and
take the expectation:
\[ E \left[ \left\{ x(i+1) - \theta \right\}^2 \right] = E \left[ \left\{ x(i) - \theta \right\}^2 \right] + \Gamma^2(i) E \left[ \left\{ f(x(i)) - \lambda \right\}^2 \right] + 2 \Gamma(i) E \left[ \left\{ x(i) - \theta \right\} \left\{ f(x(i)) - \lambda \right\} \right] \]
(5.52)

or
\[ b(i+1) = b(i) + \Gamma^2(i) e(i) - 2 \Gamma(i) d(i) \]
(5.52a)

with
\begin{align*}
\gamma(i) &= E \left[ \left\{ x(i) - \theta \right\}^2 \right] > 0 \\
e(i) &= E \left[ \left\{ f(x(i)) - \lambda \right\}^2 \right] > 0 \\
d(i) &= E \left[ \left\{ x(i) - \theta \right\} \left\{ f(x(i)) - \lambda \right\} \right] \\
\Gamma(i) &> 0
\end{align*}

For the existence of \( e(i) \) it is required that \( f(x(i)) \) is finite except possibly at a finite number of points. We require \( b(i+1) < b(i) \). So \( d(i) > 0 \).

This implies that \( f(x) \) has to be monotonic at least in the region of \( \theta \).

Writing the expressions for \( b(2), b(3), \ldots b(i+1) \) yields:
\[ b(i+1) = b(i) + \sum_{j=1}^{i} \Gamma^2(j) e(j) - 2 \sum_{j=1}^{i} \Gamma(j) d(j) \]
(5.53)

or
\[ \sum_{j=1}^{i} \Gamma(j) d(j) = \frac{1}{2} \left[ b(i) + \sum_{j=1}^{i} \Gamma^2(j) e(j) - b(i+1) \right] \leq \frac{1}{2} \left[ b(1) + \sum_{j=1}^{i} \Gamma^2(j) e(j) \right] \]
(5.54)

Under the condition that \( e(j) \) is finite and
\[ \sum_{j=1}^{\infty} \Gamma^2(j) < \infty \]
one finds
\[ \sum_{j=1}^{\infty} \Gamma(j) d(j) < \infty \]
(5.55)

If \( \Gamma(j) \) is chosen such that
\[ \sum_{j=1}^{\infty} \Gamma(j) = \infty \]
then
\[ \lim_{j \to \infty} d(j) = 0 \]
(5.56)

Consequently
\[ \lim_{i \to \infty} E \left[ \left\{ x(i) - \theta \right\} \left\{ f(x(i)) - \lambda \right\} \right] = 0 \]
(5.57)
Development of \( f(x_{i}) \) in the vicinity of \( x = \theta \) yields:

\[
\begin{align*}
\{ (x(i)) &= f(\theta) + \frac{f'(\theta)}{1!} \left\{ x(i) - \theta \right\} + o \left\{ x(i) - \theta \right\}^2 \\
\{ (x(i)) - \lambda &= f'(\theta) \left\{ x(i) - \theta \right\} + o \left\{ x(i) - \theta \right\}^2
\end{align*}
\]

or

\[
\begin{align*}
\lim_{i \to \infty} E \left[ f'(\theta) \left\{ x(i) - \theta \right\} \right] &= 0 \\
\lim_{i \to \infty} P \left\{ |x(i) - \theta| > \varepsilon \right\} &= 0
\end{align*}
\]

Sequences that obey the conditions (5.55) and (5.56) are \( \Gamma(j) = \frac{c}{f^2} \) with \( 0.5 < \alpha \leq 1. \)

As mentioned this method was developed originally for determining the root of an equation, consisting of a monotonic function. The method has been applied by Kiefer and Wolfowitz (1952) to determine the extremum of a regression function. As the regression function itself is unknown the derivative is estimated by

\[
\nabla f (x(i)) \approx \frac{f(x(i) + a(i)) - f(x(i) - a(i))}{2 a(i)}
\]

By this operation the "even" function is replaced by an "odd" function, cf. chapter 2. The algorithm for the determination of a minimum is:

\[
x(i+1) = x(i) - \Gamma(i) \nabla f (x(i))
\]

This algorithm converges if all conditions with respect to \( f(x) \) and \( \Gamma(i) \) are satisfied. Kiefer and Wolfowitz also have extended this method to more-dimensional and continuous cases. The multi-dimensional case can be formulated in the following way:

The function

\[
\begin{align*}
F(x) &= F(x_1, x_2, \ldots, x_N)
\end{align*}
\]

has to be maximized (or minimized) by a suitable choice of \( x \). Assume the maximum occurs for \( x = \hat{x} \). If \( F(x) \) is known and if there are no extremal boundary values, saddle points etc., then \( \hat{x} \) follows from

\[
\nabla x F(x) = \left\{ \frac{\partial F(x)}{\partial x_1}, \ldots, \frac{\partial F(x)}{\partial x_N} \right\}^t = 0 \quad \text{for} \quad x = \hat{x}
\]
For determining $\hat{x}$ the following algorithm is suitable:

$$x(i+1) = x(i) - \Gamma(i) \nabla_x F(x(i))$$  \hspace{1cm} (5.65)

The choice of $\Gamma(i)$ determines the iteration method. If the convergence conditions are satisfied then

$$x(i) \rightarrow \hat{x} \hspace{1cm} \text{for} \; i \rightarrow \infty$$

In cases that $F(x)$ is not known the gradient can be estimated by:

$$\nabla_x F(x) \approx \Delta F'(x, \pm a) = \left\{ \frac{F(x + ae_i) - F(x - ae_i)}{2a}, \ldots, \right.$$  

$$\ldots, \frac{F(x + ae_N) - F(x - ae_N)}{2a} \right\}$$  \hspace{1cm} (5.66)

where $e_1, \ldots, e_N$ are the basis vectors and $a$ is a small scalar value; $a$ can be made dependent on the iteration number and on the vector component $x_j$.

If in addition $F(x)$ is of a stochastic nature then this stochastic aspect can be expressed by the function $F(n|x)$ where $n$ is a stochastic vector.

The best thing to do in such cases, is to maximize (or minimize) the expected value

$$\int F(n|x)p(n) \, dn = \mathbb{E}_n \left[ F(n|x) \right]$$  \hspace{1cm} (5.67)

Then $x = \hat{x}$ follows from

$$\nabla_x \mathbb{E}_n \left[ F(n|x) \right] = 0$$

and the algorithm is

$$x(i+1) = x(i) - \Gamma(i) \nabla_x \mathbb{E}_n \left[ F(n|x(i)) \right]$$  \hspace{1cm} (5.68)

If $p(n)$ is not given beforehand and if $F(x)$ is unknown then we have to apply an estimate of the gradient as used in the next expression:

$$x(i+1) = x(i) - \Gamma(i) \frac{F(n(i)|x(i) + \xi(i)) - F(n(i)|x(i) - \xi(i))}{2\xi(i)}$$  \hspace{1cm} (5.69)

Now $x(i) \rightarrow \hat{x}$, i.e.

$$\lim_{i \rightarrow \infty} \mathbb{E} \left[ \left( x(i) - \hat{x} \right)^2 \right] = 0$$  \hspace{1cm} (5.70)
These conditions imply that \( r(i) \) has to decrease fast enough, in order that the variance disappears. On the other hand \( r(i) \) may not decrease too fast, because otherwise not a sufficient amount of data is being processed. If the process is deterministic, then one can drop the requirements with respect to \( r(i) \) and \( r \) can be chosen to be constant.

As an illustration of the capabilities of the stochastic approximation approach, the results of a simple experiment are shown in the figures

\[ \Psi \] by Mr. J.P.C.M. de Weerd, Technical University, Eindhoven, Netherlands.

5.21 and 5.22. The function \( F \) has been chosen to be one-dimensional;

\[
F\left(h(i) \mid x(i)\right) = 1 + x(i) + \lambda h(i)
\]

\( p(n) = 1 \) for \( 0 \leq n \leq 1 \)

\( a(i) = a = 0.5 \)

For a low noise level, \( \lambda = 0.1 \), the convergence to the right value, \( x = 0 \), is shown in fig. 5.21. \( r = 0.5 \) results in a one-step-approach; \( r = 0.9 \) leads to an oscillatory behaviour; \( r = 1/i \) shows a convergence in two steps.

In case of a high noise level, \( \lambda = 10 \), two approaches are given in fig. 5.22 a and b for \( r = 0.5 \) viz. \( r = 1/i \). Each figure shows the behaviour for two different noise sequences \( n(i) \). In case of the constant \( r \) there is no convergence to \( x = 0 \). The stochastic approximation algorithm with \( r = 1/i \) converges satisfactorily in spite of the same noise.

Some of these algorithms can be made continuous in time:

\[
\frac{d\tilde{x}(i)}{dt} = -\frac{r}{2} \sum_{i} F' \tilde{x}(i) \tag{5.71}
\]

for the non-stochastic case and

\[
\frac{d\bar{x}(i)}{dt} = -\frac{r}{2} \sum_{i} F' \bar{x}(i) \tag{5.72}
\]

for the stochastic case.
For more information on stochastic approximation and on different types of application the reader is referred to the excellent surveys available. See e.g. Albert and Gardner (1967) and Tsypkin (1966). Some applications are given by Sakrison (1967), Saridis and Stein (1968a, 1968b), Holmes (1968), Elliott and Sworder (1969), Neal and Bekey (1969).


Deterministic convergence schemes


Stochastic approximation.


Fig 5.1

Fig 5.2

Fig 5.3
Fig. 5.17

Gauss-Newton; μ = 0

M = ∞
Schepest descent

Fig. 5.18

Fig. 5.19
(from Rosenbrock and Storey, 1966)
fig. 5.20

\[ f(x) \]

\[ f(x) - \lambda \implies -f(x) \{ f(x) - \lambda \} \]

fig. 5.21

\[ x(i) \]

\[ \Gamma(i) = 0.5 \]
\[ \Gamma(i) = 0.9 \]
\[ \Gamma(i) = \frac{1}{i} \]
**Fig. 5.22a**

\[ \Gamma = 0.5 \]

**Fig. 5.22b**

\[ \Gamma = \frac{1}{t} \]
Chapter 6  SAMPLED SIGNALS; explicit methods

In chapter 2 the least-squares and the generalized least-squares estimators were introduced. In chapter 5 they were "derived" from the maximum likelihood approach under special conditions. This was done in order to show the relationship between the various estimation procedures.

In this chapter, these estimators are derived starting from the concepts of regression analysis. Properties such as linearity, unbiasedness and minimum-variance are discussed.

Next, the implementation of such estimators is considered, as well as the types of errors which occur in practical applications: noise error, truncation error, state-correspondence error, errors due to the implementation, sampling error.

In the estimation procedure the residuals can be used to derive more information, for example, about the noise variance and the number of process parameters; this procedure is treated in section 6.4.

An attempt to use these explicit methods on the class of generalized models leads to estimates which generally are biased. An asymptotic expression is derived for this bias. A few methods circumventing this problem are outlined.
6.1 Regression analysis.

Regression curves and planes. The elementary notions of regression analysis (Cramér, 1946; Aitken, 1952) can be given as follows using a two-dimensional case.

Let \( x \) and \( y \) be random variables with a joint probability density function \( p(x,y) \). If this function is continuous in \( y \) then the conditional probability density function of \( y \) with respect to \( x \) can be written as:

\[
p(y|x) = \frac{p(x,y)}{\int_{-\infty}^{\infty} p(x,y) \, dy} = \frac{p(x,y)}{p(x)} \quad (6.1)
\]

From this function we want to determine some functional relation, for example, the expectation of \( y \) for any value of \( x \):

\[
y_x = E[y|x] = \frac{\int_{-\infty}^{\infty} y \, p(x,y) \, dy}{p(x)} = f(x) \quad (6.2)
\]
That function is called a regression curve of y on x. It can be shown (Aitken, 1952, p. 96) that this regression curve gives the best possible estimation of y in the mean squared error sense.

In this way, one determines the functional relationship.

\[ y_x = f(x) \]

without restrictions on \( f \). We may, however, want this function to belong to a certain class, for example, the class of all linear functions or the class of all functions of a given polynomial form. Choosing such a class and minimizing the expectation in the mean squares sense leads to least squares regression curves.

Let us assume, for example, that we want to find the best linear function

\[ w = \alpha + \beta x \]  \hspace{1cm} (6.3)

This is called a linear regression model; \( \alpha \) and \( \beta \) are called the regression coefficients. Then we have to minimize

\[ E[(y-w)^2] = E[(y-\alpha-\beta x)^2] \]  \hspace{1cm} (6.4)

where the expectation operator \( E[\cdots] \) stands for \( \int \cdots \int f(x,y) \, dx \, dy \)

and the following notation will be useful, cf. appendix B:

\[
\begin{align*}
\mu_x &= E[x] \\
\mu_y &= E[y] \\
\sigma^2_x &= E[x^2] - E[x]^2 \\
\sigma^2_y &= E[y^2] - E[y]^2 \\
\sigma_{xy} &= E[xy] - E[x] \cdot E[y] \\
\rho &= \frac{\sigma_{xy}}{\sigma_x \sigma_y} \hspace{1cm} \text{correlation coefficient}
\end{align*}
\]

Differentiation of eq. (6.4) with respect to \( \alpha \) and with respect to \( \beta \) and equating to zero leads to

\[
\begin{align*}
-2E[y-\alpha-\beta x] &= -2E[y] + 2\alpha + 2\beta E[x] = 0 \\
-2E[(x/y-\alpha-\beta x)] &= -2E[x] + 2\alpha E[x] + 2\beta E[x^2] = 0
\end{align*}
\]  \hspace{1cm} (6.5)
From this it follows that:

\[ \alpha = \frac{\alpha_y - \alpha_x}{\sigma_y} \quad \frac{\sigma_{xy} L}{\sigma_x^2} \]

\[ \beta = \frac{\sigma_{xy}^2}{\sigma_x^2} \]

Substitution of the expressions for \( \alpha \) and \( \beta \) in eq. (6.3) yields the equation of the regression line:

\[ v = \frac{\alpha_y + \beta_x (x - \mu_y)}{\sigma_y} \]

These ideas can be generalized so as to find the best nonlinear regression model, for example, the polynomial functions \( w = a + \beta x + \gamma x^2 + \ldots \)

It may be useful to choose a set of orthonormal polynomials, cf. Aitken (1952), p. 115. These polynomials \( \Pi_i \) are of the degree \( i \) in \( x \) \((i = 0, 1, \ldots, m)\)

and fulfill the conditions of orthonormality:

\[ \int_{-\infty}^{\infty} \Pi_i(x) \Pi_j(x) \rho(x) \, dx = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \]  

where

\[ \rho(x) = \int_{-\infty}^{\infty} \rho(x, y) \, dy \]

is the marginal probability density function of \( x \). If now

\[ v = \sum_{i=0}^{m} y_i \Pi_i(x) \]

then we have to minimize

\[ \mathbb{E} \left[ (y - \sum_{i=0}^{m} y_i \Pi_i(x))^2 \right] \]

By differentiation with respect to \( y_i \), one obtains:

\[ \mathbb{E} \left[ \Pi_i(x) (y - \sum_{i=0}^{m} y_i \Pi_i(x))^2 \right] = \int_{-\infty}^{\infty} y \Pi_i(x) \rho(x, y) \, dx \, dy - \sum_{i=0}^{m} \int_{-\infty}^{\infty} \Pi_i(x) \rho(x, y) \, dx \, dy \]

This has to be equated to zero. As a result of the orthonormality, this reduces to:

\[ y_i = \int_{-\infty}^{\infty} y \Pi_i(x) \rho(x, y) \, dx \, dy \]

Consequently, there is no need to solve a set of linear equations. This is particularly advantageous since the linear equations often become nearly singular as the order of the model increases (Hamming)

These ideas can be extended to the more-dimensional case where a conditional probability density function of the form...
is considered.

Again the class of functions may be restricted, for example, to the linear ones.

Then the choice of

\[ w = \alpha + \beta_1 x_1 + \ldots + \beta_m x_m \]  \hspace{1cm} (6.13)

leads to a mean square regression plane for \( y \) with respect to \( x_1, \ldots, x_m \) if the following expectation is minimized:

\[ E[(y - w)^2] = E[(y - \alpha - \beta_1 x_1 - \ldots - \beta_m x_m)^2] \]  \hspace{1cm} (6.14)

If all random variables are assumed to be centered, i.e., with zero expectation, then the coefficient \( \alpha \) can be dropped.

Differentiating with respect to \( \beta_1, \ldots, \beta_m \) respectively, and equating to zero leads to:

\[ \beta_1 E[x_1^2] + \beta_2 E[x_1 x_2] + \ldots + \beta_m E[x_1 x_m] = E[x_1 y] \]

\[ \vdots \]

\[ \beta_1 E[x_m^2] + \beta_2 E[x_m x_2] + \ldots + \beta_m E[x_m^2] = E[x_m y] \]  \hspace{1cm} (6.15)

If the matrix of expectations is not singular this set of linear equations can be solved for the regression coefficients \( \beta_1, \ldots, \beta_m \). When \( w \) is chosen as a more complicated function than polynomials (e.g., a set of exponentials), then the resulting expressions for the coefficients are often quite difficult and tedious to obtain.

**Estimation from a finite number of observations.** So far we have assumed that we can determine all \( E[\,] \) operations, i.e., that the joint probability density function \( p(x_1, \ldots, x_m, y) \) is known. This will seldom be the case. Then we will have to estimate the parameters using a finite number of observations, viz. the sample values. Consequently, the estimate has to be some function of these sample values, which are actually observed values of a random variable.
This implies that the estimate is also a random variable and can be characterized by a probability density function. The quality of the estimate depends on this function, for example, its expectation and its variance.

The methods discussed have a long history. K. Gauss already used them in 1795 for the study of planetary motion. In our time they are used for example, for determining parameters of satellite orbits. In passing, it has to be noted that apart from the ordinary type of regression model

\[ y_i = \beta_0 x_{i,i} + \ldots + \beta_m x_{m,i} + n_i \]

with \( n_i \) being a random variable, in statistical literature the autoregression model

\[ y_i + \alpha_1 y_{i-1} + \ldots + \alpha_k y_{i-k} = \eta_i \]

and the generalized regression model

\[ y_i + \alpha_1 y_{i-1} + \ldots + \alpha_k y_{i-k} = \beta_0 x_{i,i} + \ldots + \beta_m x_{m,i} + n_i \]

are studied as well (Durbin, 1960; Jenkins and Watts, 1968).

**Notation.** Now we will see how estimates can be derived. The following notation will be used. Observations are made on the process output signal \( y \).

As in chapter 2, this \( y \) is composed of: - the response to the process input \( u \) and : - the noise \( n \) added in the process and in the measuring instrument.

The output at the \( j \)th sampling instant is:

\[ y(j) = y(y, b, n, j) \]  \hspace{1cm} (6.16)

with \( j = 1, \ldots, k \). The argument \( b \) of \( y \) indicates that these samples depend on the parameters \( b_0, b_1, \ldots, b_m \) which are the components of the process parameter vector \( b \).

We define

\[ y' = [y(1), y(2), \ldots, y(k)] \]  \hspace{1cm} (6.17)
The noise may be given by its mean and its covariance matrix

$$\mathbb{E}[n] = 0$$  \hspace{1cm} (6.18)

$$\mathbb{E}[n n'] = \begin{bmatrix} \mathbb{E}[n(0) n(0)] & \cdots & \mathbb{E}[n(0) n(1)] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[n(0) n(1)] & \cdots & \mathbb{E}[n(0) n(1)] \end{bmatrix} = \mathbf{N}$$  \hspace{1cm} (6.19)

The objective is to determine a coefficient vector $\beta$ which estimates $b$. This is done by using the theoretical response vector $w$, the model output, which is dependent on $\beta$, the coefficient vector, which has the components $\beta_0, \beta_1, \ldots, \beta_m$. This functional relationship can be chosen in many different ways.

The simplest case is that in which there is a linear relation between $w$ and $\beta$, a model which is linear-in-the-parameters (cf. section 4.6), i.e.,

$$w(j) = \sum_{i=0}^{m} \beta_i u_i(j) \quad j = 1, \ldots, k$$

where $u_i(j)$ are linearly independent known functions. The $w$ can be written as:

$$w = U \beta$$  \hspace{1cm} (6.20)

with

$$U = \begin{bmatrix} u_0(1) & \cdots & u_m(1) \\ \vdots & \ddots & \vdots \\ u_0(k) & \cdots & u_m(k) \end{bmatrix}$$  \hspace{1cm} (6.21)

Note again, that this choice of a linear relationship between $w$ and $\beta$ does not imply that the input-output relation of the model has to be linear. The matrix $U$ is assumed to be known perfectly, that is, it can be measured without errors. Moreover, it is assumed that there are more observations ($k$) than there are unknown parameters ($m+1$).
Linear unbiased estimators. The class of linear unbiased estimators is defined by:

\[
\hat{\beta} = Qv \quad \text{linearity}
\] (6.22)

where \( Q \) is a matrix having \( m+1 \) rows and \( k \) columns, and

\[
E[\beta] = b \quad \text{unbiasedness}
\] (6.23)

It is assumed that (6.20) may give a complete representation model of the process, that is:

\[
y = Ub + n
\] (6.24)

For the moment we will assume that \( U \) and \( n \) are statistically independent.

Now an error vector \( e \) may be defined as

\[
e = y - w = y - U\beta
\] (6.25)

An error- or loss function can be taken as a positive definite form:

\[
E = e' R e = \| e \|^2
\] (6.26)

where \( R \) is a matrix of weighting coefficients \( r_{ij} \). Without loss of generality this matrix will be assumed symmetric.

This error function can be written as:

\[
E = \| y - U\beta \|^2 = (y - U\beta)' R (y - U\beta) =
\]

\[
y' R y + [U\beta]' R [U\beta] - y' R [U\beta] - [U\beta]' R y
\] (6.27)

Since \( [U\beta]' = \beta' U' \) and since \( R \) is a symmetric matrix:

\[
E = y' R y + \beta' U' R U\beta - 2 \beta' U' R y
\] (6.28)

Differentiation of equation (6.28) with respect to \( \beta \) gives (cf. Appendix C):

\[
\frac{\partial E}{\partial \beta} = \nabla_{\beta} E = 2 U' R U\beta - 2 U' R y
\] (6.29)
This can also be written as \(-2 \mathbf{U}'\mathbf{R} \mathbf{[y-U\hat{\beta}]} \mathbf{= -2U}'\mathbf{Re}\). For some vector \(\hat{\beta}\) eq. 6.29 can be equated to zero, leading to an expression for \(\hat{\beta}\) at which an extremum of \(E\) occurs:

\[
\mathbf{U}'\mathbf{R} \mathbf{U} \hat{\beta} = \mathbf{U}'\mathbf{R} \mathbf{y}
\]  

(6.30)

This is called the "normal equation". If \(\mathbf{U}'\mathbf{R}\mathbf{U}\) is non-singular then

\[
\hat{\beta} = (\mathbf{U}'\mathbf{R}\mathbf{U})^{-1}\mathbf{U}'\mathbf{R} \mathbf{y}
\]

(6.31)

It is not difficult to show that for \(\hat{\beta}\) the error function \(E\) actually has a minimum value. This value \(E(\hat{\beta})\) is called the residual error (based on \(k\) observations).

**Note 1:** Of course eq. (6.31) can also be derived through calculus of variations:

\[
\frac{d}{d \alpha} E(\hat{\beta} + \alpha \Delta \beta) \bigg|_{\alpha=0} = 0
\]

or

\[
\Delta \beta' \mathbf{U}'\mathbf{R} \mathbf{[y-U\hat{\beta}]} = 0
\]

for arbitrary \(\Delta \beta\) (orthogonality principle).

**Note 2:** A direct proof of the minimization of \(E\) can be found by the well-known method of completing the square term in \(\beta\). From eq. (6.27):

\[
E = \mathbf{y}'\mathbf{R} \mathbf{y} + \beta'\mathbf{U}'\mathbf{R} \mathbf{U} \beta - \mathbf{y}'\mathbf{R} \mathbf{U} \beta - \beta'\mathbf{U}'\mathbf{R} \mathbf{y} =
\]

\[
= [\beta - (\mathbf{U}'\mathbf{R}\mathbf{U})^{-1}\mathbf{U}'\mathbf{R} \mathbf{y}] \mathbf{U}'\mathbf{R} \mathbf{U} [\beta - (\mathbf{U}'\mathbf{R}\mathbf{U})^{-1}\mathbf{U}'\mathbf{R} \mathbf{y}] +
\]

\[
- \mathbf{y}'\mathbf{R} \mathbf{U} [\mathbf{U}'\mathbf{R}\mathbf{U}]^{-1}\mathbf{U}'\mathbf{R} \mathbf{y} + \mathbf{y}'\mathbf{R} \mathbf{y}
\]

Evidently this is a minimum if \(\hat{\beta}\) satisfies eq. (6.31).

**Note 3:** As a mnemonic technique the reader may find it useful to remember
multiply by $U' R$

$$U' R y = U' R U b + U' R n$$

As the last term is unknown, unmeasurable and as it is assumed that $U$ and $n$ are statistically independent, that term is discarded, leaving an estimate $\hat{\beta}$ instead of the true parameter $b$ as shown in eq. (6.30). Naturally this way of deriving eq. (6.31) does not indicate in what sense the estimator is optimal.

This estimate fulfills the condition of eq. (6.22) for linearity because

$$\hat{\beta} = Q y \quad \text{with} \quad Q = [U'R U]^{-1} U' R$$

From equations (6.31) and (6.24) it follows that

$$\hat{\beta} = [U' R U]^{-1} U' R [U b + n] = b + [U' R U]^{-1} U' R n$$

As the input and noise signal are statistically independent:

$$E[\hat{\beta}] = b + E[U' R U]^{-1} E[U R n]$$

Furthermore, it was already assumed that $E[U b] = 0$, consequently, the estimate is unbiased:

$$E[\hat{\beta}] = b$$

This implies

$$E[n] = E[U \hat{\beta}] = U b$$

that is, the expectation of the "model output" equals the process output without the additive noise.

Another characteristic of the estimate, $\hat{\beta}$ of eq. (6.31), which we want to determine is the variance. One may also be interested in the correlation between the elements of the vector $\hat{\beta}$. This can be found by means of the covariance matrix:

$$\text{cov}[\hat{\beta}] = E[(\hat{\beta} - b)(\hat{\beta} - b)] = E[(U y - b)(U y - b)]$$
If it is again assumed that eq. (6.24) holds and that \( \mathbf{U} \) and \( \mathbf{n} \) are statistically independent then one finds, using (6.32):

\[
\text{Cov}[\hat{\beta}] = E \left[ (\mathbf{QU} b + \mathbf{Q} n - b)(\mathbf{QU} b + \mathbf{Q} n - b)' \right] = E \left[ (\mathbf{Q} n)(\mathbf{Q} n)' \right] = \mathbf{Q} E[\mathbf{n} n'] \mathbf{Q}' = \mathbf{Q} \mathbf{N} \mathbf{Q}'
\]

Consequently

\[
\text{Cov}[\hat{\beta}] = \left[ \mathbf{U}' \mathbf{R} \mathbf{U} \right]^{-1} \mathbf{U}' \mathbf{N} \mathbf{R} \mathbf{U} \left[ \mathbf{U}' \mathbf{R} \mathbf{U} \right]^{-1}
\]

We will find that in several cases of practical interest this expression is simplified considerably. Estimates of the variances of the parameter estimates are obtained from the main diagonal of this matrix.

**Least squares estimators.** (Levin, 1959; 1960) The least squares estimators imply a minimization of

\[
E = \mathbf{e}' \mathbf{e} = \| \mathbf{e} \|^2
\]

Consequently in (6.26) and following equations:

\[
\mathbf{R} = \mathbf{I}
\]

and from equations (6.30), (6.31) and (6.35),

\[
\mathbf{U}' \mathbf{U} \hat{\beta}_s = \mathbf{U}' \mathbf{y}
\]
or
\[ \hat{\beta}_j = (u'u)^{-1}u'y \]  
(6.38)

\[ \text{cov}[\hat{\beta}_j] = (u'u)^{-1}u'nu[u'u]^{-1} \]  
(6.39)

If \( U \) is a square matrix, i.e., if the number of samples is equal to the number of parameters to be determined, and if \( U \) has an inverse, then:
\[ (u'u)^{-1}u' = u^{-1} \]  
(6.40)

and
\[ \hat{\beta}_j = u^{-1}y \]  
(6.41)

i.e., a solution of a set of linear equations. From the engineering point of view this case is of little interest due to the disturbance by the noise.

For reduction of the influence of the noise, the number of samples chosen has to be much larger than the number of parameters.

If equation (6.37) is written in terms of \( u \)'s then:
\[
\begin{bmatrix}
\sum_{j=1}^{k} u_2^2(j) & \cdots & \sum_{j=1}^{k} u_0(j) u_2(j) \\
\cdots & \cdots & \cdots \\
\sum_{j=1}^{k} u_0(j) u_0(j) & \cdots & \sum_{j=1}^{k} u_2^2(j)
\end{bmatrix}
\begin{bmatrix}
\beta_0 \\
\vdots \\
\beta_p
\end{bmatrix}
= 
\begin{bmatrix}
\sum_{j=1}^{k} u_0(j) y(j) \\
\vdots \\
\sum_{j=1}^{k} u_m(j) y(j)
\end{bmatrix}
\]  
(6.42)

Choice of testsignals \( u_i(j) \) which are orthogonal or orthonormal may introduce desirable simplification; cf. chapter 4. In the case of orthonormality then:
\[ u'u = I, \text{ the identity matrix } \]

and
\[ \hat{\beta} = u'y \]
or
\[ \hat{\beta}_i = \sum_{j=1}^{k} u_i(j) y(j) \]
i = 0, ..., m
A simple geometrical interpretation of the least squares estimate can be given as follows; cf. fig. 6.1 for the case where $\beta$ is two-dimensional.

The length of
\[ e = y - \hat{\beta} \]
has to be minimized. This is true if $e$ is orthogonal to $u_1$ and $u_2$;
\[ u_1' e = 0 \quad \text{or} \quad u_2' e = 0 \]

Consequently
\[ u'(y - \hat{\beta}) = 0 \]
or
\[ u'\hat{\beta}_s = u'y \]
which is eq. (6.37).

Markov or generalized least squares estimators. The Markov estimators imply a minimization of
\[ E = e' N^{-1} e = ||e||^2_{N^{-1}} \]
Consequently
\[ B = N^{-1} \]
and from equations (6.30), (6.31) and (6.35):
\[ u' N^{-1} u \hat{\beta}_M = u' N^{-1} y \]
or
\[ \hat{\beta}_M = [u' N^{-1} u]^{-1} u' N^{-1} y \]
and
\[ \text{cov} [\hat{\beta}_M] = [u' N^{-1} u]^{-1} u' N^{-1} N N^{-1} u [u' N^{-1} u]^{-1} = [u' N^{-1} u]^{-1} u' N^{-1} u [u' N^{-1} u]^{-1} = [u' N^{-1} u]^{-1} \]
As the covariance matrix (6.46) is part of the estimator (6.45), one obtains both estimate and covariance at the same time. The diagonal elements of (6.46) represent the variances of the parameter estimates, the off-diagonal elements provide the correlation between the errors of the different parameters. If \( k \to \infty \) and \( \sum_{j=1}^{k} u_1^2(j) \to \infty \) the variances approach zero and consequently the estimates are consistent, cf. section 5.1.

Using Schwarz's inequality for matrices (cf. Appendix C), it can quite easily be proved that this linear unbiased estimator has the property of minimum variance.

Schwarz's inequality:

\[
\begin{bmatrix} B' \cdot B \end{bmatrix} \geq \begin{bmatrix} B' \cdot A \end{bmatrix} \begin{bmatrix} A' \cdot A \end{bmatrix}^{-1} \begin{bmatrix} A' \cdot B \end{bmatrix}
\]

Choose: \( A = N^{-1/2} U \) and \( B' = Q N^{-1/2} \), then

\[
QN'Q' = QNQ' \geq [QN^{-1/2}U][U'N^{-1/2}N^{-1/2}U]^{-1}[U'N^{-1/2}N^{-1/2}Q']
\]

Because \( Qu = I \) then

\[
QNQ' \geq [U'N^{-1/2}U]^{-1}
\]

or

\[
\text{Cov} \begin{bmatrix} \beta \end{bmatrix} \geq \text{Cov} \begin{bmatrix} \hat{\beta}_M \end{bmatrix}
\]

i.e.,

\[
\text{Cov} \begin{bmatrix} \beta \end{bmatrix} - \text{Cov} \begin{bmatrix} \hat{\beta}_M \end{bmatrix}
\]

is non-negative definite. Here \( \beta \) is any other linear estimate. This implies efficiency of the estimator, cf. section 5.1.

From the previous sections, it is clear that if the additive noise is "white" that is,

\[
N = \mathbb{E}[n n'] = \sigma_n^2 I
\]

then

\[
\hat{\beta}_S = \hat{\beta}_M
\]
and

$$\text{cov}[\hat{\beta}_M] = \sigma_n^2 [v'v]^{-1}$$  \hspace{1cm} (6.50)

As was indicated in chapter 5, for Gaussian noise the least squares and Markov estimates coincide with the maximum likelihood ones, each having the minimum variance of all unbiased estimates.
The class of linear unbiased estimates was derived by minimizing the positive definite form

\[ E = \| \varepsilon \|_R^2 \]

Subsequently, the choices \( R = I \) and \( R = N^{-1} \) led to the least squares and the Markov estimates. Other choices would be possible as well, e.g., a weighting matrix taking into account a possible non-stationary additive noise. In that case one has to minimize the error-function.

\[ E = \sum_{j} \frac{1}{\sigma^2(j)} \left( y(j) - u'(j) \beta \right)^2 \]

where \( \sigma^2(j) \) represents the noise variance at time \( j \).

6.2 Implementation of "open loop" estimation schemes.

In section 6.1 we found the following explicit expressions:

\[
\hat{\beta}_g = \left( u'u \right)^{-1} u'y
\]

\[
\text{cov}[\hat{\beta}_g] = \left( u'u \right)^{-1} u'N u \left( u'u \right)^{-1}
\]

Markov estimate

\[
\hat{\beta}_M = \left( u'N^{-1}u \right)^{-1} u'N^{-1}y
\]

\[
\text{cov}[\hat{\beta}_M] = \left( u'N^{-1}u \right)^{-1}
\]

In this section we will discuss some aspects of the implementation of these expressions.

General considerations. We will start with the simplest case, the least squares estimation, following from

\[
\min_{\beta} E \quad \text{with} \quad E = e'e = \sum_{j} e^2(j)
\]

resulting in eq. (6.52).
The implementation of this equation leads to the celebrated correlation technique. As an example let us assume that only one parameter \( \beta_1 \) has to be determined. Equation (6.52) then reduces to:

\[
\beta_i = (u'_i u'_i)^{-1} u'_i y = \frac{\sum_k u'_i(j) y(j)}{\sum_j u'_i(j)}
\]  

(6.56)

The numerator and the denominator can be recognized as the time-averaged products that approximate the respective cross and autocorrelations for \( k \to \infty \); cf. fig. 6.2. The errors that may occur will be discussed later in this section. For the multiple parameter case, the corresponding expression is:

\[
\hat{\beta} = \left[ \begin{array}{c}
\hat{\beta}_0 \\
\vdots \\
\hat{\beta}_m
\end{array} \right] = \left[ \begin{array}{ccc}
u'_0 & \cdots & u'_m \\
\vdots & \ddots & \vdots \\
u'_m & \cdots & \nu'_m
\end{array} \right]^{-1} \left[ \begin{array}{c}
u'_0 y \\
\vdots \\
u'_m y
\end{array} \right]
\]  

(6.57)

Some remarks are in order here:

- The matrix \( U'U \) is symmetric about its main diagonal because \( u'_i u'_j = u'_j u'_i \).

- Among the practical problems is the computation of the inverse matrix; cf. Westlake (1968), Peterka and Smuk (1969). Using available matrix-inversion subroutines, it is a simple matter to write a straightforward estimation procedure for the parameters. If desired, the residuals and the value of the error function can also be determined. The covariance matrix can be computed if the noise covariance matrix \( N \) is known.
If $u_i^t u_j = \delta_{ij}$, the Kronecker delta, $U^t U = I$, the identity matrix.

Such conditions of orthonormality can be approximated by a suitable choice of the input signal $u$ and appropriate transfer functions $G_i$. cf. fig.6-3. The combination $u$ = white noise, $G_i$ = timedelay elements is most frequently used. The use of orthonormal filters for $G_i$ (e.g. Laguerre filters cf. Chapter 4) is quite often proposed.

The calculation of filters that provide orthogonal components $u_j$ for a particular input signal $U$ has been studied by Douce and Roberts (1964) and by Roberts (1967).

Instead of the operation on sampled signals, one may use the corresponding operations on non-sampled or "continuous" signals; cf. Chapter 9.
Instead of operating on signals with a continuous amplitude scale one may work with quantized signals which can have only a limited number of predetermined amplitude values. The limiting case of quantization gives a binary signal; cf. Chapter 8.

If one may use a testsignal, then \( u \) or \( u(t) \) can be chosen with full emphasis on its properties. Well known choices are testsignals composed of a number of sinewaves and pseudo-random-binary-sequences. Both types of signals can be given interesting properties with respect to orthogonality and generation, cf. Chapter 10.

In the derivation of the estimate no assumption was made that the process \( P \) has to be linear. The operations \( G_i \) in fig. 6.3 may be nonlinear. The only requirements imposed by the condition \( w = U \beta \) as a model for \( y = U b \) is the linearity of \( w \) with respect to the parameters to be determined. Consequently classes of nonlinear systems may be amenable to correlation techniques: cf. Wiener's characterization of nonlinear systems, Volterra expansions, Pugashev's systems reducible to linear; cf. Chapter 4.

If \( u_i^j = \{ u_i(1), u_i(2), \ldots, u_i(k) \} \) for \( i = 0, 1, \ldots, m \) then
\[
\bar{u}_i^j = \sum_{\ell=1}^{k} u_i(\ell) u_j(\ell) = k \tilde{\psi}_{u_i^j}(0, k) \tag{6.58}
\]
where \( \tilde{\psi}(o,k) \) is an approximation of one point of a correlation function, based on \( k \) observations. If \( u_i \) and \( u_j \) are delayed versions of \( u \) then
\[
\bar{u}_i^j = \sum_{\ell=1}^{k} u_i(\ell-i) u_j(\ell-j) = k \tilde{\psi}_{u_i^j}(\ell-i, k)
\]
For ergodic (stationary) time series
\[
\tilde{\psi}_{u_i^j}(\ell, k) \rightarrow \psi_{u_i}(\ell) \quad \text{for} \ k \rightarrow \infty
\]

The next degree of complication is found in the instrumentation of the Markov estimate, following from
\[
\min_{\beta} E \quad \text{with} \quad E = \mathbf{e}'N^{-1}\mathbf{e} \tag{6.59}
\]
resulting in eq. (6.54).

The matrix \( N^{-1} \) can be separated into a lower triangular matrix \( D \) and its transpose:
\[
N^{-1} = D' D
\]
Using this notation, eq. (6.59) and (6.54) can be written as:
\[
E = [D\mathbf{e}]'D\mathbf{e}
\]
\[
\hat{\beta} = (D'U)'(D'U)^{-1}(D'U)'\mathbf{y}
\]
The matrix \( D \) represents a "noise-whitening" filter; given \( n \) as the input sequence, the output of that filter is white noise. Figure 6.4 indicates how these filters may be introduced into the instrumentation of the correlation technique. As was mentioned before, this provides the minimum variance estimate (of all linear unbiased estimates that can be found).

In spite of this interesting property and the relative simplicity of the instrumentation it seems that it has found little application. The same remarks that have been made regarding \([U'U]^{-1}\) in the least squares case, may be made as regards the matrix \([D'U]'D'U]^{-1}\).

In practice, the available a priori knowledge of the disturbances is seldom sufficient for using the Markov estimating scheme. Consequently, one may want to estimate the statistics of the disturbance in terms of autoregression of the residual. From that knowledge one may derive the filters \( D \) as shown in fig. 6.4.

One such technique can be indicated as follows. Determine
\[
\hat{\beta} = [U'U]^{-1}U'y
\]
\[
\mathbf{e} = y - U\hat{\beta}
\]
Now the residuals \( \mathbf{e} \) are considered as an approximation of the disturbances \( \mathbf{n} \). Assume that these disturbances are generated from "white" noise through
a m-th order filter. From this interpretation follows that one wants to minimize:

\[ e - E \hat{\delta} \]

or

\[
\begin{bmatrix}
 e(m+1) \\
 \vdots \\
 e(k) \\
 \end{bmatrix}
 -

\begin{bmatrix}
 e(m) & \ldots & e(1) \\
 \vdots & \ddots & \vdots \\
 e(k) & \ldots & e(k-m) \\
 \end{bmatrix}
 \begin{bmatrix}
 \delta_1 \\
 \vdots \\
 \delta_m \\
 \end{bmatrix}
\]

From this follows, via the normal equations,

\[ \hat{\delta} = (E' E)^{-1} E' e \]

With these estimates one may adjust (approximately) the model \( D \) of filter \( D \). Then a new estimate for \( \hat{\beta} \) is determined, after that a new estimate of \( \hat{\delta} \) is made, etc.; cf. section 7.2.

A brute force method of obtaining observations without correlation when the noise non-white, is by waiting long enough between successive samples. In this way, however, the observation time may become excessively long.

For other approaches to this problem the reader is referred to Durbin (1960) and Mayne (1967). Åström's work, cf. Chapter II, is directed to the estimation of both process parameters and noise statistics.

Up to this point, we did not specify the model except by the condition that its output has to be linearly related to the parameters to be determined. In the majority of cases \( u \) is chosen as a delayed version of the input signal \( u \). When using a digital computer, such a choice is of course very convenient. In that case, the parameter \( \beta_1 \) is an estimate of one point \( h(i) \) of the process impulse response. Figure 6.5 indicates this type of model, using a "delay line". In that case, we assume that the process output \( y(j) \) can be found from a convolution sum:


\[ y(j) = \sum_{i=0}^{j} u(i) \cdot h(j-i) + n(j) \quad j = 1, \ldots, k \]

or

\[ y = U \cdot \beta + n = [u_0, u_1, \ldots, u_m] \cdot \beta + n \]

In that case eq. (6.57) can be written as

\[
\hat{\beta} = \begin{bmatrix}
\tilde{\beta}_0 \\
\vdots \\
\tilde{\beta}_m
\end{bmatrix} = \begin{bmatrix}
\tilde{\psi}_{uu}(0,k) & \cdots & \tilde{\psi}_{uu}(m,k) \\
\vdots & \ddots & \vdots \\
\tilde{\psi}_{uu}(m,k) & \cdots & \tilde{\psi}_{uu}(2k)
\end{bmatrix}^{-1} \begin{bmatrix}
\tilde{\psi}_{uy}(0,k) \\
\vdots \\
\tilde{\psi}_{uy}(m,k)
\end{bmatrix}
\]

(6.60)

where each \( \tilde{\psi} \) is a summation over \( k \) products. If \( k \to \infty \) and the signal \( u \) is ergodic \( \tilde{\psi}_{uu}(j,k) \) approaches \( \psi_{uu}(j) \), the auto correlation function. For an input time series with the character of "white noise"

\[ \psi_{uu}(j) = \begin{cases} \psi_{uu}(0) = \sigma_u^2 & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \end{cases} \]

the parameter estimates follow from

\[ \hat{\beta}_j = \frac{\tilde{\psi}_{uy}(j)}{\psi_{uu}(0)} \]

6.3 Accuracy; some error causes.

For practical use it is of prime importance to recognize the possible error causes. We will distinguish:

a) noise error
b) truncation error
c) state-correspondence error
d) errors due to simplification of the implementation
e) sampling error

a) noise error

This error is due to the noise $n$, added to the process output. We assume that the model may represent the process completely, cf. fig. 6.5:

$$y = u' b + n$$

$$w = u' \beta$$

with $b$ and $\beta$ of the same dimension. For that situation we have already derived the least squares equation (6.53):

$$\text{Cov} [\hat{\beta}] = \mathbb{E} [ (\hat{\beta} - \beta) (\hat{\beta} - \beta)' ] = [u'u]^{-1} u' N u' [u'u]^{-1}$$

(6.61)

with

$$N = \mathbb{E} [n'n']$$

Note that eq. (6.61) tells the complete story with respect to the noise error, based on a priori knowledge about the noise and a posteriori knowledge (measurements) of the process input signal $u$. The experimenter can obtain a clear picture of the accuracy of his estimates and the correlation between the errors in the components of $\hat{\beta}$. If necessary, an implementation of eq. (6.61) on the digital computer together with the estimator is possible. For the sake of conceptual simplicity we will consider some cases using a few simplifying assumptions.

If the noise is white and $\mathbb{E}[n] = 0$ then

$$N = \mathbb{E}[n|u] = \sigma_n^2 I$$

and

$$\text{Cov} [\hat{\beta}] = \sigma_n^2 [u'u]^{-1}$$

(6.62)

If in addition the noise is Gaussian, then this least squares estimation is identical with the maximum likelihood estimation; cf. Chapter 5. Consequently the properties of the M.L.E. as summarized in that chapter hold, for example, $\hat{\beta}$ is asymptotically normal with mean $b$ and a covariance given by (6.62).
Now we assume further that the input \( u \) of the process has the character of stationary "white" noise with the a priori knowledge:

\[
\mathbb{E}[u] = 0
\]

\[
\mathbb{E}[u'u] = k \, \psi_u(0) \, I = k \, \sigma_u^2 \, I
\]

In this situation, the expectation operator refers to the stochastic variable \( u \). Again \( k \) samples are considered. This leads to

\[
\text{cov} \left[ \hat{\beta} \right] = \frac{\psi_{nn}(0)}{k \, \psi_u(0)} \, I = \frac{\sigma_n^2}{k \, \sigma_u^2} \, I
\]

For each parameter \( \beta_j \), one finds a standard deviation (spread):

\[
\sigma_{\beta_j} = \frac{\sigma_n}{\sqrt{k} \, \sigma_u}
\]

For \( k \to \infty \) the spread of the probability density curve goes to zero. It must be noticed, however, that due to the square root, the convergence is rather slow. \( \sigma_u^{-2} \) and \( \sigma_n^{-2} \) can be considered as a measure of the respective "powers." This expression for \( \sigma_{\beta} \) is independent of the number of parameters to be estimated.

Another way of approach to finding the error due to additive noise is the following as can be seen from eq. (6.38) and noting that \( y = U b + n \):

\[
\hat{\beta} = [U'U]^{-1} U'y = b + [U'U]^{-1} U'n
\]

or

\[
\Delta \beta = \hat{\beta} - b = [U'U]^{-1} U'n = P \, U'n \quad \text{with} \quad P = k[\beta_{ij}]
\]

Consequently

\[
\Delta \beta_j = \beta_{oj} \, \tilde{\psi}_u(o,k) + \cdots + \beta_{ij} \, \tilde{\psi}_u(i,k) + \cdots + \beta_{mj} \, \tilde{\psi}_u(m,k)
\]
If \( u \) has a white-noise character then \( p_{ij} = \tilde{\psi}_{uu}^{-1}(0, k) \approx \delta_{ij} \approx 0 \) for \( i \neq j \). Therefore the dominant term will be that with \( p_{jj} \) and
\[
\Delta \beta_j \approx \frac{\tilde{\psi}_{uu}(0, k)}{\tilde{\psi}_{uu}(0, k)} \quad (k \text{ large})
\]
(6.64)

Of course an extra noise error manifests itself when the ordinary least squares estimation is used in a situation where the additive noise is not white.

**b) truncation error**

If only part of the process parameters can be taken care of by the (conceptual) model, then for a good grasp of the problem it is advisable to consider the process to consist of two parts; cf. fig. 6.6.

The parameters (points of the impulse response) \( b_0 \ldots b_p \) can be estimated, the parameters \( b_{p+1} \ldots b_m \) are not considered.

This situation may occur:
- if the initial guess of the length of the impulse response is too restricted;
- if we are interested only in the initial (or any other restricted) part of the response;
- if the impulse response extends to infinity, i.e., if the process contains a pure integrating action.

In fig. 6.7 the model parameters \( \beta' = [\beta_0 \ldots \beta_p] \) correspond to \( \beta' = [b_0 \ldots b_p] \). The process output \( y \) consists of two parts
\[
y = y_1(b) + y_2(b)
\]
As far as the estimating procedure is concerned \( y_2 \), due to the "truncated" part of the process response, acts as additive disturbances. Clearly \( y_1 \) and \( y_2 \) may be correlated as they both originate from \( u \).

For studying the effect of the truncation we put:
\[ y = U \hat{b} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \text{ with } \hat{b}_1 = \begin{bmatrix} b_1^* \\ b_2^* \end{bmatrix} \text{ and } \hat{b}_2 = \begin{bmatrix} b_m^1 \\ b_m^2 \end{bmatrix} \]

or
\[ y = U_1 \hat{b}_1 + U_2 \hat{b}_2 \]
\[ U_1 y = U_1 U_1 \hat{b}_1 + U_1 U_2 \hat{b}_2 \]

The estimator for the "truncated" part of the process is
\[ \hat{\beta} = \left[ U_1 U_1^T \right]^{-1} U_1 y \]

Equating this expression with eq. (6.65) leads to:
\[ \Delta \beta = \hat{\beta} - \beta = \left[ U_1 U_1^T \right]^{-1} U_1 U_2 \beta_2 \]

If the input \( u \) has the character of white noise then \( \mathbb{E}[U_1 U_2^T] = 0 \)
and \( \mathbb{E}[\Delta \beta] = 0 \). For other input signals this expectation may be unequal to zero and the estimates can be biased.

For the covariance of \( \Delta \beta \):
\[ \text{cov}[\Delta \beta] = \mathbb{E}\left[ \left[ U_1 U_1^T \right]^{-1} U_1 U_2 \beta_2 \right] \left[ U_1 U_1^T \right]^{-1} U_1 U_2 \beta_2 \]

From these equations and from physical insight it is clear that \( \mathbb{E}[\Delta \beta] \)
and \( \text{cov}[\Delta \beta] \) are independent of the level ("power") of the input signal.

A case of special interest is that in which the impulse response of the process has an asymptotic value \( h(\infty) \neq 0 \), cf. fig. 6.8.

The parameters \( h_1 \) ought to be determined from
\[ \hat{\beta} = \left[ U_1 U_1^T \right]^{-1} U_1 y_1 = \left[ U_1 U_1^T \right]^{-1} U_1 (y - y_2) \]

However, \( y_2 \) is not known beforehand and \( y \) has to be used. Consequently, errors will occur. No matter how many samples are being used in the estimation, the uncertainty with respect to these estimates never converges to zero. It turns out, that for each parameter \( \beta_j \) with \( j \geq p \) holds (Melis, 1967):
\[ \sigma_j = \frac{h(\infty)}{\sqrt{2}} \]
If we are going to use $\chi$ then the last one of the estimated parameters, $\beta_p$ in fig. 6.8, may provide us with some knowledge about $\underline{h}_2$, albeit with an error due to truncation. It was assumed that $\beta_j = \beta_p$ for $j \geq p$. Therefore, $x_2$ is known with some uncertainty. This implies that we (partially) can compensate $x$ by $x_2$ and derive a better estimate for $\beta_j$, etc.

Another way to handle this situation is indicated in fig. 6.9, where an integrator (summer) is chosen to represent the "tail" of the impulse response. In an analogous way one might take care of impulse response components due to large time constants by including in the (conceptual) model a low pass filter instead of the integrator.

Note the difference between this scheme with that given by Levin (1959) who places an integrator before the delay time.

The truncation errors in the estimated parameters are highly correlated and almost equal in magnitude and sign. This holds especially for large $k$.

From this it follows that for large $k$ the shape of the impulse response is measured adequately, but that it cannot be indicated how much each estimated point differs from the corresponding point of the true impulse response (Melis, 1967).

c) state-correspondence error

When starting an observation on the input $u$ and the output $y$ of a process the state vector need not be zero. This results in contributions to the first few samples of $y$ which are not the result of the input samples $u$.

Chosing again the delay time representation, one finds

process:  \[ y = \underline{U}_p \beta + \mu \]

model :  \[ \nu = \underline{U}_m \]
If at $t = 0$ the state of the process is zero, then all elements in the triangle $A$ are zero; if not then they contribute to $y(1) \ldots y(m)$. The same holds for the end of the observation period; if we have control over the input signal $u$ and we would take $u(j) = 0$ for $j > k$, then the triangle $B$ will consist of zeros. If this control on the input signal is not available then the elements in $B$ add contributions to the last few samples of $y$, which are not relevant for our estimation. This results in an extra variance of the parameter estimates. Of course this error is only of some importance when handling small series of samples for estimating purposes.

This effect can be diminished in the following way. The least square estimation is based on minimizing

$$E = e' e = \sum e^2(u)$$

When starting the observation we want to take the first few error samples "less seriously" than the samples later on. Near the end of the observations we want to do the same with the last few error samples. This can be effected by
minimizing (Draper and Smith, 1966):

\[
E = e'R e \quad \text{with} \quad R = \begin{bmatrix}
  r_1 & 0 \\
r_2 & r_3 \\
  & \ddots \\
  0 & & \ddots & r_k
\end{bmatrix}
\]

with \( r_1 < r_2 < r_3 \ldots r_i = r_j \ldots r_{k-2} > r_{k-1} > r_k \). This results in an estimate given by

\[
\hat{\beta} = [U'U]^{-1} U'R y
\]

(6.71)

Another approach to his problem is taken by estimating the parameters and the state at \( t = 0 \) simultaneously.

d) **errors due to simplification of the implementation**.

We have seen that the estimate

\[
\hat{\beta} = [U'U]^{-1} U'y
\]

(6.72)
is unbiased. The matrix inversion and its multiplication with a matrix is unpleasant from the computational point of view. Consequently it is to know of interest whether it is permissible to simplify eq. (6.72) to

\[
\beta = \frac{U'y}{k \tilde{\psi}_u(i,k)} \quad \text{or} \quad \beta_j = \frac{\tilde{\psi}_y(j,k)}{\tilde{\psi}_u(i,k)}
\]

(6.73)
particularly as we know from section 6.11 that this expression is valid when \( u \) has the character of white noise. From eq. (6.72)
it follows that

\[ \hat{\beta}_j = \frac{\tilde{\psi}_{uv}(j,k)}{\tilde{\psi}_{uu}(0,k)} - \frac{1}{\tilde{\psi}_{uu}(0,k)} \left\{ (\hat{\beta}_{j-1} + \hat{\beta}_{j-1}) \tilde{\psi}_{uv}(j,k) + \cdots \right\} \]  

(6.74)

As \( \beta_j \) determined without using matrix inversion gives only the first term, it follows that the second term indicates the bias due to this simplified instrumentation. (Van den Boom and Melis, 1969).

Another point of interest is to compare the variance of the estimates determined with and without matrix inversion.

With matrix inversion and white additive noise we have already found that:

\[ \text{cov} [\hat{\beta}] = \left[ \mathbf{U}' \mathbf{U} \right]^{-1} \mathbf{U}' \mathbf{N} \left[ \mathbf{U}' \mathbf{U} \right]^{-1} = \sigma_n^2 \left[ \mathbf{U}' \mathbf{U} \right]^{-1} \]  

(6.75)

Without matrix inversion the estimator is:

\[ \hat{\beta} = \frac{\mathbf{U}' \mathbf{Y}}{k \tilde{\psi}_{uu}(0,k)} = \frac{\mathbf{U}' \mathbf{b} + \mathbf{U}' \mathbf{n}}{k \tilde{\psi}_{uu}(0,k)} \]

In general this will be biased, as in most cases:

\[ \mathbb{E}[\hat{\beta}] = \frac{\mathbf{U}' \mathbf{Y}}{k \tilde{\psi}_{uu}(0,k)} \neq I \hat{\beta} \]  

(6.76)

Consequently the covariance is:

\[ \text{cov} [\hat{\beta}] = \mathbb{E} [\| \hat{\beta} - \frac{\mathbf{U}' \mathbf{b}}{k \tilde{\psi}_{uu}(0,k)} \|^2] = \frac{\mathbf{U}' \mathbf{N} \mathbf{U}}{k^2 \tilde{\psi}_{uu}(0,k)} \]  

(6.77)

and for white noise:

\[ \text{cov} [\hat{\beta}] = \frac{\sigma_n^2 \mathbf{U}' \mathbf{U}}{k^2 \tilde{\psi}_{uu}(0,k)} \]  

(6.78)

If the input \( \mathbf{u} \) has the character of white noise then for large \( k \)
the covariance has the same expression as in the case with matrix inversion; cf. eq. (6.63).
Consequently, a comparison of the explicit method with and without matrix inversion results in the following list of properties:

<table>
<thead>
<tr>
<th>with matrix inversion</th>
<th>without matrix inversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) the estimate is linear $\hat{\theta} = 0 \nu$</td>
<td>same</td>
</tr>
<tr>
<td>b) the estimate is unbiased $E[\hat{\theta}] = b$</td>
<td>{the estimate is biased, unless the input signal is white and $k$ large enough; eq. (6.76).}</td>
</tr>
<tr>
<td>c) $\lim_{k \to \infty} \text{cov}[\hat{\theta}] = 0 \text{I}$</td>
<td>same</td>
</tr>
<tr>
<td>d) suited for all types of input signals, provided all modes are excited</td>
<td>requires only a matrix-vector multiplication if $u$ is white and $\tilde{\psi}_{uu}(0,k)$ is approximately known beforehand.</td>
</tr>
<tr>
<td>e) suited for all types of additive noise uncorrelated with $u$.</td>
<td>same</td>
</tr>
<tr>
<td>f) requires matrix inversion, matrix-matrix multiplication and two matrix-vector multiplications.</td>
<td>requires only a matrix-vector multiplication</td>
</tr>
<tr>
<td>g) with $u$ white, $n$ white then $\text{cov}[\hat{\theta}] \approx \frac{\sigma_n^2}{k \sigma_u^2} \text{I}$ (6.63)</td>
<td>{same}</td>
</tr>
</tbody>
</table>

**e) sampling error**

If the process under consideration is continuous (non-sampled), then one has to be aware of the influence of the choice of the sampling-interval length; cf. chapter 7.
6.4. Residuals, noise properties and model order

In the previous sections we derived estimators for the parameters and for their covariance matrices. In practice, however, there are often more aspects of interest that may not be known beforehand, e.g., the properties of the noise (disturbances) and the number of the model parameters needed.

For such type of knowledge we may study the residuals:

\[ e = y - U\hat{\beta} = U(b - \hat{\beta}) + n \]  

(6.79)

The residuals carry information on both the parameter-vector-correspondence and the additive noise.

If it is not known beforehand how many parameters are significant, then it is of interest to have available an estimation procedure that is recursive in the number of parameters. This means that, starting with an estimate of the parameter vector \( \hat{b}_m' = [b_0, \ldots, b_{m-1}] \), it is not necessary to do all the calculations again for estimating the parameter vector \( \hat{b}_p' = [b_0, \ldots, b_{p-1}] \) with \( p > m \).
Using the notation $b'_p = \left[ b'_m ; b'_{p-m} \right]$ we find

$$y = \left[ U'_m ; U'_{p-m} \right] \begin{bmatrix} b'_m \\ b'_{p-m} \end{bmatrix} + \eta \quad (6.80)$$

which leads to the following least squares equations:

$$\begin{bmatrix} U'_m U'_m & - \\ - & 0 \end{bmatrix} \begin{bmatrix} \hat{\beta}_{m,m} \\ \hat{\beta}_{m,p} \end{bmatrix} = \begin{bmatrix} U'_m y \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} U'_m U'_m & U'_m U'_{m-p} \\ U'_{m-p} U'_m & U'_{m-p} U'_{m-p} \end{bmatrix} \begin{bmatrix} \hat{\beta}_{m,p} \\ \hat{\beta}_{m,m} \end{bmatrix} = \begin{bmatrix} U'_m y \\ U'_{m-p} y \end{bmatrix}$$

$\hat{\beta}_{m,m}$ and $\hat{\beta}_{m,p}$ are the old and the new estimates of $b_m, \ldots, b_{m-p}$, respectively, that is, before and after the increase of the number of parameters from $m$ to $p$, using the same set of observations. These equations can be rewritten as:

$$U'_m U'_m \left[ \hat{\beta}_{m,p} - \hat{\beta}_{m,m} \right] + U'_m U'_{m-p} \hat{\beta}_{m,p} = 0$$

$$U'_{m-p} U'_m \left[ \hat{\beta}_{m,p} - \hat{\beta}_{m,m} \right] + U'_{m-p} U'_{m-p} \hat{\beta}_{m,p} = U'_{m-p} \left[ y - U'_m \hat{\beta}_{m,m} \right]$$

Solving for $\hat{\beta}_{m,p}$ and $\hat{\beta}_{m,p}$ leads to:

$$\hat{\beta}_{m,p} = \hat{\beta}_{m,m} - B C U'_{m,p} \left[ y - U'_m \hat{\beta}_{m,m} \right]$$

$$\hat{\beta}_{m,p} = C U'_{m-p} \left[ y - U'_m \hat{\beta}_{m,m} \right] \quad (6.81)$$

with

$$A = \begin{bmatrix} U'_m U'_m \end{bmatrix}^{-1}$$

$$B = A U'_m U'_{m-p}$$

$$C = \begin{bmatrix} U'_{m-p} U'_{m-p} - U'_{m-p} U'_m A U'_m U'_{m-p} \end{bmatrix}^{-1}$$
The term \( y - \hat{U}_m \hat{\beta}_{m,m} \) represents the residuals of the initial model with parameters \( \beta_0, \ldots, \beta_{n-1} \). Note that \( A = [\hat{U}_m' \hat{U}_m]^{-1} \) is known from the previous estimation and that the matrix to be inverted is only of the order \((m-p)\). Consequently, \( \hat{\beta}_{m;p} \) and \( \hat{\beta}_{m-p;p} \) can be computed quite easily.

For the next estimation one also needs the inverse of the augmented \( \hat{U}'\hat{U} \) matrix. This can be shown to satisfy the relationship.

\[
\begin{bmatrix}
\hat{U}'_m & \hat{U}_m \\
\hat{U}'_{m-p} & \hat{U}_{m-p}
\end{bmatrix}
\begin{bmatrix}
\hat{U}'_m & \hat{U}_m \\
\hat{U}'_{m-p} & \hat{U}_{m-p}
\end{bmatrix}^{-1} = \begin{bmatrix}
A - BC & B' - BC \\
-C'B & C
\end{bmatrix}
\] (6.82)

Consequently, it consists of the matrices that have been determined already.

In this way, it is possible to calculate the estimates for models with different numbers of parameters. We have, however, no answer yet to the question what number of parameters (what model order) actually is needed. Generally speaking, an increase of the number of model parameters will result in a decrease of the error- or loss function \( E = e'e \). Actually \( E \) may become zero if the number equals \( k \), the number of samples from input and output signals. In that case, the noise contribution is incorporated completely in the model; this, of course, is undesirable. Therefore, we have to use a statistical test for determining whether a decrease of \( E \), when increasing the number of parameters, is significant.

The idea to view the test of order as a decision problem has been discussed by Anderson (1962), it is also a standard tool in regression analysis. The following derivation of such a test is taken from Åström (1968). Assume the process to be

\[
y = \hat{U}_m b_m + \eta
\] (6.83)
where \( y \) consists of \( k \) sample values, \( b_m \) has \( m \) parameters (e.g., \( b_{0, \ldots, b_{m-1}} \))

and where \( n \) is independent normal \((0, \sigma)\). Based on a model of order \( p \) the estimate \( \hat{\beta}_p \) is:

\[
\hat{\beta}_p = \left[ U_p' U_p \right]^{-1} U_p' y = \begin{bmatrix} \hat{b}_{m} \\ \Omega_{m,m} \end{bmatrix} + \left[ U_p' U_p \right]^{-1} \hat{e}_p'
\]

(6.84)

and the residuals:

\[
\begin{align*}
E_p &= y - U_p \hat{\beta}_p \\
&= U_m \hat{b}_n + U_p \left[ \begin{array}{c} \hat{b}_n \\
\Omega_{m,m} \end{array} \right] - \left[ U_m \hat{b}_n \right] U_p' \left[ \begin{array}{c} \hat{b}_n \\
\Omega_{m,m} \end{array} \right] - U_p U_p' \hat{e}_p = \\
&= n - U_p \left[ U_p' U_p \right]^{-1} U_p' n = \left[ I - A \right] n 
\end{align*}
\]

(6.85)

\( A \) is a symmetric idempotent matrix, i.e.,

\[
A = A^2 = A^3 = \ldots
\]

the same holds for the matrix \( \left[ I - A \right] \); cf. Appendix C. Consequently:

\[
E_q = E_p' E_p = n' \left[ I_k - U_p \left[ U_p' U_p \right]^{-1} U_p' \right] n
\]

(6.86)

Now if we consider two models with \( p \) and \( q \) (\( p < q \)) as the numbers of parameters, then the following trivial identity can be considered:

\[
E_q + (E_p - E_q) + \left( n' I_n - E_q \right) = n' \left[ I_k \right]^{-1} n
\]

The quantities at the left hand side of the equation are non negative definite quadratic forms in \( e_q, e_p \) and \( n \). Using the knowledge about the corresponding matrices:

- symmetric, non negative definite;
- which, by an orthogonal transformation can be brought to diagonal form having only 0's and 1's on the diagonal:
- for which the rank = trace (invariant):
By Cochran's theorem (c.f. Kendell and Stuart, vol. 2, pag. 360 ff), each of 
\( E_q, E_p - E_q \) and \( \frac{n'r}{r} - E_p \) is distributed as \( \chi^2 \) and they are mutually independent. Consequently, if \( q > p \geq m \) then \( E_q \) and \( E_p - E_q \) are independent random variables having \( \chi^2 \) distributions with \( k-q \) and \( q-p \) degrees of freedom, respectively.

Now a test of whether the error- or loss function is significantly reduced by increasing the number of parameters from \( p \) to \( q \) can be made by

\[
\tau = \frac{(E_p - E_q)/\chi^2}{E_q/(k-q)}
\]

(6.87)

This has an \( F_{k-q, q-p}^{q-p} \) distribution, which is tabulated extensively; e.g., Fraser (1958, pag. 390 ff). From such tables, at the 5% level:

\[
\begin{align*}
F_{120}^1 &= 3.92 & F_{\infty}^1 &= 3.84 \\
F_{120}^2 &= 3.07 & F_{\infty}^2 &= 3.00 \\
F_{120}^3 &= 2.68 & F_{\infty}^3 &= 2.60
\end{align*}
\]

i.e.,

\[
\mathcal{P} \left( F_{120} < 3.92 \right) = 0.95
\]

Consequently, if \( \tau \) is greater than the \( F \) value for the case under consideration a higher-order model has to be used. In engineering situations, one is interested
in cases where $k \gg q$ and thus, it is of interest to remember that for $k \to \infty$

$$\frac{F_{q-p}}{k-q}$$

is asymptotically $\chi^2$, i.e.

$$\frac{F_{q-p}}{k-q} \Rightarrow \frac{\chi^2_{q-p}}{q-p}$$

Consequently, tables of the $\chi^2$ distribution can be used as well.
A perfect matching of model and process would result in \( e = n \). If one has a priori assumptions about the statistics of the noise then it is possible to test the residuals with respect to that assumption. If \( n \) is assumed to be "white", i.e., the sample values are statistically independent, then that property can be tested on \( e \) by correlation techniques; if \( n \) is assumed to be gaussian, then a \( \chi^2 \) test can be used as verification.

If the variance of the noise \( \sigma_n \) is not known beforehand, then this can be estimated by

\[
\sigma_n^2 = \frac{1}{k-m} e'e
\]  

(6.88)

where \( k \) is the number of samples and \( m \) is the number of parameters to be determined. This can be shown as follows (Rosenberg and Shen, 1963; Åström 1968).

From eq. (6.86) it follows that:

\[
\mathbb{E}[e'e_e] = \mathbb{E}[h'[I_k - U_m[U_m'U_m]^{-1}U_m']n] =
\]

\[
= \mathbb{E}[\text{trace } h'[I_k - U_m[U_m'U_m]^{-1}U_m']n] =
\]

\[
= \mathbb{E}[\text{trace } [I_k - U_m[U_m'U_m]^{-1}U_m']n'n'] =
\]

\[
= \sigma_n^2 \{ \text{trace } I_k - \text{trace } U_m[U_m'U_m]^{-1}U_m' \} = \sigma_n^2/(k-m)
\]

Consequently, an unbiased estimate of \( \sigma_n^2 \) is given by eq (6.88), i.e., by squaring the residuals and dividing by the number of degrees of freedom.

Note that in statistical literature expressions can also be found for confidence-intervals of parameters \( \beta_i \) and for tests of the hypothesis \( \beta_i = 0 \).
Some further points of interest. Hoppe (1965) describes a method in which the point by point measurements (samples) are replaced by the results of integration over short periods of time. This should be advantageous in situations in which the process output is corrupted by additive noise, the frequency of which is high compared to the sampling frequency.

If one has control over the input u then the interesting question arises how to choose \( U \) in order to make the expected errors as small as possible. This problem is briefly considered by Leonov (1967).

### 6.5 Extension to generalized models and processes with feedback

In view of the interesting properties and the convenient instrumentation (linearity-in-the-parameters and a quadratic error or loss function) it is tempting to try to extend the least squares or Markov estimation to the generalized model as introduced in chapter 2 and shown again in fig. 6.10.

In this case

\[
e(k) = y(k) + \sum_{j=1}^{n} \alpha_j y_j(k) - \sum_{j=1}^{m} \beta_j u_j(k)
\]

\[
e = y - \begin{bmatrix} u' \\ y \end{bmatrix} = y - \begin{bmatrix} u' \\ y \end{bmatrix} \begin{bmatrix} \beta \\ -\alpha \end{bmatrix} = y - \begin{bmatrix} u' \\ y \end{bmatrix} \begin{bmatrix} 0 \\ -\alpha \end{bmatrix}
\]

(6.89)

As the number of parameters in \( \alpha \) and \( \beta \) is not important, \( m = n \) will be used in the sequel. Each expression can easily be adapted to the situation where \( m \neq n \). For the sake of convenience, the term \( \beta_0 \) will be discarded as well.

Along the same lines as before by minimizing \( E = e'e \) the estimate is found to be:

\[
\hat{\Theta} = \left[ \Omega' \Omega \right]^{-1} \Omega' y
\]

(6.90)
We will soon notice, that this description has a severe disadvantage due to a bias in the estimates caused by the additive noise. Yet it pays to take this case a bit further, as it covers one of the most important representations of a linear, time invariant, discrete time process with one input and one output. This is obtained by choosing \( u_i(k) = u(k-i) \) and \( y_j(k) = y(k-j) \):

\[
x(k) + a_1 x(k-1) + \cdots + a_m x(k-m) = b_1 u(k-1) + \cdots + b_m u(k-m)
\]

\[
y(k) = x(k) + n(k)
\]

(6.91)

\[
y(k) + \alpha, y(k-1) + \cdots + \alpha_m y(k-m) - \beta_1 u(k-1) - \cdots - \beta_m u(k-m) = e(k)
\]

(6.92)

By introduction of the shift operator

\[
z y(k) = y(k+1)
\]

as before, then the process can be described by:

\[
A(z) x(k) = B(z) u(k)
\]

or

\[
A^*(z^{-1}) x(k) = B^*(z^{-1}) u(k)
\]

with

\[
y(k) = x(k) + n(k)
\]

For the model the corresponding expressions \( A, B, A^* \) and \( B^* \) with \( \alpha \) and \( \beta \) parameters will be used. Using the \( A^*, B^* \) notation, the generalized error can be interpreted by different kinds of models:

1) a "generalized" model:

\[
e(k) = A^*(z^{-1}) y(k) - B^*(z^{-1}) u(k)
\]

(6.93)

which is represented schematically in Fig. 6.11.
2) a "prediction" model

\[ e(k) = y(k) - w(k) \]  

(6.94)

where

\[ w(k) = \left\{ 1 - A^*(z^{-1}) \right\} y(k) + B^*(z^{-1}) u(k) \]  

(6.95)

This model is also linear-in-the-parameters. Its output can be interpreted as a prediction of \( y(k) \) based on previous observations: \( w(k) = \hat{y}(k|k-1) \); cf. fig. 6.12. Consequently, the generalized error can be considered as the difference between the actual output and its prediction according to eq. (6.95). Note that this type of description also holds for situations where there are no inputs, e.g., parametric time series analysis in the sense of fitting an autoregression. This has been discussed by Wold (1938) and Whittle (1963). Recent applications to EEG analysis have been given by Gersh (1969).

3) an "ordinary" model (and filtering of the error signal):

\[ e(k) = A^*(z^{-1}) \left[ y(k) - y_m(k) \right] \]  

(6.96)

where

\[ y_m(k) = \frac{B^*(z^{-1})}{A^*(z^{-1})} u(k) \]  

(6.97)

cf. fig. 6.13. Note that process and model can be represented as consisting of a forward part \( B^* \), viz. \( B^* \), and a backward part \( 1-A^* \), viz. \( 1-A^* \), as shown in fig. 6.14. Note also in this figure that the noise can be assumed to affect the process in different ways, e.g., as \( r(k) \) or as \( n(k) \). This implies as process descriptions:

<table>
<thead>
<tr>
<th>( r(k) \neq 0 )</th>
<th>( n(k) = 0 )</th>
<th>( r(k) = 0 )</th>
<th>( n(k) \neq 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>in ( u ) and ( x )</td>
<td>( A^* x(k) = B^* u(k) + r(k) )</td>
<td>( A^* x(k) = B^* u(k) )</td>
<td>( A^* x(k) = B^* u(k) )</td>
</tr>
<tr>
<td>( y(k) = x(k) )</td>
<td>( y(k) = x(k) + n(k) )</td>
<td>( y(k) = x(k) + n(k) )</td>
<td></td>
</tr>
<tr>
<td>in ( u ) and ( y )</td>
<td>( A^* y(k) = B^* u(k) + r(k) )</td>
<td>( A^* y(k) = B^* u(k) + A^* n(k) )</td>
<td></td>
</tr>
</tbody>
</table>
Depending on the type of model and type of application, one may prefer one of these descriptions above the other.

For the interpretations just given, the estimator:

\[
\hat{\theta} = \left[ \Omega', \Omega \right]^{-1} \Omega' y
\]  

(6.98)

consists of the following components:

\[
\hat{\theta}' = \left[ \hat{\beta}', -\hat{\alpha}' \right] = \left[ \hat{\beta}_1, \ldots, \hat{\beta}_m, -\hat{\alpha}_1, \ldots, -\hat{\alpha}_m \right]
\]

\[
y = \begin{bmatrix}
y(1) \\
y(2) \\
\vdots \\
y(k)
\end{bmatrix}, \quad \Omega = \begin{bmatrix}
u(0) & u(-1) & \ldots & u(1-m) & y(0) & y(-1) & \ldots & y(1-m) \\
u(1) & u(0) & \ldots & u(2-m) & y(1) & y(0) & \ldots & y(2-m) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
u(k-1) & \ldots & \ldots & u(k-m) & y(k-1) & y(k-2) & \ldots & y(k-m)
\end{bmatrix}
\]

and the corresponding expressions for \( \Omega' \Omega \) and \( \Omega' y \).
These expressions may be programmed in a straightforward way for a digital computer. Starting from the observations \( u \) and \( u \), such a program results in numerical values for \( \hat{\theta} \) and, possibly, for the covariance matrix. As indicated before, this is an explicit or "one shot" method where all observations are processed in one sequence.

There is a close relation between this method and the well known correlation technique for determining points of the impulse response. In that case the following (approximations of) correlation coefficients are calculated:

\[
\begin{align*}
\widehat{\eta}_{uu}(j) &= \frac{1}{k-j} \sum_{i=1}^{k-j} u(i) u(i+j) \\
\widehat{\eta}_{yy}(j) &= \frac{1}{k-j} \sum_{i=1}^{k-j} y(i) y(i+j) \\
\widehat{\eta}_{uy}(j) &= \frac{1}{k-j} \sum_{i=1}^{k-j} y(i) u(i+j) \\
\widehat{\eta}_{yu}(j) &= \frac{1}{k-j} \sum_{i=1}^{k-j} u(i) y(i+j)
\end{align*}
\]  

(6.99)

Comparing with the least squares parameter estimation we find that the elements of the matrices \( \Omega'\Omega \) and \( \Omega'\gamma \) consist essentially of auto and cross correlations, if the beginning and the end of the observation series are neglected.

Notice that the technique of this section can immediately be applied to the identification of nonlinear processes which are linear-in-the-parameters e.g.

\[ y(k) + a y(k-1) = b_1 u(k-1) + b_2 u^2(k-1) \]

The bias problem. It was mentioned previously that the least squares estimation, using a generalied model, suffers from a bias in the estimates. This is due to the fact, that in the \( \Omega \) matrix of observations the y elements are contaminated with noise:
\[ \Omega = \begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} \Omega_{y} \\ \Omega_{x} \end{bmatrix} + \begin{bmatrix} n(0) & \cdots & n(1-m) \\ \vdots & & \vdots \\ n(k-1) & \cdots & n(k-m) \end{bmatrix} \]

or in an obvious shorthand notation

\[ \Omega = \Omega_{u,y} = \Omega_{u,x} + \Omega_{0,n} \]  

(6.100)

This contamination is contrary to the assumptions under which the least squares estimator was derived. The reader may convince himself that a bias is bound to occur, except in very special cases, by considering the simple example of fig. 6.15, which is described by:

\[ \begin{align*}
    x(k+1) + a x(k) &= b u(k) \\
    y(k) &= x(k) + n(k) \\
    e(k) &= y(k) + \alpha y(k-1) - \beta u(k-1)
\end{align*} \]

process

model

The estimate follows from a (conceptual) model adjustment, giving a minimization of the loss function

\[ E = \sum_{i=1}^{k} e(i)^2 \]

with the necessary conditions

\[ \begin{align*}
    \frac{\partial E}{\partial \alpha} &= \sum_{i=1}^{k} e(i) y(i-1) = 0 \\
    \frac{\partial E}{\partial \beta} &= \sum_{i=1}^{k} e(i) u(i-1) = 0
\end{align*} \]

These can be written as

\[ \begin{align*}
    \sum \left\{ \left\{ x(i) + \alpha x(i-1) - \beta u(i-1) + n(i) + \alpha n(i-1) \right\} \left\{ x(i-1) + n(i-1) \right\} \right\} &= 0 \\
    \sum \left\{ \left\{ x(i) + \alpha x(i-1) - \beta u(i-1) + n(i) + \alpha n(i-1) \right\} \left\{ u(i-1) \right\} \right\} &= 0
\end{align*} \]

\[ (6.101) \]

These equations are fulfilled for \( \alpha = a \) and \( \beta = b \), i.e., unbiased estimates, if:
\[ n(i) = 0 \quad i = 1, \ldots, k \]

i.e., there is no noise present.

\[ \sum_{i} n(i) n(i-1) + a \sum_{i} n^2(i-1) = 0 \]

which is approximated by the respective points of the autocorrelation function:

\[ \psi_{nn}(1) + a \psi_{nn}(0) = 0 \]

i.e., there is noise which, by passing through the part of the generalized model, provides 'white' residuals.

In all other cases there will be a bias. To see how this works out consider an example:

Assume that the process is actually described by the model

\[ \dot{y}(k) - 0.5 y(k-1) = 1.0 u(k-1) + n(k) + 0.1 h(k-1) \]

where \( n(k) \) is a sequence of independent normal \((0,1)\) random variables, but that the system is identified using the least squares method under the assumption that the residuals are uncorrelated. Below we give a typical result obtained from 500 pairs of inputs and outputs.

<table>
<thead>
<tr>
<th>process parameters</th>
<th>estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a = -0.5 )</td>
<td>( \hat{a} = -0.643 \pm 0.029 )</td>
</tr>
<tr>
<td>( b = 1.0 )</td>
<td>( \hat{\beta} = 1.018 \pm 0.062 )</td>
</tr>
</tbody>
</table>

We are consequently in a very bad situation; not only is the estimate \( \hat{a} \) wrong but we have also a great deal of confidence in the wrong result. (The true value deviates from the estimate by about 5\( \sigma \).)
Now the properties of the estimate have to be studied into somewhat detail (Shaw, 1970). Substitution of the relations

\[
\frac{\Omega}{u_x} = \frac{\Omega}{u_y} - \frac{\Omega}{o_n}
\]

and

\[
\mathbf{y} = \frac{\Omega}{u_x} \mathbf{t} + \mathbf{n}
\]

where \( \mathbf{x}' = [b; - \mathbf{a}] \) is the total process parameter vector, in the expression for the estimator (6.98) leads to:

\[
\hat{\mathbf{\theta}} = \left[ \frac{\Omega}{u_y} \frac{\Omega}{u_y} \right]^{-1} \left[ \frac{\Omega}{u_y} \frac{\Omega}{o_n} \right] \mathbf{y} = \left[ \frac{\Omega}{u_y} \frac{\Omega}{u_y} \right]^{-1} \left[ \frac{\Omega}{u_y} \frac{\Omega}{o_n} \mathbf{t} - \frac{\Omega}{o_n} \mathbf{t} + \mathbf{n} \right]
\]

(6.102)

The expectation is:

\[
E \left[ \hat{\mathbf{\theta}} \right] = \mathbf{t} - E \left[ \frac{\Omega}{u_y} \frac{\Omega}{u_y} \right]^{-1} \left[ \frac{\Omega}{u_y} \frac{\Omega}{o_n} \mathbf{t} - \frac{\Omega}{o_n} \mathbf{n} \right]
\]

(6.103)

The second right hand term indicates the bias of the estimate. Now we study the asymptotic properties of this estimator for increasing number of observations; \( k \to \infty \). Consider

\[
\left[ \frac{\Omega}{u_y} \frac{\Omega}{u_y} \right]^{-1} \left[ \frac{\Omega}{u_y} \frac{\Omega}{o_n} \mathbf{t} - \frac{\Omega}{o_n} \mathbf{n} \right]
\]

where the bias term is multiplied and divided by \( k \). If \( u \) and \( n \) are well behaved stationary random processes, then the sample averages will be consistent estimators of the auto and cross correlations. Using the notion of "limit-in-probability" (p lim) \( \|$ and a theorem due to

\[\text{The non-random vector } \mathbf{c} \text{ is the p lim of } \mathbf{x}(k) \text{ for } k \to \infty, \text{ i.e.} \]

\[\text{p lim } \mathbf{x}(k) = \mathbf{c} \]

if

\[\lim_{k \to \infty} P \left( \left| \mathbf{x}(k) - \mathbf{c} \right| \geq \varepsilon \right) = 0 \quad \text{for each } \varepsilon > 0\]

This implies
\[
\lim_{k \to \infty} \mathbb{E}[x(k)] = \xi
\]

Slutsky if it follows from equation (6.102) that:

If the elements of \( A(k) \) and \( B(k) \) converge in probability then

\[
p \lim A^{-1}(k) B(k) = [p \lim A(k)]^{-1} p \lim B(k)
\]

cf., Elgerd (1967, p. 118)

\[
p \lim \hat{\Theta} = \xi - \left( p \lim \frac{\Omega'_{u,y} \Omega_{u,y}}{k} \right)^{-1} \left( p \lim \frac{\Omega'_{u,y} \Omega_{o,n}}{k} \right) = p \lim \frac{\Omega'_{u,y} \Omega_{u,y}}{k} = \varpi
\]

where:

\[
p \lim \frac{\Omega'_{u,y} \Omega_{u,y}}{k} = \varpi
\]

Assuming that the covariance matrix \( \varpi \) is positive definite:

\[
\lim_{k \to \infty} \mathbb{E}[\hat{\Theta}] = p \lim \hat{\Theta} = \xi - \varpi^{-1} \left[ \begin{array}{c|c}
\mathbb{E}[\Omega_{u,y}] & 0 \\
\hline
0 & \mathbb{E}[\Omega_{o,n}]
\end{array} \right] = \left[ \begin{array}{c}
0 \\
0 \\
\mathbb{E}[\Omega_{o,n}(1-m)] \\
\mathbb{E}[\Omega_{o,n}(0)]
\end{array} \right]
\]

(6.105)
Consequently, there is an asymptotic bias unless the term in brackets is zero.

This asymptotic bias does not exist if:

$$\frac{N}{a} = - \begin{bmatrix} \psi_n(1) \\ \psi_n(n) \end{bmatrix}$$  \tag{6.106}$$

i.e. if \( n(k) \) is generated from white noise by

$$n(k) + \alpha_1 n(k-1) + \ldots + \alpha_m n(k-m) = w(k)$$

This can be seen from multiplication by \( n(k-i) \) and taking the expectation:

$$E[n(k) n(k-1) + \alpha_1 n(k-1) n(k-2) + \ldots + \alpha_m n(k) n(k-m)] = E[w(k) n(k-i)]$$

$$E[n(k) n(k-m) + \alpha_1 n(k) n(k-m) + \ldots + \alpha_m n(k-m) n(k-m)] = E[w(k) n(k-i)]$$

Here the right hand side is zero, as \( n(k-i) \) can depend only on previous samples of \( w(k) \). Therefore, this expression is identical to eq. (6.106)

This condition for an unbiased estimate corresponds to the condition of "white residuals", as in the notation of eq. (6.93) for the right parameter values \( \alpha^*(z^{-1}) = A^*(z^{-1}) \) and \( \beta^*(z^{-1}) = B^*(z^{-1}) \):

- model:\n  \( e(k) = A^*(z^{-1}) x(k) + A^*(z^{-1}) n(k) - B^*(z^{-1}) \omega(k) \)
- process:\n  \( 0 = A^*(z^{-1}) x(k) - B^*(z^{-1}) \omega(k) \)

\[ \begin{align*}
  e(k) &= A^*(z^{-1}) n(k) \\
  &= n(k) + \alpha_1 n(k-1) + \ldots + \alpha_m n(k-m) = w(k)
\end{align*} \]

This phenomenon can be expressed by the moving average

$$n(k) + \alpha_1 n(k-1) + \ldots + \alpha_m n(k-m)$$

whitens the noise generated by

$$n(k) + \alpha_1 n(k-1) + \ldots + \alpha_m n(k-m) = w(k)$$

This is illustrated in fig. 6.15.
The condition of "white residuals" can also be approached along the following lines. Consider the representation of process and environment given in fig. 6.17
\[ A(z^-) x(k) = B(z^-) u(k) \]
\[ y(k) = x(k) + n(k) \]
Consequently,
\[ A^*(z^-) y(k) = B^*(z^-) u(k) + A^*(z^-) n(k) \] (6.107)
If \( n(k) \) is derived from a "white" noise \( w(k) \) through
\[ D^*(z^-) n(k) = F^*(z^-) w(k) \] (6.108)
then substitution of (6.108) in (6.107) with \( C^*(z^-) = A^*(z^-) F^*(z^-) \) results in:
\[ A^*(z^-) y(k) = B^*(z^-) u(k) + \frac{C^*(z^-)}{D^*(z^-)} w(k) \] (6.109)
By introducing
\[ \tilde{y}(k) = \frac{D^*(z^-)}{C^*(z^-)} y(k) \]
\[ \tilde{u}(k) = \frac{D^*(z^-)}{C^*(z^-)} u(k) \]
eq (6.107) can be written as:
\[ A^*(z^-) \tilde{y}(k) = B^*(z^-) \tilde{u}(k) + n(k) \] (6.110)
Thus, by using a generalized model operating on \( \tilde{u}(k) \) and \( \tilde{y}(k) \) the residuals are white and the estimates will be unbiased.

**Processes with feedback.** Related problems occur when the process to be identified is part of a closed loop; cf. fig. 6.18. Note that the noise is inside the loop. The characteristics of \( P \) are of interest. One has the freedom to consider either \( u \) or \( w \) as input signals.

In the first case, an estimation procedure based on \( u \) and \( y \), determines the overall characteristics of the closed loop system.
The parameter determination for the process $P$ from these overall characteristics is not a trivial problem; e.g., the error propagation from the closed loop description to the open loop one may lead to an accumulation of estimation errors.

In the second case, when the estimation procedure is based on $w$ and $y$, one has to face another problem. Due to the feedback, the input signal $w = (u+y)$ is not statistically independent of the disturbance $n$, which was recognized to be a necessary condition for obtaining unbiased estimates.

Consider as an example for the sampled case the determination of the sampled impulse response. The process is described by:

$$ x(k) = b_0 u(k) + b_1 u(k-1) + \ldots + b_n u(k-n) = \sum_{i=0}^{\infty} b_i u(k-i) $$

or

$$ x(k) = \sum_{i=0}^{\infty} b_i w(k-i) - \sum_{i=0}^{\infty} b_i y(k-i) $$

(6.111)

(6.112)

Because

$$ y(k) = x(k) + n(k) $$

the output can be found explicitly as:

$$ y(k) = \sum_{i=0}^{\infty} b_i w(k-i) - \sum_{i=1}^{\infty} b_i y(k-i) + \frac{1}{1+b_0} n(k) $$

(6.113)

where $b_i = b_i'/(1-b_0')$. The condition of "white residuals" can be studied as before.

Through the efforts to preserve the nice properties of the generalized model, that is the linearity-in-the-parameters, but to circumvent the detrimental effects of the additive noise, several techniques have been developed, viz:

a) repeated least squares
b) instrumental variables
c) tally principle
d) generalized least squares and related schemes
e) maximum likelihood
f) Levin's method
a) **Repeated least squares.** This method consists of increasing the order of the model

\[
e(k) = A(z^{-1}) y(k) - B(z^{-1}) u(k)
\]

(6.114)

in successive estimation runs. The correspondence of process and model is judged, based on the residual error. If an increase of model order does not have a significant improvement on the residual error, the estimation stops. It has been illustrated by Åström (1967) for a simple first order system, that a division of \( A(z^{-1}) \) by \( B(z^{-1}) \) results in a good model and that the rest of the division can be neglected. No formal proof of the validity of this method seems to be available yet.
b) **Instrumental variables.** A straightforward solution to the problem of bias, due to observations which are contaminated with noise, can be indicated in a simple way. From the process equation

\[ y = \alpha b + n \]  

(6.115)

through multiplication by \( \Omega' \)

\[ \Omega' y = \Omega' \alpha b + \Omega' n \]  

(6.116)

if \( \Omega \) and \( n \) are statistically independent, then the normal equation follows as

\[ \hat{\beta} = \left( \Omega' \Omega \right)^{-1} \Omega' y \]  

(6.117)

Now assume that there is a matrix \( V \) available which is correlated with \( u \) but not with \( n \):

\[ \lim [V' n] = 0 \quad \lim [V' \alpha] \quad \text{positive definite} \]

then

\[ V' y = V' \alpha b + V' n = V' \alpha \hat{\beta} \]

as \( n \) is unmeasurable, or:

\[ \hat{\beta} = \left( V' \alpha \right)^{-1} V' y \]  

(6.118)

which guarantees an unbiased estimate. The matrix \( V \) represents the so called "instrumental variables", cf. fig. 6.19. A particular application of this technique that strongly suggests itself is the case of a feedback path around the process under study; cf. fig. 6.18. As \( w \) is not correlated with \( n \) this can be used as instrumental variable. Of course such an arbitrary \( V \) does not guarantee an optimal estimate. It is also possible to find instrumental variables such that the estimate has optimal properties. Moreover, this type of estimator can be formulated in a recursive version.

The schemes proposed by Peterka and Šmuk (1969), Hsia and Landgrebe (1967) are closely related to the instrumental variable technique. Vušković, Bingulac and DJOROVIĆ (1970) indicate the use of pseudo sensitivity functions as instrumental variables.
c) **Tally principle.** This estimation procedure as described by Peterka and Halousková (1970) starts from the canonical form

\[ y(k) + a_1 y(k-1) + \ldots + a_m y(k-m) = b_1 u(k-1) + \ldots + b_n y(k-n) + \eta(k) \]  

(6.119)

where

\[ \eta(k) = \eta(k) + c_1 \eta(k-1) + \ldots + c_n \eta(k-n) \]  

(6.120)

and \{\eta(k)\} is a sequence of uncorrelated random variables. It is derived that:

\[
\begin{align*}
\mathbb{E}[\eta(i) u(i-j)] &= 0 \quad \text{for } j > 0 \\
\mathbb{E}[\eta(i) y(i-j)] &= 0 \quad \text{for } j > m
\end{align*}
\]  

(6.121)

For any estimate of the unknown parameters one finds a sequence \{\eta^*(i)\}. Now for a sequence of \( k \) observations

\[
\frac{1}{k} \sum_{i=1}^{k} \eta^*(i) u(i-j)
\]

and

\[
\frac{1}{k} \sum_{i=1}^{k} \eta^*(i) y(i-j)
\]

have to "tally" (duplicate) the corresponding expectations as well as possible in a least squares sense.

d) **Generalized least squares and related schemes;** these are discussed in chapter 7.

e) **Maximum likelihood;** this is discussed in chapter 11.

f) **Levin's method.**

A particular method for estimating the bias due to correlated residuals has been proposed by Levin (1959, 1960, 1964) for the case of a deterministic system with independent observation errors. Levin's results are based on a technique due to Koopmans and it gives the estimates in terms of an eigenvalue problem for the matrix \( \Omega^{-1} \Omega \). A careful analysis of Levin's method which includes convergence proofs and error estimates has been done by Aoki and Yue (1970). The method has also been used by Smith (1968).
6.6. Some computer procedures...
### References


see also *Control*, 8, 555


Levin, M.J. (1964). Estimation of system pulse transfer function in the presence of noise. IEEE Trans. autom. control, AC-9,


Detchmendy, D.M. and R. Shridhar ( ). On the experimental determination of the dynamical characteristics of physical systems.


Fagin, S.L. (1964). Recursive linear regression theory, optimal filter theory and error analysis of optimal systems. IEEE Convent Record, 12, part 1


Koopmans, T.C. (1950). Statistical inference in dynamic economic models. Cowless Commis- 


Watson, ( ). Serial correlation in regression analysis. Biometrika, 42, 327-


Wilks ( ). Mathematical Statistics

Chapter 7 SAMPLED SIGNALS; implicit or model-adjustment techniques.

In the previous chapter we derived the "normal equations" for the least squares and the Markov estimators. These can readily be programmed for execution on a digital computer. Some disadvantages, however, are:
- matrix inversion is a (computer-)time consuming operation.
- all sample values of input and output signals have to be stored in a fast memory in order to obtain acceptable computing time; this requires much memory space.
- this method does not offer the possibilities of "on-line" computation, i.e., taking into account new observations that may become available as time progresses and perhaps "forgetting" old observations.

Based on this third consideration the explicit method, discussed in the previous chapter, can also be characterized as an accumulative, "one-shot" or "one-step" procedure.

In many applications it is highly desirable to obtain the result of the identification as the process develops. For example it might be of interest to proceed until a specified parameter accuracy is achieved. The problem is then as follows. Assume that an estimate $\hat{\beta}(i)$ is obtained based on $i$ pairs of input-output samples. Is it necessary to repeat the whole identification procedure from the beginning, using the whole string of input/output data in order to obtain $\hat{\beta}(i+1)$ or is it possible to rearrange the computations using $\hat{\beta}(i)$? An identification scheme which does not require that the whole string of input/output data is brought in at each step is called recursive. Such schemes may be considered for applications where one wants to process observations as soon as they become available; on-line methods. Apart from being of practical interest, this point of view on identification problems will also make it possible to establish connections with other fields e.g., nonlinear filtering, stochastic
approximation, learning and adaption.

If the parameters of the process are truly time varying it is of course meaningless to do anything else but to track the parameters in real-time. This is called real-time identification.

There are many different ways to obtain algorithms for implicit estimation techniques. Practically all methods, however, will yield algorithms with a strong similarity and be of the form

\[
\hat{\beta}(i+1) = \hat{\beta}(i) + \Gamma(i) e(i)
\]

where \( e(i) \) is the (generalized) error discussed earlier and \( \Gamma(i) \) is a gain factor which can be of varying complexity.

For models which are linear-in-the-parameters it is quite simple to determine the gradient and to base the adjustment scheme on those gradients. This is discussed in section 7.1. When trying to use these methods for generalized models one meets with the problem of bias as discussed in the previous chapter. This implies the need for more extensive estimation schemes discussed in section 7.2. For models that are nonlinear-in-the-parameters special gradient-determination schemes have to be devised or direct-search methods have to be used; cf. section 7.3.
7.1. Models, linear - in - the - parameters

The recursive or iterative solution to an estimation problem can be characterized by the task of updating an old estimate, say \( \hat{\beta}(i) \). The presence of this estimate implies the need for a "model" (e.g., programmed on a digital computer) as shown in fig. 7.1a. The input is simultaneously fed to the process and to a model with adjustable parameters. The adjustable parameters are changed by an adjustment mechanism which receives the process output \( y_i \) and the model output \( \tilde{y}_i \) as inputs. This formulation of the on-line identification problem was first considered by Whitaker (1958).

The essential problem is to determine the adjustment mechanism such that the model parameters in some sense will be close to the process parameters. Note again that in this approach the \( \hat{\beta}(i) \) is imbedded in a closed loop, which implies self-correcting properties for several types of errors, cf. Chapter 2.

As indicated in chapter 4, the linearity-in-the-parameters leads to a simple case, even if the process has nonlinear dynamics. This section will be devoted to that type of model structure:

\[
\nu(k) = \sum \beta_j y_j(k) \quad \text{and} \quad \nu = \sum \beta
\]

\[
\epsilon = y - \sum \beta
\]

Iterative schemes may be derived at different levels of complexity; this will be discussed successively, viz:

a. corrections proportional to the gradient
b. least squares approach to strings of observations
c. least squares approach to single (pairs of) observations
d. stochastic approximation
e. contraction mapping
a. Corrections proportional to the gradient

Assume we are given is the diagram of fig. 7.1b. The number of parameters to be estimated is again known beforehand. \( \mathbf{u} \) and \( \mathbf{n} \) are statistically independent. The vector \( \hat{\beta}(i) \) represents the constant model-parameters during the \( i \)-th measuring interval. The next parameter vector can be chosen as the old vector corrected with a quantity proportional to the gradient of the error function:

\[
\hat{\beta}(i+1) = \hat{\beta}(i) - \frac{1}{i} \Gamma(i) \frac{\partial}{\partial \hat{\beta}} E_i
\]

with

\[
E_i = e_i' R e_i,
\]

\[
e_i = y_i - U_i \hat{\beta}(i),
\]

\[
\frac{\partial}{\partial \hat{\beta}} E_i = -2 U_i' R e_i
\]

cf. Appendix C

and where \( y_i \) and \( U_i \) have to be interpreted as the observations made during the \( i \)-th measurement interval; c.f. fig. 7.2. Consequently,

\[
\hat{\beta}(i+1) = \hat{\beta}(i) + \Gamma(i) U_i' R [y_i - U_i \hat{\beta}(i)]
\]

(7.2)

For the weighting matrix possible choices again include \( R = \mathbf{I} \) and \( R = N^{-1} \). The "gain factor" \( \Gamma(i) \) can be chosen constant or as a function of \( i \). The effects of this choice are explained in the following part of this section.
b. Least squares approach to strings of observations.

In the situation given by fig. 7.1b one might estimate the difference between the process and the model parameter vector:

$$\Delta \hat{\beta}(i) = \hat{\beta}(i) - \hat{\beta}(i) = \left[U_i'R U_i\right]^{-1} U_i'R e_i$$

i.e., by use of eq. (6.31) operating on $e_i$ instead of $y_i$. This implies that the process and the model partially compensate each other and that the parameter of the resulting dynamic system are estimated. Hence:

$$\hat{\beta}(i+1) = \hat{\beta}(i) + \Gamma(i) \Delta \hat{\beta}(i)$$

which yields:

$$\hat{\beta}(i+1) = \hat{\beta}(i) + \Gamma(i) \left[U_i'R U_i\right]^{-1} U_i'R \left[y_i - U_i \hat{\beta}(i)\right]$$  \hspace{1cm} (7.3)

Compare this result with eq. (7.2). Again $R = I$ and $R = N^{-1}$ are among the possible choices. Unless otherwise stated $R = I$ will be used in the following part.

The need of matrix inversion in each step of the iteration makes this algorithm still comparatively unpleasant for numerical solutions, although a much smaller number of samples is used per iteration than in the accumulative type of solution. The size of the $U_i'$ matrix, however, is not reduced.

If there is no noise $n$ present, then an adjustment scheme referred to as "learning identification" with good convergence properties can be used. In that case eq. (7.3) may be reduced to:

$$\hat{\beta}(i+1) = \hat{\beta}(i) + \Gamma \frac{U_i}{\|U_i\|^2} \left[y_i - U_i \hat{\beta}(i)\right]$$


In real-life estimating situations the additive noise is the limiting factor. Several properties of the algorithm eq. (7.3) have been derived by Van den Boom (1967); cf. also Van den Boom and Molis (1969).
The development in time of the expectation can be studied. Keeping in mind that $y_1 = U_1 + \eta_1$, it follows from eq. (7.3):

$$
E[\hat{\beta}(i+1)] = \Gamma(i) \cdot b + \{1 - \Gamma(i)\}E[\hat{\beta}(i)] \quad i = 1, 2, \ldots
$$

(7.4)

Consequently

$$
E[\hat{\beta}(2)] = \Gamma(1) \cdot b + \{1 - \Gamma(1)\}E[\hat{\beta}(1)]
$$

$$
E[\hat{\beta}(3)] = \Gamma(2) \cdot b + \{1 - \Gamma(2)\}\Gamma(1) \cdot b + \{1 - \Gamma(2)\}\{1 - \Gamma(1)\}E[\hat{\beta}(1)]
$$

As $\hat{\beta}(1)$ is the (known) initial model setting we may put $E[\hat{\beta}(1)] = \hat{\beta}(1)$. If no a priori knowledge about the process parameters is available a legitimate choice is $\hat{\beta}(1) = 0$.

Also of interest is the development in time of the covariance. For the sake of simplicity we choose $R = I$. As $U$ and $n$ have been assumed to be statistically independent:

$$
\hat{\beta}(i+1) = \Gamma(i) \cdot b + \{1 - \Gamma(i)\} \cdot \hat{\beta}(i) + \Gamma(i)[U_i, U_i]^T U_i, n
$$

$$
E[\hat{\beta}(i+1)] = \Gamma(i) \cdot b + \{1 - \Gamma(i)\}E[\hat{\beta}(i)] + 0
$$

$$
\hat{\beta}(i+1) - E[\hat{\beta}(i+1)] = \{1 - \Gamma(i)\}[\hat{\beta}(i) - E[\hat{\beta}(i)]] + \Gamma(i)[U_i, U_i]^T U_i, n
$$

If the noise $n$ has a wide band character then $\hat{\beta}(i)$ and $\eta_i$ can also be considered independent in most practical cases. Consequently

$$
\text{cov}[\hat{\beta}(i+1)] = \{1 - \Gamma(i)\}^2 \text{cov}[\hat{\beta}(i)] + \Gamma(i)[U_i, U_i]^T U_i, E[\hat{\beta}(i), \eta_i] U_i, E[\hat{\beta}(i), \eta_i]^T
$$

(7.5)

i.e., a vary-linear difference equation. The last term reduces to
\[ \Gamma(i) \sigma_n^2 \left[ U_i U_i^T \right]^{-1} \] for "white" additive noise n

\[ \approx \Gamma(i) \frac{\sigma_n^2}{\sigma_u^2} I \] for "white" additive noise n and "white" input signal u of sufficient length (j samples)

Again \( \hat{\beta}(1) \) is the initial model setting and is consequently known. This implies \( \text{cov} \left[ \hat{\beta}(1) \right] = 0 I \). From this initial condition the development of the covariance in time can be studied.

From equations (7.4) and (7.5) it is clear that the choice of \( \Gamma(i) \) is of crucial importance for the convergence properties. This "gain factor" may be chosen to be either a constant or a function of the adjustment number. In both cases a choice \( \Gamma(i) = 1 \) for any \( i = 1 \) implies according to eq. (7.4) that

\[ E \left[ \hat{\beta}(i+1) \right] = \beta \] (7.8)

for all \( i \geq 1 \). Consequently after that \( l \)-th adjustment the estimate is unbiased irrespective of the initial model setting.

Implementation of "closed loop" estimation schemes. Following Van den Boom and Melis (1969) we will consider the following cases:

a) \( \Gamma(i) = \frac{1}{i} \)

b) \( \Gamma(i) = c = \text{constant} \)

a) This case reminds us of the algorithms found for stochastic approximation; cf. Chapter 5. From eq. (7.5) and eq. (7.7) it follows that for "white" noise n and "white" input signal u:

\[ \text{cov} \left[ \hat{\beta}(i+1) \right] = \frac{\sigma_n^2}{\sigma_u^2 i} I \] (7.9)

and
\[
\lim_{i \to \infty} \text{cov} \left[ \hat{\beta} \left( i+1 \right) \right] = 0 \text{I} \quad (7.10)
\]

This can easily be shown by recognizing the pattern in the successive covariance expressions for \( \hat{\beta}(1), \hat{\beta}(2), \ldots \) etc. The term \( i.j \) in eq. (7.10) implies the total number of samples that have been used in all \( i \) measuring intervals together. A comparison of this expression with the one for the explicit method (eq. 6.63) shows that the covariance for both methods is equal. For rather short lengths of samples (\( j \) small) the relation

\[
\left[ U_i' U_i \right]^{-1} \approx \frac{1}{j \hat{u}^2} \text{I}
\]

does not necessarily hold. For such cases, experimental results show a shift towards greater variances; cf. fig. 7.3.

b) The case \( \Gamma(i) = c \) leads to

\[
\mathbb{E} \left[ \hat{\beta} \left( i+1 \right) \right] = c b + (1 - c) \mathbb{E} \left[ \hat{\beta} \left( i \right) \right] \quad (7.11)
\]

which can be written as:

\[
\mathbb{E} \left[ \hat{\beta} \left( i+1 \right) \right] = \left\{ 1 - (1 - c) i \right\} b + (1 - c)^i \hat{\beta}(1) \quad (7.12)
\]

Consequently,

\[
\lim_{i \to \infty} \mathbb{E} \left[ \hat{\beta} \left( i+1 \right) \right] = b \quad \text{for} \quad 0 < c < 2 \quad (7.13)
\]

i.e., the estimate is asymptotically unbiased; cf. fig. 7-4.

The successive covariance expressions for \( \hat{\beta}(2), \hat{\beta}(3), \ldots \) etc. lead to

\[
\lim_{i \to \infty} \text{cov} \left[ \hat{\beta} \left( i+1 \right) \right] = \frac{c}{2 - c} \frac{\sigma_n^2}{\hat{\sigma}_u^2} \text{I} \quad (7.14)
\]

For this case of a constant "gain factor" the covariance does not reduce to zero even for an infinite observation interval.

Apparently method a) is usable if the parameters to be determined are constant. Such parameters can be found with any desired accuracy if the observation interval is of sufficient length. Note that because \( \Gamma(1) = 1/i \) new observations add less and less to the estimate as \( i \) increases.
The method b) gives to new observations as much weight as to old observations. In cases where the parameters are slowly varying (parameter tracking) one would care for gradually-forgetting-the-past. The implementation of that feature is discussed further on in this section.

Just as in chapter 6 we may ask what will be the effect of simplifying the instrumentation and specially what will happen if the matrix inversion is discarded. It turns out that:

- for \( \Gamma(i) = 1/i \) the covariance is about the same as in the case when matrix inversion is used, with some deviation for small values of i.j, the total number of samples, cf. fig. 7.5.

- for \( \Gamma(i) = c \) for large values of i.j the covariance has the same asymptotic value as with matrix inversion given by eq. (7.14).

c. Least squares approach to single (pairs of) observations.

For our next iterative scheme we change the notation of \( U \) to \( U_k \), indicating the last observation incorporated (Lee, 1964; Aoki, 1967).

\[
U_k = \begin{bmatrix} u_o(0) & \ldots & u_m(0) \\ \vdots & \ddots & \vdots \\ u_o(k) & \ldots & u_m(k) \end{bmatrix} \quad U_{k+1} = \begin{bmatrix} u_o(0) & \ldots & u_m(0) \\ \vdots & \ddots & \vdots \\ u_o(k) & \ldots & u_m(k) \\ u_o(k+1) & \ldots & u_m(k+1) \end{bmatrix} = \begin{bmatrix} U_k \\ u'(k+1) \end{bmatrix}
\]

(7.15)

where \( u'(k+1) \) represents the new observations. We define

\[
P_k^{-1} = U'_k U_k \quad \text{nonsingular}
\]

\[
P_{k+1}^{-1} = U_{k+1}' U_{k+1} = \begin{bmatrix} U_k' & u(k+1) \end{bmatrix} \begin{bmatrix} U_k \\ u'(k+1) \end{bmatrix}
\]

(7.16)

or

\[
P_{k+1}^{-1} = P_k^{-1} + u(k+1) u'(k+1)
\]

(7.17)
By using the matrix identities given in Appendix C one finds:

\[ P_{k+1} = P_k - P_k u(k+1) \left[ u'(k+1) P_k u(k+1) + 1 \right]^{-1} u'(k+1) P_k = \]

\[ P_k - \gamma(k) P_k u(k+1) u'(k+1) P_k \]  

(7.18)

where

\[ \gamma(k) = \left[ u'(k+1) P_k u(k+1) + 1 \right]^{-1} \]

is a scalar. Using \( y_k = [y(1), \ldots, y(k)] \) we find the estimator for \( b \) from eq. (6.31):

\[ \hat{\beta}(k) = P_k U_k' y_k \]  

(7.19)

\[ \hat{\beta}(k+1) = P_{k+1} U_{k+1} y_{k+1} = P_{k+1} \left[ U_k' y_k + u(k+1) y(k+1) \right] \]  

(7.20)

Substituting eq. (7.18) into (7.20) yields:

\[ \hat{\beta}(k+1) = \frac{P_k U_k' y_k}{P_k - \gamma(k) P_k u(k+1) u'(k+1) P_k} + \gamma(k) P_k u(k+1) y(k+1) \]

\[ + \frac{P_k u(k+1) y(k) y'(k) y(k+1)}{-P_k u(k+1) y(k) u'(k+1) P_k} \left[ U_k' y_k + u(k+1) y(k+1) \right] \]  

(7.21)

where the form \( \gamma(k) \cdot \gamma'(k) = 1 \) is introduced to facilitate further manipulation. Using (7.19) this can be reworked to the following expressions:

\[ \hat{\beta}(k+1) = \hat{\beta}(k) + \gamma(k) P_k u(k+1) \left\{ y(k+1) - u'(k+1) \hat{\beta}(k) \right\} \]  

(7.22)

or

\[ \hat{\beta}(k+1) = \hat{\beta}(k) + P_{k+1} u(k+1) \left\{ y(k+1) - u'(k+1) \hat{\beta}(k) \right\} \]  

(7.23)

The equivalence of the last two expressions can be shown by direct substitution of \( \frac{P_{k+1} u(k+1)}{\gamma(k+1)} \) or by use of the matrix identities given in Appendix C.

The estimate after using the new observations \( u(k+1) \) and \( y(k+1) \) equals the old estimate plus a linear correction term. Note that \( \gamma \) acts as a normalizing factor. The term \( \{y(k+1) - u'(k+1) \hat{\beta}(k)\} \) represents that part of the new output-observation that is not accounted for by the input observations and the last parameter estimate. Note that there is no need for matrix inversion and that only "old" knowledge and new observations are used for the new estimate. The pair \( (\hat{\beta}, P) \) thus represents the smallest number of variables, characterizing the input/output data, which are necessary to
carry along in the computations. This is indicated in fig. 7.6. Each estimation iteration is given by the equations (7.18) and (7.22) or (7.23).

Figure 7.7 shows schematically the operations that are needed for each such iteration. $P_k$ is a symmetrical matrix, hence only a triangular part needs to be updated. Notice that $\hat{\beta}(k+1)$, as given by eq. (7.22) or (7.23), is the least square estimate; the convergence of the iterative scheme follows directly from the consistency proof of the least squares estimate.

Another way to write eq. (7.22) is:

$$\hat{\beta}(k+1) = \left\{ 1 - \gamma(k) P_k U(k+1) U'(k+1) \right\} \hat{\beta}(k) + \gamma(k) P_k U(k+1) Y(k+1)$$

(7.24)

which indicates that the new estimate consists of a linear combination of the old estimate and an estimate following from the new observations only.

The "weighting factors" in this linear combination represent the "credibility" of the respective parts.

As our estimator is just an iterative version of the accumulative type derived in chapter 6 we know that

$$\mathbb{E} \left[ \hat{\beta}(k) \right] = b$$

$$\text{cov} \left[ \hat{\beta}(k) \right] = \left[ U_k U_k' \right]^{-1} U_k' N U_k \left[ U_k U_k' \right]^{-1} = P_k U_k' N U_k P_k$$

(7.25)

(7.26)

If $\mathbf{n}$ is white noise with $\mathbf{N} = \sigma_n^2 \mathbf{I}$

$$\text{cov} \left[ \hat{\beta}(k) \right] = \sigma_n^2 \left[ U_k U_k' \right]^{-1} = \sigma_n^2 P_k$$

(7.27)

Consequently $P_k$ and $P_{k+1}$ are proportional to the covariance matrices of the estimates $k$ and $k+1$ respectively.

It can be shown for many type of input signals that $P_k$ is positive definite and that

$$\lim_{k \to \infty} P_k = 0$$

An easy grasp of the significance of the different terms is provided by the simplification to a one-dimensional case.
\[ \hat{\beta}(k+1) = \hat{\beta}(k) + P_{k+1} u(k+1) \{ y(k+1) - \hat{y}(k+1) \hat{\beta}(k) \} \]  
\[ P_{k+1}^{-1} = P_k^{-1} + u^2(k+1) = P_0^{-1} + \sum_{j=1}^{k+1} u^2(j) \]  
\[ (7.28) \]
\[ (7.29) \]

If at the start little a priori knowledge is available on \( \beta \), then a proper choice for \( P_0^{-1} \) is 0. If \( u \) is a white stationary signal with \( \mathbb{E}[u] = 0 \), and \( \text{var}[u] = \sigma_u \), then

\[ \mathbb{E}[P_{k+1}] = \frac{1}{(k+1) \sigma_u^2} \]  
\[ (7.30) \]

It is interesting to compare this weighting factor with the condition derived for the convergence in probability for stochastic approximation; cf. section 5.4.

By a straightforward manipulation along the lines given in this section the following relations can be derived as well:

\[ \hat{\beta}(k+1) = \hat{\beta}(k) + P_{k+1} u(k+1) \{ y(k+1) - \hat{y}(k+1) \hat{\beta}(k) \} \]  
\[ P_{k+1} = P_k - P_k u(k+1) \left[ u'(k+1) P_k u(k+1) + r_{k+1} \right]^{-1} u(k+1) P_k \]  
\[ (7.31) \]
\[ (7.32) \]

where \( r_{k+1} \) is the corresponding element of the diagonal matrix \( R \) in the error function \( e'R e \).

For a more rigorous discussion of the iterative least squares estimators the reader is referred to the paper by Albert and Sittler (1965).

Starting the estimation procedure. This estimation scheme, using an iteration after each new observation, can be started by several techniques; cf. Klinger (1968).

1) By using possibly available a priori knowledge about \( \beta \) and \( P \). If one knows that the input signal \( u \) has the character of stationary "white" noise
with variance $\sigma_u^2$ then this knowledge can be used in choosing:

$$P_\ell = E \left[ U_\ell' U_\ell \right]^{-1} = \frac{I}{\ell \sigma_u^2}$$

Then the iterative procedure starts with iteration number $\ell+1$.

2) By determination from the explicit method:

$$P_\ell = \left[ U_\ell' U_\ell \right]^{-1}$$

$$\beta(t) = P_\ell U_\ell' x_\ell$$

using $\ell$ samples together for an initial estimate; after this calculation the iterative scheme can be used. This way of starting has the disadvantage that it needs a matrix inversion.

3) By using $P_0 = aI$, where $a$ is a very large scalar and $\hat{\beta}(0)$ is chosen arbitrarily. After one iteration:

$$P_1^{-1} = P_0^{-1} + U_1' u_1(y)$$

$$\hat{\beta}(1) = \hat{\beta}(0) + P_1 U_1' \left\{ y_1(y) - u_1'(y) \hat{\beta}(0) \right\} =$$

$$= P_1 \left\{ y_1(y) + P_0^{-1} \hat{\beta}(0) \right\}$$

cf. eq. (7.17) and (7.23). In the same way follows:

$$P_\ell^{-1} = P_0^{-1} + U_\ell' U_\ell$$

$$\hat{\beta}(\ell) = P_\ell \left\{ U_\ell x_\ell + P_0^{-1} \hat{\beta}(0) \right\}$$

If $P_0^{-1} = \frac{1}{a} I$ and $a \rightarrow \infty$ then

$$P_\ell^{-1} \rightarrow U_\ell' U_\ell$$

$$\hat{\beta}(\ell) \rightarrow P_\ell U_\ell' x_\ell$$

i.e., the expressions corresponding to the explicit estimation method.

Now the iterative estimation procedure can start with the first observation samples.
Exponential weighting of past data. As mentioned before the algorithms just given are iterative versions of the accumulative (one shot) type of estimator. This implies that all $k$ samples processed for the estimate $\hat{\beta}(k)$ get an equal weight, irrespective of the "age" of those samples. For the estimation of (slowly) changing parameters this obviously is undesirable; one needs a method to "forget" gradually the past. This can be done as follows:

$$U'_k, U'_{k+1} = [c U_k]' [c U_k] + U(k+1) U'(k+1)$$

or

$$c^{-2} P^{-1}_{k+1} = P^{-1}_k + U(k+1) c^{-2} U'(k+1)$$

(7.33)

with $0 < c \leq 1$. With $\rho = c^2$ this leads to the following equations; cf. Appendix C:

$$\hat{\beta}(k+1) = \hat{\beta}(k) + \gamma(k) P_k U(k+1) \left\{ y(k+1) - U(k+1) \hat{\beta}(k) \right\}$$

(7.34)

$$P_{k+1} = \rho^{-1} \left[ P_k - \gamma(k) P_k U(k+1) U'(k+1) P_k \right]$$

(7.35)

$$\gamma(k) = \left\{ U'(k+1) P_k U(k+1) + \rho \right\}^{-1}$$

(7.36)

Recognizing that the weighting factors on the previous samples change with the sample number $j$ as $c^{k-j}$, with $j \leq k$, it is quite simple to determine the number of samples that effectively contribute to the estimate for a given value of $c$. For noisy observations in general $c$ will have to be close to one, in order to obtain an appreciable effective observation length.
d. Stochastic approximation.

This method was discussed in section 5.3. The algorithm that may
be used is

$$\hat{\beta}(k+1) = \hat{\beta}(k) - \frac{1}{2} \Gamma(k) \nabla_{\beta} E$$

(7.37)

where $\Gamma(k)$ has to fulfill the conditions

$$\Gamma(k) \geq 0 \quad \sum_{k=0}^{\infty} \Gamma(k) = \infty \quad \sum_{k=0}^{\infty} \Gamma^2(k) < \infty$$

For an iteration after each observation the following holds:

$$E = e^2(k+1)$$

$$e(k+1) = y(k+1) - u'(k+1) \hat{\beta}(k)$$

Consequently eq. (7.37) can be rewritten as:

$$\hat{\beta}(k+1) = \hat{\beta}(k) + \Gamma(k) u(k+1) \{ y(k+1) - u'(k+1) \hat{\beta}(k) \}$$

(7.38)

This estimate converges in probability to the true parameter values; c.f.
Albert and Gardner (1967), Holmes (1968). The conditions of $\Gamma(k)$ can be ful-
filled by the sequence:

$$\Gamma(k) = \frac{c}{k^\alpha} \quad \text{with} \quad c > 0 \quad 0.5 < \alpha \leq 1$$

Comparing eq. (7.23) with eq. (7.38), one notices that the only difference is
that instead of $P_{k+1}$ we have $\Gamma(k)$; that is the $P$ matrix, that is closely re-
lated to the covariance matrix of the estimate, cf. eq. (7.27), is replaced by
a diagonal matrix $\Gamma(k) I$. Consequently stochastic approximation provides a
simplified algorithm that needs less computational effort because the $P$ matrices
need not to be calculated. On the other hand as the $P$ matrix is not available,
the stochastic approximation method does not provide a measure of the accuracy
of the estimate. In the least-squares error sense it is suboptimal compared to
the algorithms (7.23) and (7.18). Some references on applications of the sto-
chastic approximation procedure are: Sakrison (1967), Saridis and Stein (1968a,
A technique of constructing recursive algorithms has been suggested by Oza and Jury (1968, 1969). We will explain the technique in connection with the least squares problem. Instead of solving the equation
\[ U_k' U_k \hat{\beta}(k) = U_k' y_k \]
for each \( k \) and showing that \( \hat{\beta}(k) \) satisfies a recursive equation, Oza and Jury introduce the mapping
\[ \mathcal{J}_k [\beta] = \beta - \gamma [ U_k' U_k \beta - U_k' y_k ] \]
where \( \gamma \) is a scalar. It is then shown that the sequence
\[ \beta(k+1) = \mathcal{J}_k [\beta(k)] \]
under suitable conditions converges to the true parameters as \( k \to \infty \). When applied to the ordinary least squares problems the algorithm (7.39) is not efficient in contrast with the recursive least squares method. To obtain an efficient algorithm it is necessary to make \( \gamma \) a matrix.

With the choice
\[ \gamma = \left[ U_k' U_k \right]^{-1} \]
the algorithm becomes equivalent to the least squares.

The method of Oza and Jury can be applied to more general cases than the least squares. It was actually proven for the case where there are errors in the measurements of both inputs and outputs, provided the covariance function of the measurement errors are known. The assumption of known covariances of the measurement errors severely limits the practical applicability of the method.
7.2 Extension to generalized and state space models.

The iterative schemes given in the previous section can also be applied to the generalized model given by eq. (6.83) and fig. 7.8:

\[
\hat{e} = y + \hat{y} - \frac{U \hat{y}}{\beta} = y - \left[ U \hat{y} \right] \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix} = y - \Omega \hat{\theta} \quad (7.41)
\]

Note that this model is linear-in-the-parameters. Of course we will meet with the same problems as when using the explicit method for this case, i.e., a bias of the parameters due to the noise components in \( \hat{y} \); cf. section 6.5.

The same generalized model can be given by fig. 7.9, a pulse-transfer-function representation, where:

\[
A^*[z^{-i}; i] = 1 + \alpha_1(i) z^{-1} \ldots + \alpha_m(i) z^{-m} \\
B^*[z^{-i}; i] = \beta_1(i) z^{-1} \ldots + \beta_n(i) z^{-n} \quad (7.42)
\]

Here the symbol \( i \) refers to the parameter values at \( i \)-th iteration. For such a scheme Kalman (1958) proposed a simple estimation scheme leading to asymptotically biased estimates.

Techniques to circumvent the detrimental effects of additive noise have been indicated in chapter 6. The discussion of one of those techniques, generalized least squares and related schemes, was deferred to this chapter, as it implies model-adjustment procedures. In this category the following approaches will be discussed:

- successive linear regression and filtering
- generalized-least-squares estimation
- extended matrix method.

Successive linear regression and filtering. Instead of choosing the generalized error according to fig. 7.9, as:
\[ e(k) = A^*(z^{-1}; i) y(k) - B^*(z^{-1}; i) u(k) \]  

(7.43)

A proposal by Steiglitz and McBride (1965) implies the use of the error expression:

\[ e'(k) = \frac{e(k)}{A^*(z^{-1}; i-1)} = \frac{A^*(z^{-1}; i)}{A^*(z^{-1}; i-1)} y(k) - \frac{B^*(z^{-1}; i)}{A^*(z^{-1}; i-1)} u(k) = \]

\[
= \left\{ y(k) - \frac{B^*(z^{-1}; i)}{A^*(z^{-1}; i)} \right\} \frac{A^*(z^{-1}; i)}{A^*(z^{-1}; i-1)} 
\]

(7.44)
We still have a generalized model, linear with parameters $\alpha(i)$ and $\beta(i)$, but the term between brackets represents an ordinary model, fig. 7.10. For further improvement of the estimate it is also suggested that, after the convergence has proceeded far enough, the scheme be switched from this generalized model to an ordinary model:

$$e(k) = y(k) - \frac{B^*(z^{-i};i)}{A^*(z^{-i};i)} u(k)$$

(7.45)

This has the advantage that the additive noise does not cause biased estimates. Near the optimum adjustment the nonlinearity-in-the-parameters of $A^*$ hardly has a detrimental effect on the convergence. The nonlinearity-in-the-parameters can also be handled by means of quasilinearization, cf. Schulz (1968).

The iterative procedure is as follows: As the first step, a linear regression gives

$$\hat{\theta}(2) = \left[ \begin{array}{c} \hat{\beta}(1) \\ -\frac{\hat{\alpha}(1)}{\hat{\gamma}(1)} \end{array} \right] = \left[ \begin{array}{c} \hat{\alpha}^T(z;1) \\ \hat{\beta}^T(z;1) \end{array} \right]^{-1} \hat{\Omega} \hat{\gamma}(1) y$$

(7.46)

where $\alpha(1)$ and $\beta(1)$ are the parameters of $A^*(z^{-i};i)$ and $B^*(z^{-i};i)$. The next step is to filter process input and output signals with the filter $A^*(z^{-i};i)$. Based on these filtered signals an estimate $\hat{\theta}(2)$ is found, etc. No proof is known for the convergence of this scheme.

Generalized-least-squares (Markov) estimation method. The basic idea is as follows. Let the process be governed by

$$A^*(z^{-i}) y(k) = B^*(z^{-i}) u(k) + n(k)$$

(7.47)

where $A^*$ and $B^*$ are polynomials and $\{n(k)\}$ a sequence of correlated random variables. Suppose that the correlations of $n(k)$ are known. Say that they can be represented as
where \( \{w(k)\} \) is a sequence of uncorrelated random variables and \( G^* \) a pulse transfer function. The equation describing the process can be then written as

\[
A^*(z^{-1}) y(k) = B^*(z^{-1}) u(k) + G^*(z^{-1}) w(k)
\]

(7.49)

where \( A^* \) and \( B^* \) characterize the process and \( G^* \) characterizes the "environment", or:

\[
A^*(z^{-1}) \tilde{y}(k) = B^*(z^{-1}) \tilde{u}(k) + w(k)
\]

(7.50)

where

\[
\tilde{y}(k) = \frac{1}{G^*(z^{-1})} y(k) \quad \text{and} \quad \tilde{u}(k) = \frac{1}{G^*(z^{-1})} u(k)
\]

Note that if the number of terms of \( G^*(z^{-1}) \) is quite limited, the number of significant terms of \( G^*(z^{-1})^{-1} \) may be very high. This implies that many samples of \( u \) are needed for \( \tilde{u} \).

Hence if the signals \( \tilde{u} \) and \( \tilde{y} \) are considered as the inputs and outputs we have an ordinary least squares problem. Compare with (6.110). We thus find that the generalized least squares can be interpreted as a least squares identification problem where the criterion is chosen as

\[
\min \sum \varepsilon^2(k)
\]

with the generalized error defined as:

\[
\varepsilon(k) = \frac{G^*(z^{-1})}{G^*(z^{-1})} y(k) - \frac{B^*(z^{-1})}{G^*(z^{-1})} u(k) = \left( G^*(z^{-1}) \right) ^{-1} \frac{y(k)}{G^*(z^{-1})} - \frac{B^*(z^{-1})}{G^*(z^{-1})} \frac{u(k)}{G^*(z^{-1})}
\]

(7.51)

Compare with the block diagram of Fig. 7.11. This shows how the generalized error can be obtained from the process input and output and the model parameters. \( \star \)
and \( \beta \) in the generalized least squares method.

The correlation of the residuals and the pulse transfer function \( G^* \) are seldom known in practice. Clarke (1967) has proposed an iterative procedure to determine \( G^* \) which has been tested on simulated data as well as on practical measurements (distillation column identification). The procedure consists of the following steps.

(In order to accentuate the link between pulse-transfer-function representation and the matrix notation we will use both in parallel)

1) Make an ordinary least squares fit to the process input and output sequence

\[
\mathbf{A}^*(z^{-1}; i) y(k) = (\mathbf{G}^*(z^{-1}; i) u(k) + r(k)
\]

2) Now the error sequence is an approximation of the residuals. Analyze those residuals and fit an auto regression:

\[
\mathbf{D}^*(z^{-1}; i) e(k) = \hat{f}(k)
\]

in such way that

\[
\mathbf{D}^*(z^{-1}; i) e(k) = \hat{f}(k)
\]

where \( \hat{f}(k) \) is discrete time white noise,

3) Filter the process input and output sequence:

\[
\tilde{u}(k) = \mathbf{D}^*(z^{-1}; i) u(k)
\]

\[
\tilde{y}(k) = \mathbf{D}^*(z^{-1}; i) y(k)
\]
4) Make a new least squares fit to the filtered process input and output sequence, i.e., repeat step 1.

5) repeat from step 2.

This recursive procedure is indicated in fig. 7.12. The stopping is governed by a comparison of the respective values of the loss function.

If \( A^*(i) \rightarrow A^* \), \( B^*(i) \rightarrow B^* \), and \( D^*(i) \rightarrow \{ \mathbf{g}^* \}^{-1} \), then we will obtain

\[
A^*(z^{-1}) \{ \mathbf{g}^*(z^{-1}) \}^{-1} y(k) = B^*(z^{-1}) \{ \mathbf{g}^*(z^{-1}) \}^{-1} u(k) + v(k)
\]

(7.52)

i.e., a description with uncorrelated residuals. There do not seem to be available any convergence proofs as yet. It has, however, been shown to work very well in specific examples with reasonable ad hoc choices for the order of the process. The procedure has the disadvantage that there are no systematic rules (yet) for the choice of model order and for the order of the auto regression.

The procedure just described is not recursive in the process input/output data; a whole sequence of observations is handled in a one-shot manner.

Strictly speaking therefore this method could have been discussed in chapter 6. It has found its place in this chapter for the sake of continuity in the discussion.

A recursive version of this procedure has been described by Hasting-James and Sage (1969). Now the estimates of \( \beta \) and \( \delta \) are updated when new observations become available.

In the beginning, the estimates \( \delta \) will be poor and consequently the resulting samples \( u(k) \) and \( y(k) \) after filtering will not be optimal for estimating \( \beta \). Therefore, one has to introduce some kind of weighting in order to diminish the influence of older input–output data on the estimation procedure. As the exponential weighting of past data has been discussed in the previous section, we only summarize the algorithms for this recursive estimation scheme:
\[ \Omega_k = \begin{bmatrix} U_k & Y_k \end{bmatrix} \]

\[ \Omega_{k+1} = \begin{bmatrix} \Omega_k \\ \omega_{k+1} \end{bmatrix} = \begin{bmatrix} U_k & Y_k \\ U_{k+1} & Y_{k+1} \end{bmatrix} \]

\[ \tilde{y}(k+1) = y(k+1) + \omega'(k) \hat{\delta}(k) \]

\[ \tilde{\omega}(k+1) = \omega(k+1) + \hat{\omega}'(k) \hat{\delta}(k) \]

\[ \hat{\omega}(k+1) = \omega(k+1) + \Omega_k \hat{\delta}(k) \]

\[ \hat{f}(k+1) = y(k+1) - \omega'(k+1) \hat{\delta}(k) \]

\[ \hat{f}(k+1) = \begin{bmatrix} \hat{f}(k), \hat{f}(k-1), \ldots, \hat{f}(k-p+1) \end{bmatrix} \]

\[ \hat{\delta}(k+1) = \hat{\delta}(k) + y(k) P_k \tilde{\omega}(k+1) \{ \tilde{y}(k+1) - \tilde{\omega}'(k+1) \hat{\delta}(k) \} \]

\[ P_{k+1} = \rho^{-1} \left[ P_k - y(k) \tilde{\omega}'(k) \tilde{\omega}(k+1) P_k \right] \]

\[ y(k) = \{ \tilde{\omega}'(k) P_k \tilde{\omega}(k+1) + \rho \}^{-1} \]

\[ \hat{\delta}(k+1) = \hat{\delta}(k) + \lambda(k) \hat{\delta}_k \hat{f}(k+1) \{ \hat{f}(k+1) - \hat{f}'(k+1) \hat{\delta}(k) \} \]

\[ \hat{Q}_{k+1} = \nu^{-1} \left[ \hat{Q}_k - \lambda(k) \hat{\delta}_k \hat{f}(k+1) \hat{f}'(k+1) \hat{Q}_k \right] \]

\[ \lambda(k) = \{ \hat{f}'(k+1) \hat{Q}_k \hat{f}(k+1) + \nu \}^{-1} \]

A choice has to be made for the weighting coefficients \( \rho \) and \( \nu \). When starting the estimation procedure the coefficients could be low, e.g., 0.975 in order to obtain a good convergence rate; in due course \( \rho \) and \( \nu \) could approach 1.0 for obtaining a good accuracy (small variance).
No general theory is available yet that dictates the optimal choice of both $\rho$ and $\nu$. Some simulation results are given in section 7.3.

Extended matrix method. The previous method was based on the assumption that the noise can be represented by

$$ r(k) = \frac{1}{\mathbf{D}^*(z^{-1})} w(k) $$

where \( \{w(k)\} \) is a sequence of uncorrelated random variables (discrete "white" noise). Depending on the situation at hand, the number of parameters to characterize the noise can be quite large. Consequently, a minimum number of parameters for such a characterization could be obtained by using:

$$ r(k) = \frac{C^*(z^{-1})}{\mathbf{D}^*(z^{-1})} w(k) $$

Consequently, it is of importance to develop an estimation scheme where the "environment" of the process is characterized by estimates of the $c$ and $d$ parameters; cf. fig. 7.13.

The process and the environment are described by

$$ \sum_{i=0}^{n} a_i x(k-i) = \sum_{i=1}^{n} b_i u(k-i) \quad \text{process dynamics} $$

$$ y(k) = x(k) + n(k) \quad \text{process output} $$

Consequently,

$$ \sum_{i=0}^{n} a_i y(k-i) - \sum_{i=1}^{n} b_i u(k-i) = \sum_{i=0}^{n} a_i n(k-i) = r(k) $$
where the output noise is represented by \( \{r(k)\} \) and \( r(k) \) may be assumed to be derived from "white" noise samples \( w(k) \) by:

\[
    r(k) + \sum_{i=1}^{b} d_i r(k-i) = w(k) + \sum_{i=1}^{q} c_i w(k-i)
\]  

(7.62)

Choosing again \( a_0 = 1 \) one finds:

\[
    y(k) = \sum_{i=1}^{m} b_i u(k-i) - \sum_{i=1}^{b} a_i y(k-i) + \\
    + \sum_{i=1}^{q} c_i w(k-i) - \sum_{i=1}^{b} d_i r(k-i) + w(k)
\]  

(7.63)

and in the same way as before:

\[
    y = \begin{bmatrix} \hat{U} & \hat{Y} & \hat{W} & \hat{R} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{c} \\ \hat{d} \end{bmatrix} + \hat{w}
\]  

(7.64)

This, again, is a model that is linear in the process parameters and in the noise parameters, and is disturbed by white equation noise. As before, a consistent estimator is:

\[
    \begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{c} \\ \hat{d} \end{bmatrix} = \left[ \begin{bmatrix} \hat{U} & \hat{Y} & \hat{W} & \hat{R} \end{bmatrix}' \begin{bmatrix} \hat{U} & \hat{Y} & \hat{W} & \hat{R} \end{bmatrix} \right]^{-1} \begin{bmatrix} \hat{U} & \hat{Y} & \hat{W} & \hat{R} \end{bmatrix}' y
\]  

(7.65)

As we do not know, nor can measure, the elements of \( \hat{R} \) and \( \hat{W} \) these elements have to be replaced by their estimates \( \hat{R} \) and \( \hat{W} \):
\[
\begin{align*}
\hat{R}(k) &= \hat{y}(k) + \sum_{i=1}^{n} \alpha_i \hat{y}(k-i) - \sum_{i=1}^{m} \beta_i \hat{u}(k-i) \\
\hat{\omega}(k) &= \hat{\omega}(k) + \sum_{i=1}^{p} \delta_i \hat{\omega}(k-i) - \sum_{i=1}^{q} \gamma_i \hat{\omega}(k-i)
\end{align*}
\] 

(7.66)

Consequently, after each estimation of the parameter vector, the elements of \( \hat{R} \) and \( \hat{\omega} \) are calculated. As in the beginning of the estimation procedure, the accuracy will be poor, so that again a weighting factor has to be introduced in order to de-emphasize older information. This algorithm too can be written in recursive form along the same lines as before. With proper interpretation, e.g.,

\[
\begin{align*}
\tilde{\theta}'(k+1) &= \left[ \tilde{\beta}'(k+1) \mid -\tilde{\alpha}'(k+1) \mid \tilde{\gamma}'(k+1) \mid -\tilde{\delta}'(k+1) \right] \\
\omega^*(k+1) &= \left[ \tilde{\omega}'(k+1) \mid \tilde{\gamma}'(k+1) \mid \tilde{\hat{\omega}}'(k+1) \mid \tilde{\hat{\omega}}'(k+1) \right]
\end{align*}
\] 

(7.67)

etc. the algorithm given by eq. (7.56) can be used. This method is discussed by Smets (1970) for \( \gamma = 0 \), by Young (1968) for \( \delta = 0 \) and by Talmon (1971) for the general case. Some simulation results taken from the last reference, are given in section 7.3.

Another approach to circumvent the bias problems of the generalized model is by means of the maximum-likelihood method, cf. chapter 11.

The instrumental variable method (cf. chapter 6) can also easily be arranged as an on-line iterative procedure; cf. Peterka and Smuk (1969). An approximate on-line version of the maximum likelihood methods has been proposed by Panuska (1968). A discussion of on-line methods is also given in Leathrum (1969).
State space models. The previous methods can also be applied for the estimation of the parameters, used in the state space description of the process.

A discussion of the estimation of $A$ in

$$
\begin{align*}
    x(k+1) &= A x(k) \\
    y(k) &= x(k) + v(k)
\end{align*}
$$

is given in Ho and Whalen (1963). The situation where in addition, the input signal is measurable can be indicated as follows; cf. fig. 7.14

process:

$$
\begin{align*}
    x(k+1) &= A x(k) + B u(k) \\
    y(k) &= x(k) + v(k)
\end{align*}
$$

model:

$$
\varepsilon(k) = y(k+1) - A(k) y(k) - B(k) u(k)
$$

The criterion is chosen to minimize $E = \varepsilon'(k) \varepsilon(k)$, leading to

$$
\begin{align*}
    \frac{\partial E}{\partial A(k)} &= -2 \varepsilon(k) y'(k) \\
    \frac{\partial E}{\partial B(k)} &= -2 \varepsilon(k) u'(k)
\end{align*}
$$

Consequently, the iterative estimation procedure can be chosen as:

$$
\begin{align*}
    A(k+1) &= A(k) + \Gamma \varepsilon(k) y'(k) \\
    B(k+1) &= B(k) + \Gamma \varepsilon(k) u'(k)
\end{align*}
$$

The coefficient $\Gamma$ may again be chosen:

- to be constant, if $A$ and $B$ are changing in time
- to be $1/k$ if $A$ and $B$ are constant; stochastic approximation.
Because of the additive noise the estimates will be biased.

Anderson et al. (1969) describe a procedure in order to obtain consistent estimates of the elements of the $A$ matrix in

$$
\begin{align*}
    x(k+1) &= A x(k) + v(k) \\
    y(k) &= x(k) + n(k)
\end{align*}
$$

(7.73)

where $v(k)$ and $n(k)$ are white noises.

7.3 Some computer procedures and results

As the methods discussed in section 7.2 are promising with respect to practical application, the results of a number of experiments will be given to illustrate their properties. These illustrations are taken from Talmon (1971).

Process. In all cases given, the process has the pulse-transfer function.

$$
H(z^{-1}) = \frac{B^*(z^{-1})}{A^*(z^{-1})} = \frac{z^{-1} + 0.5 z^{-2}}{1 - 1.5 z^{-1} + 0.7 z^{-2}}
$$

(7.74)

and consequently:

$$
\begin{bmatrix}
    b^t \\
    b_0 & b_1 & b_2 & -a_1 & -a_2
\end{bmatrix} = \begin{bmatrix}
    0, 1, 0.5, 1.5, -0.7
\end{bmatrix}
$$

It has poles for $z = 0.75 \pm 0.36 j$ and a zero at $z = -0.5$. The pole-zero pattern is given in fig. 7.15, and the impulse response is shown in fig. 7.16.

In all programs, the input is chosen as a white noise signal with a rectangular
distribution between -1 and +1.

If the white noise input signal has a power of $\sigma_u^2$ then the power of the output, $\sigma_x^2$, is given by:

$$
\sigma_x^2 = \frac{\sigma_u^2}{2\pi j} \int_{|z|=1} H(z) H(z^{-1}) \frac{dz}{z}
$$

(7.75)

cf. Jury (1964). Åström, Jury & Agniel (1970) give in their paper a fast method to calculate this integral. For the given process the ratio $\frac{\sigma_x^2}{\sigma_u^2} = 18.8$.

The results given in the following sections are the averages of the estimates over ten runs of 1000 iterations. In addition, an average of the standard deviation, based on those ten runs, is given.

Notes:

1) In all tables, first the averages of the estimated values of the parameters over ten runs are given, unless otherwise indicated. Their standard deviations are given immediately below the averages.

2) Consider a stochastic process $z(k) = \mu + x(k)$, in which $\mu$ is constant and $x(k)$ is a stochastic variable with $E[x(k)] = 0$.

The estimate of $\mu$ is given by $\bar{z} = \frac{1}{N} \sum_{i=1}^{N} z(i)$

The confidence interval for $\mu$ is given by

$$
\bar{z} - \frac{t_\nu \left( \frac{1}{\sqrt{N}} \right)}{\sqrt{N}} < \mu < \bar{z} + \frac{t_\nu \left( \frac{1}{\sqrt{N}} \right)}{\sqrt{N}}
$$

(7.76)
where $s^2$ is the estimate of the variance of $z$ and where $v$ is the degree of freedom.

In our case, $z(i)$ is the estimated value of a parameter in the i-th run.

We have 10 runs, so $v = 9$. When we want the confidence interval, with a chance of 5% that $\mu$ lies outside this interval, then $t_9(\frac{1}{2}) = 2.26$.

Thus, we obtain:

$$\bar{z} - 0.7s < \mu < \bar{z} + 0.7s$$

3) The programs are written in ALGOL 60. The listings and the papertapes of the programs are available at the Eindhoven University of Technology, Group Measurement and Control.

The recursive generalized-least-squares estimation method. The algorithms used are those given by equations (7.56) and (7.57). The difference equation describing process and disturbances is:

$$y(k) = \sum_{i=0}^{2} b_i y(k-i) - \sum_{i=1}^{2} a_i y(k-i) - \sum_{i=1}^{2} c_i r(k-i) + \lambda \nu(k)$$

(7.77)

$$b_0 = 0$$

$$b_1 = 1$$

$$a_1 = -1.5$$

$$a_2 = -1.0$$

$$b_2 = 0.5$$

$$a_1 = 0.7$$

$$d_2 = 0.2$$

cf. fig. 7.17 a and b. $\{\nu(k)\}$ is a white noise sequence with a rectangular amplitude distribution between -1 and +1. The transfer function that colours the white noise is:

$$\frac{1}{D^*(z^{-1})} = \frac{1}{1 - z^{-1} + 0.2z^{-2}}$$

(7.78)

with poles at $z = 0.725$ and $z = 0.275$. For this case $\frac{\sigma^2}{\sigma_\nu^2} = 98.52$
Five process parameters and two noise parameters are estimated. After each estimation of \(\theta\) a new vector \(\hat{\omega}(k+1)\) and a new value \(\hat{r}(k+1)\) are calculated. After each new estimation \(\hat{\theta}\) of the parameter vector a new vector \(\hat{F}(k+1)\) and a new value of \(\hat{r}(k+1)\) are calculated.

The following tables of experimental results are given:

<table>
<thead>
<tr>
<th>table</th>
<th>(\lambda)</th>
<th>signal ratio noise at process output (\frac{S}{N} = \frac{\sigma_x^2}{\sigma_n^2})</th>
<th>weighting coefficient (\rho = \nu =)</th>
<th>number of samples in which the influence of model output errors is reduced to 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>0.25</td>
<td>3.2</td>
<td>+5 dB</td>
<td>0.9913</td>
</tr>
<tr>
<td>7.2</td>
<td>0.5</td>
<td>0.8</td>
<td>-1.2 dB</td>
<td>same</td>
</tr>
<tr>
<td>7.3</td>
<td>1</td>
<td>0.2</td>
<td>-7 dB</td>
<td>same</td>
</tr>
<tr>
<td>7.4</td>
<td>2</td>
<td>0.05</td>
<td>-13 dB</td>
<td>same</td>
</tr>
<tr>
<td>7.5</td>
<td>4</td>
<td>0.0125</td>
<td>-19 dB</td>
<td>same</td>
</tr>
<tr>
<td>7.6</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>1.0</td>
</tr>
<tr>
<td>7.7</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>0.995</td>
</tr>
<tr>
<td>7.8</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>0.990</td>
</tr>
<tr>
<td>7.9</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>0.985</td>
</tr>
<tr>
<td>7.10</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>0.980</td>
</tr>
<tr>
<td>7.11</td>
<td>same</td>
<td>same</td>
<td>same</td>
<td>0.975</td>
</tr>
</tbody>
</table>
INiber of
itaratiOIII

a2

ao

a1

Bt

h

lz

-1.556
.056

.752
.082

+,049
,088

1,077

.SZI
.049

-o.t95
,161

.270
,1.57

zoo

-1.529
.056

.730
,079

+,038
.076

-1.060

300

-1.519
.042

.720
,061

400

-1.514
.028

•?15
.038

+,020
.067
+.025
. ,.048

500

-1.509
.017

.708
.024

.284
.131
.286
.086
.235
,090
.190

600

-1.499
.017

.700
.020

700

-1.502

.oil

.703
·,017

800

-1.498
.022

,702
,022

900

-1.498
.017

1000

-1.502
.016

.708
.019
,702
.016

,514
.022
.505
.013
,504
.012
.501
.017
,504
,018
.501
.016
.489
.020
.493
.CIIII
.498
.• 016

. at

100

Table 7 .I

-Ns = 3.2 (+SdB)

-.

~

-

- ....

.....

.il98
1;048
.082
1.030
.069
1.028
.046
1.015
.034
1.004
.027
1,008
.017
0.990
.• 031
0.998
.018
1.007
.018

+'.014
,030
;.';001
.024
+.001
.016
-,011
.024
--.003
,017
+.001
,016

. .

..

----

-

-1.068
.092
-1.003
.066
-o.971
,101

.080

-o.985
.076

.201
.061

-0.950
.077

,172
.058

-o.960
.057·

.158
.052

-o.977
.066

.204
.052

-o.985
.090

.196
.087

...

l!z

~2

1.053
,085

.510
.080

-o.H.5
.132

.244
.152

+.022
.033

1.038
.049

.516
.043

-1.000
.085

.272
.127

.698
.030

-.001
,032

1.020
.035

,027

.509

-1.015
,070

.270
.052

-1.513
.026

.696
.028

+,022
.028

-1.517
.026

.ioo
.034

+. 006
.030

1,024
.036
1.012
.043

.509
.028
.505
.034

-o.962
.059
-o.948
.092

.234
.080
.193
.062

600

-1.504
,025

,697
,034

-.010
.030

.515
.Q35

700

-1.505
,020

.700
.029

- •.po6
.023

1.007
.042.
1.010
.024

.506
.030

-o.971
.077
-o.945
,080

.207
.066
.182
.072

1!00

-1.502
.031

.702
.036

-.023
.040

0.983
,051

.• 482
,039

-o.957
.061

.166
. • 062

900

-1.500
,030

,714
,034

-.005
,031

0,999
.033

.489
.038

-o.973
. ,o62

.206
.052

1000

-1.504
,028

• 704
.028

+,001
,Q31

1,014
.o38

.496
.035

-0.913
.083

.196
.080

az

Bo

lit

100

-1.565
,056

.722
,056

+,022
.054

200

-1.538
.032

.706
.036

300

-1.525
.028

Table 7.2

400

s = 0.8
N

500

-

'·

...

--

....

a\aber of
itaratioae

---

--......,...·--:---------.....-----.....-

..

Q)

az

llo

l!t

az

4t

6.a

100

-1.624
.047

.762
.053

+,035
.114

1,105
.190

.507
,171

-o.849
,102

.232
,158

200

-1.592
.034

• 737
.0411

+.043
.068

1.084
.119

.526
.088

-o.938
.078

.274
,141

300

-1.578
.028
-1.540
.034
-1.551
,035

.726
.032

1.050
.068
1.057
.066
1,027
.089

.510
.051
.516
.061

-o.962
.082
-o.911!
.061

.274
.072
,252
.096

.710
.048

-.003
.071
+,050
.058
+,009
.065

.507
.075

-o.909
,096

.205
.062

-1.533
.023
-1.523
,035

.704
.042
• 700
.038

-.021
.062
-.011
,051

1.025
.082
1.027
.050

-o.943
.077
-0.918
,076

.225
,082
,198
,091

800

-1.524
.034

.704
,045

0.975
.094

-o.936
.063

,181
.075

900

-1.517
.046
-1.517
.039

. 719
.048

-.045
.082
-,003
.065

.534
.079
.513
.060
.467
.079

1.013
.068

.486
.084

-0,957
,056

.226
,055

• 705
.041

+.007
.068

1.038
.081

.496
.079

-o.969
,082

.208
.082

Table 7.3

400

s = 0,2 (-7dB)
N

500
600

700

1000

• 702
.046

·.

- ----

-·- ··

6t

Qj

aWiber of
iteratioaa

(-ldB)

..

.u.s


Table 7.4

\[ S/N = 0.05 \text{ (-13dB)} \]
(averaged over runs)

\[
\begin{array}{ccccccc}
\text{number of iterations} & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\
100 & -1.672 & .806 & +.043 & 1.106 & .456 & -0.805 & .245 \\
 & .049 & .042 & .225 & .226 & .258 & .096 & .167 \\
200 & -1.649 & .785 & +.068 & 1.093 & .494 & -0.886 & .279 \\
 & .037 & .024 & .143 & .198 & .150 & .088 & .139 \\
300 & -1.641 & .775 & +.007 & 1.076 & .497 & -0.897 & .273 \\
 & .032 & .024 & .127 & .132 & .080 & .082 & .070 \\
400 & -1.599 & .740 & +.111 & 1.125 & .511 & -0.869 & .260 \\
 & .028 & .039 & .100 & .124 & .117 & .057 & .092 \\
500 & -1.606 & .744 & +.026 & 1.045 & .500 & -0.859 & .219 \\
 & .037 & .042 & .128 & .186 & .154 & .090 & .065 \\
600 & -1.588 & .734 & -.024 & 1.066 & .547 & -0.893 & .250 \\
 & .026 & .031 & .119 & .160 & .152 & .065 & .087 \\
700 & -1.575 & .729 & -.006 & 1.071 & .520 & -0.873 & .222 \\
 & .045 & .032 & .094 & .094 & .120 & .075 & .097 \\
800 & -1.576 & .724 & -.050 & 0.997 & .445 & -0.878 & .196 \\
 & .037 & .035 & .112 & .148 & .144 & .054 & .071 \\
900 & -1.572 & .742 & +.006 & 1.038 & .462 & -0.917 & .257 \\
 & .044 & .040 & .114 & .125 & .159 & .045 & .066 \\
1000 & -1.564 & .722 & +.036 & 1.098 & .490 & -0.926 & .243 \\
 & .041 & .042 & .116 & .140 & .149 & .069 & .091 \\
\end{array}
\]

Table 7.5

\[ S/N = 0.0125 \text{ (-19dB)} \]

\[
\begin{array}{ccccccc}
\text{number of iterations} & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\
100 & -1.691 & .816 & +.157 & 1.508 & .636 & -0.776 & .240 \\
 & .052 & .047 & .492 & 1.122 & .769 & .101 & .155 \\
200 & -1.667 & .790 & +.154 & 1.370 & .658 & -0.857 & .278 \\
 & .040 & .036 & .279 & .830 & .481 & .099 & .122 \\
300 & -1.662 & .787 & +.002 & 1.245 & .579 & -0.872 & .275 \\
 & .040 & .034 & .242 & .529 & .296 & .071 & .063 \\
400 & -1.614 & .746 & +.201 & 1.262 & .582 & -0.842 & .273 \\
 & .041 & .048 & .202 & .266 & .293 & .052 & .081 \\
500 & -1.629 & .757 & +.035 & 1.097 & .526 & -0.835 & .232 \\
 & .040 & .039 & .251 & .382 & .308 & .081 & .060 \\
600 & -1.616 & .751 & -.091 & 1.084 & .612 & -0.872 & .259 \\
 & .038 & .032 & .252 & .324 & .298 & .054 & .081 \\
700 & -1.596 & .725 & -.051 & 1.098 & .542 & -0.852 & .236 \\
 & .049 & .033 & .209 & .203 & .231 & .058 & .086 \\
800 & -1.602 & .739 & -.192 & 0.889 & .358 & -0.867 & .220 \\
 & .039 & .027 & .334 & .368 & .275 & .048 & .070 \\
900 & -1.589 & .744 & -.036 & 1.025 & .435 & -0.893 & .272 \\
 & .052 & .046 & .241 & .236 & .303 & .042 & .064 \\
1000 & -1.587 & .731 & +.023 & 1.142 & .470 & -0.905 & .254 \\
 & .041 & .043 & .243 & .260 & .272 & .087 & .093 \\
\end{array}
\]
\[ \frac{S}{N} = 0.0125 \]  

### Table 7.6

**\( \rho = \nu = 1.0 \)**

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( \theta_0 )</th>
<th>( \theta_1 )</th>
<th>( \theta_2 )</th>
<th>( \delta_1 )</th>
<th>( \delta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.698</td>
<td>0.820</td>
<td>+.142</td>
<td>1.103</td>
<td>.354</td>
<td>-0.802</td>
<td>.301</td>
</tr>
<tr>
<td>200</td>
<td>-1.673</td>
<td>0.796</td>
<td>+.073</td>
<td>1.109</td>
<td>.368</td>
<td>-0.834</td>
<td>.287</td>
</tr>
<tr>
<td>300</td>
<td>-1.667</td>
<td>0.787</td>
<td>+.083</td>
<td>1.120</td>
<td>.350</td>
<td>-0.838</td>
<td>.282</td>
</tr>
<tr>
<td>400</td>
<td>-1.664</td>
<td>0.785</td>
<td>+.020</td>
<td>1.066</td>
<td>.449</td>
<td>-0.853</td>
<td>.288</td>
</tr>
<tr>
<td>500</td>
<td>-1.655</td>
<td>0.778</td>
<td>+.024</td>
<td>1.059</td>
<td>.450</td>
<td>-0.854</td>
<td>.288</td>
</tr>
<tr>
<td>600</td>
<td>-1.651</td>
<td>0.775</td>
<td>+.026</td>
<td>1.046</td>
<td>.460</td>
<td>-0.853</td>
<td>.285</td>
</tr>
<tr>
<td>700</td>
<td>-1.649</td>
<td>0.772</td>
<td>+.001</td>
<td>1.042</td>
<td>.468</td>
<td>-0.856</td>
<td>.281</td>
</tr>
<tr>
<td>800</td>
<td>-1.644</td>
<td>0.769</td>
<td>+.002</td>
<td>1.047</td>
<td>.453</td>
<td>-0.855</td>
<td>.279</td>
</tr>
<tr>
<td>900</td>
<td>-1.642</td>
<td>0.767</td>
<td>-.007</td>
<td>1.029</td>
<td>.446</td>
<td>-0.856</td>
<td>.272</td>
</tr>
<tr>
<td>1000</td>
<td>-1.640</td>
<td>0.764</td>
<td>-.010</td>
<td>1.024</td>
<td>.458</td>
<td>-0.858</td>
<td>.275</td>
</tr>
</tbody>
</table>

### Table 7.7

**\( \rho = \nu = 0.995 \)**

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( \theta_0 )</th>
<th>( \theta_1 )</th>
<th>( \theta_2 )</th>
<th>( \delta_1 )</th>
<th>( \delta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.676</td>
<td>0.817</td>
<td>+.138</td>
<td>1.102</td>
<td>.449</td>
<td>-0.806</td>
<td>.300</td>
</tr>
<tr>
<td>200</td>
<td>-1.665</td>
<td>0.787</td>
<td>+.062</td>
<td>1.100</td>
<td>.508</td>
<td>-0.843</td>
<td>.286</td>
</tr>
<tr>
<td>300</td>
<td>-1.653</td>
<td>0.774</td>
<td>+.084</td>
<td>1.124</td>
<td>.461</td>
<td>-0.848</td>
<td>.280</td>
</tr>
<tr>
<td>400</td>
<td>-1.643</td>
<td>0.766</td>
<td>-.026</td>
<td>1.018</td>
<td>.454</td>
<td>-0.865</td>
<td>.289</td>
</tr>
<tr>
<td>500</td>
<td>-1.625</td>
<td>0.754</td>
<td>-.001</td>
<td>1.012</td>
<td>.454</td>
<td>-0.869</td>
<td>.289</td>
</tr>
<tr>
<td>600</td>
<td>-1.621</td>
<td>0.753</td>
<td>+.003</td>
<td>0.990</td>
<td>.472</td>
<td>-0.866</td>
<td>.279</td>
</tr>
<tr>
<td>700</td>
<td>-1.618</td>
<td>0.749</td>
<td>-.058</td>
<td>1.003</td>
<td>.496</td>
<td>-0.876</td>
<td>.270</td>
</tr>
<tr>
<td>800</td>
<td>-1.606</td>
<td>0.738</td>
<td>-.034</td>
<td>1.030</td>
<td>.447</td>
<td>-0.874</td>
<td>.268</td>
</tr>
<tr>
<td>900</td>
<td>-1.602</td>
<td>0.734</td>
<td>-.050</td>
<td>0.978</td>
<td>.433</td>
<td>-0.876</td>
<td>.267</td>
</tr>
<tr>
<td>1000</td>
<td>-1.595</td>
<td>0.731</td>
<td>-.036</td>
<td>0.983</td>
<td>.490</td>
<td>-0.886</td>
<td>.264</td>
</tr>
</tbody>
</table>

### Table 7.8

**\( \rho = \nu = 0.990 \)**

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( \theta_0 )</th>
<th>( \theta_1 )</th>
<th>( \theta_2 )</th>
<th>( \delta_1 )</th>
<th>( \delta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.667</td>
<td>0.790</td>
<td>+.285</td>
<td>1.335</td>
<td>.580</td>
<td>-0.865</td>
<td>.334</td>
</tr>
<tr>
<td>200</td>
<td>-1.639</td>
<td>0.765</td>
<td>+.167</td>
<td>1.193</td>
<td>.584</td>
<td>-0.857</td>
<td>.286</td>
</tr>
<tr>
<td>300</td>
<td>-1.633</td>
<td>0.758</td>
<td>+.052</td>
<td>1.080</td>
<td>.463</td>
<td>-0.883</td>
<td>.253</td>
</tr>
<tr>
<td>400</td>
<td>-1.630</td>
<td>0.757</td>
<td>-.024</td>
<td>1.063</td>
<td>.385</td>
<td>-0.899</td>
<td>.291</td>
</tr>
<tr>
<td>500</td>
<td>-1.594</td>
<td>0.728</td>
<td>+.104</td>
<td>1.062</td>
<td>.477</td>
<td>-0.873</td>
<td>.285</td>
</tr>
<tr>
<td>600</td>
<td>-1.602</td>
<td>0.748</td>
<td>+.024</td>
<td>1.097</td>
<td>.560</td>
<td>-0.868</td>
<td>.285</td>
</tr>
<tr>
<td>700</td>
<td>-1.595</td>
<td>0.741</td>
<td>-.075</td>
<td>1.145</td>
<td>.665</td>
<td>-0.885</td>
<td>.263</td>
</tr>
<tr>
<td>800</td>
<td>-1.577</td>
<td>0.720</td>
<td>-.022</td>
<td>1.084</td>
<td>.504</td>
<td>-0.883</td>
<td>.242</td>
</tr>
<tr>
<td>900</td>
<td>-1.587</td>
<td>0.742</td>
<td>-.123</td>
<td>0.938</td>
<td>.414</td>
<td>-0.908</td>
<td>.246</td>
</tr>
<tr>
<td>1000</td>
<td>-1.574</td>
<td>0.731</td>
<td>-.093</td>
<td>1.028</td>
<td>.551</td>
<td>-0.926</td>
<td>.272</td>
</tr>
</tbody>
</table>

\[ S/N = 0.0125 \]  

\( \rho = \nu = 1.0 \)
\[
\frac{S}{N} = 0.0125 \quad (-1948)
\]

### Table 7.9

<table>
<thead>
<tr>
<th>(\rho = v = 0.985)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of iterations</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>200</td>
</tr>
<tr>
<td>300</td>
</tr>
<tr>
<td>400</td>
</tr>
<tr>
<td>500</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>700</td>
</tr>
<tr>
<td>800</td>
</tr>
<tr>
<td>900</td>
</tr>
<tr>
<td>1000</td>
</tr>
</tbody>
</table>

### Table 7.10

<table>
<thead>
<tr>
<th>(\rho = v = 0.980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of iterations</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>200</td>
</tr>
<tr>
<td>300</td>
</tr>
<tr>
<td>400</td>
</tr>
<tr>
<td>500</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>700</td>
</tr>
<tr>
<td>800</td>
</tr>
<tr>
<td>900</td>
</tr>
<tr>
<td>1000</td>
</tr>
</tbody>
</table>

### Table 7.11

<table>
<thead>
<tr>
<th>(\rho = v = 0.975)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of iterations</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>200</td>
</tr>
<tr>
<td>300</td>
</tr>
<tr>
<td>400</td>
</tr>
<tr>
<td>500</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>700</td>
</tr>
<tr>
<td>800</td>
</tr>
<tr>
<td>900</td>
</tr>
<tr>
<td>1000</td>
</tr>
</tbody>
</table>
Consequently, from the experimental results one may judge:

a) the influence of increasing noise level; table 7.1 - 7.5

b) the effect of the weighting coefficients; table 7.6 - 7.11

a). One notices that, with increasing noise power:

- the standard deviations of $\beta_0$, $\beta_1$, and $\beta_2$ grow linearly with $\lambda$, i.e., proportionally with the square root of S/N;
- the standard deviation of $\alpha_1$ and $\alpha_2$ go to a constant value;
- the standard deviation of $\delta_1$ and $\delta_2$ remain nearly constant; cf. fig. 7.18.

With increasing noise power $b_1 - \beta_1(k)$ becomes larger for increasing $k$.

In addition, it seems that with increasing noise power the speed of convergence decreases.

b). One notices that, with decreasing weighting factors:

- the means of the estimated values of the $\beta$'s vary more and more around the true values, whereas their standard deviations increase;
- $a_i - \alpha_i$ decrease constantly; the standard deviations diminish until $\rho = \nu = 0.990$ and then grow again;
- the means of the $\delta$'s improve, while the standard deviations become a little bit larger.

These experimental results suggest that for low noise levels greater weighting factors can be chosen than for high noise levels. For such situations where the parameters remain constant if seems appropriate to increase the weighting factors gradually to one in course of the estimation procedure in order to obtain both rapid convergence and small standard deviations.
The extended matrix method. The process description and the estimation algorithms used are those indicated by equations (7.60) to (7.67). Using for the process description the same parameters as before one finds:

\[ y(k) = \sum_{i=0}^{2} b_i u(k-i) - \sum_{i=1}^{2} a_i y(k-i) + r(k) \]  
(7.79)

\[
\begin{align*}
    b_0 &= 0 \\
    b_1 &= 1 \\
    b_2 &= 0.5
\end{align*}
\]

\[
\begin{align*}
    a_1 &= -1.5 \\
    a_2 &= 0.7
\end{align*}
\]

For the disturbances now a description with c and d parameters is chosen:

\[ r(k) = \lambda \sum c_i w(k-i) - \sum d_i r(k-i) + \lambda w(k) \]  
(7.80)

\[
\begin{align*}
    c_1 &= 0.3 \\
    c_2 &= 0.0 \\
    c_3 &= 0.0 \\
    c_4 &= 0.0
\end{align*}
\]

\[
\begin{align*}
    d_1 &= -0.5 \\
    d_2 &= 0.0 \\
    d_3 &= 0.0
\end{align*}
\]

Again the factor \( \lambda \) controls the noise level. The noise filter has a "power gain" \( \sigma_n^2 / \sigma_w^2 = 43.62 \).

Now one may study the effect on the process parameter estimates when using the wrong number (and kind) of noise parameters.

In the experimental results given, the weighting factor is increased to one in an exponential way:

\[ \rho(k+1) = \rho(k) + (1 - \rho(k)) \Delta \rho = (1 - \Delta \rho) \rho(k) + \Delta \rho \]  
(7.81)

\[ \rho(0) = 0.9912 \]

The following tables and figures of experimental results are given:
| Table | \( \lambda \) | Signal-to-
noise ratio | Estimated process parameters: \( \alpha_1, \alpha_2, \beta_0, \beta_1, \beta_2 \) | Estimated noise parameters: | Fig. |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7.12</td>
<td>1.000</td>
<td>0.432 dB</td>
<td>0.001 none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.13</td>
<td>same</td>
<td>same same same</td>
<td>same ( \delta_1 )</td>
<td></td>
<td>7.20</td>
</tr>
<tr>
<td>7.14</td>
<td>same</td>
<td>same same same</td>
<td>same ( \delta_1, \delta_2 )</td>
<td></td>
<td>7.21</td>
</tr>
<tr>
<td>7.15</td>
<td>same</td>
<td>same same same</td>
<td>same ( \delta_1, \delta_2, \delta_3 )</td>
<td></td>
<td>7.22</td>
</tr>
<tr>
<td>7.16</td>
<td>same</td>
<td>same same same</td>
<td>same ( \gamma_1 )</td>
<td></td>
<td>7.23</td>
</tr>
<tr>
<td>7.17</td>
<td>same</td>
<td>same same same</td>
<td>same ( \gamma_1, \gamma_2 )</td>
<td></td>
<td>7.24</td>
</tr>
<tr>
<td>7.18</td>
<td>same</td>
<td>same same same</td>
<td>same ( \gamma_1, \gamma_2, \gamma_3 )</td>
<td></td>
<td>7.25</td>
</tr>
<tr>
<td>7.19</td>
<td>same</td>
<td>same same same</td>
<td>same ( \gamma_1, \gamma_2, \gamma_3, \gamma_4 )</td>
<td></td>
<td>7.26</td>
</tr>
<tr>
<td>7.20</td>
<td>0.250</td>
<td>6.92 +8.3 dB</td>
<td>same ( \delta_1, \gamma_1 )</td>
<td></td>
<td>7.27</td>
</tr>
<tr>
<td>7.21</td>
<td>1.000</td>
<td>0.432 dB</td>
<td>same same</td>
<td></td>
<td>7.28</td>
</tr>
<tr>
<td>7.22</td>
<td>4.000</td>
<td>0.0269 -15.7 dB</td>
<td>same same</td>
<td></td>
<td>7.29</td>
</tr>
<tr>
<td>7.23</td>
<td>16.000</td>
<td>0.0017 -27.7 dB</td>
<td>same same</td>
<td></td>
<td>7.30</td>
</tr>
<tr>
<td>7.24</td>
<td>1.000</td>
<td>0.432 dB</td>
<td>0.001 same</td>
<td></td>
<td>7.31</td>
</tr>
<tr>
<td>7.25</td>
<td>same</td>
<td>same same same</td>
<td>0.001 same</td>
<td></td>
<td>7.32</td>
</tr>
<tr>
<td>7.26</td>
<td>same</td>
<td>same same same</td>
<td>0.005 same</td>
<td></td>
<td>7.33</td>
</tr>
<tr>
<td>7.27</td>
<td>same</td>
<td>same same same</td>
<td>0.025 same</td>
<td></td>
<td>7.34</td>
</tr>
</tbody>
</table>
Table 7.12
Only process parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(b_0)</th>
<th>(b_1)</th>
<th>(b_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.709</td>
<td>0.859</td>
<td>+0.024</td>
<td>1.042</td>
<td>0.386</td>
</tr>
<tr>
<td>200</td>
<td>-1.701</td>
<td>0.857</td>
<td>+0.012</td>
<td>1.029</td>
<td>0.386</td>
</tr>
<tr>
<td>300</td>
<td>-1.690</td>
<td>0.850</td>
<td>+0.014</td>
<td>1.026</td>
<td>0.394</td>
</tr>
<tr>
<td>400</td>
<td>-1.690</td>
<td>0.853</td>
<td>+0.059</td>
<td>1.057</td>
<td>0.398</td>
</tr>
<tr>
<td>500</td>
<td>-1.693</td>
<td>0.851</td>
<td>+0.042</td>
<td>1.035</td>
<td>0.392</td>
</tr>
<tr>
<td>600</td>
<td>-1.687</td>
<td>0.849</td>
<td>-0.004</td>
<td>0.968</td>
<td>0.357</td>
</tr>
<tr>
<td>700</td>
<td>-1.695</td>
<td>0.852</td>
<td>-0.011</td>
<td>0.974</td>
<td>0.375</td>
</tr>
<tr>
<td>800</td>
<td>-1.689</td>
<td>0.853</td>
<td>-0.029</td>
<td>0.968</td>
<td>0.390</td>
</tr>
<tr>
<td>900</td>
<td>-1.688</td>
<td>0.849</td>
<td>-0.011</td>
<td>0.974</td>
<td>0.380</td>
</tr>
<tr>
<td>1000</td>
<td>-1.689</td>
<td>0.846</td>
<td>+0.012</td>
<td>0.998</td>
<td>0.410</td>
</tr>
</tbody>
</table>

Table 7.13
Only one backward parameter estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(b_0)</th>
<th>(b_1)</th>
<th>(b_2)</th>
<th>(c_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.558</td>
<td>0.742</td>
<td>+0.005</td>
<td>1.038</td>
<td>0.482</td>
<td>-0.625</td>
</tr>
<tr>
<td>200</td>
<td>-1.523</td>
<td>0.719</td>
<td>-0.012</td>
<td>0.998</td>
<td>0.468</td>
<td>0.657</td>
</tr>
<tr>
<td>300</td>
<td>-1.503</td>
<td>0.715</td>
<td>+0.015</td>
<td>1.004</td>
<td>0.470</td>
<td>-0.696</td>
</tr>
<tr>
<td>400</td>
<td>-1.488</td>
<td>0.708</td>
<td>+0.014</td>
<td>1.006</td>
<td>0.483</td>
<td>-0.726</td>
</tr>
<tr>
<td>500</td>
<td>-1.482</td>
<td>0.707</td>
<td>-0.006</td>
<td>0.997</td>
<td>0.496</td>
<td>-0.721</td>
</tr>
<tr>
<td>600</td>
<td>-1.482</td>
<td>0.708</td>
<td>-0.020</td>
<td>0.960</td>
<td>0.472</td>
<td>-0.706</td>
</tr>
<tr>
<td>700</td>
<td>-1.482</td>
<td>0.713</td>
<td>-0.044</td>
<td>0.939</td>
<td>0.476</td>
<td>-0.702</td>
</tr>
<tr>
<td>800</td>
<td>-1.463</td>
<td>0.704</td>
<td>-0.003</td>
<td>0.967</td>
<td>0.482</td>
<td>-0.719</td>
</tr>
<tr>
<td>900</td>
<td>-1.479</td>
<td>0.715</td>
<td>-0.012</td>
<td>0.965</td>
<td>0.506</td>
<td>-0.703</td>
</tr>
<tr>
<td>1000</td>
<td>-1.481</td>
<td>0.718</td>
<td>-0.011</td>
<td>0.981</td>
<td>0.509</td>
<td>-0.700</td>
</tr>
</tbody>
</table>

Table 7.14
Two backward noise-parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(b_0)</th>
<th>(b_1)</th>
<th>(b_2)</th>
<th>(c_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.623</td>
<td>0.827</td>
<td>-0.019</td>
<td>0.983</td>
<td>0.438</td>
<td>-0.661</td>
</tr>
<tr>
<td>200</td>
<td>-1.569</td>
<td>0.741</td>
<td>+0.023</td>
<td>1.004</td>
<td>0.431</td>
<td>-0.735</td>
</tr>
<tr>
<td>300</td>
<td>-1.566</td>
<td>0.737</td>
<td>+0.004</td>
<td>0.977</td>
<td>0.442</td>
<td>-0.745</td>
</tr>
<tr>
<td>400</td>
<td>-1.547</td>
<td>0.723</td>
<td>-0.007</td>
<td>1.007</td>
<td>0.472</td>
<td>-0.762</td>
</tr>
<tr>
<td>500</td>
<td>-1.533</td>
<td>0.710</td>
<td>-0.022</td>
<td>0.980</td>
<td>0.471</td>
<td>-0.751</td>
</tr>
<tr>
<td>600</td>
<td>-1.515</td>
<td>0.698</td>
<td>-0.031</td>
<td>0.952</td>
<td>0.470</td>
<td>-0.763</td>
</tr>
<tr>
<td>700</td>
<td>-1.510</td>
<td>0.695</td>
<td>-0.015</td>
<td>0.960</td>
<td>0.467</td>
<td>-0.768</td>
</tr>
<tr>
<td>800</td>
<td>-1.511</td>
<td>0.699</td>
<td>-0.016</td>
<td>0.958</td>
<td>0.481</td>
<td>-0.762</td>
</tr>
<tr>
<td>900</td>
<td>-1.517</td>
<td>0.701</td>
<td>-0.007</td>
<td>0.982</td>
<td>0.507</td>
<td>-0.765</td>
</tr>
<tr>
<td>1000</td>
<td>-1.512</td>
<td>0.696</td>
<td>+0.005</td>
<td>0.998</td>
<td>0.507</td>
<td>-0.770</td>
</tr>
</tbody>
</table>
Table 7.15
3 backward noise
parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.583</td>
<td>.756</td>
<td>.025</td>
<td>1.006</td>
<td>.419</td>
<td>-.720</td>
<td>.238</td>
</tr>
<tr>
<td>200</td>
<td>-1.558</td>
<td>.735</td>
<td>.017</td>
<td>.994</td>
<td>.436</td>
<td>-.670</td>
<td>.223</td>
</tr>
<tr>
<td>300</td>
<td>-1.551</td>
<td>.737</td>
<td>.008</td>
<td>.999</td>
<td>.557</td>
<td>-.752</td>
<td>.209</td>
</tr>
<tr>
<td>400</td>
<td>-1.535</td>
<td>.722</td>
<td>-.014</td>
<td>.989</td>
<td>.666</td>
<td>-.764</td>
<td>.205</td>
</tr>
<tr>
<td>500</td>
<td>-.525</td>
<td>.719</td>
<td>.053</td>
<td>.945</td>
<td>.642</td>
<td>-.791</td>
<td>.207</td>
</tr>
<tr>
<td>600</td>
<td>-.503</td>
<td>.697</td>
<td>-.019</td>
<td>.956</td>
<td>.662</td>
<td>-.776</td>
<td>.224</td>
</tr>
<tr>
<td>700</td>
<td>-.502</td>
<td>.701</td>
<td>-.026</td>
<td>.952</td>
<td>.668</td>
<td>-.778</td>
<td>.215</td>
</tr>
<tr>
<td>800</td>
<td>-.510</td>
<td>.708</td>
<td>-.012</td>
<td>.975</td>
<td>.505</td>
<td>-.771</td>
<td>.218</td>
</tr>
<tr>
<td>900</td>
<td>-.514</td>
<td>.707</td>
<td>.003</td>
<td>.995</td>
<td>.507</td>
<td>-.776</td>
<td>.241</td>
</tr>
<tr>
<td>1000</td>
<td>-.511</td>
<td>.710</td>
<td>.010</td>
<td>1.008</td>
<td>.503</td>
<td>-.783</td>
<td>.232</td>
</tr>
</tbody>
</table>

Table 7.16
One forward noise-
parameter estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.643</td>
<td>.804</td>
<td>-.009</td>
<td>1.026</td>
<td>.429</td>
<td>.642</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>-1.621</td>
<td>.784</td>
<td>-.008</td>
<td>1.001</td>
<td>.418</td>
<td>.600</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>-1.620</td>
<td>.786</td>
<td>.028</td>
<td>1.023</td>
<td>.419</td>
<td>.673</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>-1.615</td>
<td>.781</td>
<td>.029</td>
<td>1.033</td>
<td>.436</td>
<td>.702</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>-1.611</td>
<td>.776</td>
<td>.006</td>
<td>1.016</td>
<td>.439</td>
<td>.693</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>-1.612</td>
<td>.776</td>
<td>-.016</td>
<td>.975</td>
<td>.414</td>
<td>.666</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>-1.609</td>
<td>.776</td>
<td>-.040</td>
<td>.954</td>
<td>.421</td>
<td>.659</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>-1.598</td>
<td>.769</td>
<td>-.010</td>
<td>.980</td>
<td>.437</td>
<td>.672</td>
<td></td>
</tr>
<tr>
<td>900</td>
<td>-1.607</td>
<td>.772</td>
<td>-.014</td>
<td>.972</td>
<td>.461</td>
<td>.648</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>-1.602</td>
<td>.767</td>
<td>-.014</td>
<td>.995</td>
<td>.465</td>
<td>.677</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.17
Two forward noise-
parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.625</td>
<td>.794</td>
<td>-.029</td>
<td>.973</td>
<td>.435</td>
<td>.683</td>
<td>.237</td>
</tr>
<tr>
<td>200</td>
<td>-1.578</td>
<td>.751</td>
<td>.023</td>
<td>1.003</td>
<td>.434</td>
<td>.746</td>
<td>.306</td>
</tr>
<tr>
<td>300</td>
<td>-1.579</td>
<td>.750</td>
<td>.007</td>
<td>.961</td>
<td>.438</td>
<td>.747</td>
<td>.306</td>
</tr>
<tr>
<td>400</td>
<td>-1.572</td>
<td>.745</td>
<td>-.006</td>
<td>1.011</td>
<td>.465</td>
<td>.751</td>
<td>.299</td>
</tr>
<tr>
<td>500</td>
<td>-1.559</td>
<td>.732</td>
<td>-.026</td>
<td>.977</td>
<td>.458</td>
<td>.738</td>
<td>.305</td>
</tr>
<tr>
<td>600</td>
<td>-1.555</td>
<td>.728</td>
<td>-.033</td>
<td>.951</td>
<td>.451</td>
<td>.741</td>
<td>.322</td>
</tr>
<tr>
<td>700</td>
<td>-1.547</td>
<td>.724</td>
<td>-.015</td>
<td>.958</td>
<td>.448</td>
<td>.766</td>
<td>.300</td>
</tr>
<tr>
<td>800</td>
<td>-1.542</td>
<td>.727</td>
<td>-.018</td>
<td>.956</td>
<td>.462</td>
<td>.736</td>
<td>.315</td>
</tr>
<tr>
<td>900</td>
<td>-1.533</td>
<td>.728</td>
<td>-.007</td>
<td>.990</td>
<td>.487</td>
<td>.741</td>
<td>.297</td>
</tr>
<tr>
<td>1000</td>
<td>-1.552</td>
<td>.726</td>
<td>.005</td>
<td>.997</td>
<td>.487</td>
<td>.741</td>
<td>.297</td>
</tr>
</tbody>
</table>
### Table 7.18

3 forward noise-parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.591</td>
<td>0.764</td>
<td>+0.039</td>
<td>1.020</td>
<td>4.424</td>
<td>7.604</td>
<td>.316</td>
<td>.124</td>
</tr>
<tr>
<td>200</td>
<td>-1.561</td>
<td>0.760</td>
<td>+0.015</td>
<td>1.001</td>
<td>4.444</td>
<td>7.777</td>
<td>.351</td>
<td>.087</td>
</tr>
<tr>
<td>300</td>
<td>-1.555</td>
<td>0.760</td>
<td>+0.009</td>
<td>1.002</td>
<td>4.577</td>
<td>7.533</td>
<td>.350</td>
<td>.121</td>
</tr>
<tr>
<td>400</td>
<td>-1.543</td>
<td>0.725</td>
<td>-0.008</td>
<td>0.992</td>
<td>4.666</td>
<td>7.677</td>
<td>.368</td>
<td>.141</td>
</tr>
<tr>
<td>500</td>
<td>-1.533</td>
<td>0.717</td>
<td>-0.019</td>
<td>0.947</td>
<td>4.500</td>
<td>7.698</td>
<td>.359</td>
<td>.149</td>
</tr>
<tr>
<td>600</td>
<td>-1.517</td>
<td>0.699</td>
<td>-0.017</td>
<td>0.961</td>
<td>4.559</td>
<td>7.744</td>
<td>.347</td>
<td>.171</td>
</tr>
<tr>
<td>700</td>
<td>-1.518</td>
<td>0.704</td>
<td>-0.025</td>
<td>0.956</td>
<td>4.633</td>
<td>7.688</td>
<td>.358</td>
<td>.156</td>
</tr>
<tr>
<td>800</td>
<td>-1.521</td>
<td>0.707</td>
<td>-0.016</td>
<td>0.975</td>
<td>5.023</td>
<td>7.699</td>
<td>.355</td>
<td>.173</td>
</tr>
<tr>
<td>900</td>
<td>-1.525</td>
<td>0.704</td>
<td>-0.001</td>
<td>0.997</td>
<td>5.031</td>
<td>7.788</td>
<td>.332</td>
<td>.160</td>
</tr>
<tr>
<td>1000</td>
<td>-1.544</td>
<td>0.707</td>
<td>+0.007</td>
<td>1.009</td>
<td>4.988</td>
<td>7.799</td>
<td>.346</td>
<td>.151</td>
</tr>
</tbody>
</table>

### Table 7.19

4 forward noise-parameters estimated.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.621</td>
<td>0.788</td>
<td>-0.018</td>
<td>0.966</td>
<td>4.111</td>
<td>7.739</td>
<td>.241</td>
<td>.052</td>
</tr>
<tr>
<td>200</td>
<td>-1.587</td>
<td>0.757</td>
<td>+0.003</td>
<td>0.951</td>
<td>4.602</td>
<td>7.378</td>
<td>.304</td>
<td>.067</td>
</tr>
<tr>
<td>300</td>
<td>-1.565</td>
<td>0.744</td>
<td>-0.019</td>
<td>0.996</td>
<td>4.448</td>
<td>7.670</td>
<td>.307</td>
<td>.092</td>
</tr>
<tr>
<td>400</td>
<td>-1.547</td>
<td>0.729</td>
<td>-0.014</td>
<td>0.969</td>
<td>4.647</td>
<td>7.670</td>
<td>.327</td>
<td>.151</td>
</tr>
<tr>
<td>500</td>
<td>-1.523</td>
<td>0.708</td>
<td>-0.019</td>
<td>0.938</td>
<td>4.448</td>
<td>7.779</td>
<td>.368</td>
<td>.170</td>
</tr>
<tr>
<td>600</td>
<td>-1.509</td>
<td>0.701</td>
<td>-0.023</td>
<td>0.947</td>
<td>4.477</td>
<td>7.769</td>
<td>.381</td>
<td>.175</td>
</tr>
<tr>
<td>700</td>
<td>-1.515</td>
<td>0.706</td>
<td>-0.020</td>
<td>0.975</td>
<td>4.999</td>
<td>7.799</td>
<td>.361</td>
<td>.182</td>
</tr>
<tr>
<td>800</td>
<td>-1.517</td>
<td>0.705</td>
<td>-0.009</td>
<td>0.987</td>
<td>5.101</td>
<td>7.863</td>
<td>.363</td>
<td>.199</td>
</tr>
<tr>
<td>900</td>
<td>-1.537</td>
<td>0.702</td>
<td>+0.017</td>
<td>1.007</td>
<td>5.053</td>
<td>7.888</td>
<td>.352</td>
<td>.177</td>
</tr>
<tr>
<td>1000</td>
<td>-1.546</td>
<td>0.706</td>
<td>+0.003</td>
<td>1.006</td>
<td>4.955</td>
<td>7.936</td>
<td>.362</td>
<td>.175</td>
</tr>
</tbody>
</table>

### Table 7.20

\[ N = 6.92 \text{ (8.3 dB)}. \]
Table 7.21

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b_0 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( y_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.601</td>
<td>.777</td>
<td>-.018</td>
<td>.983</td>
<td>.432</td>
<td>-.366</td>
<td>.332</td>
</tr>
<tr>
<td>200</td>
<td>-1.546</td>
<td>.729</td>
<td>+.023</td>
<td>1.008</td>
<td>.456</td>
<td>-.455</td>
<td>.311</td>
</tr>
<tr>
<td>300</td>
<td>-1.544</td>
<td>.729</td>
<td>-.003</td>
<td>.974</td>
<td>.650</td>
<td>-.474</td>
<td>.293</td>
</tr>
<tr>
<td>400</td>
<td>-1.533</td>
<td>.721</td>
<td>-.005</td>
<td>1.019</td>
<td>.488</td>
<td>-.483</td>
<td>.301</td>
</tr>
<tr>
<td>500</td>
<td>-1.522</td>
<td>.709</td>
<td>-.023</td>
<td>.980</td>
<td>.479</td>
<td>-.474</td>
<td>.305</td>
</tr>
<tr>
<td>600</td>
<td>-1.509</td>
<td>.702</td>
<td>-.027</td>
<td>.955</td>
<td>.476</td>
<td>-.516</td>
<td>.273</td>
</tr>
<tr>
<td>700</td>
<td>-1.499</td>
<td>.698</td>
<td>-.020</td>
<td>.960</td>
<td>.464</td>
<td>-.503</td>
<td>.288</td>
</tr>
<tr>
<td>800</td>
<td>-1.504</td>
<td>.703</td>
<td>-.018</td>
<td>.964</td>
<td>.488</td>
<td>-.499</td>
<td>.285</td>
</tr>
<tr>
<td>900</td>
<td>-1.510</td>
<td>.706</td>
<td>-.011</td>
<td>.978</td>
<td>.503</td>
<td>-.481</td>
<td>.301</td>
</tr>
<tr>
<td>1000</td>
<td>-1.508</td>
<td>.703</td>
<td>+.002</td>
<td>1.000</td>
<td>.504</td>
<td>-.482</td>
<td>.308</td>
</tr>
</tbody>
</table>

\( S_N = 0.432 \) (-3.7 dB).

Table 7.22

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b_0 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( y_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.627</td>
<td>.793</td>
<td>-.077</td>
<td>.915</td>
<td>.415</td>
<td>-.327</td>
<td>.349</td>
</tr>
<tr>
<td>200</td>
<td>-1.564</td>
<td>.740</td>
<td>+.098</td>
<td>.997</td>
<td>.347</td>
<td>-.425</td>
<td>.337</td>
</tr>
<tr>
<td>300</td>
<td>-1.557</td>
<td>.744</td>
<td>+.011</td>
<td>.901</td>
<td>.364</td>
<td>-.432</td>
<td>.323</td>
</tr>
<tr>
<td>400</td>
<td>-1.558</td>
<td>.741</td>
<td>-.025</td>
<td>1.030</td>
<td>.439</td>
<td>-.440</td>
<td>.329</td>
</tr>
<tr>
<td>500</td>
<td>-1.542</td>
<td>.726</td>
<td>-.092</td>
<td>.912</td>
<td>.412</td>
<td>-.441</td>
<td>.317</td>
</tr>
<tr>
<td>600</td>
<td>-1.526</td>
<td>.714</td>
<td>-.124</td>
<td>.797</td>
<td>.378</td>
<td>-.502</td>
<td>.269</td>
</tr>
<tr>
<td>700</td>
<td>-1.520</td>
<td>.712</td>
<td>-.058</td>
<td>.834</td>
<td>.359</td>
<td>-.460</td>
<td>.317</td>
</tr>
<tr>
<td>800</td>
<td>-1.516</td>
<td>.710</td>
<td>-.074</td>
<td>.826</td>
<td>.430</td>
<td>-.490</td>
<td>.281</td>
</tr>
<tr>
<td>900</td>
<td>-1.517</td>
<td>.710</td>
<td>-.044</td>
<td>.906</td>
<td>.533</td>
<td>-.470</td>
<td>.308</td>
</tr>
<tr>
<td>1000</td>
<td>-1.515</td>
<td>.706</td>
<td>+.005</td>
<td>.969</td>
<td>.517</td>
<td>-.473</td>
<td>.311</td>
</tr>
</tbody>
</table>

\( S_N = 0.0269 \) (-15.7 dB).

Table 7.23

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b_0 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( y_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.623</td>
<td>.788</td>
<td>-.329</td>
<td>.603</td>
<td>+.304</td>
<td>-.342</td>
<td>.330</td>
</tr>
<tr>
<td>200</td>
<td>-1.562</td>
<td>.737</td>
<td>+.386</td>
<td>.553</td>
<td>-.052</td>
<td>-.426</td>
<td>.338</td>
</tr>
<tr>
<td>300</td>
<td>-1.568</td>
<td>.702</td>
<td>1.302</td>
<td>1.956</td>
<td>1.875</td>
<td>1.155</td>
<td>1.80</td>
</tr>
<tr>
<td>400</td>
<td>-1.561</td>
<td>.743</td>
<td>+.093</td>
<td>1.111</td>
<td>+.331</td>
<td>-.434</td>
<td>.333</td>
</tr>
<tr>
<td>500</td>
<td>-1.546</td>
<td>.729</td>
<td>-.370</td>
<td>.634</td>
<td>+.196</td>
<td>-.435</td>
<td>.320</td>
</tr>
<tr>
<td>600</td>
<td>-1.530</td>
<td>.718</td>
<td>-.497</td>
<td>.179</td>
<td>+.240</td>
<td>-.495</td>
<td>.272</td>
</tr>
<tr>
<td>700</td>
<td>-1.525</td>
<td>.716</td>
<td>-.230</td>
<td>.335</td>
<td>-.043</td>
<td>-.451</td>
<td>.320</td>
</tr>
<tr>
<td>800</td>
<td>-1.519</td>
<td>.712</td>
<td>-.297</td>
<td>.299</td>
<td>+.235</td>
<td>-.485</td>
<td>.283</td>
</tr>
<tr>
<td>900</td>
<td>-1.518</td>
<td>.711</td>
<td>-.180</td>
<td>.613</td>
<td>+.645</td>
<td>-.470</td>
<td>.308</td>
</tr>
<tr>
<td>1000</td>
<td>-1.516</td>
<td>.707</td>
<td>-.013</td>
<td>.863</td>
<td>+.576</td>
<td>-.471</td>
<td>.312</td>
</tr>
</tbody>
</table>

\( S_N = 0.0017 \) (-27.7 dB).
Table 7.21

<table>
<thead>
<tr>
<th>S/N</th>
<th>0.432 (-3.7 dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>𝑞_1 𝜔_2 𝜖_0 𝜖_1 𝜖_2 𝜖_3 𝜀_1</td>
</tr>
<tr>
<td>100</td>
<td>-1.601 0.777 -0.018 0.983 0.433 -3.366 0.322</td>
</tr>
<tr>
<td>200</td>
<td>-1.546 0.729 +0.03 1.008 0.656 -4.555 0.311</td>
</tr>
<tr>
<td>300</td>
<td>-1.544 0.729 -0.003 0.974 0.650 -4.744 0.293</td>
</tr>
<tr>
<td>400</td>
<td>-1.533 0.721 -0.005 1.019 0.488 -4.833 0.301</td>
</tr>
<tr>
<td>500</td>
<td>-1.522 0.709 -0.023 0.980 0.479 -4.744 0.301</td>
</tr>
<tr>
<td>600</td>
<td>-1.509 0.702 -0.027 0.955 0.476 -5.161 0.273</td>
</tr>
<tr>
<td>700</td>
<td>-1.499 0.698 -0.020 0.960 0.664 -5.033 0.288</td>
</tr>
<tr>
<td>800</td>
<td>-1.504 0.703 -0.018 0.964 0.488 -4.999 0.285</td>
</tr>
<tr>
<td>900</td>
<td>-1.510 0.706 -0.011 0.978 0.503 -4.811 0.301</td>
</tr>
<tr>
<td>1000</td>
<td>-1.508 0.703 +0.002 1.000 0.504 -4.821 0.308</td>
</tr>
</tbody>
</table>

Table 7.22

<table>
<thead>
<tr>
<th>S/N</th>
<th>0.0269 (-15.7 dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>𝑞_1 𝜔_2 𝜖_0 𝜖_1 𝜖_2 𝜖_3 𝜀_1</td>
</tr>
<tr>
<td>100</td>
<td>-1.627 0.793 -0.077 0.915 0.415 -3.27 0.349</td>
</tr>
<tr>
<td>200</td>
<td>-1.564 0.740 +0.098 0.997 0.347 -4.425 0.337</td>
</tr>
<tr>
<td>300</td>
<td>-1.557 0.744 +0.011 0.901 0.364 -4.332 0.323</td>
</tr>
<tr>
<td>400</td>
<td>-1.558 0.741 -0.025 1.030 0.439 -4.400 0.329</td>
</tr>
<tr>
<td>500</td>
<td>-1.542 0.726 -0.092 0.912 0.412 -4.441 0.317</td>
</tr>
<tr>
<td>600</td>
<td>-1.526 0.714 -0.124 0.797 0.378 -4.502 0.269</td>
</tr>
<tr>
<td>700</td>
<td>-1.520 0.712 -0.058 0.834 0.359 -4.660 0.317</td>
</tr>
<tr>
<td>800</td>
<td>-1.516 0.710 -0.074 0.826 0.430 -4.490 0.281</td>
</tr>
<tr>
<td>900</td>
<td>-1.517 0.710 -0.044 0.906 0.533 -4.707 0.308</td>
</tr>
<tr>
<td>1000</td>
<td>-1.515 0.706 +0.005 0.969 0.517 -4.733 0.311</td>
</tr>
</tbody>
</table>

Table 7.23

<table>
<thead>
<tr>
<th>S/N</th>
<th>40017 (-27.7 dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>𝑞_1 𝜔_2 𝜖_0 𝜖_1 𝜖_2 𝜖_3 𝜀_1</td>
</tr>
<tr>
<td>100</td>
<td>-1.623 0.788 -0.329 0.603 +0.304 -3.342 0.330</td>
</tr>
<tr>
<td>200</td>
<td>-1.562 0.737 +0.386 0.553 -0.052 -4.266 0.338</td>
</tr>
<tr>
<td>300</td>
<td>-1.568 0.702 1.302 1.956 1.875 0.155 0.180</td>
</tr>
<tr>
<td>400</td>
<td>-1.561 0.743 -0.098 1.111 +0.331 -4.393 0.333</td>
</tr>
<tr>
<td>500</td>
<td>-1.546 0.729 -0.370 0.636 +0.196 -4.435 0.320</td>
</tr>
<tr>
<td>600</td>
<td>-1.530 0.718 -0.497 0.179 +0.404 -4.957 0.272</td>
</tr>
<tr>
<td>700</td>
<td>-1.525 0.716 -0.230 0.335 -0.024 -4.511 0.320</td>
</tr>
<tr>
<td>800</td>
<td>-1.519 0.712 -0.297 0.299 +0.235 -4.858 0.283</td>
</tr>
<tr>
<td>900</td>
<td>-1.518 0.711 -0.180 0.613 +0.645 -4.708 0.308</td>
</tr>
<tr>
<td>1000</td>
<td>-1.516 0.707 -0.013 0.863 +0.576 -4.471 0.312</td>
</tr>
<tr>
<td>number of iterations</td>
<td>$c_1$</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>100</td>
<td>1.605</td>
</tr>
<tr>
<td>200</td>
<td>1.548</td>
</tr>
<tr>
<td>300</td>
<td>1.543</td>
</tr>
<tr>
<td>400</td>
<td>1.531</td>
</tr>
<tr>
<td>500</td>
<td>1.515</td>
</tr>
<tr>
<td>600</td>
<td>1.497</td>
</tr>
<tr>
<td>700</td>
<td>1.488</td>
</tr>
<tr>
<td>800</td>
<td>1.496</td>
</tr>
<tr>
<td>900</td>
<td>1.516</td>
</tr>
<tr>
<td>1000</td>
<td>1.509</td>
</tr>
</tbody>
</table>

Table 7.24

$\rho_o = 0.9913$. $\Delta \rho = 0.0$.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.601</td>
<td>0.777</td>
<td>-0.018</td>
<td>0.983</td>
<td>0.433</td>
<td>-0.366</td>
<td>0.332</td>
</tr>
<tr>
<td>200</td>
<td>1.546</td>
<td>0.729</td>
<td>+0.023</td>
<td>1.008</td>
<td>0.456</td>
<td>-0.435</td>
<td>0.311</td>
</tr>
<tr>
<td>300</td>
<td>1.544</td>
<td>0.729</td>
<td>-0.003</td>
<td>0.974</td>
<td>0.450</td>
<td>-0.474</td>
<td>0.293</td>
</tr>
<tr>
<td>400</td>
<td>1.533</td>
<td>0.721</td>
<td>-0.005</td>
<td>0.979</td>
<td>0.488</td>
<td>-0.483</td>
<td>0.301</td>
</tr>
<tr>
<td>500</td>
<td>1.522</td>
<td>0.709</td>
<td>-0.023</td>
<td>0.980</td>
<td>0.479</td>
<td>-0.474</td>
<td>0.305</td>
</tr>
<tr>
<td>600</td>
<td>1.509</td>
<td>0.702</td>
<td>-0.027</td>
<td>0.980</td>
<td>0.476</td>
<td>-0.516</td>
<td>0.273</td>
</tr>
<tr>
<td>700</td>
<td>1.499</td>
<td>0.698</td>
<td>-0.020</td>
<td>0.960</td>
<td>0.464</td>
<td>-0.503</td>
<td>0.288</td>
</tr>
<tr>
<td>800</td>
<td>1.504</td>
<td>0.703</td>
<td>-0.018</td>
<td>0.964</td>
<td>0.489</td>
<td>-0.499</td>
<td>0.285</td>
</tr>
<tr>
<td>900</td>
<td>1.510</td>
<td>0.706</td>
<td>-0.011</td>
<td>0.978</td>
<td>0.503</td>
<td>-0.481</td>
<td>0.301</td>
</tr>
<tr>
<td>1000</td>
<td>1.508</td>
<td>0.703</td>
<td>+0.002</td>
<td>1.000</td>
<td>0.504</td>
<td>-0.482</td>
<td>0.308</td>
</tr>
</tbody>
</table>

Table 7.25

$\rho_o = 0.9913$. $\Delta \rho = 0.001$.

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.603</td>
<td>0.778</td>
<td>-0.018</td>
<td>0.984</td>
<td>0.434</td>
<td>-0.366</td>
<td>0.328</td>
</tr>
<tr>
<td>200</td>
<td>1.556</td>
<td>0.736</td>
<td>+0.019</td>
<td>1.006</td>
<td>0.447</td>
<td>-0.466</td>
<td>0.305</td>
</tr>
<tr>
<td>300</td>
<td>1.551</td>
<td>0.733</td>
<td>-0.001</td>
<td>0.978</td>
<td>0.444</td>
<td>-0.468</td>
<td>0.286</td>
</tr>
<tr>
<td>400</td>
<td>1.543</td>
<td>0.727</td>
<td>-0.001</td>
<td>1.007</td>
<td>0.448</td>
<td>-0.475</td>
<td>0.286</td>
</tr>
<tr>
<td>500</td>
<td>1.533</td>
<td>0.720</td>
<td>-0.010</td>
<td>0.990</td>
<td>0.466</td>
<td>-0.471</td>
<td>0.294</td>
</tr>
<tr>
<td>600</td>
<td>1.528</td>
<td>0.715</td>
<td>-0.014</td>
<td>0.974</td>
<td>0.467</td>
<td>-0.484</td>
<td>0.284</td>
</tr>
<tr>
<td>700</td>
<td>1.523</td>
<td>0.712</td>
<td>-0.013</td>
<td>0.974</td>
<td>0.464</td>
<td>-0.484</td>
<td>0.284</td>
</tr>
<tr>
<td>800</td>
<td>1.522</td>
<td>0.712</td>
<td>-0.013</td>
<td>0.974</td>
<td>0.474</td>
<td>-0.485</td>
<td>0.282</td>
</tr>
<tr>
<td>900</td>
<td>1.522</td>
<td>0.711</td>
<td>-0.011</td>
<td>0.977</td>
<td>0.483</td>
<td>-0.481</td>
<td>0.287</td>
</tr>
<tr>
<td>1000</td>
<td>1.519</td>
<td>0.709</td>
<td>-0.005</td>
<td>0.985</td>
<td>0.486</td>
<td>-0.482</td>
<td>0.290</td>
</tr>
</tbody>
</table>

Table 7.26

$\rho_o = 0.9913$. $\Delta \rho = 0.005$.
Table 7.27

\( p_0 = 0.9913 \), \( \Delta p = 0.025 \).

<table>
<thead>
<tr>
<th>number of iterations</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \gamma_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.612</td>
<td>.785</td>
<td>-.019</td>
<td>.092</td>
<td>.447</td>
<td>-.343</td>
</tr>
<tr>
<td></td>
<td>.043</td>
<td>.050</td>
<td>.091</td>
<td>.113</td>
<td>.124</td>
<td>.179</td>
</tr>
<tr>
<td>200</td>
<td>-1.578</td>
<td>.753</td>
<td>+.012</td>
<td>1.001</td>
<td>.434</td>
<td>-.410</td>
</tr>
<tr>
<td></td>
<td>.040</td>
<td>.047</td>
<td>.072</td>
<td>.087</td>
<td>.069</td>
<td>.142</td>
</tr>
<tr>
<td>300</td>
<td>-1.571</td>
<td>.747</td>
<td>+.001</td>
<td>.983</td>
<td>.437</td>
<td>-.434</td>
</tr>
<tr>
<td></td>
<td>.032</td>
<td>.040</td>
<td>.076</td>
<td>.087</td>
<td>.068</td>
<td>.094</td>
</tr>
<tr>
<td>400</td>
<td>-1.562</td>
<td>.760</td>
<td>-.003</td>
<td>.996</td>
<td>.449</td>
<td>-.442</td>
</tr>
<tr>
<td></td>
<td>.029</td>
<td>.036</td>
<td>.057</td>
<td>.087</td>
<td>.063</td>
<td>.088</td>
</tr>
<tr>
<td>500</td>
<td>-1.554</td>
<td>.734</td>
<td>-.007</td>
<td>.989</td>
<td>.451</td>
<td>-.445</td>
</tr>
<tr>
<td></td>
<td>.027</td>
<td>.033</td>
<td>.046</td>
<td>.085</td>
<td>.062</td>
<td>.085</td>
</tr>
<tr>
<td>600</td>
<td>-1.548</td>
<td>.729</td>
<td>-.011</td>
<td>.978</td>
<td>.454</td>
<td>-.463</td>
</tr>
<tr>
<td></td>
<td>.029</td>
<td>.035</td>
<td>.050</td>
<td>.082</td>
<td>.057</td>
<td>.093</td>
</tr>
<tr>
<td>700</td>
<td>-1.544</td>
<td>.726</td>
<td>-.009</td>
<td>.978</td>
<td>.456</td>
<td>-.450</td>
</tr>
<tr>
<td></td>
<td>.023</td>
<td>.029</td>
<td>.050</td>
<td>.076</td>
<td>.048</td>
<td>.085</td>
</tr>
<tr>
<td>800</td>
<td>-1.542</td>
<td>.725</td>
<td>-.010</td>
<td>.977</td>
<td>.462</td>
<td>-.458</td>
</tr>
<tr>
<td></td>
<td>.020</td>
<td>.025</td>
<td>.049</td>
<td>.077</td>
<td>.063</td>
<td>.080</td>
</tr>
<tr>
<td>900</td>
<td>-1.540</td>
<td>.723</td>
<td>-.009</td>
<td>.981</td>
<td>.472</td>
<td>-.454</td>
</tr>
<tr>
<td></td>
<td>.021</td>
<td>.027</td>
<td>.041</td>
<td>.069</td>
<td>.066</td>
<td>.077</td>
</tr>
<tr>
<td>1000</td>
<td>-1.538</td>
<td>.721</td>
<td>-.005</td>
<td>.985</td>
<td>.475</td>
<td>-.458</td>
</tr>
<tr>
<td></td>
<td>.022</td>
<td>.025</td>
<td>.038</td>
<td>.060</td>
<td>.065</td>
<td>.071</td>
</tr>
</tbody>
</table>
Table 7.12 represents cases where only process parameters are estimated. As expected, there is an appreciable bias in some of the estimates such that the true value does not fall in the one-standard-deviation interval.

Tables 7.13-7.15 represent cases where the model interprets the noise by means of an autoregression expression ("backward" parameters), i.e., as if it were generated by:

\[ r(k) = - \sum a_i r(k-i) + \lambda w(k) \]

or

\[ r(k) = \frac{\lambda w(k)}{D^*(z^{-1})} \] \hspace{1cm} (7.82)

The true noise generation mechanism can be approximated in this way:

\[ r(k) = \frac{C^*(z^{-1})}{D^*(z^{-1})} \lambda w(k) = \frac{1 + 0.3 z^{-1}}{1 - 0.5 z^{-1}} \lambda w(k) = \]

\[ = \frac{\lambda w(k)}{1 - 0.8 z^{-1} + 0.24 z^{-2} - 0.072 z^{-3} + \ldots} \] \hspace{1cm} (7.83)

Increasing the number of parameters will result in a better representation of the noise and the process parameters; a comparison with Fig. 7.19 indicates the improvement with respect to the bias. Note the true parameter values with respect to the estimated values ± the standard deviation.

Tables 7.16-7.19 represent cases where the model interprets the noise by means of a moving-average model ("forward" parameters), i.e., as if it were generated by:
The true noise generation mechanism can be approximated in this way:

\[ r^*(k) = \lambda \left\{ \sum c_i w(k-i) + w(k) \right\} \]

or

\[ r''(k) = \lambda C(z^{-1}) w(k) \]  \hspace{1cm} (7.84)

Again increasing the number of parameters will result in a better representation of the noise and the process parameters; cf. with fig. 7.19.

It turns out that one has to estimate more \( c \) parameters than \( d \) parameters in order to obtain results of the same quality. This is due to the fact that the values of \( d \) decline more rapidly with increasing \( i \) than the values of \( c \).

Tables 7.20-7.23 represent the case where the model interprets the noise by means of one \( c \) and one \( d \) parameter, i.e., a model that has the right topology and number of parameters. Table 7.21/fig. 7.28 has the same \( S/N \) ratio as the examples discussed before; the process and noise parameters are estimated very well.

These tables and figures give the results of experiments with decreasing signal to noise ratios at the process output; in each step the \( S/N \) is deteriorated 4-fold; in fig. 7.30 note that \( S/N = 0.0017 \) !
Study of these results indicates that:
- the standard deviations of the $\beta$ parameters grow linearly with $A$, i.e., proportionally with the square root of $S/N$;
- the standard deviations of the $\alpha$ parameters tend to a constant value;
- the standard deviations of the $\gamma$ and $\delta$ parameters remain constant.

Tables 7.24-7.27 represent cases where the speed of change with $k$ of the weighting factor $\rho$ changes. The cases chosen:

$$\Delta \rho = 0 \quad 0.001 \quad 0.005 \quad 0.025$$

correspond to a reduction of the difference between its value and 1 to half the initial value of that difference after the following number of iterations:

$$\infty \quad 698 \quad 128 \quad 27$$

From the tables, one notices that with increasing $\rho$ the standard deviations tend to diminish. When increasing $\Delta \rho$ then in general the means of the estimates become worse, while the standard deviation tends to some lower bound.

For a complete insight into the properties of the different estimation schemes, more theoretical work and experiments have to be done. This holds particularly for the choice of the sequence of weighting factors. Of vital importance remains the problem of choosing or deriving the right number of $\alpha, \beta, \gamma$ and $\delta$ parameters.
7.4 Models, nonlinear-in-the-parameters.

In chapter 4, the importance of linearity-in-the-parameters has been stressed. The methods in the previous sections were devoted to such processes. For the simplicity of the estimating schemes this type of linearity proved to be of more importance than the linearity-of-the-dynamics. For special classes of nonlinear systems, where the linear dynamics and the nonlinearity-without-memory can be separated, the identification technique by means of an adjustable model is sometimes feasible; c.f. Butler and Bohn (1966), Narendra and Gallman (1966).

Let us now consider the general case of nonlinearity-in-the-parameters. The model output may be written as

\[ w = f(U, \beta) \]  
(7.86)

the error and error (loss) function may again be given as

\[ E = e' e \quad \text{with} \quad e = y - w = y - f(U, \beta) \]  
(7.87)

From this point on the techniques mentioned in section 5.2 for minimization of a function can be used. Note that linearity-in-the-parameters resulted in an error function which is quadratic-in-the-parameters. This neat property is lost in the cases that are considered now. Two approaches may serve as illustrations:

a) Gauss or Gauss-Newton method, i.e., an expansion of \( f \) in a Taylor series:

\[ w = f(U, \beta, \Delta \beta) \approx f(U, \beta) + \frac{\partial f}{\partial \beta} \Delta \beta \]

or

\[ w \approx f^* + M \Delta \beta \]  
(7.88)

Now \( w \) is again linear in \( \Delta \beta \); consequently, \( \Delta \beta \) can be found by the standard least squares techniques discussed before. The model is adjusted according to \( \Delta \beta \) and the same procedure is followed in the next iteration.
b) **Gradient method.** The gradient of the error function is determined:

\[ \nabla_\beta E = \left[ \frac{\partial E}{\partial \beta_1}, \ldots, \frac{\partial E}{\partial \beta_m} \right]' \tag{7.89} \]

and the model adjustment is chosen as \( \Gamma \nabla_\beta E \).

In section 2.4 the determination of partial derivatives or gradients has been discussed. It may be done by evaluating the effects of parameter perturbations of the "model". For such situations, a fast analog computer model, combined with a digital machine, that is, a hybrid computer, may be a useful implementation.
Generally speaking, method a) may suffer from divergence, method b) may converge very slowly in the vicinity of the optimum. In Marquardt (1963) both methods are combined.

The reason for using a model nonlinear-in-the-parameters may be: - the type of process-representation needed for the application in mind; - the properties of the estimates.

If our attention is focussed on the pulse transfer function representation, then the generalized model is an adequate choice. It has the disadvantage, however, that the additive noise causes bias of some parameters. Therefore, it may be necessary to consider the model according to fig. 7-31 and perform a minimization of

$$E = \Sigma e^2(k)$$

with

$$e(k) = y(k) - \frac{B(z^{-1})}{A(z^{-1})} u(k)$$

Using quasilinearization, Bellman and Kalaba (1965) have presented a solution for the nonlinear optimization problem. Interesting applications of this technique are found in Buel, Kagiwada and Kalaba (1967), Schulz (1968), Buel and Kalaba (1969), Vaneček and Fessl (1970).

Another method for solving the nonlinear optimization problem has been given by Taylor, Iliff and Powers (1969) in connection with application to in-flight determination of stability derivatives.

Again the error criterion can be given a probabilistic interpretation if it is assumed that the only disturbances are white noise measurement errors. A technique which admits the measurement errors to be a stationary process with unknown rational spectral density has been proposed by Åström, Bohlin and Wensmark (1965). Due to specific assumptions that are made concerning
the model structure, one might expect that serious mistakes can be made if these assumptions are not true. Results which prove or disprove this are not known.


ieh, H.C. (1964). Least square estimation of linear and nonlinear weighting function matrices. Information and Control, 7, 84-


Oza, K.G. and E.I. Jury (1968) System identification and the principle of contraction mapping. SIAM J. control, 6, 244-257.


Müller, J.A. (1968). Analysis of control loops by means of adaptive models (in German) Z. Messen, Steuern, Regeln, 11, 78-80, 146-152


c.f. also same journal AC-13, 207-208.


fig. 7.1

fig. 7.2
fig. 7.8

fig. 7.9

fig. 7.10

fig. 7.11

fig. 7.12
fig. 7.13

fig. 7.14
fig. 7.25

fig. 7.26

fig. 7.27
Chapter 8 CONTINUOUS SIGNALS; explicit methods.

In this chapter correlation techniques are studied for signals which are continuous (nonsampled) time functions. In chapter 2 it was stated that as regards to amplitude, one can distinguish between "analog", quantized and binary signals. For a quantized signal the amplitude is restricted to a limited number of prefixed levels; for a binary signal this number is two.

The correlation techniques for "analog" signals are discussed in section 8.1. The next section is devoted to the description of the (nonlinear) quantizing operation. As a result of quantizing, the correlation-multiplier can be of a simpler construction than the one used for analog signals. This holds even more in the case of binary signals where a still simpler relay or logic circuit may be used for multiplication. This is discussed in section 8.3.

Figure 8.1 gives a schematic representation of these cases mentioned.

The differential approximation method, discussed in section 8.4, is closely related to the correlation techniques. In this case, attention is focussed on determination of the coefficients of differential equations.
8.1 Operations on "analog" signals.

General properties of correlation measurements. As a start it seems worthwhile to recall that correlation techniques have found application for a wide variety of purposes: e.g.,

- stochastic signal characterization
- system characterization
- signal detection

and in a wide variety of fields: e.g.,

- communication
- radioastronomy
- radar and sonar receivers
- engineering process studies
- engineering measurement techniques
- mechanical vibration studies
- seismology
- acoustics
- brain research

Although the importance of correlation measurements was recognized for many years, it took a long time before commercial instruments became available. This may be attributed to the rather difficult and critical nature of some of the elements which constitute a correlator. Particularly this is true for the timedelay unit, the multipliers usable over an extensive range of frequencies and the integrators for long integration intervals. The advent of digital and hybrid techniques has alleviated many of these problems.

Correlation techniques may provide estimates of the unknown parameter in the sense of a minimum mean squared error; cf. chapter 2. This can easily be shown for continuous signals. We consider a linear process \( P \), stationary signals (ergodic random processes), and want to estimate the impulse response of \( P \), cf. fig. 3.2.
Due to linearity, the convolution integral holds:

\[ y(t) = \int_{-\infty}^{\infty} h(\theta) u(t-\theta) \, d\theta + n(t) = h(t) * u(t) + n(t) \]  

(8.1)

The signals \( u(t) \) and \( n(t) \) are stochastic; consequently, \( y(t) \) is stochastic too. Stationarity of \( u(t) \) implies that for processes with a "memory" (impulse response \( h(t) \) of finite duration \( y(t) \) will be stationary also.

Correlation technique. Multiplying both sides with \( u(t-\tau) \) and taking the mathematical expectation we find:

\[ E\left[ u(t-\tau) y(t) \right] = \int_{-\infty}^{\infty} h(\theta) E\left[ u(t-\tau) u(t-\theta) \right] \, d\theta + E\left[ u(t-\tau) n(t) \right] = \]

(8.2)

or using the notation of Appendix B:

\[ \Psi_{uy}(\tau) = \int_{-\infty}^{\infty} h(\theta) \Psi_{uu}(\tau-\theta) \, d\theta + 0 = h(\tau) * \Psi_{uu}(\tau) + 0 \]

(8.3)

if \( n(t) \) and \( u(t) \) are independent. Equation (8.1) relates the stochastic signals \( u(t) \) and \( y(t) \); eq. (8.3) is a relation between deterministic functions, describing characteristics of those stochastic signals.

Minimization of the averaged squared error. One wants to determine a function \( h(t) \) as an estimate of the impulse response. Again the average squared error is considered:

\[ E = \frac{1}{T} \int_{0}^{T} e^{2} \, dt \]

(8.4)

with

\[ e(t) = y(t) - \int_{0}^{\infty} h(\tau) u(t-\tau) \, d\tau \]

(8.5)
Assume that $\hat{f}(\theta)$ is the function that minimizes $E$. Then any $h(\theta)$ can be represented as

$$h(\theta) = \hat{f}(\theta) + \alpha h_a(\theta)$$

(8.6)

where $h_a$ is an arbitrary function. By definition $E$ will be larger for $\alpha \neq 0$ than for $\alpha = 0$. Consequently,

$$\frac{\partial E}{\partial \alpha} \bigg|_{\alpha = 0} = 0$$

or

$$\frac{2}{T} \int \left\{ \int y(t) - \int \hat{f}(\theta) u(t-\theta) \, d\theta \right\} \left\{ \int h_a(\tau) u(t-\tau) \, d\tau \right\} \, dt = 0$$

This can be written as:

$$\int h_a(\tau) \left\{ \frac{1}{T} \int y(t) u(t-\tau) \, dt + \int \hat{f}(\theta) \, d\theta \frac{1}{T} \int u(t-\theta) u(t-\tau) \, dt \right\} \, d\tau = 0$$

(8.7)

As $h_a(\tau)$ is an arbitrary function, we can infer that

$$\frac{1}{T} \int y(t) u(t-\tau) \, dt = \int \hat{f}(\theta) \, d\theta \frac{1}{T} \int u(t-\theta) u(t-\tau) \, dt$$

(8.8)

or

$$\tilde{\psi}_{uy}(\tau, \tau) = \int \hat{f}(\theta) \tilde{\psi}_{uu}(\tau-\theta, \tau) \, d\theta$$

(8.9)

In practical situations the correlation measurement has to be over a finite time interval, which results in an approximation of the true correlation function. For $T \to \infty$ and using the assumption of ergodicity:
\[ \psi_{uy} (\tau) = \int_{0}^{\infty} h(\theta) \psi_{uu} (\tau - \theta) d\theta \] 

Wiener-Hopf equation (8.10)

Compare this result with eq. (8.3)!

By studying \( \frac{\partial^2 E}{\partial \alpha^2} \) one has to verify that \( E \) really is minimized.

Note that no restrictions have been introduced with respect to \( \hat{h}(\theta) \); it is clear that for physical realizability of the process

\[ \hat{h}(\theta) = 0 \quad \text{for} \quad \theta < 0 \]

The determination of \( h(\theta) \) from the correlation functions \( \psi_{uy} \) and \( \psi_{uu} \) (deconvolution) can be done by means of analog or digital techniques and along the lines of time domain or frequency domain operations.

In the time domain the situation is quite simple if the bandwidth of \( u(t) \) is large compared to the bandwidth of the process. For such practical cases of a "white" input signal the following approximation can be used:

\[ \psi_{uu} (\tau - \theta) \approx c \delta (\tau - \theta) \quad \text{with} \quad c = \int_{-\infty}^{+\infty} \psi_{uu} (\tau) d\tau \]

and, using the properties of Dirac "functions"

\[ \psi_{uy} (\tau) \approx c h(\tau) \quad \text{or} \quad h(\tau) \approx \frac{1}{c} \psi_{uy} (\tau) \]

In such a case, the crosscorrelation function is a good approximation of the process impulse response; c.f. fig. 8.3.

If the input signal can not be considered to be "white", then one has to face the problem of solving for \( h(\theta) \) from eq. (8.10).
One type of solution is to adjust a model in such a way that $\psi_{uy}(t)$ is its output signal if $\psi_{uu}(t)$ is its input; cf. Wallman (1950); Goodman and Reswick (1956); Chang, Goodman and Reswick (1956); Goodman (1957).

Another approach is by approximating the convolution integral $(8.10)$ as a convolution sum:

$$
\psi_{uy}(t) = \int_{t_0}^{t} \psi_{uu}(\tau) d\tau = \int_{t_0}^{t} h(\theta) \psi_{uu}(t-\theta) d\theta + \sum_{l=1}^{n} \psi_{uu}(lT) e^{lT}
$$

or, as indicated in fig. 8.4:

$$
\psi_{uy}(kT_5) \approx \frac{T_5}{2} \psi_{uu}(0) + T_5 \sum_{l=1}^{\infty} \frac{1}{T_5} \psi_{uu}(kT - lT_5)
$$

(8.11)

As an example, eq. (8.11) is written in matrix form for three points of the impulse response:

$$
\begin{bmatrix}
\psi_{uy}(0) \\
\psi_{uy}(T_5) \\
\psi_{uy}(2T_5)
\end{bmatrix} \approx
\begin{bmatrix}
\psi_{uu}(0) & \psi_{uu}(-T_5) & \psi_{uu}(-2T_5) \\
\psi_{uu}(T_5) & \psi_{uu}(0) & \psi_{uu}(-T_5) \\
\psi_{uu}(2T_5) & \psi_{uu}(T_5) & \psi_{uu}(0)
\end{bmatrix}
\begin{bmatrix}
\frac{1}{2} h(0) \\
h(T_5) \\
h(2T_5)
\end{bmatrix}
$$

Because $\psi_{uu}(T) = \psi_{uu}(T - \tau)$, the matrix is symmetrical. This formulation is closely related to those given in Chapter 6. Of course one can easily develop other and better approximations of the convolution integral, e.g., using linear interpolations between sample points of the h and $\psi$ functions.
A third type of approach towards solving for $h(\theta)$ is by means of the frequency domain. By taking the Fourier transform of eq. (8.10), one obtains:

$$\tilde{\phi}_{xy}(j\omega) = H(j\omega) \tilde{\phi}_{uv}(j\omega)$$  \hspace{1cm} (8.12)

from which one can determine

$$H(j\omega) = \frac{\tilde{\phi}_{xy}(j\omega)}{\tilde{\phi}_{uv}(j\omega)}$$

Due to the fact that only estimates of the power spectral density functions will be available, $H(j\omega)$ will be an estimate also. Numerical Fourier transformation can be performed by the celebrated Fast-Fourier Transform method; c.f. Brigham and Morrow (1967) as well as the special issue of IEEE Trans. on Audio and Electroacoustics, June 1967. For the computational aspects of power spectral density curves, including the use of lag windows / spectral windows (block, triangular, Hanning, Hamming) for obtaining smoothed density curves the reader is referred to Blackman and Tukey (1959) and Richards (1967).

The use of orthogonal filters. So far, the discussion of correlation techniques was restricted to instrumentations using a timedelay element. This, however, is not necessary. Another interesting choice of linear dynamic operators is that of orthogonal filters (Lampard, 1955; Kitamori, 1960). In this way, one circumvents the need for building an analog delay line, still an expensive and critical component of a correlator. Yet, one may estimate the impulse response. The discussion of orthogonality properties can be done both in the time- and in the frequency domain. The first approach can be compared with the discussion in section 3.2. Here we consider the frequency domain.
The (conceptual) situation is that of fig. 8.5. For the mean squared error one finds:

$$e^2(t) = \mathcal{Y}(\omega) =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left\{ H'(j\omega) - G'(j\omega) \right\} \{ H(j\omega) - G(j\omega) \} \phi_{uu}(j\omega) \, d\omega + \mathcal{Y}_{nn}(\omega)$$  \hfill (8.13)

where $H$ and $G$ represent the dynamic properties of process $P$ and "model" $M$. Apart from the noise contribution, this integral indicates the squared difference between $H$ and $G$, weighted with respect to $\phi_{uu}$. Now the model consists of some filters $F_i(j\omega)$, i.e.:

$$G(j\omega) = \sum_{i=0}^{m} \beta_i \cdot F_i(j\omega)$$  \hfill (8.14)

The coefficients $\beta_0, ..., \beta_m$ have to be "adjusted" such that $e^2(t)$ is a minimum. Substitution of eq. (8.14) in (8.13), differentiation with respect to $\beta_k$ and equating this to zero for $\beta_k = \hat{\beta}_k$ yields:

$$\frac{1}{2\pi} \sum_{i=0}^{m} \hat{\beta}_i \int_{-\infty}^{+\infty} F_k^*(j\omega) F_i(j\omega) \phi_{uu}(j\omega) \, d\omega =$$

$$= \frac{1}{2} \int_{-\infty}^{+\infty} \left\{ F_k^*(j\omega) H(j\omega) + F_k(j\omega) H^*(j\omega) \right\} \phi_{uu}(j\omega) \, d\omega$$  \hfill (8.15)

or with symbols $I$ instead of the integrals:

$$\sum_{i=0}^{m} \hat{\beta}_i \cdot I_{ki} = I_{kkH}$$  \hfill (8.16)
From this equation follows:

\[
\hat{\beta}_k = \frac{1}{I_{kk}} \left\{ I_{kk} - \sum_{i \neq k} \hat{\beta}_i I_{ki} \right\} \quad (8.17)
\]

Consequently \( \hat{\beta}_k \) is dependent on all other parameters; after adjustment of some other parameters the parameter \( \hat{\beta}_k \) has to be corrected again. This nuisance can be prevented by choosing

\[
I_{ki} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{\phi_{uu}(j\omega)} \phi_{uu}(j\omega) \ d\omega = \delta_{ki} = \begin{cases} 1 & k = i \\ 0 & k \neq i \end{cases}
\]

(8.18)

i.e., the \( F_k \) form an orthonormal system with respect to \( \phi_{uu}(j\omega) \). In that case

\[
\hat{\beta}_k = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\phi_{uu}(j\omega)} \phi_{uu}(j\omega) \ d\omega}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi_{uu}(j\omega) \ d\omega} \quad (8.19)
\]

which gives \( \hat{\beta}_k \) as an explicit relation, without interdependence of the parameter values. The \( F_k \)'s have to be orthonormal with respect to \( \phi_{uu} \).

If \( u \) represents white noise, then \( \phi_{uu}(j\omega) \) is constant and as impulse responses of the filters \( F_k \) the well known Laguerre functions can be used, c.f. Chapter 3 and fig. 4.17.

Note that numerator and denominator of eq. (8.19) do not change if under the integral signs is included \( e^{-\tau} \) with \( \tau = 0 \). Consequently, these integrals may be recognized as inverse \( F \)-transforms, i.e., correlation functions (with argument \( \tau = 0 \)).
\[ \beta_k = \frac{i}{\psi_{uu}(0)} \left( \psi_{uy}(0) + \psi_{yx}(0) \right) = \frac{\psi_{uy}(0)}{\psi_{yx}(0)} \]

\( k = 0, 1, \ldots, m \) \hspace{1cm} (8.20)

This can be implemented according to fig. 8.6.

**Statistical errors of correlation measurements.** The averaging operation over an infinite time interval is, of course, highly impractical. In actual applications one has to restrict the correlation time to a finite interval \([0, T]\). Consequently, one has to be satisfied with approximations of the correlation functions denoting them by \( \tilde{\psi}_{uu}(\tau, T) \) and \( \tilde{\psi}_{uy}(\tau, T) \). These experimental "correlation functions" show a variance (uncertainty) with respect to the theoretical correlation functions.

To discuss this aspect we consider fig. 8.7 where \( x \) and \( y \) are ergodic (stationary) stochastic signals and where the switch \( S \) is closed at \( t = 0 \). The initial condition \( r(0) = 0 \). In appendix E it is shown that for this situation the following holds:

\[ \mu_r(t) \overset{def}{=} \mathcal{E}[r(t)] = t \psi_{xy}(0) \] \hspace{1cm} (8.21)

and if \( x \) and \( y \) are jointly Gaussian:

\[ \sigma_r^2(t) \overset{def}{=} \mathcal{E}\left[ \left( r(t) - \mu_r(t) \right)^2 \right] = 2 \int_0^t (t - \nu) R(\nu) d\nu \] \hspace{1cm} (8.22)
with

\[ R(\nu) = \gamma_{xx}(\nu) \gamma_{yy}(\nu) + \gamma_{xy}(\nu) \gamma_{yx}(\nu) \]  \hspace{1cm} (8.23)

By comparing figures 8.3 and 8.7 one notices that these cases correspond if

\[ x(t) \Rightarrow u(t - \tau) \]
\[ y(t) \Rightarrow y(t) = x(t) + n(t) \]

Consequently,

\[ \gamma_{xy}(\tau) \Rightarrow \gamma_{uy}(\tau) = \gamma_{ux}(\tau) + \gamma_{yn}(\tau) \]

and

\[ \gamma_{xx}(\nu) \Rightarrow \gamma_{uu}(\nu) \]
\[ \gamma_{yy}(\nu) \Rightarrow \gamma_{yy}(\nu) = \gamma_{xx}(\nu) + \gamma_{nn}(\nu) + \gamma_{xy}(\nu) + \gamma_{yx}(\nu) \]
\[ \gamma_{xy}(\nu) \Rightarrow \gamma_{uy}(\nu + \tau) = \gamma_{ux}(\nu + \tau) + \gamma_{yn}(\nu + \tau) \]
\[ \gamma_{yx}(\nu) \Rightarrow \gamma_{yu}(\nu - \tau) = \gamma_{ux}(\nu - \tau) + \gamma_{yn}(\nu - \tau) \]

where several terms drop out as the input and noise signals are assumed to be independent. This results in

\[ \mu_{r}(t) = t \gamma_{ux}(\tau) \]  \hspace{1cm} (8.24)

The (independent) additive noise does not contribute to the expected value.

For the variance, one can split \( R(\nu) \) into two parts:

\[ R(\nu) = R_{1}(\nu) + R_{2}(\nu) \]

in such a way that

\[ R_{1}(\nu) = \gamma_{uu}(\nu) \gamma_{xx}(\nu) + \gamma_{ux}(\nu + \tau) \gamma_{xu}(\nu - \tau) \]  \hspace{1cm} (8.25)
\[ R_2 (\nu) = \psi_{uu} (\nu) \psi_{nn} (\nu) \]  

(8.26)

The contribution due to \( R_1 \) is present even if there is no additive noise; this might be called the \underline{inherent statistical uncertainty}. The contribution due to \( R_2 \) is the \underline{uncertainty due to the additive noise}.

This discussion of statistical errors can be applied to the experimentally determined points of the correlation function \( \psi_{ux} (\tau) \), estimated over the finite time interval \([0, T]\):

\[ \tilde{\psi}_{ux} (\tau, T) = \frac{1}{T} \int_0^T x(t-\tau) x(t) \, dt \]  

(8.27)

The \underline{inherent statistical uncertainty} can be expressed by the signal-to-noise ratio \( S/N_1 \) of the correlator output:

\[ S = E[\tilde{\psi}_{ux} (\tau, T)] = \frac{1}{T} \mu_x (\tau) = \psi_{ux} (\tau) \]  

(8.28)

\[ N_1^2 = E[\tilde{\psi}_{ux}^2 (\tau, T)] - \psi_{ux}^2 (\tau) = \frac{1}{T^2} O^2 (H) = \tau = \frac{2}{T} \int_0^T (T-\nu) R_1 (\nu) \, d\nu \]  

(8.29)

i.e., the expectation of the integrator output is the correct value of the desired correlation function independent of the length of the observation interval \( T \). Of course the accuracy (variance or standard deviation) of the determined value does depend on the length of the observation interval.

Because

\[ N_1^2 = \frac{2}{T^2} \left\{ \int_0^T R_1 (\nu) \, d\nu - \int_0^T \psi_{ux} (\nu) \psi_{ux} (\nu) \, d\nu \right\} \]  

(8.30)

this can be simplified for large \( T \) if the second term can be neglected, i.e., if \( R_1 (\nu) \) approaches rapidly to zero. In that case

\[ N_1^2 = \frac{2}{T} \int_0^T R_1 (\nu) \, d\nu = \frac{2}{T} \int_0^T \left\{ \psi_{uu} (\nu) \psi_{xx} (\nu) + \psi_{nx} (\nu) \psi_{ux} (\nu-\tau) \right\} \, d\nu \]  

(8.31)
If for large $\tau$, $\psi_{ux}(\tau) \to 0$ then under that condition

$$N^2 = \frac{2}{T} \int_0^T \psi_{uu}(\nu) \psi_{xx}(\nu) \, d\nu$$

(8.32)

As an example, the measurement of the auto-correlation function according to fig. 8.8 is considered, where $u(t) = x(t)$.

The white noise $n(t)$ has a power density $k^2$

the RC-filter has a transfer function $H(s) = \frac{\alpha}{s + \alpha}$

the power density of the signal $x(t)$ is $\varphi_{xx}(s) = k^2 \frac{\alpha^2}{\alpha^2 - s^2}$

the correlation function of $x(t)$ is $\psi_{xx}(\nu) = \psi_{xx}(0) \, e^{-\alpha |\nu|}$

with $\psi_{xx}(0) = \frac{\alpha k^2}{2}$

Consequently:

$$S' = \psi_{xx}(\nu) = \psi_{xx}(0) \, e^{-\alpha |\nu|}$$

$$N^2 = \frac{2}{T^2} \int_0^T (T - \nu) \, R_i(\nu) \, d\nu$$

with

$$R_i(\nu) = \psi_{xx}^2(\nu) + \psi_{xx}(\nu + \tau) \psi_{xx}(\nu - \tau) = \psi_{xx}^2(0) \left[ e^{-2\alpha |\nu|} + e^{-\alpha |\nu|} + e^{-\alpha |\nu|} \right]$$

c.f. fig. 8.9 which leads to

$$N^2 = \psi_{xx}^2(0) \frac{2 \alpha T - 1 + 2 \epsilon^{-2\alpha T} + [(2\alpha T + 1)(2\alpha T - 1) - 2(\alpha T)^2] \epsilon^{-2\alpha T}}{2(\alpha T)^2}$$

(8.33)

We introduce the dimensionless quantities $\alpha \tau = \beta$ and $\alpha T = \gamma$; then for $\beta \geq 0$

$$S' = \psi_{xx}(0) \, e^{-\beta}$$

$$N^2 = \psi_{xx}^2(0) \frac{2 \gamma - 1 + 2 \epsilon^{-2\gamma} + [(2\beta + 1)(2\gamma - 1) - 2\beta^2] \epsilon^{-2\beta}}{2(\gamma)^2}$$

(8.34)
Consider some special cases, where $\gamma \to \infty$ indicating long correlation intervals:

\[
\beta = 0 \quad \gamma \to \infty \quad N^2 = \frac{\psi^{2}_x(0)}{\gamma^2} \quad (8.35)
\]

\[
\beta \to \infty \quad \gamma \to \infty \quad N^2 \to \psi^{2}_x(0) \frac{2}{\gamma} \quad (8.36)
\]

Consequently, the variance of the measurements of points of the correlation function is inversely proportional to $\gamma$. Because $\gamma = \alpha T$, this means that the variance becomes smaller if $\alpha$ (the bandwidth of the filter) and/or $T$ (the observation interval) increase.

For the example given above, the signal-to-noise ratio is, in the case of a "long" correlation interval:

\[
\frac{S}{N} \approx \epsilon^{-\beta} \sqrt{\frac{\gamma}{2}} \quad (8.38)
\]

In terms of accuracy, one might specify e.g., $N_1 = 0.05 S$, i.e., a standard deviation equal to 5% of the measurement value. For $\beta = 0$ this results in $\gamma = \alpha T = 800$. Note that this means, that the correlation interval is 800 times the time constant of the RC-filter under consideration. Yet this rather long interval holds for a mediocre accuracy and for the most favourable point of the correlation function ($\beta = 0$). For large values of $\beta = \alpha \tau$, i.e., for points of the correlation function representing large arguments $\tau$, the S/N-ratio will be much worse. Correlation intervals of such lengths are dictated by the inherent statistical uncertainty only. Note that this S/N-ratio is independent of the level of the input signals to the correlator.
The uncertainty due to possible additive noise will put an even higher demand on correlation time. This uncertainty due to additive noise is of a simpler nature and can be expressed by an additional "noise" term in the correlator output, viz:

\[ N_z^2 = \frac{2}{T^2} \int_0^T (T - \nu) R_2(\nu) d\nu \]  
\[ (8.39) \]

with eq. (8.26)

\[ R_2(\nu) = \psi_{uu}(\nu) \psi_{nn}(\nu) \]  
\[ (8.40) \]

Of course, for this type of error, the signal-to-noise ratio \( S/N_2 \) improves linearly with an increasing level of the input signal \( u \).

As an example, let us consider the same case as before, given by fig. 8.8, but now assuming that to one channel only additive noise is added. Let this noise \( n(t) \) have the same bandwidth as the signal \( x(t) \), so:

\[ \psi_{nn}(\tau) = \psi_{nn}(0) e^{-a/|\tau|} \]  
\[ (8.41) \]

This leads to

\[ N_z^2 = \psi_{xx}(0) \psi_{nn}(0) \frac{2\gamma - 1 + e^{-2\gamma}}{2\gamma^2} \]  
\[ (8.42) \]

and for large \( \gamma \):

\[ N_z^2 \rightarrow \psi_{xx}(0) \psi_{nn}(0) \frac{\gamma}{\gamma} \]  
\[ (8.43) \]

Now the total signal to noise ratio is found to be for long correlation intervals

\[ \frac{S}{N} = \frac{S}{\sqrt{N_z^2 + N_2^2}} \approx \varepsilon^{-\beta} \sqrt{\frac{\gamma}{2 + \frac{\psi_{nn}(0)}{\psi_{xx}(0)}}} \]  
\[ (8.44) \]
Up to now, only the determination of one parameter (viz. one point of the impulse response) has been discussed. A new element enters the discussion if a number of parameters has to be determined. After one quantity has been estimated this estimate can be considered as being a priori knowledge for the next estimation cycle, etc. Intuitively, it is clear that the use of a priori knowledge may result in a more efficient estimation procedure. A crucial question, however, is in what way such knowledge can be incorporated into the instrumentation. This can be done by using a model of the process that is well adapted to the parameter-description of the process. For example, if the attention is focussed on the impulse response, then a timedelay circuit with taps and potentiometers is adequate. Figure 8.10 gives an illustration of this situation. In order to obtain some insight into the effect of the model, the relations (8.21) through (8.25) will be applied to an arbitrary simple example, viz. a process with impulse response:

\[ h(\tau) = \begin{cases} \frac{6}{\pi^2} \sin^2 \frac{\pi \tau}{T} & \text{for } \tau \geq 0 \\ 0 & \text{for } \tau < 0 \end{cases} \]

We will consider the determination of one point of this function, viz. for \( \tau = 0.225 \) (fig. 8.11a) using a "white" noise input signal.

If the switch \( S \) in fig. 8.10 is open this method is identical to the correlation scheme of fig. 8.3. In that case we find for the standard deviation with respect to the expected value

\[ \frac{1}{T} \mu_r = \psi_{uy}(\tau) \quad \text{with} \quad \tau = 0.225 \]

the relation given in fig. 8.12, curve a. From this diagram, the correlation interval needed for a certain accuracy (standard deviation) of the estimated parameter can be found.
Now we will assume that in fig. 8.10 the parameters $\beta_1$ through $\beta_6$ have already been estimated. The determined values are used for adjusting those potentiometers, the other potentiometers are set equal to zero and switch $S$ is closed. The adjusted part of the model, with the hold-circuit included, has the impulse response $m(\tau)$ shown in fig. 8.11. Consequently, the error signal $e$ is determined by the difference between $h(\tau)$ and $m(\tau)$ as shown in fig. 8.11c; yet, in this impulse response, $h(\tau) - m(\tau)$, the parameter to be estimated still has its original value. Now turning our attention to equation (8.25) we notice that in this case the index $x$ has to be replaced by the index $e$. This implies that the corresponding terms are smaller than the original ones, resulting in a smaller variance (or a reduction of the correlation interval that is needed for obtaining the same variance). For the example given we find the reduction of the correlation interval given by fig. 8.12, curve b; now the same standard deviation is obtained in about one tenth of the original time interval.

Fig. 8.13a and b give a number of actual recordings of the integrator output $r(t)$ for the cases indicated above; $S$ open viz. closed. The fluctuations represent the inherent statistical uncertainty of that correlator without and with a model. The influence of additive noise has been neglected; it can be taken into account using equation (8.25).

For further discussions the reader is referred to Davenport and others (1952), Laning and Battin (1956), Green (1957), Solodovnikov (1960), Lee (1960) Bendat and Piersol (1966).

There is an extensive literature on correlation techniques; only a very small part of it can be mentioned as references (e.g: Rake (1968), Welfounder (1969) Buchta (1969), Hayashi (1969), Raibman et al (1970)) or as additional literature. Quite a number of correlator types is commercially available nowadays, some of them equipped with additional features such as probability-density-function meters.
Another quite popular type of additional instrumentation is that of Fourier transformation of the correlation functions to power spectral density curves. Often in such an instrument an appreciable number of points of the correlation function (eq. 50 or 100) are determined simultaneously by parallel implementation. Another development is that of a continuously changing time delay \( \tau \) and associated correcting filter; cf. Ball (1963).

8.2 Operations on quantized signals.

Amplitude quantizing. As indicated earlier, the quantizing operation \( Q[ \cdot ] \) on "analog" signals \( x \) and \( y \) is a nonlinear operation, i.e., if

\[
x'_2 = Q[x]
\]

and

\[
y'_2 = Q[y]
\]

then

\[
\alpha x'_2 + \beta y'_2
\]

and

\[
Q[\alpha x + \beta y]
\]

are not necessarily equal for all values of \( x \) and \( y \). Consequently we will meet with the problems inherent to nonlinear systems; that is they fail to obey the superposition principle.

For a discussion of the quantizing operation and its use in correlation determination the reader is referred to the papers by Widrow (1956) and Watts (1962). The following results are taken from Watts' paper.

An "analog" input signal \( x \) is transformed by the general equi-interval quantizer to a quantized signal \( x'_2 \). The input-output characteristics may be that given in fig. 8.14 a. This general quantizer \( Q \) can be derived by means of a unit quantizer \( Q_1 \), the characteristics of which are shown in fig. 8.14b, as indicated in fig. 8.15.
The corresponding relationships are the following:

**Gain**

\[ X_i = \frac{1}{q} x \]  \hspace{1cm} (8.47)

**Shift**

\[ X_2 = x_i - a \]  \hspace{1cm} (8.48)

**Unit quantizer**

\[ X_q = i \quad \text{for} \quad (i - \frac{1}{2}) \leq x_2 \leq (i + \frac{1}{2}) \]  \hspace{1cm} (8.49)

**Shift**

\[ X_q' = X_q + c \]  \hspace{1cm} (8.50)

**Gain**

\[ X_2' = r X_q' \]  \hspace{1cm} (8.51)

It is shown by Watts that the relationship between the probability density functions at input \( p(x) \) and output \( p_q(x_q) \) is given by

\[
p_q(x_q) = \frac{1}{r} \sum_{i=-\infty}^{\infty} \left[ \int_{-\infty}^{+\infty} p(x) f_i \left( \frac{x q - c - \left( \frac{X q}{q} - a \right)}{r} \right) dx \right] \delta \left( \frac{x_q}{r} - c - i \right) \hspace{1cm} (8.52)
\]

with

\[ \delta(x) \text{ Dirac function} \]

\[ f_i(x) = \begin{cases} 1 & \text{for } -\frac{1}{2} < x \leq \frac{1}{2} \\ 0 & \text{elsewhere} \end{cases} \]

For \( a = c = 0, q = r = 1 \) this reduces to

\[
p_q(x_q') = \sum_{i'=-\infty}^{\infty} \left[ \int_{-\infty}^{+\infty} p(x) f_{i'}(x_q' - x) dx \right] \delta(x_q' - i) \hspace{1cm} (8.53)
\]
i.e., a convolution of the probability density function with the window function $f_1(x)$, the result "sampled" in the $x'$ domain.

For further insight and ease of manipulations it pays to take the Fourier transform, by which the convolution changes into a multiplication; i.e., we determine the characteristic function

$$
\mathcal{P}_{x_2}(s) = \mathcal{E}\left[ \exp j s x_2 \right] = \int \mathcal{P}_2(x) \exp j s x_2 \, dx
$$

and

$$
\mathcal{P}_{x_1 x_2}(s_1, s_2) = \mathcal{E}\left[ \exp j (s_1 x_1 + s_2 x_2) \right] = \int \mathcal{P}_2(x_1, x_2) \exp j (s_1 x_1 + s_2 x_2) \, dx_1 \, dx_2
$$

(8.54)

(8.55)

By differentiation of the characteristic function with respect to $s_1$ and $s_2$ one obtains the moment

$$
\mathcal{E}\left[ x_1 y_2 \right] = \langle \rangle \frac{\partial^2 \mathcal{P}_{x_1 y_2}(s_1, s_2)}{\partial s_1 \partial s_2} \bigg|_{s_1 = s_2 = 0}
$$

(8.56)

which is the correlation coefficient between these signals. In this way the correlation coefficient for the general form of a multiplier correlator can be found. One way to interpret the results is by considering the quantized signals as the original signal plus an additive quantization noise:

$$
x_q = x + n_x \quad y_q = y + n_y
$$

(8.57)

$$
\mathcal{E}[x_q y_q] = \mathcal{E}[xy] + \mathcal{E}[n_x y] + \mathcal{E}[x n_y] + \mathcal{E}[n_x n_y]
$$

(8.58)

From this general form, the different types of correlators can be derived, including the ones operating with extreme quantization, resulting in binary signals.
Types of correlators. From the previous discussion it may be clear, that
the quantizing operation can be characterized by the following parameters:

- a: shift along x axis
- c: shift along x₂ axis
- q: step size along x axis
- r: step size along x₂ axis

Special cases are:

- \( a = c = 0; \ q = r = 1 \) \hspace{1cm} \text{unit quantizer}
- \( a = c = 0; \ q = r \rightarrow 0 \) \hspace{1cm} \text{analog channel, no quantizer}
- \( a = c = 0; \ q = r \) \hspace{1cm} \text{general (symmetric) quantizer}
- \( a = \frac{1}{2}; \ c = \frac{1}{2}; \ q \rightarrow \infty; \ r = 2 \) \hspace{1cm} \text{binary quantizer}

Depending on the types of quantization used in either or both channels of a
 correlator one may distinguish, cf. fig. 8.16:

1) "analog" correlator
2) "digital" correlator
3) polarity-coincidence correlator
4) relay or ring-modulator correlator
5) "Stieltjes" correlator
A comparison of advantages and disadvantages of these correlators is given by Watts (1961). The reader is referred to that paper. Some remarks may be in order here.

1) "analog" correlator. For this type of correlator, \( x_\Omega = x \) and \( y_\Omega = y \), i.e., there is no quantization noise and:

\[
\mathcal{E} \left[ x_\Omega y_\Omega \right] = \mathcal{E} \left[ xy \right] = \mathcal{E} \left[ \tilde{\psi}_{xy}(0,T) \right] = \psi_{xy}(0) \tag{8.59}
\]

where as before

\[
\psi_{xy}(0,T) = \frac{1}{T} \int_0^T x(t) y(t) \, dt \tag{8.60}
\]

This case was discussed extensively in section 8.1.

2) "digital" correlator. The signals in both channels are quantized. It turns out that even for rather coarse quantization there is an acceptable correspondence between \( \mathcal{E} \left[ x_\Omega y_\Omega \right] \) and \( \mathcal{E} \left[ xy \right] \). This implies that for digital correlators even 3 or 4 bit word lengths may suffice. Chang and Moore (1970) discuss the use of auxiliary random noise signals to unbias the output of this type of correlator, i.e., to obtain relatively high accuracy with a small number of quantization levels; cf. also section 8.3.

3) polarity-coincidence correlator; cf. section 8.3.

4) relay or ring-modulator correlator. In this case the multiplication is very simple to implement, e.g., \( x \cdot \text{sgn} \, y \) with \( \text{sgn} \, y = 1 \) for \( y \geq 0 \) and \( \text{sgn} \, y = -1 \) for \( y < 0 \), which can be done by means of a mechanical or electronic relay. Of course the information about the power level of \( y \) is lost. For interpretation of the measurement results, the probability density function of the input signals has to be known.
5) "Stieltjes" correlator. This method was introduced by Watts. Its name refers to the fact that the correlation integral for the analog signal \( x \) and the quantized signal \( y \) can be recognized as a Stieltjes integral. It has interesting properties with respect to both theory and implementation, which can be quite simple. In spite of a rather coarse quantization, say 4 or 5 levels, it turns out that the error in the (expectation of the) measured correlation function is remarkably small. Figure 8.17 shows a way to implement this principle. For digital computer calculation of correlation functions, this principle is worth considering.
8.3 Operations on binary signals

It was recognized early on (Faran and Hills, 1952) that for signals with a joint Gaussian probability density function, the correlation measurement could well be performed using bivalent or binary versions of these signals. This makes it possible to use purely digital means for building a correlator. As only the polarity of the signals comes into play, this principle is also called **polarity coincidence correlation**.

Assume the signals \( x(t-\tau) \) and \( y(t) \) have a Gaussian probability density function with \( \xi[x] = \xi[y] = 0 \), i.e.:

\[
\varphi(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp \left[ \frac{-1}{2(1-\rho^2)} \left\{ \left( \frac{x}{\sigma_x} \right)^2 - 2\rho \left( \frac{x}{\sigma_x} \right) \left( \frac{y}{\sigma_y} \right) + \left( \frac{y}{\sigma_y} \right)^2 \right\} \right]
\]

with \( \rho = \rho(\tau) \). The correlation function is defined as

\[
\Phi_{xy}(\tau) = \mathcal{E}[xy] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot \varphi(x,y) \, dx \, dy
\]

(8.62)

For the density function given by eq. (8.61) this results in

\[
\Phi_{xy}(\tau) = \sigma_x \sigma_y \rho(\tau)
\]

(8.63)

When using binary signals \( \text{sgn } x \) and \( \text{sgn } y \) instead of the original signal, one finds:

\[
\Phi_{xy}(\tau) = \mathcal{E}[\text{sgn } x \cdot \text{sgn } y] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \text{sgn } x \cdot \text{sgn } y \cdot \varphi(x,y) \, dx \, dy
\]


Substitution of eq. (8.61) for the Gaussian density function leads to:

\[ \psi'_{xy}(\tau) = \frac{2}{\pi} \arcsin \rho(\tau) \]  

(8.65)

Comparing \( \psi \) and \( \psi' \) according to eq.'s (8.63) and (8.65) one notices that:

- as expected from the clipping operation, the "powers" of the signals have disappeared;
- the \( \psi' \) is a monotonic function of \( \rho(\tau) \) because \(-1 \leq \rho \leq 1\)

It can be shown that the variance of this correlation function is only slightly worse than that of normalized correlation function calculated by means of an "analog" correlator. (Veltman and Kwakernaak, 1961)

The use of auxilliary signals. The reasoning given above holds for signals with a Gaussian distribution. In practical situations, the amplitude distribution may be unknown. In those cases the relationship between \( \psi' \) and \( \psi \) is not known. There has been a number of attempts to alleviate this drawback of polarity-coincidence correlation. All these attempts are based on the addition of auxilliary signals to the correlator input signals; cf. fig. 8.18. An analysis of this type of implementation indicates that the measured correlation functions correspond quite well with the proper correlation functions, in spite of the polarity multiplication. The auxilliary signals have been chosen as:

a. random numbers that are realizations of two statistically independent random processes, uniformly distributed. (Veltman and Kwakernaak, 1961). In addition, other distributions can be used (Chang and Moore, 1970).
b. random sawtooth signals that are realizations of two independent processes, each of which has a uniform probability density function. (Jespers, Chu and Fettweis, 1962)

c. deterministic functions which, however, in a sense are mutually independent over a finite time interval, even under time-shift operations; "shift-invariant independent functions" (Peek, 1967).

The difference between the latter method and the methods mentioned under a and b is the use of systematic instead of random auxiliary signals.

A kind of independence of functions is defined by:

The set of functions \( \{ f_i(t) \} \), \( i = 1, 2, \ldots, n \) is "shift-invariant independent" in the interval \((0, T)\) if and only if

\[
\frac{1}{T} \int_0^T f_i(t) f_j(t + \tau_j) \, dt = \frac{1}{T} \sum_{\tau_i=0}^{T} \int_0^T f_i(t) f_j(t + \tau_j) \, dt \quad (8.66)
\]

for every value of \( \tau_j, \ldots, \tau_n \) and all positive integers \( k_1, \ldots, k_n \)

A particularly interesting set of functions that exhibit this type of independence are the bivalent Rademacher functions defined by

\[
r_i(t) = \text{sgn} \left\{ \sin \frac{2\pi t}{T_0} \right\} \quad i = 1, 2, 3, \ldots
\]

These are orthonormal. The set is not complete, which is of no importance in this type of use.

Figure 8.19 shows the first four elements of the set of Rademacher functions. They can be generated by means of a binary counter fed with periodic pulses. Linear combinations of these functions are also "shift-invariant independent" if they do not contain a common Rademacher function. In this way signals with more than two levels may be constructed. Figure 8.20 shows two such three-level signals, having the desired independence property. Signals of such kind are added to the correlator input signals as indicated in fig. 8.18. The resulting sum -signals
are fed to hard limiters. It can be shown that the correlation function of these binary signals is the same as the correlation function of quantized input signals as discussed in section 8.2. Consequently, the same analysis for distortion errors of the correlation function holds. Due to the periodic nature of the auxiliary signals their contribution to the correlator output noise, if measured over an integral multiple of periods, is favourable compared to random auxiliary signals. In Peek (1967) the output signal-to-noise ratio for several types of correlators is derived.

8.4 Differential approximation method

Under the name differential approximation, a method has been indicated that actually represents an extension of the correlation techniques discussed. Assume the situation sketched in fig. 8.21. It is assumed that all state variables \( x(t) \) and all input variables \( u(t) \) are measurable and that there is no noise present. The process dynamics are described by

\[
\dot{x} = f(x, u, \beta)
\]

For a value of the unknown parameter vector \( \beta \) we may write

\[
\dot{x} - f(x, u, \beta) = e
\]

Again the criterion is to minimize \( E(\beta) \) where

\[
E(\beta) = \text{some functional quadratic in } e, \text{ e.g.,}
\]

\[
E(\beta) = \int_0^T (e^T \cdot e) \, dt
\]

or generally

\[
E(\beta) = \int_0^T \| \dot{x} - f(x, u, \beta) \|^2 \, dt
\]
The necessary condition for the minimum:

\[ \frac{\partial}{\partial \beta} E(\beta) = 0 \]  \hspace{1cm} (8.71)

leads to:

\[ \int_0^T \left[ \frac{\partial f_i}{\partial \beta_j} \right] \dot{x} \, dt = \int_0^T \left[ \frac{\partial f_i}{\partial \beta_j} \right] f(x, u, \beta) \, dt \]  \hspace{1cm} (8.72)

where

\[ \left[ \frac{\partial f_i}{\partial \beta_j} \right] = \begin{bmatrix} \frac{\partial f_1}{\partial \beta_1} & \cdots & \frac{\partial f_1}{\partial \beta_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \beta_1} & \cdots & \frac{\partial f_m}{\partial \beta_m} \end{bmatrix} \]  \hspace{1cm} (8.73)

From the set of equations (8.72) the m unknown parameters \( \beta \) have to be solved.

If \( f \) is linear in \( \beta \) then again this procedure is quite simple.

Neither is it unavoidable to use the state space description, nor does a non-linear process behaviour necessarily imply additional complications. (As pointed out in chapter 4, linearity-in-the-parameters is of paramount importance).

The following simple example may be illustrative. The process is described by

\[ \ddot{x} + a_1 \dot{x} + a_2 x^3 = b u \]  \hspace{1cm} (8.74)

and

\[ \dddot{x} + a_1 \dot{x} + a_2 x^3 - \beta u = e \]  \hspace{1cm} (8.75)
Consequently,
\[
\frac{\partial e}{\partial \alpha_i} = \dot{x} \quad \frac{\partial e}{\partial \alpha_c} = x^3 \quad \frac{\partial e}{\partial \beta} = -u
\]  
(8.76)

Using \( \min \int_0^T e^t \, dt \) this leads to the set of equations, linear in the unknown parameters:
\[
\int_0^T e \, \dot{x} \, dt = 0 \quad \int_0^T e \, x^3 \, dt = 0 \quad \int_0^T e \, u \, \omega_0 \, dt = 0
\]  
(8.77)

or:
\[
\begin{align*}
\alpha_1 \int_0^T x^2 \, dt &+ \alpha_0 \int_0^T x^3 \, \dot{x} \, dt - \beta \int_0^T u \, \dot{x} \, dt = -\int_0^T x \, \dot{x} \, dt \\
\alpha_1 \int_0^T x \, x^3 \, dt &+ \alpha_0 \int_0^T x^6 \, dt - \beta \int_0^T u \, x^3 \, dt = -\int_0^T x \, x^3 \, \dot{x} \, dt \\
\alpha_1 \int_0^T x \, u \, dt &+ \alpha_0 \int_0^T x^3 \, u \, dt - \beta \int_0^T u^2 \, dt = -\int_0^T x \, u \, \dot{x} \, dt
\end{align*}
\]  
(8.78)

So far it has been assumed that \( x \) can be measured without additive noise. If only a contaminated output is available, e.g. \( y(t) = x(t) + n(t) \) then one is tempted to apply the equations (8.76) using \( y \) and \( \dot{y} \) instead of \( x \) and \( \dot{x} \). The reader will note the intimate relationship with the generalized-model-notion. From this relationship it will be clear that in this case as well, additive noise will lead to biased estimates of the parameters. Such a bias can be prevented by the so-called instrumental variable approach; cf. sect. 9.4.

This leads to a set of equations:
\[
\int_0^T e \, \dot{i}_1 \, dt = 0 \quad \int_0^T e \, \dot{i}_2 \, dt = 0 \quad \int_0^T e \, \dot{i}_3 \, dt = 0
\]  
(8.79)
where \( i_1, i_2, i_3 \) are the instrumental variables, i.e., time functions related to \( \omega(t) \) but not contaminated with noise.

For more details the reader may refer to the additional literature on this topic given at the end of this chapter.

8.5 Higher-order correlation functions

In the previous sections attention was restricted to the correlation functions for two time functions; these might be called second-order correlation functions. In some cases, however, it is of interest to consider correlation functions of the order \( n \), defined for ergodic signals by:

\[
\Psi_{x_1 \ldots x_n}(t, \ldots, t_n) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x_1(t - t_1) \cdots x_n(t - t_n) \, dt
\]

These play a role in the characterization of non-linear signal and process relations, e.g., in the Wiener characterization of non-linear processes. As these methods still have found very little application in practice the reader is referred to the additional literature on this topic.
References


Additional literature

8.1. Operations on "analog" signals.


Levin, M.J. (1964). Estimation of a system pulse transfer function in the presence of noise. IEEE Trans. autom. control, AC-9,


8.2. Operations on quantized signals.


Additional literature

1.3. Operations on binary signals.

Anderson, G.W. and J.E. Murrin. (1956). Auto-
correlation analysis of binary noise signals. RE WESCON conv. rec.

Anderson, G.W. and others (1958). A self-
djusting system for optimum dynamic perfor-
ance. IRE natn. conv. rec. part. 4, 182-190.

of the Self-Adaptive Flight Control Systems
ympp., WADC Techn. Report 59-49. ASTIA
ocument AD 209389: Wright-Patterson Air
orce Base, Ohio, 349-406.

Anderson, G.W. (1959). Use of crosscorrela-
tion in an adaptive control system. Proc.


ystems. IRE natn. conv. rec., part 4,
2-44, 1959 IRE Trans. autom. control.,
C-4, 30-42.

dynamics in computer controlled adaptive
ystems. Proc. first IFAC congress, Moscow,
96-603.

of drift corrections schemes for periodic

determining the parameters of physical
ystems. (correspondence). Proc. IEEE, 53,
05-206.

Iamessis, J.E. (1965). A method for deter-
mining the parameters of certain time-varying
ystems. (correspondence). Proc. IEEE, 53,
96-397.

(1965). On the moments of shifted Gaussian
probability density functions. CCMC report
5-2, T.H. Delft, (Netherlands).


"analog" signals

quantized signals

binary signals
fig. 8.10

fig. 8.11

fig. 8.12

fig. 8.13
As in chapter 7, the model-adjustment techniques may be interpreted either as a control scheme for the parameters of a physical (analog) model or as a computation scheme in which the "mathematical model" can be clearly indicated. In chapter 2 a "generalized model" and "error" have been introduced. In this case there is the advantage of a linear relationship between the error and the parameters to be adjusted. Some considerations of a general nature are given in section 9.1. Some of the properties of these model-adjustment schemes are discussed. A distinction is made with respect to the use of an instantaneous error criterion, e.g.,

$$E_1 = e^t (\beta, t)$$

(cf. section 9.2)

and a time-average criterion e.g.,

$$E_2 = \frac{1}{T} \int_{T-T}^{T} e^2 (\beta, t) \, dt$$

(cf. section 9.3)

In the last equation a certain averaging time T is introduced. In practice this time must be chosen long enough to satisfy requirements as to accuracy of the estimate. The instantaneous criterion can be applied only when the corresponding adjustment loop is sufficiently slow. In this case, the adjustment loop, which always will be of a low pass character, will provide the necessary averaging effect to obtain a certain accuracy.

Section 9.4 is devoted to the parameter-sensitivity functions and their use for continuous and for intermittent adjustment schemes.

The approximation of the sensitivity functions through the use of two models with parameters $\beta$ and $(\beta + \delta \beta)$ respectively is the subject of section 9.5.

The final section 9.6 is devoted to the use of the relation

$$\frac{dE}{dt} = \frac{\partial E}{\partial \beta} \frac{d\beta}{dt} + \frac{\partial E}{\partial t}$$

where $\beta$ is varied, e.g., sinusoidally, and $dE/dt$ can be measured. If $\partial E/\partial t$ can be neglected then $\partial E/\partial \beta$ can be found. This provides us the direction to which the parameters have to be adjusted.
9.1 Models, linear in the parameters.

Generalized error. The generalized error was defined in chapter 2 as

\[ e = \sum_{i=0}^{n} \alpha_i y_i + \sum_{j=0}^{m} \beta_j u_j \]  

(9.1)

where \( u_j \) and \( y_i \) are the outputs of the dynamic operators.

The purpose of the arrangement of fig. 9.1 is to adjust the potentiometers \( \alpha_0, \ldots, \alpha_n \) and \( \beta_0, \ldots, \beta_m \) in such a way that some functional of the generalized error \( e \) is minimized. One of the potentiometer settings can be chosen freely as reference, e.g. \( \beta_0 = -1 \). With this choice:

\[ e = \sum_{i=0}^{n} \alpha_i y_i + \sum_{j=1}^{m} \beta_j u_j - u \]  

(9.2)

As has already been pointed out, the essential feature of such a generalized model is the linear relation between \( e \) and the parameters \( \alpha_i \) and \( \beta_j \). This implies that the partial derivatives (parameter-sensitivity functions; cf. section 9.4) are given in a very simple way:

\[ \frac{\partial e}{\partial \alpha_i} = y_i \quad \frac{\partial e}{\partial \beta_j} = u_j \]  

(9.3)

Consequently, no extra models are necessary to instrument the sensitivity equations and all parameter-sensitivity functions can be obtained from the one (generalized) model.

For the moment, let us choose as the error criterion a minimization of the squared error over an interval of length \( T \). In order to eliminate the stochastic aspect we will take the expectation \( \bar{e} \), thus:

\[ E = \bar{e} \left[ \int_{t-T}^{t} e^2 \, dt \right] \]  

(9.4)

For this minimization we differentiate with respect to the parameters that have to be determined

\[ \frac{1}{2} \zeta_i = \frac{1}{2} \frac{\partial E}{\partial \alpha_i} = \bar{e} \left[ \int_{t-T}^{t} e y_i \, dt \right] \overset{\text{def}}{=} \langle e, y_i \rangle \]  

(9.5a)

\[ i = 0, 1, \ldots, n \]
\[
\frac{1}{2} \eta_j = \frac{1}{2} \frac{\partial E}{\partial \theta_j} = \mathcal{E} \left[ \sum_{t=1}^{T} e u_j dt \right] \overset{\text{def}}{=} \langle e, u_j \rangle
\]

This leads to \( m + n + 1 \) equations. The symbol \( \langle \cdot, \cdot \rangle \) defined by equations (9.5a) and (9.5b) has been introduced for the sake of convenience. In the next section a wider interpretation will be given to it. For optimality \( 0_1, \ldots, 0_n \) and \( \eta_1, \ldots, \eta_m \) all have to be equal to zero. At the moment we will assume \( \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \) to be constant. If one substitutes the expression (9.2) for the generalized error \( e \) in the equations (9.5a and b) then one will find in matrix notation:

\[
\begin{bmatrix}
\zeta_0 \\
\vdots \\
\zeta_n \\
\eta_1 \\
\vdots \\
\eta_m \\
\end{bmatrix} =
\begin{bmatrix}
\langle y_0, y_0 \rangle & \cdots & \langle y_0, y_n \rangle & \langle u_1, y_0 \rangle & \cdots & \langle u_m, y_0 \rangle \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\langle y_n, y_0 \rangle & \cdots & \langle y_n, y_n \rangle & \langle u_1, y_n \rangle & \cdots & \langle u_m, y_n \rangle \\
\langle y_0, u_1 \rangle & \cdots & \langle y_0, u_n \rangle & \langle u_1, u_1 \rangle & \cdots & \langle u_m, u_1 \rangle \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\langle y_n, u_1 \rangle & \cdots & \langle y_n, u_n \rangle & \langle u_1, u_n \rangle & \cdots & \langle u_m, u_n \rangle \\
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\
\vdots \\
\alpha_n \\
\beta_1 \\
\vdots \\
\beta_m \\
\end{bmatrix} =
\begin{bmatrix}
\langle u_0, y_0 \rangle \\
\vdots \\
\langle u_n, y_n \rangle \\
\langle u_1, u_1 \rangle \\
\vdots \\
\langle u_m, u_m \rangle \\
\end{bmatrix}
\]

or renaming the matrices in this order:

\[
\frac{1}{2} \nabla E = M \Theta - \Psi
\]

The vector \( \nabla E \) with elements \( \zeta_i \) and \( \eta_j \) will be called the gradient vector, the vector \( \Theta \) with elements \( \alpha_i \) and \( \beta_j \) is the parameter vector, the vector \( \Psi \) with elements \( \langle u, v_i \rangle \) and \( \langle u, u_j \rangle \) will be denoted by reference vector.

On the matrix \( M \) one can make observations analogously to those made in chapter 6:

- from the definition of \( M \) it follows that \( \langle u_j, y_i \rangle = \langle y_i, u_j \rangle \), consequently \( M \) is symmetric about its main diagonal. In addition, it is clear that the elements on the main diagonal are non-negative.
- the elements of the matrix depend only on the input signal u, the process 
and the linear operators of the model; they do not depend on the coefficient 
vector if, as it was assumed, this vector is constant.

- in principle, each component of the gradient vector is dependent on all para-
metrics; this leads to parameter interaction.

- by use of orthogonality relations a number of matrix elements can be made 
zero for certain classes of input signals, cf. Chapters 3 and 4.

- if M is not singular then, by multiplication of the matrix equation with 
$M^{-1}$, in the right hand side of equation (9.6) becomes: $M^{-1}M = I$ (with $I$ = identity matrix). This means that if $M^{-1}$ can be instrumented in the esti-
mating scheme then there is no interdependence in the determination of the 
different coefficients.

The general expression of equation (9.6) looks rather formidable. One has to 
remember, however, that it governs the determination of not less than $m + n + 1$ 
unknown coefficients. Moreover, it will become clear that in the instrumentation 
of the problem it will not be necessary to determine all elements of matrix 
M explicitly. Starting from equation (9.6) there are two routes:

- Select an operational relation between the gradient vector $\nabla E$ and the 
parameter vector $\theta$

$$\theta = \mathcal{G}[\frac{1}{2} \nabla E]$$

(9.7)

where $\mathcal{G}$ is called the policy operator. Substituting equation (9.7) in 
equation (9.6) leads to:

$$\theta = \mathcal{G}[M \theta - \varphi]$$

(9.8)

A proper choice of $\mathcal{G}$ will give a set of equations so that in time:

$$[\zeta_0, \ldots, \zeta_m, \eta_1, \ldots, \eta_m] \rightarrow [\sigma, \ldots, \sigma, \sigma, \ldots, \sigma]$$
Except for pathological cases this implies:
\[ [\alpha_0, \ldots, \alpha_n, \beta_1, \ldots, \beta_m] \rightarrow [a_0, \ldots, a_n, b_1, \ldots, b_m] \]
where \([a_i, b_j]\) are the parameters of the process.

As a result of the condition which requires the coefficient vector to be constant, the policy operator has to be one of the intermittent type, i.e., measuring over an interval with constant coefficients, correcting the model coefficients after the interval according to the results of the measurements, measuring again, correcting again, etc.

Consequently, if the argument \(i\) denotes the \(i\)-th measuring and adjustment cycle then at the end of the measuring interval one finds for the error vector:
\[ \frac{1}{2} \nabla E(i) = M \Theta(i) - \Psi(i) \]  

(9.9)

In the next adjustment interval the coefficient vector may be changed according to:
\[ \Theta(i+1) = \Theta(i) - \Gamma(i) \nabla E(i) \]  

(9.10)

where \(\Gamma(i)\) may be a matrix of real numbers, representing "gain factors" of the adjustment loop. With this new value of \(\Theta\) over the \((i+1)\)th measuring interval the error vector is determined again:
\[ \frac{1}{2} \nabla E(i+1) = M \Theta(i+1) - \Psi(i+1) \]

This procedure leads to a set of difference equations.

In case we choose as a policy operator
\[ \dot{\Theta} = -\frac{1}{2} \Gamma \nabla E \]  

(9.11)

then we are led to a differential equation:
\[ \dot{\Theta} + \Gamma M \Theta = \Gamma \Psi \]  

(9.12)

Strictly speaking, this does not hold as now \(\Theta\) is changing whereas equation (9.6) was derived assuming \(\Theta\) to be constant. In cases where the convergence is relatively slow, however, equation (9.12) leads to an approximation of the dynamic behaviour.
The instrumentation of these equations leads to the class, mentioned in chapter 2, "using a model-adjustment technique"; the model adjustment gives an approximation of the process characteristics.

Put \( \mathbf{y} = 0 \) where \( \mathbf{0} \) is the zero-vector. Now equation (9.7) reduces to:

\[
M \hat{\mathbf{y}} = \mathbf{y} \quad \text{and} \quad \hat{\mathbf{y}} = M^{-1} \mathbf{y}
\]

Provided \( M \) is not singular, this matrix equation can be solved for the unknown elements of the vector \( \hat{\mathbf{y}} \) as all elements of \( M \) and \( \mathbf{y} \) can be determined by a proper instrumentation. This leads us as to the other class of chapter 2 called: "using an explicit mathematical relation".

If the model can describe the dynamics of the process completely and if there is no additive noise then there is a set of parameters \( [a_0, \ldots, a_n, b_1, \ldots, b_m] \) for which the following holds:

\[
a = \sum_{i=0}^{n} a_i y_i + \sum_{j=1}^{m} b_j u_j - u
\]

Subtraction of eq. (9.14) from (9.2) leads us to

\[
e = \sum_{i=0}^{n} (\alpha_i - a_i) y_i + \sum_{j=1}^{m} (\beta_j - b_j) u_j
\]

Consequently, the error is linearly related to the deviation of the model parameters from their optimal values. Using again as the error criterion the minimization of (9.4) and differentiation according to (9.5a) and (9.5b), then this results in

\[
\begin{bmatrix}
\zeta_0 & \langle y_0, y_0 \rangle & \langle y_0, y_0 \rangle & \langle u_0, y_0 \rangle & \langle u_0, y_0 \rangle \\
\frac{1}{2} \zeta_n & \langle y_n, y_n \rangle & \langle u_n, y_n \rangle & \langle u_n, y_n \rangle & \langle u_n, y_n \rangle \\
\end{bmatrix}
\begin{bmatrix}
\alpha_0 - a_e \\
\alpha_n - a_n \\
\beta_1 - b_1 \\
\beta_m - b_m \\
\end{bmatrix}
\]

or renaming the vectors in this order:
\[ \frac{1}{2} \nabla E = M \lambda \]  

(9.16)

where \( \nabla E \) and \( M \) are the same as defined for eq. (9.7) and where the vector \( \lambda \) with elements \( \alpha_i - a_i \) and \( \beta_j - b_j \) is the parameter-error vector.

As before different kinds of policy operators \( \mathcal{P} \) can be used, e.g.

\[ \Delta \lambda = - \frac{1}{2} \Gamma \nabla E \]  

(9.17)

or

\[ \Delta \lambda = - \Gamma M \lambda \]

This leads to a vector difference equation

\[ \Delta \lambda = \lambda (i+1) - \lambda (i) = - \Gamma M \lambda (i) \]

or

\[ \lambda (i+1) = (I - \Gamma M) \lambda (i) \]  

(9.18)

or

\[ \lambda (i) = (I - \Gamma M)^i \lambda (0) \]  

(9.18a)

Again the "gain factors" \( \Gamma \) may be made dependent on the iteration number \( i \).

As before, we also may choose

\[ \dot{\lambda} = - \frac{1}{2} \Gamma \nabla E \]

which leads to

\[ \dot{\lambda} = - \Gamma M \lambda \]  

(9.19)

or

\[ \lambda (t) = \lambda (0) \exp [ - \Gamma M t ] \]  

(9.19a)

where \( \lambda (0) \) is the parameter-error vector at \( t = 0 \). As \( M \) is a real symmetric matrix its characteristic roots are real. According to the theory of matrix exponentials \( \lambda (t) \rightarrow 0 \) for \( t \rightarrow \infty \) if all characteristic roots of \( M \) have real positive parts.
The choice of the error criterion and the type of information processing. Let us consider the case where the parameter $b(t)$ in the process $P$ is the quantity that changes according to an arbitrary, unknown time-function and $\beta(t)$ is the parameter in the model $M$ that has to follow that time-function (parameter tracking). The quantities $b$ and $\beta$ may be coefficients of differential equations, of difference equations, of transfer functions or of sampled data transfer functions. They may also represent the locations of poles and zeros of the transfer functions, coordinates of the impulse response functions, moments of the impulse response functions, coordinates of amplitude- or phase characteristics.

The error in the correspondence of the transfer characteristics of $P$ and $M$ can be defined in many ways. One might propose to use as an error criterion the minimization of

$$\int_{t-T}^{t} \left\{ b(t) - \beta(t) \right\}^2 dt$$

(9.20)

or some corresponding relationship.

A complete solution for $\beta(t)$ is explicitly obtainable only for the most simple cases. Furthermore, $b(t)$ is not known in practical applications. Consequently, the error criterion of equation (9.20) can be used in simulation studies only.

Therefore another type of error, $e = y - \psi$, was defined. As an error criterion, the squared error over an interval of time was used. There is, however, a considerable freedom in choosing error criteria, e.g.,

$$E_1 = \mathcal{E} \left[ \int_{t-T}^{t} c^2 dt \right] \quad \text{def} \quad \mathcal{E} \langle e, e \rangle \quad (9.21a)$$

$$E_2 = \mathcal{E} \left[ \int_{t-T}^{t} |e| dt \right] \quad \text{def} \quad \mathcal{E} \langle |e|, 1 \rangle \quad (9.21b)$$

$$E_3 = \mathcal{E} \left[ \int_{0}^{\infty} e^{(t-\theta)} h(\theta) d\theta \right] \quad \text{def} \quad \mathcal{E} \langle e, e \rangle_h \quad (9.21c)$$

$$E_4 = \mathcal{E} \left[ \int_{0}^{\infty} |e| e^{(t-\theta)} h(\theta) d\theta \right] \quad \text{def} \quad \mathcal{E} \langle |e|, 1 \rangle_h \quad (9.21d)$$
In this order, the criteria may be called:

- \( E_1 \) integral squared error
- \( E_2 \) integral absolute error
- \( E_3 \) time weighted squared error
- \( E_4 \) time weighted absolute error

In \( E_1 \) and \( E_2 \) the integration interval is "sliding" along the t-axis.

Another possibility is the use of intervals that are prefixed on the t-axis (intermittent schemes). In \( E_3 \) and \( E_4 \) there is a weighting of the error with a function that puts less emphasis on older information, i.e., "the past is gradually forgotten". This can be achieved by choosing a low pass filter with impulse response \( h(t) \to 0 \) for \( t \to \infty \) instead of an integrator.

As before, the bracket symbols defined by equation (9.21) are introduced for the sake of convenience. Differentiation of equations (9.21) with respect to \( \alpha_i \) and \( \beta_j \) analogously to equation (9.5) leads to:

\[
\begin{align*}
\frac{1}{2} \zeta_i &= \frac{1}{2} \frac{\partial E_i}{\partial \alpha_i} = \langle e, y_i \rangle \\
\eta_j &= \frac{1}{2} \frac{\partial E_i}{\partial \beta_j} = \langle e, y_j \rangle \\
\zeta_i &= \frac{\partial E_2}{\partial \alpha_i} = \langle \frac{e}{|e|}, y_i \rangle \\
\eta_j &= \frac{\partial E_2}{\partial \beta_j} = \langle \frac{e}{|e|}, y_j \rangle \\
\frac{1}{2} \zeta_i &= \frac{1}{2} \frac{\partial E_3}{\partial \alpha_i} = \langle e, y_i \rangle _h \\
\eta_j &= \frac{1}{2} \frac{\partial E_3}{\partial \beta_j} = \langle e, y_j \rangle _h \\
\zeta_i &= \frac{\partial E_4}{\partial \alpha_i} = \langle \frac{e}{|e|}, y_i \rangle _h \\
\eta_j &= \frac{\partial E_4}{\partial \beta_j} = \langle \frac{e}{|e|}, y_j \rangle _h
\end{align*}
\]

Of course, the error criterion used determines the implementation. The term

\[
\frac{e}{|e|} = \text{sgn } e = \pm 1
\]

is the signum function.
in equation (9.22) leads to a binary operation; for example a multiplication \( x \cdot \text{sgn} \, e \) can be performed by a type of relay driven by \( e \) and switching between \(+x\) and \(-x\).

In chapter 3, a distinction was made with regard to the type of signal used. The names "continuous" (a better term would be: non-sampled) and sampled refer to the information-representation as a function of time. As a function of amplitude a distinction can be made as to: analog, quantized and binary. In studying the following sections, it has to be kept in mind that many schemes can be implemented for the use of each of these signal types mentioned.

The choice of the dynamic operators. In the previous sections, nothing has been said about the dynamic operators \( F_0, \ldots, F_m \) and \( G_0, \ldots, G_n \) that are a part of the generalized model in fig. 9.1.

It is clear that the choice of these operators influences the properties of the parameter estimating scheme, e.g., the kind of information produced, the speed of convergence to the solution, the possible bias of the solution, the influence of additive noise, the possibilities of simplification of the implementation using orthogonality relations.

The following are a few possible choices for the linear operators:

\[
\begin{align*}
F_j & G_i \quad \text{differentiators} \\
F_j & G_i \quad \text{high pass filters} \\
F_j & G_i \quad \text{integrators} \\
F_j & G_i \quad \text{low pass filters} \\
F_j & G_i \quad \text{band pass filters} \\
F_j & \quad \text{pure time delay units; } \alpha_0 = 1, \quad G_0 = 1 \quad \text{and } G_i = 0 \text{ for } i > 0 \\
F_j & \quad \text{orthogonal filters; } \alpha_0 = 1, \quad G_0 = 1 \quad \text{and } G_i = 0 \text{ for } i > 0
\end{align*}
\]
The last two choices lead to a scheme (cf. fig. 2.3) where the model is no longer a "generalized" one. In such a case, the signal \( e = y - w \) can be called the error-in-the-dependent-variable, as both \( y \) and \( w \) are dependent on the input variable \( u \).

It has to be noted of course that a "mixed" set of operators can also be used. A linear operator may consist also of a multiplier with the signal \( u \) (or \( y \)) as one input and an independent time function as the other input. This arrangement may be used in development of the signal \( u \) (or \( y \)) in orthogonal components.

The choice of the linear operators hinges on the completeness and the simplicity of the process representation. It also depends on the type of information that is needed from the process (e.g., time domain or frequency domain information).

It has to be remembered that the best-fit-in-some-sense to the observed data depends on the assumptions that are used for a particular type of description.

The propagation of errors or uncertainty from one description to another one is also often difficult to analyze.

The following table summarizes some of the properties of the different types of models (cf. fig. 2.13)

<table>
<thead>
<tr>
<th>model:</th>
<th>usable a priori knowledge?</th>
<th>output linear in parameters?</th>
<th>interaction of the parameter adjustment?</th>
<th>bias due to noise?</th>
</tr>
</thead>
<tbody>
<tr>
<td>timedelay</td>
<td>little (1)</td>
<td>yes</td>
<td>little</td>
<td>no</td>
</tr>
<tr>
<td>orthogonal filter</td>
<td>little (1)</td>
<td>yes</td>
<td>little</td>
<td>no</td>
</tr>
<tr>
<td>generalized model</td>
<td>much (2)</td>
<td>yes</td>
<td>some</td>
<td>yes</td>
</tr>
<tr>
<td>topologically identical model</td>
<td>much (2)</td>
<td>no</td>
<td>much</td>
<td>no</td>
</tr>
</tbody>
</table>

1) the process is (approximately) linear;
   the approximate length of the impulse response.

2) order and form of the differential equation;
   order of magnitude of the parameter values that have to be estimated;
   known values of some of the parameters.
Analogous considerations hold for nonlinear models. The restriction that the error has to be linear in the parameters limits the choice of such nonlinear dynamic operators to a special class; cf. chapter 4.

Apart from the previous considerations the interaction of the model adjustments may be important to consider. For a minimal parameter interaction, one may choose orthogonal filters according to section 3.2. and 3.3. In that case

$$\sum_{j=0}^{T} u_j \, dt = \langle u_i, u_j \rangle = c \delta_{ij} \tag{9.23}$$

With the error criterion and the error

$$E = \int_0^T e^2 \, dt = <e, e> \quad \text{and} \quad e = y - \sum_{j=0}^{m} \beta_j u_j$$

this leads to the simplified relationship:

$$-\frac{1}{2} \frac{dE}{\beta_j} = c \beta_j - <y, u_j> \tag{9.24}$$

i.e., the gradient with respect to one parameter is now independent of all other parameters. The optimal value of the parameter is

$$\frac{\partial E}{\partial \beta_j} = c \quad \text{or} \quad \beta_j = \frac{<y, u_j>}{c}$$

From

$$E = \int_0^T e^2 \, dt = \int_0^T (y - \sum_{j=0}^{m} \beta_j u_j)^2 \, dt$$

it follows that the residual part of the error criterion is:

$$E_{res} = \int_0^T y^2 \, dt - c \sum_{j=0}^{m} \beta_j^2 \tag{9.25}$$
This shows the contribution of each model parameter $\beta_j$ to the minimization of the error criterion. It gives us a means for judging the number of model parameters needed. Extending the number of model parameters with $m+1, \ldots$ does not influence the optimal settings of the parameters $\beta_1, \ldots, \beta_m$.

The problem remains as to how one chooses the impulse responses of the dynamic operators in such a way that their output signals $u_j(t)$ are orthogonal. Assume we want

$$\int_0^T u_i(t) u_j(t) \, dt = 0$$

with

$$u_i(t) = \int_0^\infty u(t - \tau) g_i(\tau) \, d\tau$$
$$u_j(t) = \int_0^\infty u(t - \tau) g_j(\tau) \, d\tau$$

This leads to the condition

$$c_i j = \int_0^T \int_0^\infty \int_0^\infty u(t - \tau) g_i(\tau) g_j(\tau) \, dt \, d\tau \, dw = 0$$

Consequently the impulse responses of the dynamic operators have to satisfy this condition in order to derive the orthogonal output signals $u_j(t)$ from $u(t)$.

Therefore, these impulse responses are dependent on the input signal characteristics. For "white" noise

$$E \left[ \tilde{u}_i(t) \tilde{u}_j(t) \right] = \sigma_i^2 \delta(t - \tau) \delta(t - \tau)$$

taking the expectation leads to the simple condition

$$\int_0^\infty \delta(t - \tau) \, d\tau = \frac{1}{\sigma_i^2} \delta_i^2$$

For arbitrary input signals it is not a simple task to find the desired impulse responses; cf. section 3.2. and Kitamori (1960), reference of chapter 8.
It is evident that for a particular application one orthogonal series is "better" than some other series. This connotation "better" may be based on the number of terms needed for the approximation (with predetermined accuracy) of the process dynamics. This leads to the proposal of adjusting separately some parameter of the orthogonal filters in such a way that, given a limited number of filters, the residual error is minimized. For orthogonal dynamic operators for multivariable and nonlinear systems cf. chapter 4.
9.2 Minimization of the instantaneous error.

In the previous section, a number of error criteria has been mentioned. The mathematical expectation was taken in order to remove the stochastic aspect. In a hardware realization of a model-adjustment system, one has only one ensemble element to operate on. In this section, the minimization of the instantaneous error is discussed; the minimization of an integral error is the topic of the next section.

Basic equations. The generalized model is shown in fig. 9.1. For the moment the additive noise is neglected. The generalized error is defined before as:

\[ e = \sum_{i=0}^{n} \alpha_i y_i + \sum_{j=1}^{m} \beta_j u_j - u \]  
(9.27)

The error criterion is chosen as minimizing:

\[ E = e^2 \]  
(9.28)

From this it follows that

\[ \left\{ \begin{array}{l}
\frac{1}{2} \zeta_i = \frac{1}{2} \frac{\partial E}{\partial \alpha_i} = e \frac{\partial e}{\partial \alpha_i} = e y_i \\
\frac{1}{2} \eta_j = \frac{1}{2} \frac{\partial E}{\partial \beta_j} = e \frac{\partial e}{\partial \beta_j} = e u_j 
\end{array} \right. \]  
(9.29)

and the vector equation is

\[ \frac{1}{2} \Delta E = M \theta - \psi \]  
(9.30)

with

\[ \Delta E = [\zeta_0, \ldots, \zeta_m, \eta_1, \ldots, \eta_m]' \]
\[ \theta = [\alpha_0, \ldots, \alpha_n, \beta_1, \ldots, \beta_m]' \]
\[ \psi = [uy_0, \ldots, uy_n, uu_1, \ldots, uu_m]' \]

and

\[ M = \begin{bmatrix}
y_0 y_0 & \cdots & y_n y_0 & u_0 y_0 & \cdots & u_m y_0 \\
y_0 y_1 & \cdots & y_n y_1 & u_0 y_1 & \cdots & u_m y_1 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
y_0 y_n & \cdots & y_n y_n & u_0 y_n & \cdots & u_m y_n \\
y_0 u_1 & \cdots & y_n u_1 & u_0 y_1 & \cdots & u_m y_1 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
y_0 u_m & \cdots & y_n u_m & u_0 u_m & \cdots & u_m u_m 
\end{bmatrix} \]
where ' indicates the transpose of the vector.

The same reasoning holds for the error criterion
\[ E = |e| \]  \hspace{1cm} (9.28a)

Now
\[ \zeta_i = \frac{\partial E}{\partial \alpha_i} = \frac{e}{|e|} \frac{\partial e}{\partial \alpha_i} = \frac{e}{|e|} y_i \]
\[ \eta_j = \frac{\partial E}{\partial \beta_j} = \frac{e}{|e|} \frac{\partial e}{\partial \beta_j} = \frac{e}{|e|} u_j \]
and a vector equation results of the form
\[ \nabla E = \frac{M \theta - \Psi}{|e|} \]  \hspace{1cm} (9.31)

where \( \nabla E, \theta, \Psi, M \) are the same as given before.

It is interesting to note, that with these error criteria there is no need
to stipulate that \( \theta \) has to be constant as was done in section 9.1; equations
(9.30) and (9.31) also hold exactly for varying \( \theta \). So one may choose the
policy-operator in the equation
\[ \dot{\theta} = \mathcal{P}[\frac{1}{2} \nabla E] \]  \hspace{1cm} (9.32)
for example as in eq. (9.11)
\[ \dot{\theta} = -\frac{i}{2} \sum \nabla E \]  \hspace{1cm} (9.33)

With eq. (9.30) this leads to the matrix-differential equation
\[ \dot{\theta} + \Gamma M \theta = \Gamma \Psi \]  \hspace{1cm} (9.34)

If the input \( u \) is stochastic then the elements of \( M \) are also stochastic. This
differential equation can be simplified again by assuming that the model is
capable of an ideal representation; see Miller (1962). This implies that
\[ \dot{\theta} = \sum_{i=0}^{n} \alpha_i y_i + \sum_{j=1}^{m} b_j u_j - u \]  \hspace{1cm} (9.35)
and
\[ \dot{e} = \sum_{i=0}^{n} \left( \alpha_i - \alpha_i^* \right) y_i + \sum_{j=1}^{m} \left( b_j - b_j^* \right) u_j \]  \hspace{1cm} (9.36)
leads to
\[ \dot{\lambda} = -\Gamma M \lambda \]  \hspace{1cm} (9.37)
cf. equations (9.14), (9.15) and (9.19).
For a stationary stochastic input, signal the elements of the matrix $M$ are stationary stochastic as well.

The vector equations (9.30) and (9.32) govern the behaviour of the adjusting system. For the instrumentation scheme the equations (9.27) and (9.33) are sufficient:

$$
\dot{\theta} = -\frac{1}{2} \sum \psi \varepsilon
$$

or

$$
\begin{bmatrix}
\dot{x}_i \\
\dot{\beta}_j
\end{bmatrix} = -\sum \begin{bmatrix}
ey_i \\
\varepsilon u_j
\end{bmatrix}
$$

(9.38)

Using a simple "gain matrix" $\sum = \gamma I$ then:

$$
\begin{cases}
\dot{x}_i = -\gamma \varepsilon y_i \\
\dot{\beta}_j = -\gamma \varepsilon u_j
\end{cases}
$$

(9.38a)

As an example we will assume a process with a simple linear transfer function

$$
H(s) = \frac{b}{1+as}
$$

where $a$ and $b$ have to be estimated. Written in Laplace notation and neglecting the initial conditions, the process obeys the equation

$$
b U(s) - a s Y(s) - Y(s) = 0
$$

or

$$
W(s) \left\{ b U(s) - a s Y(s) - Y(s) \right\} = 0
$$

or

$$
b U_0(s) - a Y_1(s) - Y_0(s) = 0
$$

where

$W(s)$ is an arbitrary (filter) transfer function

$$
\begin{align*}
U_0(s) &= W(s) U(s) \\
Y_0(s) &= W(s) Y(s) \\
Y_1(s) &= W(s) s Y(s)
\end{align*}
$$

By means of a suitable choice of $W(s)$ the transfer function $W(s)s$ may be more practically-realizable than without this filter, e.g.,
9 - 18

\[ W(s) = \frac{1}{1+\tau s} \quad W(s)s = \frac{s}{1+\tau s} \]

Now the model equation is chosen as:

\[ \beta U_0(s) - \alpha Y_1(s) - Y_0(s) = \mathcal{L} \left[ e(t) \right] \tag{9.39} \]

This is represented in fig. 9.2 for that particular choice of \( W(s) \).

Now one may choose from a variety of error criteria.

For \( E = e^2(t) \) one finds:

\[ \frac{\partial E}{\partial \alpha} = \frac{\partial e^2}{\partial \alpha} = 2e \frac{\partial e}{\partial \alpha} = 2e(-Y_1) \]
\[ \frac{\partial E}{\partial \beta} = \ldots \ldots = 2e u_0 \]

and for the gradient equations

\[ \frac{d\alpha}{dt} = -\gamma \frac{\partial E}{\partial \alpha} = -2\gamma e(-Y_1) \]
\[ \frac{d\beta}{dt} = \ldots \ldots = -2\gamma e u_0 \]

From this it follows that fig. 9.2 has to be completed using the instrumentation of fig. 9.3a. Choosing the error criterion \( E = |e| \), one finds the instrumentation of fig. 9.3b, where a couple of analog multipliers has been replaced by simple relays.

**Convergence properties.** The convergence and stability properties of the parameter-estimation schemes, given in previous sections, are often difficult to establish. Therefore, another approach, where the basic requirement is not some type of error criterion but a monotonic model-convergence, is given in this section; see Graupe (1961).

The process is given by

\[ g \left( t, b_1, \ldots, b_m \right) = c \]

where \( b_j \) are unknown parameters. This equation holds for a period of time, where the time dependence is given by several functions \( x_i(t) \) and \( y_j(t) \), which
may include derivatives. The model is given by

\[ g(t; \beta_1, \ldots, \beta_m) = e \quad (9.43) \]

where \( e \) is the error that has to be minimized. Now the sense in which it will be minimized is not, however, determined in advance but follows from the following type of reasoning. It generally holds that:

\[ \frac{de}{dt} = \sum_{j=1}^{m} \frac{de}{\partial \beta_j} \frac{d\beta_j}{dt} + \frac{de}{dt} \quad (9.44) \]

Each term of this equation is multiplied by an odd function \( -f \) of \( e \); \( f(-e) = -f(e) \). Now assume that the following condition is fulfilled:

\[ -f(e) \sum_{j=1}^{m} \frac{de}{\partial \beta_j} \frac{d\beta_j}{dt} > \left| -f(e) \frac{de}{dt} \right| \quad \text{for} \quad 0 < e_{\text{min}} < |e| \quad (9.45) \]

In that case

\[ -f(e) \frac{de}{dt} > 0 \quad \text{for} \quad 0 < e_{\text{min}} < |e| \quad (9.46) \]

The last inequality assures monotonic convergence of the error towards zero as long as the condition posed by equation (9.45) is fulfilled. After the error is diminished in such a way that this condition no longer holds, then the error will fluctuate in the interval \(-e_{\text{min}} < e < e_{\text{min}}\).

For a certain choice of \( f(e) \) there are several possible choices for \( \frac{d\beta}{dt} \) in order to fulfill the requirements of equation (9.45). A few examples for the case of one parameter adjustment are given below. Take \( f(e) = e^{-1} \), then

\[
\begin{align*}
\frac{d\beta}{dt} & = -\gamma \frac{de}{\partial \beta} \\
& \quad - \gamma e \frac{de}{\partial \beta} \\
& \quad - \gamma e \text{sgn} \left[ \frac{de}{\partial \beta} \right] \\
& \quad - \gamma \text{sgn} [e] \frac{de}{\partial \beta} \\
& \quad - \gamma \text{sgn} [e] \text{sgn} \left[ \frac{de}{\partial \beta} \right] \\
\frac{1}{e} \frac{de}{\partial \beta} & \quad \frac{\frac{de}{\partial \beta}}{|e|} \quad \frac{\frac{de}{\partial \beta}}{|e|} \quad \text{for} \quad \gamma > 0 \quad (9.47)
\end{align*}
\]
For one of these cases and for the simultaneous adjustment of $n$ parameters, fig. 9.4 shows the instrumentation schematically. The gains $\gamma_k$ of the integrators have to be chosen with respect to noise filtering ($\gamma_k$ low) and convergence speed ($\gamma_k$ high).

Taking as a simple example a steepest descent instrumentation for each parameter:

$$\frac{d\theta_i}{dt} = - \gamma \frac{\partial E}{\partial \theta_i}$$  \hspace{1cm} (9.48)

then for

$$E = e^2 \Rightarrow \frac{\partial E}{\partial \theta_i} = 2e \frac{\partial e}{\partial \theta_i} \quad \text{and} \quad \frac{d\theta_i}{dt} = - 2\gamma e \frac{\partial e}{\partial \theta_i}$$  \hspace{1cm} (9.49a)

$$E = |e| \Rightarrow \frac{\partial E}{\partial \theta_i} = \text{sgn} \, [e] \frac{\partial e}{\partial \theta_i} \quad \text{and} \quad \frac{d\theta_i}{dt} = - \gamma \text{sgn} \, [e] \frac{\partial e}{\partial \theta_i}$$  \hspace{1cm} (9.49b)

Thus, equation (9.47) guarantees the monotonic convergence of the steepest descent adjustment using the error criterion given above for one parameter.

Using the (generalized) error of equation (9.23), one finds for the condition of equation (9.45) if $E = e^2$:

$$2 \gamma \left\{ \sum_{j=1}^m u_j^2 + \sum_{i=1}^n y_i^2 \right\} > \left| - \frac{\partial e}{\partial t} \right| \quad \text{for} \quad 0 \leq e_{\min} < |e|$$  \hspace{1cm} (9.50)

For the $E_e|e|$ criterion one finds for the condition of equation (9.45):

$$\frac{\gamma}{|e|} \left\{ \sum u_j^2 + \sum y_i^2 \right\} > \left| - \frac{\partial e}{\partial t} \right|$$

or

$$\gamma \left\{ \sum u_j^2 + \sum y_i^2 \right\} > \left| \frac{\partial e}{\partial t} \right| \quad \text{for} \quad 0 \leq e_{\min} < |e|$$  \hspace{1cm} (9.51)

If the input signal is stationary then the $u_j$ and $y_i$'s are stationary too.

The coefficient $\gamma$ gives a contribution to the "loop-gain".

An intuitive type of reasoning about the convergence of this type of system is given in the paper by Potts, Ornstein and Clymer (1961); the following example is taken from that reference.
The process dynamics equation is given by:
\[ \dot{y} + a_0y = u \]
where \( a_1 \) and \( a_0 \) have to be determined. For comparison with the previous discussion put \( \dot{y} = y_1 \); \( y = y_0 \) and \( u = u_o \). The generalized error is:
\[ e = \alpha_1 \dot{y} + \alpha_0 y - u = \alpha_1 y_1 + \alpha_0 y_0 - u_o \]
where \( \alpha_1 \) and \( \alpha_0 \) are the computed values of the coefficients. Taking the absolute magnitude
\[ |e| = |\alpha_1 y_1 + \alpha_0 y_0 - u_o| \]
gives the plane in the \( |e(\alpha_1, \alpha_0)| \) space shown in fig. 9.5. This surface is tangent to the \( (\alpha_1, \alpha_0) \) plane in line \( L \) given by:
\[ \alpha_1 y_1 + \alpha_0 y_0 - u_o = 0 \]
The slopes of the surface along the \( \alpha_1 \) and \( \alpha_0 \) axis can be found as
\[ \frac{\partial |e|}{\partial \alpha_i} = \frac{e}{|e|} y_i \]
\[ \frac{\partial |e|}{\partial \alpha_o} = \frac{e}{|e|} y_o \]
The right hand sides form the components of the gradient vector. If the coordinates of the running point are driven at rates proportional to the negatives of these slopes then the running point moves towards the line \( L \) along a path normal to that line.

As the values of \( y_1, y_0 \) and \( u_o \) change with time, the line \( L \) will rotate about the fixed point \( (a_1, a_0) \). The running point will converge uniformly to the point \( (a_1, a_0) \). The instrumentation of this case has already been given in fig. 9.3.

Some discussion of convergence properties can be found in the paper by Miller (1962), where the second method of Lyapunov is used. Define \( R^2(t) \) as:
\[ R^2(t) = \sum_{i=0}^{n} (\alpha_i - \alpha_i)^2 + \sum_{j=1}^{m} (\beta_j - b_j)^2 \geq 0 \] (9.52)
Differentiation of equation (9.52) with respect to time yields:
\[ \frac{dR^2}{dt} = -2Ye \left\{ \sum (\alpha_i - \alpha_i) y_i + \sum (\beta_j - b_j) u_j \right\} = -2Ye^2 \]
as a result of equations (9.38a) and (9.36). Thus, \( R^2(t) \) will be a Lyapunov function.
Influence of additive noise. There has not yet been a discussion concerning the influence of additive noise. Consider, however, the simple example of instrumentation given in fig. 9.6. Assume the process is described by

\[ \dot{x} + a_0 x = u \]  \hspace{1cm} \text{(9.53)}

Without the additive noise, the generalized error is:

\[ e = \dot{x} + \alpha_0 x - u \]  \hspace{1cm} \text{(9.54)}

Now assume that \( \alpha_0 = a_0 \) and that the adjustment loop is cut open at \( B \). The signal and noise are not related; \( \mathcal{E}[u,n] = 0 \) and \( \mathcal{E}[u,h] = 0 \). If \( n \neq 0 \), then the error signal \( e \) has two components due to the noise signal, viz.:

\( \alpha_0 n \) and \( h \). The (negative) output of the multiplier \( M_1 \) consists of:

\[ (x + n) \left( u - \dot{x} - a_0 x - h - a_0 n \right) = - (x + n) (h + a_0 n) \]

Now

\[ \mathcal{E}[x,n] = 0 \quad \text{and} \quad \mathcal{E}[x,h] = 0 \]

due to the assumption that the signal and noise are not related. Further it can be proved that

\[ \mathcal{E}[n,h] = 0 \]

The fourth term, however, gives

\[ \alpha_0 \mathcal{E}[n,n] > 0 \quad \text{for} \quad n \neq 0 \]  \hspace{1cm} \text{(9.55)}

This amounts to a d.c. component at the output of the multiplier. So when the loop is closed at \( B \) the system is not in equilibrium. Loosely speaking, one would, therefore, expect an offset of the coefficient \( \alpha_0 \) such that:

\[ \alpha_0 \mathcal{E}[n,n] = - (\alpha_0 - \alpha_c) \mathcal{E}[x,x] \]  \hspace{1cm} \text{(9.56)}

Because \( x.x \) is the only combination of the inputs of \( M_1 \) that can provide an expectation \( \neq 0 \), this means that a bias of the coefficient will be found which is related to the noise signal ratio.
The input-combination of $M_1$ which does not contribute to the expectation of the output does, however, add to the variance of that output and thus to the uncertainty of $\sigma$.

Examples of simulations, implementations and applications. Examples of these model-adjustment schemes where the instantaneous error is minimized can be found in several references.

In pioneering papers, Clymer (1958, 1959) introduces the idea of "system synthesis by dynamic implicit computation". This line of thought has its origin in the implicit circuit technique of analog computation, an example of which is the implementation of a divider circuit by the use of a multiplier and a high gain amplifier. "Implicit synthesis is thus a distinct mode of the application of computers. It differs from the usual mode of application, in which one knows the differential equation including all characteristics, and the excitation, while one wishes to determine the solution to the equation, which is the dependent variable as a function of the independent variable. By contrast, in implicit synthesis one knows the solution and the excitation, but one wishes to determine one or more unknown characteristics in the differential equation".

Some computer results for simple systems are given, as well as an illustration based on chemical reaction rate equations.

Further information on the development of this line of thought can be found in Potts, Ornstein, Clymer (1961), where application of the methods for the determination of aerodynamic coefficients and for the determination of human coefficients are indicated. Also cf. Ornstein (1963).
Miller (1962) gives computer results for the simultaneous determination of the coefficients $b_1$ and $a_0$ in the differential equation.

$$b_1 \dot{x} + x = a_0 u$$

The influence of additive noise is not reported on. See also Schoenemann (1965).
References


fig. 9.4

fig. 9.5
fig. 9.6
Fundamental for the estimation problem is the fact that the process has to be excited. It is often advantageous to use test signals, introduced from the outside if this is allowed; observations from spontaneous fluctuations alone give less accurate results. The knowledge about the system can be used to find an optimal test signal, which can be defined as the signal with a prescribed maximum amplitude (allowable disturbance) that yields the required knowledge with a specified accuracy in a minimum length of time. There is much choice in test signals, but if little a priori information is available about the system, the choice is seldom critical. For an extensive discussion of testsignals the reader is referred to the excellent monograph by Strobel (1968).

In this chapter, several types of testsignals are discussed. Generally speaking, the type of input signal determines the kind of knowledge that becomes available. Assume the linear process behaviour to be described by the equation:

$$a_2 \dot{y}(t) + a_1 \ddot{y}(t) + a_0 y(t) = b_1 \dot{u}(t) + b_0 u(t)$$

(10.1)

where $u(t)$ is the input and $y(t)$ is the (uncorrupted) output signal.

Description in the time domain leads to the convolution integral:

$$y(t) = \int_{-\infty}^{+\infty} h(\theta) u(t-\theta) \, d\theta$$

and description in the frequency domain leads to:

$$h(t) = 0 \quad \text{for } t<0$$

$$h'(t) = 0 \quad \text{for } t<0$$

(10.2)
\[ Y(s) = H(s) U(s) \]

with
\[ H(s) = \frac{b_1 s + b_0}{a_2 s^2 + a_1 s + a_0} \]

and \( s = j\omega \) for sinusoidal signals. Generally speaking, the measurement of transients is most useful for the first type of description whereas for the second type of description periodic signals are mostly used.

The relations between time and frequency domain description have been indicated in chapters 3 and 4. For practical purposes, it is necessary to know how to go from the time domain description to the frequency domain by means of Fourier transform methods.

The frequency transfer function can be obtained as the cross spectrum divided by the input power spectrum. The use of disturbances, present already in the process, may suffer from the following disadvantages:

- the power spectrum of the disturbances may be too narrow or it may contain some dominant frequencies, which require either prewhitening or periodic rejection in order to make the input spectrum approach that of white noise.
- because only finite records of input and output signals are available, the accuracy of correlation functions is limited. This necessitates the use of lag windows of a certain width. The wider the window, the greater the resolution of the spectrum, but the less the accuracy of the estimates.

This calls for a compromise.

Some aspects of the influence of testsignal properties on the accuracy of the estimates are discussed in section 11.4.
10.1 Impulse and step signals.

For information in the time domain there is the obvious solution of testing with impulse or step signals. Of theoretical interest is an input signal that approaches the properties of the Dirac impulse function, defined by:

\[ u(t) = c \delta(t) \quad \begin{cases} \delta(t) = 0 & \text{for } t \neq 0 \\ \int_{-\infty}^{\infty} \delta(t) \, dt = 1 \end{cases} \]  

(10.4)

If this is substituted in the convolution integral eq. (10.2) then \( y(t) = c \, h(t) \). A somewhat more practical input test signal approximates a step signal:

\[ u(t) = c \, r(t) \quad \begin{cases} r(t) = 0 & \text{for } t < 0 \\ r(t) = 1 & \text{for } t \geq 0 \end{cases} \]  

(10.5)

From the convolution integral follows

\[ y(t) = c \int_{0}^{t} h(\tau) \, d\tau \]  

(10.6)

and by subsequent differentiation:

\[ \frac{dy}{dt} = c \, h(t) \]  

(10.7)

Step signals have decreasing power at high frequencies where processes usually show attenuation.

The impulse and step signals should, of course, not endanger the process, and not drive it into nonlinear behaviour if a linear description of the behaviour around a set point condition is needed. In addition, ramp signals have some relevance in process testing.

From the response of the process to a step input, a number of practical parameters can be determined immediately, e.g., (cf. fig. 10.1)

- \( \tau \) dead time
- \( T_d \) delay time
- \( T_b \) build-up time.
For a more complete survey and nomenclature of such definitions, also in cases where overshoot occurs, the reader is referred to the ASA standards (ASA C 85.1 - 1963). A rather common situation is that where the transfer function can be expected to have poles on the real axis only. Consequently,

$$H(s) = \frac{k e^{-s\tau}}{s(s\tau + 1)}$$

If one may assume that the time constants are equal then from the step response curve the number \( n \) of those time constants and the numerical value of each of those time constants can be determined directly. For more general situations such rules for determining time constants have been reported, e.g., Schwarze (1960 - 1964), Strobel (1968).

The influence of additive noise. Now we have to consider the effect of additive noise \( n(t) \). Instead of an uncorrupted output \( x(t) \) there is available only \( y(t) = x(t) + n(t) \).

Assume \( n(t) \) to be characterized by:

$$\mathbb{E}[n(t)] = 0 \quad \mathbb{E}[n(t)^2] = \sigma^2$$

The uncertainty in one measurement introduced by the noise is given by the standard deviation \( \sigma \). This uncertainty can be diminished by performing the same test a number (say \( k \)) of times, assuming the deterministic part of the test to be reproducible.

The \( k \) time functions are added and this sum is divided by \( k \); cf. fig. 10.2. The resulting time function has a smaller standard deviation, a fact which can be shown as follows:

Assume the successive tests start at \( t = t_1, t_2, \ldots, t_k \). Consider the samples of the output signal \( \theta \) seconds after the start of each test and call

$$y(t_i + \theta) = x(t_i + \theta) + n(t_i + \theta)$$

or

$$y_i = x_i + n_i$$
Now the average over the \( k \) tests is:

\[
\bar{Y}_k = \frac{1}{k} \sum_{i=1}^{k} Y_i = x + \frac{1}{k} \sum_{i=1}^{k} n_i
\]

(10.8)
because \( x_i = x \) for all \( i \) due to the reproducible deterministic part of the test.

Consequently,

\[
\mathbb{E}[\bar{Y}_k] = x
\]

\[
\mathbb{E}[(\bar{Y}_k - x)^2] = \frac{\sigma^2}{k}
\]

The latter relation holds only if \( n_i \) and \( n_j \) are uncorrelated for \( i \neq j \).

This condition is sufficiently approximated if the autocorrelation function of the noise satisfies:

\[
\rho_{nn}(t_i - t_j) \approx 0
\]

that is, if the successive tests are sufficiently separated in time.

Under that condition, the standard deviation diminishes inversely proportional to the square root of the number of tests; Clark (1958); M.I.T. (1959); Clynes and Kohn (1961); Clark (1961).

This principle is instrumented in commercially available equipment, e.g., one called CAT (Computer of Average Transients), developed by Mmometron Corp.

This instrument extracts repetitive signals from noise by calculating the average response. Its applications are in the field of the study of evoked brain potentials, nerve potentials, retiograms, cardiograms, etc. Furthermore, smaller digital computers with analog-to-digital converters are being used for this purpose. Another quantity that may be of interest to estimate is the variance \( \sigma^2 \). One might propose to do this as follows from \( k \) observations:

\[
\frac{(\bar{Y} - \bar{Y}_k)^2}{\sigma^2}_k = \frac{1}{k} \sum_{i=1}^{k} (Y_i - \bar{Y}_k)^2
\]

(10.9)

For the expectation of this quantity we find:

\[
\mathbb{E}[(\bar{Y} - \bar{Y}_k)^2] = \frac{1}{k} \sum_{i=1}^{k} (Y_i - x + x - \bar{Y}_k)^2 = \frac{1}{k} \mathbb{E}[(\Sigma Y_i - x)^2 + 2 \Sigma (Y_i - x)(x - \bar{Y}_k) + \Sigma (x - \bar{Y}_k)^2] = \]

\[= \sigma^2 - \frac{2}{k} \sigma^2 + \frac{k}{k} \sigma^2 = \frac{k-1}{k} \sigma^2
\]
Hence as an unbiased estimate for $\sigma^2$ can be chosen:

$$s^2 = \frac{k}{k-1} \frac{(y_j - \bar{y})^2}{k} = \frac{1}{k-1} \frac{k}{k} \frac{(y_j - \bar{y})^2}{k}$$

(Note the case $k=1$).

The use of correlating filters. For testing with impulse or step functions the input signal has to be of a considerable amplitude. This implies that one may drive the process into a region of nonlinear operation for which the simple (impulse) response does not provide useful information. The process characteristics found in this way may differ from those under normal conditions. The problem may be circumvented by spreading the test signal energy over a longer time interval in such a way that the amplitude remains small. At any particular instant this can be arranged by using a so-called correlating filter (CF) and a corresponding test signal (cf. fig. 10.3a); Lichtenberger (1961). The test signal $u(t)$ and the impulse response function $g(t)$ are related to each other as $g(t) = u(T-t)$. Consequently, the output of the filter with this signal at its input is:

$$v(t) = \int_{-\infty}^{\infty} u(t-\theta) g(\theta) d\theta = \int_{-\infty}^{\infty} u(t-\theta) u(T-\theta) d\theta$$

(10.11)

This gives the autocorrelation function of the signal $u$. By a proper choice of $u(t)$ and the related $g(t)$ the signal $v(t)$ may approximate the impulse "function" $c \delta(t)$ arbitrarily well. In that case the output $y(t)$ of the process $P$ is an approximation of its impulse response. If $P$ is linear, then CF and $P$ may be rearranged according to fig. 10.3b without changing $y(t)$ whereas now the energy of the signal $v(t) = c \delta(t)$ is spread over a time interval of length $T$. Again the effect of additive noise can be diminished by repeating the experiment and by averaging over the responses.

The freedom of choice with respect to $g(t) = u(T-t)$ can be used to satisfy the criterion that the maximum amplitude of $u$ has to remain within acceptable bounds. This problem is identical with that of shaping radar signals for high resolution cf. Woodward (1955).
10.2 **Sinusoidal test signals.**

Information in the frequency domain is most easily obtained by using sinusoidal or other periodic test signals. This method of testing processes has had a long history (Angström, 1861). There is a host of publications on this subject; Fuchs (1954), Cowley (1957). Sinusoidal signals have many advantages. One can limit the measurements to the frequencies of interest; orthogonality properties of periodic signals give labelling facilities; linearity assumptions are easily verified; the duration of the test can be chosen quite arbitrarily; the generation and processing of such signals has been studied extensively; in "classical" feedback theory, stability studies are formulated in terms of sinusoidal signals. A disadvantage may be the possibility of dangerously exciting a resonance frequency of the process.

The principles of this method can be summarized using the following example.

Assume the process to be described by the differential equation

$$a_n y^{(n)} + \ldots + a_1 y + a_0 y = b_m w^{(m)} + \ldots + b_i w + u$$

(10.12)

The input signal is chosen as:

$$u(t) = u_0 \sin \omega t = \frac{u_0}{\omega} \left( e^{i\omega t} - e^{-i\omega t} \right)$$

(10.13)

Due to the linearity of the process its output consists of the sum of the responses of both input components after the transients have disappeared:

$$y(t) = \frac{u_0}{2j} \left( H(j\omega) e^{i\omega t} - H(j\omega) e^{-i\omega t} \right)$$

(10.14)

with

$$H(j\omega) = \frac{b_m (j\omega)^m + \ldots + b_1 (j\omega) + 1}{a_n (j\omega)^n + \ldots + a_1 (j\omega) + a_0}$$

(10.15)

and

$$H(j\omega) = \left| H(j\omega) \right| e^{j\theta(j\omega)}$$

$$H(j\omega) = \left| H(j\omega) \right| e^{-j\theta(j\omega)}$$

(10.16)

Consequently,
\[ \frac{y(t)}{u(t)} = \frac{u_0}{|H(j\omega)|} \sin \left( \omega t + \phi(\omega) \right) = u_0 \left\{ c(\omega) \sin \omega t + d(\omega) \cos \omega t \right\} \]  

(10.17)

with

\[ c(\omega) = \frac{1}{|H(j\omega)|} \cos \phi(\omega) \]
\[ d(\omega) = \frac{1}{|H(j\omega)|} \sin \phi(\omega) \]

(10.18)

and

\[ |H(j\omega)| = \left\{ c^2(\omega) + d^2(\omega) \right\}^{1/2} \]
\[ \phi(\omega) = \arctan \left\{ \frac{d(\omega)}{c(\omega)} \right\} \]

(10.19)

The process parameters to be determined are \( c(\omega_i) \) and \( d(\omega_i) \) for a number of frequencies \( \omega_i \). These can be determined by exciting the process with \( u(t) = u_0 \sin \omega_i t \) and by applying "Fourier filtering" to the output signal \( y(t) \). This filtering procedure consists of the multiplication of \( y(t) \) respectively with \( \sin \omega_i t \) and \( \cos \omega_i t \) and averaging over an integer number of periods:

\[ c(\omega_i) = \frac{2}{u_0 T} \int_0^T y(t) \sin \omega_i t \, dt \]
\[ d(\omega_i) = \frac{2}{u_0 T} \int_0^T y(t) \cos \omega_i t \, dt \]

(10.20)

with \( T = k T_i \)

and \( T_i = \frac{2\pi}{\omega_i} \)

Comparison with chapter 8 will illustrate the close relation to the correlation technique. In the same way, the harmonic content of \( y(t) \) can be measured if \( u(t) = u_0 \sin \omega_i t \) and the output is multiplied with frequencies \( l\omega_i \);

\[ l = 2, 3, 4, \ldots \]

The next problem is to derive, from the measured Nyquist plot, the parameters \( a_i \) and \( b_j \) in the transfer function given by eq. (10.15). Cf. Strobel (1968).

The influence of additive noise. (Sins, 1967). Thus far a situation is considered in which there are no disturbances. If \( y(t) \) is contaminated with a "noise signal" \( n(t) \) then the result of the Fourier filtering is
\[ \tilde{c}(\omega, T) = c(\omega) + \frac{2}{u_0 T} \int_0^T n(t) \sin \omega t \, dt \]

and

\[ \tilde{d}(\omega, T) = d(\omega) + \frac{2}{u_0 T} \int_0^T n(t) \cos \omega t \, dt \]

Now \( \tilde{c} \) and \( \tilde{d} \) are approximations which depend on \( T = kT \), the time interval over which integration is performed. Because \( n(t) \) is a stochastic signal, \( c \) and \( d \) will be of a stochastic nature. In the following we will assume \( \mathbb{E}[n(t)] = 0 \).

By considering the expectation we find:

\[ M_c(T) = \mathbb{E}[\tilde{c}(\omega, T)] = \mathbb{E}[c(\omega)] + \frac{2}{u_0 T} \int_0^T \mathbb{E}[n(t)] \sin \omega t \, dt = c(\omega) \]

\[ M_d(T) = \mathbb{E}[\tilde{d}(\omega, T)] = d(\omega) \]

Consequently, the numerical results of the measurements are not biased.

The variance of the measurements is also of interest

\[ \sigma_c^2(T) = \mathbb{E}\left[ (\tilde{c}(\omega, T) - c(\omega))^2 \right] = \mathbb{E}\left[ \left( \frac{2}{u_0 T} \int_0^T n(t) \sin \omega t \, dt \right)^2 \right] = \]

\[ = \frac{4}{u_0^2 T^2} \int_0^T \int_0^T \mathbb{E}[n(t) n(v)] \sin \omega t \sin \omega v \, dt \, dv \]

where

\[ \mathbb{E}[n(t) n(v)] = \Psi_{nn}(v-u) = \Psi_{nn}(v) \]

In Appendix C is derived that

\[ \left( \sigma_c^2(T) \right) = \frac{4}{u_0^2 T} \int_0^T \Psi_{nn}(\tau) \left( 1 - \frac{\tau}{T} \right) \cos \omega_0 \tau + \frac{1}{\omega_0 T} \sin \omega_0 \tau \right) \, d\tau \]

(10.24)

If the additive noise is "white" then

\[ \Psi_{nn}(t) = \delta(t) \]

and one finds \( \Psi_{nn}(t) \).

\( \Psi_{nn}(t) \) due to the relationship

\[ \int_{-\infty}^{\infty} \delta(t) \, dt = \frac{1}{2} \int_{-\infty}^{\infty} \delta(t) \, dt = \frac{1}{2} \]
\[ \sigma_c^2(T) = \sigma_d^2(T) = \frac{2c}{\omega^2 T} = \frac{c}{\mu_{uu}(t) T} \]  

(10.25)

Because \( \psi_{uu}(0) = u_0^2/2 \) for \( u = u_0 \sin \omega t \). As in cases discussed in previous chapters the standard deviation \( \sigma \) is inversely proportional to \( \sqrt{T} \) or \( \sqrt{k} \).

If the additive noise \( n(t) \) can be considered as white noise \( w(t) \), processed by low pass filter with transfer function

\[ H(s) = \frac{\alpha}{s + \alpha} \]

then

\[ \phi_{nn}(s) = \frac{c}{\alpha^2 x^2 s^2} \Rightarrow \mu_{nn}(t) = \frac{c}{2} \alpha e^{-\omega_l t} \]  

(10.26)

Substitution of this autocorrelation function in eq. (10.26) leads to:

\[ \sigma_c^2(T) = \frac{c}{\mu_{uu}(t)} \frac{\alpha}{T} \left\{ \frac{\alpha}{\omega_1^2 + \alpha^2} + \frac{2}{T} \frac{\omega_1^2}{(\omega_1^2 + \omega_2^2)(1 - e^{-\omega_1^2 T})} \right\} \]

(10.27a)

\[ \sigma_d^2(T) = \frac{c}{\mu_{uu}(t)} \frac{\alpha}{T} \left\{ \frac{\alpha}{\omega_1^2 + \alpha^2} - \frac{2}{T} \frac{\omega_1^2}{(\omega_1^2 + \omega_2^2)(1 - e^{-\omega_1^2 T})} \right\} \]

(10.27b)

The behaviour for \( T \to \infty \) is found to be

\[ \sigma_c^2(T) \approx \sigma_d^2(T) \approx \frac{c}{\mu_{uu}(t) T} \frac{\alpha^2}{\omega_1^2 + \alpha^2} \]

(10.28)

which for \( \alpha \to \infty \) ("white noise") reduces to eq. (10.25). For white noise, using a certain interval, the variance is independent of \( \omega_1 \); for "red colored" noise that variance improves with increasing \( \omega_1 \); cf. fig. 10.4.

It has been shown by Sins (1967) that the difference between equations (10.27a) and (10.27b) for \( \sigma_c^2 \) and \( \sigma_d^2 \) can be attributed to the particular choice of integration interval, viz. \( (0, T) = (0, kT_1) \). One may as well use \( \left\{ \frac{T_1}{2n}, (k + \frac{1}{2n})T_1 \right\} \). If \( \phi \) is stochastic and uniformly distributed over the interval \( 0 \leq \phi < 2 \pi \) then

\[ \mathcal{E}_\phi [\sigma_c^2(T)] = \mathcal{E}_\phi [\sigma_d^2(T)] \]

Not only \( \sigma_c \) and \( \sigma_d \) are of interest, but also the correlation between the errors in the c and d direction. This can be expressed by

\[ \sigma_{cd}^2(T) = \mathcal{E} \left[ \left( c(\omega_i, T) - c(\omega_i) \right) \left( c(\omega_d, T) - c(\omega_d) \right) \right] \]

(10.29)

It is not difficult to prove that \( \sigma_{cd}^2(T) \equiv 0 \) irrespective of \( T, \omega_i \) and the noise correlation function \( \psi_{nn}(0) \).
As an example, let us consider the measurement of the Nyquist or polar plot of a simple first-order process

\[ H(j\omega) = \frac{\alpha}{j\omega + \beta} \]

under the influence of additive noise given by eq. (10.26); cf. fig. 10.5.

Assume \( w(t) = n(t) \) and \( \beta = \alpha \). Then \( \sigma_c \) and \( \sigma_d \) can be calculated for different values of \( \omega_1 \) which, together with our knowledge \( \sigma_{cd} = 0 \), leads to fig. 10.6. Here the circles around a point indicate the one-sigma-regions for different lengths of integration intervals.

For further aspects of Fourier filtering see: Balchen (1959), Balchen and Blandhol (1962), Mesch (1964), Leonov and Lipatov (1959), Dwyer (1965), Werner (1962), Hennig (1963), Seifert (1962), Lange (1962).

Related procedures (Sins, 1967). Considering the detrimental influence of the additive noise one may ask whether there are better techniques available. In this respect "better" may be defined as: a method giving a smaller variance, using the same observation or integration interval, than the Fourier filter.

One may think of:

- the use of an extra bandpass filter with the Fourier filter in order to reduce the noise signal that interferes with our measurements; cf. fig. 10.7.

Indeed, along these lines the correlation function \( \psi_{nn}(\tau) \) in eq. (10.24) can be influenced in a favourable way. There are, however, several disadvantages, viz.: - the center frequency of the bandpass filter has to be slaved with the measuring frequency \( \omega_1 \); - the adjustment of the bandpass filter with respect to \( \omega_1 \) is quite critical as a slight detuning manifests itself in a phase contribution \( \neq 0 \) which introduces errors in \( c(\omega_1) \) and \( d(\omega_1) \); - when switching the sinusoidal signal to the bandpass filter there is a transient, during which the Fourier filter can not yet be used. This transient time interval has to be added to the integration interval of the Fourier filter.
in order to obtain a fair comparison with the unmodified Fourier filter.

It turns out that these disadvantages outweigh the possible advantages.

- the use of a compensation (model adjustment) method; cf. fig. 10.8. This can be compared with the chapters 7 and 9. In this case, the model is very simple because the only parameters are c and d. Again some type of error criterion containing $e(t)$ has to be minimized. It can be shown that the integral-squared-error-criterion is very closely related to the Fourier filter discussed already.


- the use of samplers. This can be illustrated in the following way. Let

$$y(t) = \frac{c_0}{|H(j\omega)|} \sin \omega t + \varphi(\omega) = \frac{c_0}{|H(j\omega)|} \sin \omega t \cos \varphi(\omega) + \frac{c_0}{|H(j\omega)|} \cos \omega t \sin \varphi(\omega)$$

If $\omega t_i = \frac{\pi}{2} + i \pi$:

$$y(t_i) = \frac{c_0}{|H(j\omega)|} \cos \varphi(\omega) = c_0 \cos(\omega)$$

If $\omega t_j = j \pi$:

$$y(t_j) = \frac{c_0}{|H(j\omega)|} \sin \varphi(\omega) = c_0 \sin(\omega)$$

Consequently, the parameters to be determined can be found by sampling at the instant $t_i$ and $t_j$. Therefore the method is extremely simple; fig. 10.9.

It is clear, however, that this method is very sensitive to the presence of higher harmonics. The influence of additive noise can be diminished again by averaging over a number of samples $t_i \; ; \; i = 0, 1, 2, \ldots, k$ respectively $t_j \; ; \; j = 0, 1, 2, \ldots, k$.

A comparison of the variance obtained by this method with that of Fourier filtering leads to the result that the latter is (much) better. Some improvement can be obtained, e.g., by sampling two times per period ($t_j = j \pi$) and taking care of inverting the proper observations. In spite of such added complications, the conclusion given remains valid; cf. Schussler (1961).
the use of a **band filter**. In order to remove harmonics and noise, the output signal $y(t)$ is filtered by a bandpass filter and then used for making a Lissajous plot on a scope or x-y plotter; cf. fig. 10.10. From this plot, both parameters can be determined as indicated in fig. 10.11. In principle the method is simple but, as with the previous application of a bandpass filter, there are some problems. From the theoretical consideration it follows that the Fourier filter is preferable; cf. Mesch (1964).

- the use of an **optimal filter**. The problem can be described as estimating the unknown amplitude and phase (or real and imaginary component) of a signal of known frequency amidst contaminating noise. There is a multitude of publications on such types of optimal filter problems, that can be solved by using linear filters. From that point of view, the problem is sketched in fig. 10.12, where $g_r(t)$ and $g_i(t)$ indicate the impulse responses of the optimal filters for the real and imaginary component respectively. It turns out that for white additive noise the optimal filters are identical with the Fourier filter.

A disadvantage of the use of sinewaves is the small amount of information that becomes available per frequency of the testsignal. A means of diminishing the drawback of the sinewave testing is the addition of several harmonics into one periodic test signal. In this way by an intelligent choice of phase relations of the sinewaves, the periodic signal can be made binary (i.e., having only amplitude levels $+B$ and $-B$) without much disturbance of the original energy spectrum, cf. Jensen (1959).

If permitted, such a binary signal can easily be applied to many types of linear processes. The process output is filtered with respect to the harmonics which constitute the input signal. In this way, information regarding the amplitude transference of the process for several frequencies is obtained simultaneously. Another approach to determining time dependent transfer-characteristics (human operator) using the sum of a number of sine waves is described by Sheridan (1960).
There still is a need for a theory that enables us to construct on-off signals with a prescribed power spectrum. Trial and error methods, but also optimization procedures with on-off signals, can give good approximations to certain desired power spectra and may give results superior to pseudo-random sequences.

If no a priori information is present, a pseudo-random sequence, sect. 10.3, is very good for obtaining an overall picture of the response so that sharper measuring strategies can be devised. The foundation of optimum test signals is given in decision theory. As more knowledge is obtained from the measurements, sharper measuring strategies can be applied. If this is done with an automatic control system, the term "dual control" is used. (Fel'dbaum, 1960).
10.3 Binary sequences.

In section 8.1 it was shown that "white" noise has desirable properties with respect to the determination of the impulse response of a linear process. Its autocorrelation function, approximating a Dirac "function", eliminates the need for deconvolution; the crosscorrelation function of the process input and the output signal gives us directly the process impulse response function. For sampled signals, the same idea holds; for "white" noise, the matrix \[
\begin{bmatrix}
U'U
\end{bmatrix}
\] in the normal equations (section 6.1) reduces to a diagonal matrix if the observation interval is of sufficient length. This alleviates the need for matrix inversion, corresponding to the deconvolution operation in the case of nonsampled signals. Therefore it is not surprising that a number of research workers have devoted time and energy to find signals with autocorrelation properties corresponding to those of white noise. In spite of the fact that the same type of requirements occurs in the field of radar and sonar signal modulation. (Woodward, 1955) there seems to have been only little interaction between this field and that of process parameter estimation.

The signal or time sequence that we want to construct may be a periodic or an aperiodic one. Examples of the latter type are given by Huffman (1956). In the class of periodic signals, the subclass of binary testsignals shows some interesting engineering properties:
- they can be generated using relay or digital type of instrumentation;
- they reduce multiplication with an analog signal to a relay-type of operation;
- of all signals with amplitude limitation -a≤u≤a the binary signal \[u = \pm a\] transfers the maximum amount of "power" to the process under test.

The following diagram shows some relationships:
The random interval (telegraph) signal may have a Poisson distribution for
the number of switches from one amplitude to the other in a given time
interval. It may be generated by using a radioactive source and a radiation
detector, initiating a change of polarity at each detected particle. Its
autocorrelation function is given by
\[ \psi_{uv}(\tau) = \psi_{uv}(0) \exp\left(-2\lambda\tau\right) \]
where \(\lambda\) is a real constant.

The switching intervals may be discretized by the use of a clock pulse with
period \(\theta\). In that case the change of signal polarity may occur only at the
end of an integer number of clock pulse periods. Such a signal has an auto-
correlation function
\[ \psi_{uv}(\tau) = \begin{cases} \psi_{uv}(0) \left(1 - \frac{\tau}{\theta}\right) & \tau \leq \theta \\ 0 & \tau > \theta \end{cases} \]

This correlation function can arbitrary-closely approximate the Dirac "function".
This type of signal has the added advantage that it can easily be delayed by
using a shift register driven by the same clock signal.
M-Sequences. Due to the random nature of the signals mentioned, one has to face the inherent statistical uncertainty in the correlation operation as explained in section 6.1. Therefore, there is an advantage in using a deterministic signal that still has the desired autocorrelation function; pseudo random binary sequence (P.R.B.S). One of the early papers reporting on process dynamics estimation of this type is the one by Anderson, Buland and Cooper (1959).
Here the generation of the binary test signal and the multiplication with the same time function were effected by using a codedisk.

Soon it was recognized, a.o. by Van Rede and Poortvliet (Poortvliet, 1962), that the so called maximum-length-sequences (m-sequences), generated by fed back shift registers, have desirable autocorrelation characteristics. They have the following properties:

- their autocorrelation function, if determined over an integer number of sequence periods, has no stochastic element (uncertainty) in it;
- they can easily be generated by using digital components; cf. fig. 10-13 as an example;
- moreover the signal, delayed over any integer number of clock signal periods $\theta$, can easily be generated;
- the sequence period $T$ can easily be changed according to $T = (2^n-1)\theta$, where $n$ = number of shift register elements in use. Petersen (1961) gives a tabulation of feedback connections around such shift registers up to $n = 33$;
- the "power" of the sequence is $\psi_{\mu\nu}(0) = a^2$;
- the d.c. level of the sequence is only $a/N$; consequently, for large $N$ the normal operating point of the process need hardly be disturbed;
- the autocorrelation function is shown in fig. 10-14; by choosing $N$ and $\theta$ a Dirac function can be well approximated;
- the signal, consisting of amplitudes $+a$ and $-a$, can easily be applied to a variety of processes. Such signals represent the maximum "power" under a constraint of amplitude.

A disadvantage, due to the periodic character of these signals, is that the correlation functions necessarily also will be periodic.

The best way to illustrate these points is by using the simple example of fig. 10-15. The table shows the "state" of the shift register during the
intervals of time, marked by the clock pulses at \( t = k\theta \). If the state of register element \( \alpha \) is transformed to a testsignal through \( 1 \rightarrow +a \) and \( 0 \rightarrow -a \), then results the time function represented as \( \psi(t) \).

Obviously, \( \psi \) and \( \psi - \alpha \) are delayed one, viz. two clockpulse intervals. By direct multiplication the points of the autocorrelation function for \( r = 10 \) can easily be checked. For the intervals between those points, it is advantageous to prove first, that in each interval the correlation function consists of a linear interpolation between the end points of that interval.

The small negative value of the correlation function is due to the fact that of the \( 2^n \) possible states there is only one not being used, viz. \((0, 0, 0)\). Consequently, in averaging over the \( N = 2^n - 1 \) terms of the sequence there will remain one positive element uncompensated. A remedy may be the choice of slightly different values for positive and negative signal amplitudes, leading to correlation functions as sketched in fig. 10-16.

Another aspect which may be recognized as a disadvantage is the lack of freedom to distribute the signal power over the frequency range. Due to the triangular shape of the autocorrelation function, the power spectrum has an envelope of the form (cf. fig. 10-17):

\[
\theta \left( \frac{\sin \frac{\omega \theta}{2}}{\omega \theta} \right)^2
\]

Due to the periodicity of the correlation function with a period \( T = N\theta \), the frequency spectrum consists of lines at distances \( \frac{2\pi}{T} \).

(It is left to the reader as an exercise to derive \( \Phi_{\psi}(\omega) \) from \( \psi(t) \) either:

- by finding the coefficients of the Fourier series for the periodic function or
- by recognizing that
\[ \psi_{uu}(\tau) = \frac{1}{\theta} f(\tau) * f(\tau) * \sum_{k=-\infty}^{\infty} \delta(\tau-k\tau) \]

\[ f(\tau) = \begin{cases} 1 & |\tau| \leq \frac{\theta}{2} \\ 0 & |\tau| > \frac{\theta}{2} \end{cases} \]

*meaning convolution, which becomes multiplication in the timedomain.

Consequently,

\[ \Phi_{uu}(\omega) = \frac{1}{\theta} F(\omega) \cdot F(\omega) \cdot \sum_{k=-\infty}^{\infty} \delta(\omega-k\frac{2\pi}{T}) \]

cf. fig. 10-18).

Hence one has considerable freedom in determining the bandwidth of the test-signal (by choosing the clockpulse frequency of the shift-register m-sequence generator) and the frequency resolution of its spectrum (by choosing \( T = N\theta \), i.e., the length of and the feedback around the shift register).

The use of an m-sequence as a testsignal can be compared with the use of a large number of sinusoidal components and measuring at all frequencies simultaneously.

Errors when using an m-sequence testsignal. An important reason for choosing the m-sequence was the shape of its autocorrelation function. This may approximate the Dirac "function" of white-noise testsignals, eliminating the need for deconvolution when determining points of a process impulse response. Consequently, we start from the true relationship between the correlation functions

\[ \gamma_{yy}(\tau) = \int_{0}^{\infty} h(t) \psi_{uu}(\tau-t) \, dt \]

and compare the results of our approximation. The following types of error can
be recognized (While, 1968):

a) periodicity of the signal
b) d.c. component of the signal
c) use of a "nonwhite" (bandlimited) signal
d) additive noise
e) transducer and measurement error

a) **periodicity of the signal.** Because the m-sequence is periodic, its autocorrelation function will have the same property. This implies that when determining the impulse response $h(\tau)$ of a process, one has to be sure that $h(\tau) = 0$ for $\tau > T$, the sequence period. If that were not the case, then the effect of overlapping contributions due to different peaks of the $\nu_{uu}(\tau)$ (aliasing) will occur. Consequently, this technique cannot be used on processes including a pure integration.

When using m-sequences as process input signals for least squares estimation one has to keep in mind that in spite of the periodicity of that signal the matrix $[u'u]$ has to be nonsingular, i.e., its columns must not be completely dependent.

Because the expression for the variance of the estimate contains the term this variance can be reduced by either increasing the amplitude of the test-signal or by observing over a large number of periods of the sequence.

b) **d.c. component of the signal.** The d.c. component in the test-signal leads to a constant term in the autocorrelation function $\nu_{uu}(\tau)$, viz. $-\alpha/N$. This will influence the crosscorrelation. To study this influence $\nu_{uu}(\tau)$ is split up into two parts, viz.
triangular functions with area \( \frac{N+1}{2} \alpha^2 \theta \)
a constant term \(- \frac{\alpha^2}{N} \)

The latter contributes to the crosscorrelation function

\[- \frac{\alpha^2}{N} \int_0^{N_0} h(t) dt \]

Provided \( h(t) \to 0 \) for \( t \to T \) this contribution can be recognized from \( y(t) y(t) \)
at large positive \( (t > T) \) or small negative values of \( t \). This provides means to remove this error contribution.

c) use of a "nonwhite" (bandlimited) signal. The use of this testsignal was stimulated by its autocorrelation function, which may approximate a Dirac "function". Whether an approximation is fair enough depends on the situation at hand. As the fine details in a transfer function are caused by the smallest timeconstant \( \tau_{\text{min}} \) of the process under study, it appears that \( \theta \approx \tau_{\text{min}} / 5 \) is an acceptable choice. The amount of error depends of course on the situation at hand, viz. on \( h(t) \) of the process to be studied.

d) additive noise. Due to the deterministic nature of these testsignals, there is no inherent statistical uncertainty. The study of the uncertainty due to additive noise can be done as in previous cases. One finds for the variance in the estimate due to additive noise:

\[ \sigma^2 = \frac{\sigma_n^2}{iN\theta \alpha^2} \]

where \( \sigma_n^2 \) is the variance of the additive noise and \( i \) is the (integer) number of sequence periods over which correlation has been performed (compare this with eq. 6.63). For a discussion of this case cf. Briggs and others (1964).
e) transducer and measurement error. Godfrey and Murgatroyd (1965) have studied the problem of the errors, introduced by the finite response of transducers used for introducing m-sequence signals into processes.

Based on the previous considerations one has to choose the signal parameters $\theta$, $N$ and $a$. Because the signal is binary, the multiplication operation for such correlators can simply be instrumented. Apart from instruments built along these lines in many laboratories and the use of small digital computers, there is at least one such instrument available commercially (automatic servo-tester, Sperry).

An important extension is towards the use of such testsignals for multi-input/ multi-output processes. In the case where there are $p$ inputs and $q$ outputs, then the process can be characterized by $pq$ impulse responses. It is simple to show that the "correlation" technique indicated in the previous lines applies also to this more general case and that the impulse response $h_{ij}$ from input $i$ to output $j$ can be found from

$$\psi_{ij}(\tau) = \int_{-\infty}^{\infty} h_{ij}(\theta) \psi_{ui}(\tau - \theta) d\theta$$

if $u_i$ and the input signals at the other $p-1$ inputs are statistically independent. If this condition is not fulfilled, then the result is a set of simultaneous integral equations in the unknown $h_{ij}$ which is difficult to solve.

The independence criterion for the $u$'s may be relaxed to: orthogonality over the $\tau$ interval of interest (which may be infinite for processes containing a pure integration).
It has been shown by Godfrey and Briggs (1966) that the crosscorrelation between two m sequences with the same $\Theta$ and with periods of $M$ and $N$ (coprime) respectively, can be negligibly small for all shifts only if correlation is performed over the interval $MN\Theta$.

For further information on the use of these testsignals the reader is referred to Nikiforuk and Gupta (1969) and Godfrey (1970).
10.4 Other approaches.

Now consider an engineering situation where the amount of noise can be neglected. In such a case it would be of interest to have a simple and straightforward method of obtaining an estimate of the coefficients of the process differential equation. Such a method is described by Strejc (1961). His method uses parts of the record of process input and output signals. In the following, a brief outline of this method is given.

Assume the differential equation is known to be

\[ a_2 \dot{y}(t) + a_1 \dot{y}(t) + a_0 y(t) = b_2 \ddot{u}(t) + b_1 \dot{u}(t) + u(t) \]  

(10.35)

where \( u \) and \( y \) are the input and the output signal respectively. The part of the recording used is that over the interval \((t_1, t_2)\). The method is based on repeated integration of eq. (10.35) over the interval \((t, t_2)\).

One integration yields:

\[ a_2 \left[ y(t_2) - y(t) \right] + a_1 \left[ y(t) - y(t_1) \right] + a_0 \left[ \int_{t_1}^{t_2} y(t) \, dt \right] = b_2 \left[ \ddot{u}(t_2) - \ddot{u}(t) \right] + b_1 \left[ \dot{u}(t_2) - \dot{u}(t) \right] + \left[ \int_{t_1}^{t_2} \dot{u}(t) \, dt \right] \]  

(10.36)

Integrated again:

\[ a_2 \left[ y(t_2) \{ t_2 - t \} - y(t) + y(t_1) \} + a_1 \left[ y(t) \{ t_2 - t \} - \int_{t}^{t_2} \dot{y}(t) \, dt \right] + \right. \]

\[ + a_0 \left[ \int_{t}^{t_2} y(t) \, dt \right] = \ldots \ldots \]  

(10.37)

and integrated again:

\[ a_2 \left[ \ddot{y}(t_2) \left\{ t_2 \left( t_2 - t \right) - \frac{1}{2} \left( t_2^2 - t^2 \right) \right\} - y(t_2) \left\{ t_2 - t \right\} + \int_{t}^{t_2} \dot{y}(t) \, dt \right] + \right. \]

\[ + a_1 \left[ \ddot{u}(t_2) \left\{ t_2 - t \right\} - \frac{1}{2} \left( t_2^2 - t^2 \right) \right\} - y(t_2) \left\{ t_2 - t \right\} + \int_{t}^{t_2} \dot{u}(t) \, dt \right] + \right. \]

\[ + a_0 \left[ \int_{t}^{t_2} y(t) \, dt \right] = \ldots \ldots \]  

(10.38)

By substituting \( t = t_1 \) in each of these equations the expressions in square brackets can be evaluated for the signal records over \((t_1, t_2)\). This can be done by numerical or graphical way. Consequently, one obtains a set of linear equations in the unknown coefficients:

\[ \sum_{i=0}^{2} a_i \, Y_{ij} = \sum_{i=0}^{2} b_i \, U_{ij} \]  

(10.39)

where the index \( j \) refers to the \( j \)-th equation. One needs the same number of independent equations as the number of unknown coefficients, in our case five.

One may obtain this number by further integrations. In view of accuracy
requirements, however, one would better take more signal recordings and integrate at most two times.

The best way to determine $u(t_1), u(t_2), y(t_1), y(t_2)$ is at extremes or inflection points of the recording, where some of these derivatives are zero. This may be used as a guideline in choosing the interval(s) $(t_1, t_2)$ on the signal records.

A comparatively simple case (Strejc, 1958, 1959) is obtained if the process is described by

$$a_n y^{(n)}(t) + a_{n-1} y^{(n-1)}(t) + \ldots + a_1 y(t) + a_0 y(0) = u(t)$$

(10.40)

and if the input and output are constant for $t < 0$ and for $t \to \infty$ such that

$$y(0) = y'(0) = \ldots = y^{(n-2)}(0) = 0$$

$$y'(\infty) = y''(\infty) = \ldots = y^{(n-2)}(\infty) = 0$$

(10.41)

Example: if in addition $y(0) = y(\infty)$ and $u(0) = u(\infty)$, then by repetitive integration one can show that

$$a_0 = c \int \limits_0^\infty u(t) \, dt$$

$$a_1 = - c \int \limits_0^\infty \left\{ u(t) - a_0 y(t) \right\} \, dt$$

$$a_2 = c \int \limits_0^\infty \left[ \int \limits_t^\infty \left\{ u(\xi) - a_0 y(\xi) \right\} \, d\xi + a_1 y(t) \right] \, dt$$

(10.42)

with $c^{-1} = \int \limits_0^\infty y(t) \, dt$

Due to unavoidable measurement errors, this method is limited to differential equations of the order $\leq 3$. Calculation of the coefficients $a_0, a_1, a_2$ of an $n$-th order differential equation, however, is sufficient to determine an approximation of that dynamics by supposedly $n$ equal time constants. For that situation the following holds in the Laplace notation:

$$a_0 (sT + 1)^n X = a_0 \left\{ s^n T^n + \binom{n}{1} s^{n-1} T^{n-1} + \ldots + \binom{n}{2} s T^2 + \binom{n}{1} s T + 1 \right\} X$$

Consequently:

$$nT = \frac{a_1}{a_0}$$

and

$$\binom{n}{2} T^2 = \frac{a_2}{a_0}$$

(10.43)

from which $n$ and $T$ follow; $n$ has to be an integer number.
References


Dwyer, B. (1965). A feedback transfer function analyser, Control, 676-678.

Fel'dbaum, A.A. (1960). cf. references Chapter I.


Schwarze, G. (1964). Algorithmic determination of the order and time constants of P-, I- and D-action elements with two different time constants and delays up to the 6th order. (in German). Z. Messen, Steuern, Regeln, 7, 10-18.


Additional literature

Impulse and step signals.


El Harès, H. (1965). Determination of the transferfunction based on a

Neue Technik, 7, 395-400.


Sinusoidal test signals


Darovskikh, L.N. (1962). Determination of frequency characteristics with the assistance of the computing unit of the electronic model. *Automn. remote control*, 23.


Isermann, R. (1963). Frequency-response measurement on controlled plants by means of square wave signals. (in German). Regelungstechnik, 3,


Reinisch, K. (1965). A method for the measurement of continuous and sampled control loops including dead time and some nonlinearities. (in German). Habilitationsschrift, Technische Universität, Dresden.


Binary testsignals


Boardman, K.D. (1966). The measurement of the dynamic behaviour of a superheater by random signal testing. IEE Conference on Automatic Control in Electricity Supply, Manchester, 198-


10.4. Other approaches.


$f(t)$

$F(\omega)$

$\text{fig. 10.18}$
Chapter II  BAYES' AND MAXIMUM LIKELIHOOD ESTIMATION.

In chapter 5 the relation between different types of estimates was indicated. It was shown that Markov and least squares estimation have the advantage of requiring little or no a priori knowledge. It was also shown that for additive Gaussian noise these estimates are the same as the maximum likelihood ones.

Depending on the point of view, the strength of the least squares estimator may turn into a weakness if a priori knowledge is available. Intuitively, it is clear that the use of the proper type of a priori knowledge may result in a "better" estimate. In the least-squares model-adjustment approach it is possible to set the initial values of the model parameters according to our prior knowledge or guess. It is, however, not possible in a simple way to bring into the estimation scheme the amount of confidence that we have in our initial value. This is where the Bayes' estimation may be useful. This is discussed in section II.1.

Section II.2 is devoted to the maximum likelihood estimation, which has the advantage of a well-developed theoretical background in statistics. Of great importance is the problem of maximum achievable accuracy. This is also dealt with in section II.2. Some instrumentation schemes for these types of estimates are discussed in II.3.
11.1 Bayes' estimation.

The central aspects of the problem have been discussed in section 5.1. It was shown that the crucial point in the estimation procedure is the determination of the (a posteriori) conditional probability density function \( p(b|y) \), given the observations \( y \).

For a certain sequence of observed process output variables \( y_k = \{y(1), \ldots, y(k)\} \), this probability density function provides all the knowledge that is of interest to the experimenter.

From the development of \( p(b|y_k) \), with \( k \) increasing, one may learn about the characteristics of the estimate, such as bias, consistency, asymptotic normality, cf. fig. 5.2. It would be possible to determine \( p(b|y) \) and then let the experimenter decide on the basis of that knowledge what he considers to be the "best" estimate \( \hat{b} \) of \( b \).

Several attempts to determine the development of conditional probability functions in time can be found in literature. There are approaches along the lines of solving Focker-Planck types of partial differential equations (Stratonovitch, 1962; Kushner, 1964; Bucy, 1965) or expansion of the process output in eigenfunctions (Kroy, 1968).

Objections to this procedure of first determining \( p(b|y) \) and then having the experimenter decide on the "best" estimate may be of the following nature:

- for a situation where \( \mathbf{b} \) is an \( m \)-vector with \( m > 2 \) it is not a simple task to present \( p(b|y) \) to the experimenter in a form suitable for decision;
- "best" has to be defined in some way or another, requiring the choice of a type of cost or loss function.

The choice of such a cost function remains more or less arbitrary; only very few problems clearly suggest a particular choice. Cost functions in connection with probability density functions that are discussed in the literature include:
- $C(\beta, b) = (\beta - b)^2 c$ a quadratic function, leading to the conditional expectation or mean of $p(b | y)$ as in the one dimensional case

$$\min \int c(\beta - b)^2 p(b | y) \, db$$

through differentiation with respect to $\beta$ and equating to zero results in

$$\beta = \int b \, p(b | y) \, db$$

i.e. the mean of the probability density functions.

- $C(\beta, b) = |\beta - b| c$ an absolute value function which results in the median of $p(b | y)$

$$\beta = \max_b p(b | y)$$

i.e., the most probable value, leading to the mode of $p(b | y)$.

In fig. 11.1 this is illustrated for several types of probability density functions. Figure 11.2 indicates these and other types of cost functions.

For symmetric unimodal probability functions the mean, median and mode coincide. For further discussion on the choice of cost functions cf. Deutsch (1965).

In section 5.1 it was shown how the Bayes' estimate $\beta$ depends on the a priori knowledge:

- the conditional probability density function of the output variable : $p(y | b)$

- the (a priori) probability density function of the parameter values $b$ : $q(b)$

- the cost or loss function : $C(\beta, b)$

through the average risk

$$R(\beta) = \mathbb{E}_b \left[ \mathbb{E}_y \left[ C(\beta, b) | y \right] \right] = \int \int C(\beta, b) p(y, b) \, dy \, db$$ (11.1)

which has to be minimized. For a given observation $y$ this reduces to

$$\min_{\beta} \int C(\beta, b) p(b | y) \, db$$ (11.2)

or the necessary condition

$$\frac{\partial}{\partial \beta} \int C(\beta, b) p(b | y) \, db = 0$$ (11.3)
For one variable $\beta = \beta$ and $C(\beta, b) = (\beta - b)^2$ we found

$$\hat{\beta} = \int b \ p(b \mid y) \ db$$

(11.4)

In order to clarify the ideas presented we will consider the estimation of one parameter in the simple situation indicated by fig. 11.3:

$$y = u \ b + n$$

where

- $p_b(b)$ is Gaussian, $E[b] = \mu_b$, $\text{var}[b] = \sigma_b^2$
- $p_n(n)$ is Gaussian, $E[n] = 0$, $\text{cov}[n] = \sigma_n^2 I$

$b$ and $n$ are statistically independent.

From this knowledge we also easily derive the probability density function of $y$:

$$p(y) \ \text{is Gaussian, } E[y] = u \mu_b, \ \text{cov}[y] = \sigma_b^2 u u' \ + \sigma_n^2 I$$

where the expression for the covariance follows from:

$$\text{cov} [y] = E \left[ \{y - u \mu_b \} \{y - u \mu_b \}' \right] =$$

$$= E \left[ \{u (b - \mu) + n \} \{u (b - \mu) + n \}' \right] =$$

$$= E \left[ (b - \mu_b)^2 \right] u u' + E \left[ uu' \right]$$

(11.5)

Our interest is focussed on

$$p(b \mid y) = \frac{p(y \mid b)}{p(y)} = \frac{p_b(b) \ p_n(y - u b)}{p(y)}$$

as with $n = y - u b$

(11.6)

Consequently,

$$p(b \mid y) = C \exp \left[- \frac{1}{2} \left\{ \frac{(b - \mu_b)^2}{\sigma_b^2} + \frac{(y - u b)'(y - u b)}{\sigma_n^2} \right\} \right] \left(y - u \mu_b\right)' \left(\sigma_b^2 u u' + \sigma_n^2 I\right)^{-1}$$
with $C_1$ = normalizing constant.

From this point on, a student may easily spend an excessive amount of time trying to evaluate the integral of eq. (II.4), keeping track of all parts of the exponent and the normalizing constant. The trouble may be drastically reduced by recalling some aspects of the Gaussian probability density function:

\[
p(x) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp \left[ -\frac{1}{2} \left\{ \frac{x-\mu_x}{\sigma_x} \right\}^2 \right]
\]

\[\begin{align*}
\int p(x) \, dx &= 1 \quad \text{and} \quad \int (x-\mu_x) \, p(x) \, dx = 0
\end{align*}\]

As we integrate with respect to $b$ the only thing of importance is to shape the exponent into a square in $b$:

\[
-\frac{1}{2} \left\{ \frac{b^2}{\sigma_b^2} + \frac{u' u}{\sigma_n^2} \right\} - 2 \alpha \left( \frac{\mu_b}{\sigma_b^2} + \frac{u' y}{\sigma_n^2} \right) + \ldots \ldots \ldots \ldots = 0
\]

\[
= -\frac{1}{2} \left\{ \frac{b^2}{\sigma_b^2} + \frac{u' u}{\sigma_n^2} \right\} \left\{ (b-\alpha)^2 + \ldots \ldots \ldots + \ldots \ldots \right\}
\]

(11.8)

with

\[
\alpha = \frac{\mu_b}{\sigma_b^2} + \frac{u' y}{\sigma_n^2}
\]

(11.9)

Consequently

\[
\hat{\beta} = \int b \, p(b|y) \, db = \int (b-\alpha + \alpha) \, p(b|y) \, db = \int (b-\alpha) \, p(b|y) \, db + \alpha \int p(b|y) \, db = \alpha + \alpha
\]

and

\[
\hat{\beta} = \alpha = \frac{\sigma_n^2 \mu_b + \sigma_b^2 u' y}{\sigma_n^2 + \sigma_b^2 u' u}
\]

(11.10)

Let us look at the two limiting cases:

\[
\begin{align*}
\sigma_n^2 &\neq 0 \quad \sigma_b^2 = 0 \quad \text{then} \quad \hat{\beta} = \mu_b \\
\sigma_n^1 &\neq 0 \quad \sigma_b^2 \neq 0 \quad \text{then} \quad \hat{\beta} = \frac{u' y}{u' u}
\end{align*}
\]

In the first case there is no a priori uncertainty about the value of $b$, and our noisy measurements cannot improve on that knowledge. In the second case of noiseless observations we can obtain a perfect estimate from our measurements.
In the general case $\sigma_n^2 \neq 0, \sigma_b^2 \neq 0$, we use both the a priori knowledge and the knowledge obtained by the measurement for estimation. The variances $\sigma_n^2$ and $\sigma_b^2$ determine the relative weight given to each of these pieces of knowledge.

Note that

\[ \mathbb{E} [u' y] = k \Psi_{uy} \]
\[ \mathbb{E} [u' u] = k \Psi_{uu} \]  \hspace{1cm} (11.11)

which implies that the weight given to the estimation measurements increases with the number of signal samples processed. If

\[ \mathbb{E} [b] \neq \mathbb{E} [b] \]

then for $\sigma_n^2 \neq 0, \sigma_b^2 \neq 0$ the estimate is initially biased but becomes asymptotically unbiased.

As an exercise it is recommended to consider the same problem for the case that the noise $\text{n}$ is not white, but is coloured according to a covariance matrix $\text{N}$. Also the reader is advised to derive the equivalent of eq. (11.10) for the case of a parameter vector $\beta$.

A still more general case is given in the interesting attempt of using decision theory to process parameter estimation by Maslov (1963).

The problem discussed is characterized by fig. 11.4 and

- sampled input and output signals;
- both input and output are contaminated with noises $\text{v}$ and $\text{w}$;
- $u, v, w, b$ have known probability density functions. They are stationary and independent;
- the input-output operator is given by

\[ x(k) = F(u, b, k) \]

the way the noises are added to the signal is known. These channels have no memory; $p(z(k) | u(k))$ and $p(y(k) | x(k))$ can be determined.

It is required to determine a "best" estimate of $b$: 
Again statistically the best thing one can do is to determine the a posteriori density

\[ p(\beta \mid z, y) \]  

for all \( \beta \). Taking into account a risk function, a decision rule

\[ \tilde{\beta} = \Gamma (z, y) \]  

has to be constructed such that it fulfills the minimum risk criterion.

This rule may be either deterministic or stochastic. A stochastic rule is necessary to minimize the risk in case of a "game" with an intelligent opponent. For the estimation situation with "nature" as "opponent" a deterministic rule has to be used.

Maslov discusses the cases: 1) where \( z \) and \( y \) are used for estimation, and: 2) where only \( y \) is being used. We restrict our summary to the first case. Now the conditional risk for given observations \( z \) and \( y \) is defined as:

\[ \mathcal{E} (C(\beta, b) \mid z, y) = \int C(\beta, b) p(\beta, b \mid z, y) \, d\Omega (\beta, b) \]  

The average risk for different experiments is given as

\[ R(\beta) = \int C(\beta, b) p(\beta, b \mid z, y) p(z, y) \, d\Omega (\beta, b, z, y) \]  

Using \( \beta \) yes' rule, the independence of the noises and the physical realizability of the process Maslov derives a somewhat more general expression for the average risk. He shows that the optimal estimate \( \tilde{\beta} \) is found by means of a deterministic decision rule. For further details and some examples the reader is referred to the paper cited.
Rather few other publications have appeared on the use of Bayes' method in parameter estimation techniques; Galtieri (1963), Kerr (1970). This is probably due to the computational problems when evaluating the conditional expectations. McGhee and Walford (1968) propose a solution by using a Monte Carlo approach.
11.2 Maximum Likelihood estimation (M.L.E.)

Essential aspects of the M.L.E. have been summarized in sections 2.2 and 5.2. A condition for its application is the availability of a priori knowledge of the probability density function for the observable sample variables \( y \). This functional relationship will be denoted by

\[
p\{y; \theta\}
\]

if \( \theta \) is known and our attention is focussed on \( p(y) \)

\[
L\{c; \theta\}
\]

if \( y \) is known (observed), \( y = c \), and an estimate \( \hat{\theta} \) has to be found.

The M.L. estimates \( \hat{\theta} \) of \( \theta \) are those for which the likelihood function \( L \) is maximized, i.e. for which the necessary condition is fulfilled:

\[
\frac{\partial}{\partial \theta} \ln L\{y; \theta\} \bigg|_{\theta = \hat{\theta}} = 0
\]

(11.17)

One has to find the absolute maximum of \( L \).

A simple example may illustrate this technique. Let \( \theta \) be a parameter to be estimated and \( n(i) \) additive noise samples which are independent and given by

\[
p(n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma} \exp \left[ -\frac{1}{2} \frac{n^2}{\sigma^2} \right]
\]

(11.18)

\[
E[n] = 0 \quad E[n^2] = \sigma^2
\]

The observed sample values are:

\[
y(i) = \theta + n(i)
\]

This leads to a likelihood function:

\[
L\{y; \theta\} = \prod_{i=1}^{k} p\{y(i); \theta\} = \frac{1}{(2\pi)^{\frac{k}{2}} \sigma^k} \prod_{i=1}^{k} \exp \left[ -\frac{1}{2} \frac{(y(i) - \theta)^2}{\sigma^2} \right]
\]

(11.19)

and by taking the logarithm:

\[
\ln L\{y; \theta\} = C - \frac{1}{2\sigma^2} \sum_{i=1}^{k} (y(i) - \theta)^2
\]

(11.20)

Differentiation and equating to zero yields:
or
\[
\hat{\beta} = \frac{1}{k} \sum_{i=1}^{k} y(i) \tag{11.22}
\]
as anticipated for this simple case. Just as an exercise and as a comparison with Bayes' method it pays to rework the example of the previous section as given by fig. 11.3 for the M.L.E.

In section 5.2 the relation between the Markov and the least squares estimators and the M.L. estimate has been indicated.

The achievable accuracy. The accuracy obtainable from M.L.E. in terms of variances and covariances can be derived along the following lines (Kendall and Stuart, 1961). We consider the likelihood function of \( k \) observations in case of a one-dimensional parameter \( b \):

\[
L \{ y(1), \ldots, y(k); b \} \tag{11.23}
\]

and

\[
\int L \, dy(1) \ldots dy(k) = \int L \, dy(k) = 1 \tag{11.24}
\]

Differentiation with respect to \( b \) leads to:

\[
\frac{\partial}{\partial b} \int_L d^k y = 0 \tag{11.25}
\]

which can be rewritten as

\[
\int \left( \frac{L}{\partial b} \right) L \, d^k y = 0 \tag{11.25a}
\]

Differentiating again yields:

\[
\int \left\{ \left( \frac{\partial}{\partial b} \frac{\ln L}{\partial^2} \right)^2 + \frac{\partial}{\partial b} \frac{\partial^2 L}{\partial^2 b} \right\} L \, d^k y = 0 \tag{11.26}
\]
or

\[
\mathbb{E} \left[ \left( \frac{\partial \ln L}{\partial b} \right)^2 \right] = - \mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial b^2} \right] \tag{11.27}
\]

Now we consider \( \hat{\beta} = \hat{\beta}(y) \) as an estimator for \( b \):

\[
\mathbb{E} \left[ \hat{\beta} \right] = \int \hat{\beta} L dy = b + \lambda(b) \tag{11.28}
\]

where \( \lambda(b) \) represents the bias that might be present in the estimate.

Upon differentiation

\[
\int \hat{\beta} \frac{\partial \ln L}{\partial b} L dy = 1 + \frac{d \lambda(b)}{db} \tag{11.29}
\]

or, using eq. (11.29) and (11.28):

\[
\int (\hat{\beta} - f(b)) \frac{\partial \ln L}{\partial b} L dy = 1 + \frac{d \lambda(b)}{db} \tag{11.30}
\]

From the Cauchy-Schwarz inequality this can be written as

\[
\left\{ \int (\hat{\beta} - f(b))^2 L dy \right\} \left\{ \int \left( \frac{\partial \ln L}{\partial b} \right)^2 L dy \right\} \geq (1 + \frac{d \lambda(b)}{db})^2 \tag{11.31}
\]

or using eq. (11.28 and 11.27):

\[
\text{Var} \left[ \hat{\beta} \right] = \mathbb{E} \left[ (\hat{\beta} - f(b))^2 \right] \geq \frac{(1 + \frac{d \lambda(b)}{db})^2}{\mathbb{E} \left[ \left( \frac{\partial \ln L}{\partial b} \right)^2 \right]} = - \frac{(1 + \frac{d \lambda(b)}{db})^2}{\mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial b^2} \right]} \tag{11.32}
\]

This is called the Cramér-Rao (in)equality (Rao, 1945; Cramér, 1946).

For unbiased estimators \( \lambda(b) = 0 \) and:

\[
\mathbb{E} \left[ \hat{\beta} \right] = f(b) = b \quad \text{Var} \left[ \hat{\beta} \right] \geq - \frac{1}{\mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial b^2} \right]} \tag{11.32a}
\]

for biased estimators \( \lambda(b) \neq 0 \) and:

\[
\mathbb{E} \left[ \hat{\beta} \right] = b + \lambda(b) \quad \text{Var} \left[ \hat{\beta} \right] \geq - \frac{(1 + \frac{d \lambda(b)}{db})^2}{\mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial b^2} \right]} \tag{11.32b}
\]

In the more dimensional case of unbiased estimates one finds:

\[
\text{cov} \left[ \hat{\beta} \right] = \mathbb{E} \left[ (\hat{\beta} - b)(\hat{\beta} - b)' \right] \geq \mathbf{J}^{-1} \tag{11.32c}
\]

with

\[
\mathbf{J} = \mathbb{E} \left[ \left( \frac{\partial \ln L}{\partial b} \right)' \left( \frac{\partial \ln L}{\partial b} \right) \right] = - \mathbb{E} \left[ \frac{\partial^2 \ln L}{\partial b \partial b'} \right] \tag{11.33}
\]

\( \mathbf{J} \) is called the information matrix (Fisher).
As an application of the Cramér-Rao inequality consider the case given by

\[ y = \mathbf{b} + \mathbf{n} \]

\[ p(\mathbf{n}) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{N}|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} \mathbf{n}' \mathbf{N}^{-1} \mathbf{n} \right] \]

which leads to the likelihood function

\[ L \{ y; \mathbf{b} \} = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{N}|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} \left( y - \mathbf{u} \right)' \mathbf{N}^{-1} \left( y - \mathbf{u} \right) \right] \]

\[ \ln L \{ y; \mathbf{b} \} = c - \frac{1}{2} \left( y - \mathbf{u} \right)' \mathbf{N}^{-1} \left( y - \mathbf{u} \right) \]

\[ \frac{\partial \ln L}{\partial \mathbf{b}} = \mathbf{u}' \mathbf{N}^{-1} \left( y - \mathbf{u} \right) \]

\[ \frac{\partial}{\partial \mathbf{b}} \left( \frac{\partial \ln L}{\partial \mathbf{b}} \right) = -\mathbf{u}' \mathbf{N}^{-1} \mathbf{u} \]

\[ \mathbf{A} = \mathbf{E} \left[ \mathbf{u}' \mathbf{N}^{-1} \mathbf{u} \right] = \mathbf{u}' \mathbf{N}^{-1} \mathbf{u} \]

Consequently, the covariance is restricted to

\[ \mathbf{cov} \left[ \mathbf{A} \right] \geq \left[ \mathbf{u}' \mathbf{N}^{-1} \mathbf{u} \right]^{-1} \]

A comparison of this expression with equation (6.46), the covariance of the Markov estimator, is interesting. Considering the minimum variance properties of the Markov estimator and the correspondence between Markov and M.L.E. for Gaussian noise, this does not surprise us.

The \( \mathbf{E} [ \cdot ] \) in eq. (11.35) has been dropped on the assumption that \( \mathbf{u} \) can be measured without error. Consequently, \( \mathbf{A} \) is known a posteriori. If one wants to have an a priori guess of \( \mathbf{A} \) then the expectation operator cannot be dropped if \( \mathbf{u} \) is (a priori) a stochastic signal. If \( \mathbf{n} \) is white noise then

\[ \mathbf{N}^{-1} = \mathbf{\frac{1}{\mathbf{c}_n^2}} \mathbf{I} \]

and

\[ \mathbf{A} = \mathbf{\frac{1}{\mathbf{c}_n^2}} \mathbf{E} \left[ \mathbf{u}' \mathbf{u} \right] \]
We may continue this discussion of obtainable accuracy with a simple example viz. the estimation of two parameters $\hat{\beta}_0 = \{\beta_0, \beta_1\}$ for the case where $u_0$ and $u_1$ represent the same sequence of samples as $u_0$, shifted in time over one sampling interval. In that case, for ergodic $u_0$:

$$E[u'u] = k \begin{bmatrix} \Psi_{u0}(o) & \Psi_{u0}(1) \\ \Psi_{u1}(1) & \Psi_{u1}(o) \end{bmatrix} = k \sigma_u^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$  (11.38)

and the minimum covariance is given by

$$\text{cov} [\hat{\beta}] = \frac{\sigma_n^2}{k \sigma_u^2} \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} = C \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$  (11.39)

For $\rho \neq 0$ the off-diagonal terms represent a correlation between the errors in $\beta_0$ and $\beta_1$. By means of a simple transformation one may study the variance of the vector components $^T\alpha$, e.g.

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \beta_0 + \beta_1 \\ -\beta_0 + \beta_1 \end{bmatrix} = \chi$$

Consequently

$$\text{cov} [\hat{\chi}] = \frac{C}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -\rho & 1 \\ 1 & -\rho \end{bmatrix} = C \begin{bmatrix} 1 - \rho & 0 \\ 0 & 1 + \rho \end{bmatrix}$$  (11.40)

Now the off-diagonal terms are zero and the components of $\hat{\chi}$ are not correlated.

This can be illustrated by fig. 11.5, where the so-called concentration ellipsoids are given. These provide us with a pictorial presentation of the accuracy with which the parameters can be determined, taking into account the correlation of the errors. With increasing number of samples $k$, the concentration ellipsoids shrink to the point $(\beta_0, \beta_1)$ as $k$ is contained in the denominator of eq. (11.39).
Some of the most important properties. For the case that $k$, the number of samples available, is constant an unbiased estimate is called efficient if in the Cramer-Rao relation (eq. 11.32) the equal sign holds. The efficiency is defined as:

$$
\eta(\beta) = \frac{\min \text{var} [\hat{\beta}]}{\text{var} [\beta]}
$$

(11.41)

An estimate $\hat{\beta}$ is called sufficient if the following holds:

$$
P \{ y(1), \ldots, y(k) \}; b \} = q(\hat{\beta} ; b) \cdot \mathcal{L} \{ y(1), \ldots, y(k) \}
$$

(11.42)

From Cramer (1946, p 499):

- if an efficient estimate $\hat{\beta}$ of $b$ exists, then the likelihood equation will have an unique solution equal to $\hat{\beta}$;
- if a sufficient estimate $\hat{\beta}$ of $b$ exists, then any solution of the likelihood equation will be a function of $\hat{\beta}$.

Not all maximum-likelihood estimates are either unbiased or efficient.

For the case that $k$, the number of samples available, increases one can study the asymptotic behaviour of $\hat{\beta} \{ y(1), \ldots, y(k) \}$. If $\hat{\beta}$ converges in probability to $b$ as $k \to \infty$ then $\hat{\beta}$ is called a consistent estimate:

$$
\lim_{k \to \infty} P \left[ \left| \hat{\beta} \{ y(1), \ldots, y(k) \} - b \right| \leq \varepsilon \right] = 1
$$

(11.43)

with $\varepsilon$ arbitrarily small. Each M.L.E. is consistent. In Åström and Bohlin (1966), a process is called identifiable if the estimate is consistent. A necessary condition is that the information matrix, associated with the estimation problem, is positive definite.

If the efficiency $\eta(\hat{\beta}) \to 1$ as $k \to \infty$ then the estimate is called asymptotically efficient. It can be proved that the maximum-likelihood estimates are asymptotically efficient. In addition, they have the property of being asymptotically normal, i.e., they approximate the normal distribution with mean $b$ and variance $\sigma^2$.
11.3 Some implementation schemes

An interesting application of maximum-likelihood estimation to process parameter estimation is given by Åström and Böhlín (1966). Their process is characterized by a linear, time invariant difference equation. The process is of the single input, single output type:

\[ A^*(z^{-1})y(k) = B^*(z^{-1})u(k) + \lambda C^*(z^{-1})n(k) \]  

(11.44)

with

- \( z \) - shift operator: \( z^{-1} x(k) = x(k-1) \)
- \( u(k) \) - process input
- \( y(k) \) - process output
- \( n(k) \) - additive noise, independent, with a Gaussian distribution, \( \sigma^2_n = 1 \)
- \( \lambda \) - level of the noise signal.

\[ A^*(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n} \]
\[ B^*(z^{-1}) = b + b_1 z^{-1} + \ldots + b_n z^{-n} \]
\[ C^*(z^{-1}) = 1 + c_1 z^{-1} + \ldots + c_n z^{-n} \]

In fig. 11.6a this relation is represented as a block diagram. The polynomials \( A^* \), \( B^* \), and \( C^* \) are assumed to have no common factors; some coefficients may be zero; the zeroes of \( A^* \) and \( C^* \) all lie inside the unit circle.

As a "model" one may choose the relation

\[ r^*(z^{-1}) e(k) = C^*(z^{-1})y(k) - B^*(z^{-1}) u(k) \]  

(11.45)

which may be represented by fig. 11.6b, c or d. The blocks \( A^*/r^* \) may be recognized as noise-whitening filters.
It is worth noting that the equations 11.44 and 11.45 also can be written in vector form:

\[ a \gamma = b u + \lambda c h \]  
(11.46)

\[ a' \gamma = \beta u + \sigma e \quad \text{or} \quad a' e = a' \gamma - \beta u \]  
(11.47)

In passing it may be mentioned that this scheme includes quite a number of distinguishable problems (Åström) for example:

\[ y(k) = b_0 u(k) + \lambda h(k) \]
regression

\[ a y = \lambda h \]
autoregression

\[ y = \lambda c h \]
estimation of a parameter in a moving average

\[ a y = \lambda c h \]
estimation of a rational power spectrum

\[ a y = b u + h \]
least squares estimation

\[ a y = b u + \lambda c u \]
Markov estimation

\[ a(y - \lambda h) = b u \]
estimation of a noise-free process with measurement errors

Based on the knowledge that the noise \( h \) is Gaussian distributed, the likelihood function for \( e^t = \{e(1), \ldots, e(k)\} \) can be found as:

\[
L(e; u, \lambda, \beta, X, \lambda) = \frac{1}{(2\pi)^{k/2} \lambda^k} e^{- \frac{e'e}{2\lambda^2}}
\]  
(11.48)

or

\[
\ln L = -\frac{k}{2} \ln 2\pi - k \ln \lambda - \frac{e'e}{2\lambda^2}
\]  
(11.49)
The M.L. estimate $\hat{\lambda}$ follows from

$$\frac{\partial \ln L}{\partial \lambda} = \left[ -\frac{k}{\lambda} + \frac{e'e}{\lambda^3} \right] \bigg|_{\lambda = \hat{\lambda}} = 0 \quad \text{or} \quad \hat{\lambda}^2 = \frac{e'e}{k} \quad (11.50)$$

For the following discussion there is an advantage in putting

$$\hat{\theta} = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix}, \quad \hat{e} = \hat{e}(\hat{\theta}) \quad \text{and} \quad \hat{e}'e = 2E(\hat{e})$$

where $\hat{\theta}$ represents the vector of parameters to be determined. From eq. (11.47) and from fig. 11.6 it is clear that $\hat{e}$ is linear in $\hat{\alpha}$ and $\hat{\beta}$ and nonlinear in $\hat{\gamma}$. Consequently, finding the maximum of $L$ or $\ln L$ by differentiation and equating to zero is not necessarily a simple procedure. Åström has chosen the Newton-Raphson procedure:

$$\hat{\theta}(i+1) = \hat{\theta}(i) - \left[ \frac{\partial^2 E(\hat{\theta})}{\partial \hat{\theta}' \partial \hat{\theta}} \right]^{-1} \frac{\partial E(\hat{\theta})}{\partial \hat{\theta}} \hat{\theta} = \hat{\theta}(i) \quad (11.51)$$

For further details the reader is referred to the paper cited.


The maximum likelihood procedure can also be interpreted as finding the coefficients of the prediction model

$$y_{\hat{\lambda}}(k) = \hat{y}(k|k-1) = \frac{B^*(z^{-1})}{\Gamma^*(z^{-1})} u(k) - \frac{A^*(z^{-1}) - \Gamma^*(z^{-1})}{\Gamma^*(z^{-1})} y(k) \quad (11.52)$$
in such a way that the criterion

$$E = \sum_{l=1}^{K} \left\{ y(l) - y_H(l) \right\}^2 = \sum_{l=1}^{K} \varepsilon^2(l) \quad (11.53)$$

is as small as possible.

Notice, that (11.44) can also be written as

$$A^*(z^{-1}) \left[ \frac{1}{C^*(z^{-1})} y(k) \right] = B^*(z^{-1}) \left[ \frac{1}{C^*(z^{-1})} \nu(k) \right] + \lambda \eta(t) \quad (11.54)$$

This means that the maximum likelihood method can also be interpreted as a generalized least squares method where the filter function $G = 1/C$ is determined from the model (11.40); cf. Åström (1970), (1967). The maximum likelihood method has been applied extensively to industrial measurements. See e.g., Åström (1967), (1964), and Gustavsson (1970). In the last mentioned paper, comparisons with other techniques such as correlation methods and generalized least squares are also given. The maximum likelihood method has also been applied to time series analysis (put $B = 0$). The maximum likelihood estimate is a strongly non-linear function of the parameters. Since time series analysis is concerned mostly with quadratic functions, such as covariances and spectral densities, one might expect that the estimates can be expressed as non-linear functions of the sample covariances. Estimates of this nature which are asymptotically equivalent to the maximum likelihood estimates for parametric time series analysis have been given by Zetterberg (1969). The generalization of the maximum likelihood method to the multi-variable case has been done by Woo (1970) and Caines (1970).
11.4 The effect of the process-input signal.

In section 11.2, the information matrix proved to be an important measure for the accuracy obtainable by the estimation method, cf. eq. (11.33). From the equations (11.37) and (11.38) it is clear how for the given process description the process-input signal \( u \) affects that accuracy. The question can be asked as to what conditions \( u \) has to satisfy in order to get consistent estimates. This problem has been studied by Råström and Bohlin (1966), Råström (1968), Aoki and Staley (1969a). It follows that for those conditions, the limits

\[
\bar{u} = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} u(i)
\]

and

\[
R_{uu}(j) = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} \left\{ u(i) - \bar{u} \right\} \left\{ u(i+j) - \bar{u} \right\}
\]  

(11.55)

have to exist and the matrix \( A_n \) defined by

\[
A_n = \begin{bmatrix}
R_{uu}(j-\theta) & \ldots & 0 \\
R_{uu}(j-\theta) & \ldots & 0 \\
\vdots & \ddots & \vdots \\
0 & \ldots & R_{uu}(j-\theta) 
\end{bmatrix}
\]

has to be positive definite. These conditions are called persistent excitation (of the order \( n \)).

One might perhaps dare to conjecture that a condition of this nature will be required in general.

Notice, that if the "mean square fourier transform" of the input, i.e.,

\[
F_u(\omega) = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} \left\{ u(i) - \bar{u} \right\} e^{j\omega i}
\]  

(11.57)

exists, then

\[
R_{uu}(\tau) = \int \left| F_u(\omega) \right|^2 e^{j\omega \tau} d\omega
\]  

(11.58)
and the matrix \( A_n \) given by (11.52) is automatically nonnegative definite for arbitrary \( n \). Moreover, if

\[
\hat{f}_u(\omega) > 0
\]  

(11.59)

the matrix \( A_n \) is also positive definite for arbitrary \( n \). The condition (11.59) is thus sufficient to guarantee that an input signal is persistently exciting of arbitrary order. Notice as well that if the input signal is an ergodic stationary stochastic process then its spectral density is given by

\[
\phi_{uu}(\omega) = |\hat{f}_u(\omega)|^2
\]  

(11.60)

The condition (11.59) then implies that the spectral density of the input signal does not vanish for any \( \omega \), a condition which is well known for correlation analysis.

Apart from persistent excitation many applications will require that the output is kept within specified limits during the experiment. The problem of designing input signals, energy and time constrained, which are optimal e.g., in the sense that they minimize the variances of the estimates, has been discussed by Levadi (1966), Aoki and Staley (1969a), (1969b), and Nahi (1969). The same problem is also discussed in Rault et al. (1969). It is closely related to the problem of optimal signal design in communication theory; see e.g., Middleton (1960).


Bucy, R.S. (1965). Nonlinear filtering theory. IEEE Trans. autom. control, AC-12, 198-


Fig. II.1

Fig. II.3

Fig. II.4
\[ \rho = 0 \quad \rho > 0 \]

```
\includegraphics[width=\textwidth]{fig115.png}
```

```
\includegraphics[width=\textwidth]{fig116.png}
```
In chapter 1 it was mentioned that the problem of parameter estimation is closely related with that of optimal filtering and state estimation. In the next chapter methods will be discussed which provide parameter and state estimates together. In this chapter we summarize some state estimation (Kalman filter) methods in order to put the next chapter in the proper perspective. Data processing schemes which separate desired signals from unwanted disturbances (noise) are referred to as filters.

The Kalman filter can be described as a method (implementation, computer program) for combining measurement data, that are contaminated with noise; they have to be combined in such a way that an "optimal" estimate will be obtained of some desired variable.

It can be applied for different types of problems:
- filtering; interpolation, smoothing, extrapolation (prediction) of time series or signals; cf. sect. 12.1.
- state estimation as needed for implementation of optimal-control algorithms in cases where not all, or even no elements of the process state vector can be measured directly.
- combination of measurement data on some desired variable(s), provided by different measuring devices, each of which has its own type of errors (e.g. in navigation problems: gyroscopes, accelerometers, doppler radar, etc.).
- parameter estimation, by restating the problem with a parameter in stead of a state vector.
For a highly interesting discussion of the historic perspective of these techniques the reader is referred to Sorenson (1970), who traces back the use of the least-squares method to applications made by K.F. Gauss (1795) and shows its intimate relationship with Kalman filtering.


During the second world war Wiener did his fundamental and, by now, classical studies on interpolation, extrapolation and smoothing of time series (Wiener, 1949). For our purpose we can restrict the attention to the situation sketched in fig. 12.1.

<table>
<thead>
<tr>
<th>continuous case</th>
<th>sampled case</th>
</tr>
</thead>
<tbody>
<tr>
<td>x = information carrying signal</td>
<td>x(t)</td>
</tr>
<tr>
<td>n = disturbing signal (observation noise)</td>
<td>n(t)</td>
</tr>
<tr>
<td>x_d = desired signal: interpolated</td>
<td>x(t-τ)</td>
</tr>
<tr>
<td>filtered</td>
<td>x(t)</td>
</tr>
<tr>
<td>extrapolated, predicted</td>
<td>x(t+τ)</td>
</tr>
<tr>
<td>ξ &gt; 0</td>
<td>m &gt; 0</td>
</tr>
</tbody>
</table>

\[ \hat{x}_d = \text{estimation of } x_d, \text{ obtained by means of the filter } F \]

\[ \hat{x}_d(t) \]

\[ \hat{x}_d(k) \]

\[ e = x_d - \hat{x}_d = \text{error signal that, by the choice of } F, \text{ has to be minimized in the least-squares sense.} \]

\[ e(t) \]

\[ e(k) \]

Of course the choice of F has to be based on the a priori knowledge about both x and n.
Wiener assumed the signals $x(t)$ and $n(t)$ to be stationary, i.e. their statistical properties do not change in time. The filter $F$ is restricted to be linear, time invariant and physically realizable. Consequently it can be characterized by an impulse response $h(t)$. Assuming that $x(t)$ and $n(t)$ are sample functions of an ergodic stochastic process, a suitable criterion of "optimality" is $\min E[e^{2\eta}]$. For $x_d(t) = x(t)$ and $\hat{x}_d(t) = \hat{x}(t)$ this results in a minimization of

$$E[e^{2\eta}] = E[\{\hat{x}(t) - x(t)\}^2]$$

with

$$\hat{x}(t) = \int_{-\infty}^{\infty} h(\theta)x(t-\theta) \, d\theta$$

(12.1)

The minimization proceeds along the same lines as given before in section 8.11 and it results in the necessary relationship expressed by the Fredholm integral equation:

$$\psi_{xy}(t) = \int_{-\infty}^{\infty} h(\theta) \psi_{xy}(t-\theta) \, d\theta$$

(12.2)

For physical realizability of the filter one has to restrict the impulse response to $h(t) = 0$ for $t < 0$. With this condition we have again the Wiener-Hopf (W.H) equation. The solution of this equation for $h(\theta)$ is generally not simple. Mostly it is done by transformation to the frequency domain as discussed before.

The Wiener filter thus found is optimal for signals that have a Gaussian distribution. For non-Gaussian signals the Wiener filter is the best linear filter.

12.2 The Kalman-Bucy filter.

The Kalman (or K.-Bucy) filter provides the solution for the same problem, described in terms of state variables. Unlike the Wiener type of description it can easily be adapted to the case of non-stationary signals.
Usually the resulting filter is a time-varying one. In addition the K.-B. filter has a recursive form that is well suited for programming on a digital computer. The new estimate is found by updating the old one based on new observations; there is no need to store the whole string of observation data, consequently the storage and computing requirements are modest.

Even if the interest is focussed purely on signals it pays to assume these (stochastic) signals as being derived from a "white" noise by means of a so called "shaping filter". In the sequel also such a filter will be referred to as "process".

In the general case the dynamic behaviour of a process can be given by:

\[ \dot{x} = f(x, u, v, t) \]  
\[ y = g(x, u, u, t) \]

(12.3) (12.4)

cf. fig. 12.2. In the case of a linear(ized) process this reduces to

\[ \dot{x} = A \dot{x} + B u + v \]
\[ y = C \dot{x} + D u + n \]

cf. fig. 12.3. The functional relationship between the state variables and the other "signals" are assumed to be known, including the numerical values of the parameters. Based on the observation of \( y \) over a time interval \((0, T)\) we want to find some estimate for \( x(t) \), denoted by \( \hat{x}(t|T) \). As before there are three cases:

\[ t > T \] prediction problem
\[ t = T \] filtering problem
\[ t < T \] smoothing problem

In line with chapter 5 we can approach the problem by starting from the minimum risk or Bayes estimation, the maximum likelihood estimation or the least squares estimation. For this survey we will take the first one as being the most general.
Again from the statistical point of view the crucial information is provided by the conditional probability density function

\[ p(x \mid y) \]

where \( y \) are the observed sample values of the process output. As in chapter 5 from this probability function one can derive different estimates. Using Bayes' rule

\[
p(x \mid y) = \frac{p(x, y)}{p(y)} = \frac{p(y \mid x) p(x)}{p(y)} \tag{12.5}
\]

one can use the a priori knowledge about \( x \) and, after the observation of \( y \), come up with a new a posteriori probability density function for \( x \). This leads quite naturally to an iterative estimation procedure, in which the outcome of the previous calculation is the a priori knowledge for the next calculation, using the new observations.

First we will consider the simple one-stage filter and after that the multistage version.

One-stage filter. Given is the following a priori knowledge

process \( y = Cx + n \), cf. fig. 12.4:

\[
p(x) \quad \text{Gaussian; } \mathcal{E}[x] = x_0 ; \quad \text{cov}[x] = P_0
\]

\[
p(n) \quad \text{Gaussian; } \mathcal{E}[n] = 0 ; \quad \text{cov}[n] = N
\]

\( x \) and \( n \) statistically independent.

From this knowledge we can easily derive the probability density function of \( y \), although actually we will not need it:

\[
p(y) \quad \text{Gaussian; } \mathcal{E}[y] = Cy_0 ; \quad \text{cov}[y] = C P_0 C' + N
\]

as

\[
\text{cov}[y] = \mathcal{E} \left[ \left\{ y - Cy_0 \right\} \left\{ y - Cy_0 \right\}' \right] = \\
= \mathcal{E} \left[ \left\{ C (x - x_0) + n \right\} \left\{ C (x - x_0) + n \right\}' \right]
\]

Now we want to derive an expression for \( p(x \mid y) \):

\[
p(x \mid y) = \frac{p(x, y)}{p(y)} = \frac{p(x, n) \frac{\partial p_n}{\partial y}}{p(y)} = \\
= \frac{p(x)p(n)}{p(y)} = \frac{p(x)p_n(y - Cy)}{p(y)} \tag{12.6}
\]
The expressions for each of the probability density functions can be substituted. Then one has to decide on what kind of estimate one wants by choosing a cost function. It has been pointed out in chapter 5 already that in the case of a Gaussian distribution the use of any even cost function and the maximization of the density function leads to the same estimate. We will chose the last approach.

In order to save as much manipulation as possible it is worth while to note that \( p(y) \) is known completely from our a priori knowledge. Consequently for our purpose it is considered as being part of the normalizing constant only and

\[
p(x|y) = \text{const.} \exp \left[ -\frac{1}{2} \left( (x-x_0)^T P_0^{-1} (x-x_0) + \frac{1}{2} (y-Cx)^T N^{-1} (y-Cx) \right) \right]
\] (12.7)

This expression has to be maximized with respect to \( x \), which can be done by completing the square of the exponent; c.f. Aoki (1967). Another way is by differentiating \( \ln p(x|y) \) and equating to 0 for the optimal estimate:

\[
\left\{ P_0^{-1} (x-x_0) - C' N^{-1} (y-Cx) \right\} \bigg|_{x = \hat{x}} = 0
\] (12.8)

or

\[
\left[ P_0^{-1} + C' N^{-1} C \right] \hat{x} = P_0^{-1} x_0 + C' N^{-1} y
\] (12.9)

or the following expressions:

\[
\hat{x} = P_0^{-1} x_0 + P_0 C' N^{-1} y
\] (12.10a)

\[
\hat{x} = x_0 + P_0 C' N^{-1} \left\{ y - C x_0 \right\}
\] (12.10b)

\[
\hat{x} = x_0 + P_0 C' \left[ C P_0 C' + N \right]^{-1} \left\{ y - C x_0 \right\}
\] (12.10c)

where

\[
P_0 = P_0 - P_0 C' \left[ C P_0 C' + N \right]^{-1} C P_0
\] (12.11)

These expressions follow from application of the matrix identities summarized in Appendix C. Eq. (12.10) can be represented by fig. 12.5.

Some remarks are in order here:

- Note that \( y \) are the only observations; all other quantities are given a priori. The "gain factor" \( \Gamma = P_0 C' \left[ C P_0 C' + N \right]^{-1} \) does not depend on the observation and can be determined beforehand.
- It is easily shown that $\mathcal{E}[\hat{x}] = x_0$ and

$$\mathcal{E}[(\hat{x} - x_0)(\hat{x} - x_0)'] = P_o C' \left[ C P_o C' + N \right]^{-1} C P_o$$  \hspace{1cm} (12.12)

which suggests an interpretation of eq. (12.11).

- Some more insight is provided if we simplify these equations to a one dimensional case: $x = x, y = y$ such that $P, N$ and $C$ become scalars:

If $C = 1$

$$\hat{x} = x_0 + \frac{P_o}{P_o + N} \left\{ y - x_0 \right\}$$  \hspace{1cm} (12.13)

$$P_i = P_o - \frac{P_o^2}{P_o + N} = \frac{P_o N}{P_o + N} \text{ or } \frac{1}{P_i} = \frac{1}{P_o} + \frac{1}{N}$$  \hspace{1cm} (12.14)

The terms $P_o$ and $N$ remind us of signal and noise powers in communication problems. If $N$ is comparatively large then $y - x_0$ has little influence.

- Some other simplifications that may aid to the insight:

$$C = I \quad \hat{x} = x_0 + P_o \left[ P_o + N \right]^{-1} \left\{ y - x_0 \right\}$$

or, if $C^{-1}$ exists:

$$P_i = P_o - P_o \left[ P_o + N \right]^{-1} P_o$$  \hspace{1cm} (12.15)

$$N = 0 \quad \hat{x} = x_0 + C^{-1} C P_o C' \left[ C P_o C' \right]^{-1} \left\{ y - C x_0 \right\} = x_0 + C^{-1} \left\{ y - C x_0 \right\} = C^{-1} y$$  \hspace{1cm} (12.16)

**Multi-stage filter.** Given the following a priori knowledge:

$$\begin{align*}
\chi(k+1) &= A \chi(k) + \nu(k) \\
y(k) &= C \chi(k) + \eta(k)
\end{align*}$$  \hspace{1cm} (12.17)

The notation $\hat{x}(k|k)$ apparently refers to the estimate of $x(k)$, based on the $k$ observations $\chi(1), \ldots, \chi(k)$. This type of notation can also be used for the prediction and the smoothing problem. As our attention is limited to the filtering problem it suffices to simplify the notation to $\hat{x}(k)$. 

The notation $\mathcal{P}(x|y)$ apparently refers to the estimate of $x(k)$, based on the $k$ observations $\chi(1), \ldots, \chi(k)$. This type of notation can also be used for the prediction and the smoothing problem. As our attention is limited to the filtering problem it suffices to simplify the notation to $\hat{x}(k)$. 

The notation $\mathcal{P}(x|y)$ apparently refers to the estimate of $x(k)$, based on the $k$ observations $\chi(1), \ldots, \chi(k)$. This type of notation can also be used for the prediction and the smoothing problem. As our attention is limited to the filtering problem it suffices to simplify the notation to $\hat{x}(k)$. 

The notation $\mathcal{P}(x|y)$ apparently refers to the estimate of $x(k)$, based on the $k$ observations $\chi(1), \ldots, \chi(k)$. This type of notation can also be used for the prediction and the smoothing problem. As our attention is limited to the filtering problem it suffices to simplify the notation to $\hat{x}(k)$.
From this follows:
\[
p \{ x(k+1) \mid Y_k \} \quad \mathcal{E}[x(k+1) \mid Y_k] = A \hat{x}(k) \quad \text{cov}[x(k+1) \mid Y_k] = \begin{bmatrix} A P_k & A' & Y \end{bmatrix} = \begin{bmatrix} C & N' & C' + N \end{bmatrix}.
\]

Also we can derive the probability density function for \( x(k+1) \) but, as before, we will not need it:
\[
p \{ \gamma(k+1) \mid Y_k \} \quad \mathcal{E}[\gamma(k+1) \mid Y_k] = C A \hat{x}(k) \quad \text{cov}[\gamma(k+1) \mid Y_k] = \begin{bmatrix} C & Q_{k+1} & C' + N \end{bmatrix}.
\]

Due to the dependence on \( Y_k \) the formulation of Bayes' rule needed is a little more general:
\[
p(a, b \mid c) p(c) = p(a, b, c) = p(a \mid b, c) p(b \mid c) p(c)
\]
from which follows:
\[
p(a \mid b, c) = \frac{p(a, b \mid c)}{p(b \mid c)}
\]
Hence we notice a direct parallel with the previous case:

one stage filter \[ p(x \mid Y) \quad \text{multistage filter} \quad p\{ x(k+1) \mid \gamma(k+1), Y_k \} \]

and for Gaussian distributions:
\[
p\{ x(k+1) \mid \gamma(k+1), Y_k \} = \text{const.} \exp \left\{ \frac{-1}{2} \left[ \| x(k+1) - A \hat{x}(k) \|^2_{Q_{k+1}^{-1}} + \| \gamma(k+1) - C x(k+1) \|^2_{N'^{-1}} \right] \right\}
\]
Differentiation of the logarithm of this function and equating it to 0 for the optimal estimate \( \hat{x}(k+1) \) yields:
\[
Q_{k+1}^{-1} \left[ \hat{x}(k+1) - A \hat{x}(k) \right] - C' N'^{-1} \left[ \gamma(k+1) - C \hat{x}(k+1) \right] = 0
\]
or
\[
\begin{bmatrix} Q_{k+1}^{-1} + C' N'^{-1} C' \\ C' N'^{-1} \end{bmatrix} \hat{x}(k+1) = Q_{k+1}^{-1} A \hat{x}(k) + C' N'^{-1} C \gamma(k+1)
\]
or
\[
\hat{x}(k+1) = \Delta \hat{x}(k) + Q_{k+1} C' [C Q_{k+1} C' + N]^{-1} [y(k+1) - CA \hat{x}(k)]
\]
which is closely parallel to eq. (12.10c). Both process and filter are shown in fig. 12.6. Note the similarity of process and part of the filter.

If a measurable input vector \( y(k) \) is added:
\[
\begin{align*}
\hat{x}(k+1) &= A \hat{x}(k) + B u(k) + \nu(k) \\
y(k) &= C \hat{x}(k) + \omega(k)
\end{align*}
\]

one finds in exactly the same way:
\[
\begin{align*}
\hat{x}(k+1) &= A \hat{x}(k) + B u(k) + \Gamma(k+1) [y(k+1) - CA \hat{x}(k) - CB u(k)] \\
\Gamma(k+1) &= Q_{k+1} C' [C Q_{k+1} C' + N]^{-1} = P_{k+1} C' N^{-1} \\
Q_{k+1} &= A P_k A' + \Gamma \\
P_{k+1} &= [Q_{k+1}^{-1} + C' N^{-1} C]^{-1} = Q_{k+1} - Q_{k+1} C' [C Q_{k+1} C' + N]^{-1} C Q_{k+1}
\end{align*}
\]

where \( \hat{x}(0) \) and \( P_0 \) are given.

In this scheme one may recognize:
\[
\begin{align*}
A \hat{x}(k) + B u(k) & \quad \text{extrapolated state} \\
C A \hat{x}(k) + C B u(k) & \quad \text{observation of extrapolated state} \\
Q_{k+1} & \quad \text{a priori covariance matrix of} \ \hat{x}(k+1) \ \text{based on} \ k \ \text{observations} \\
P_{k+1} & \quad \text{a posteriori covariance matrix of} \ \hat{x}(k+1) \ \text{based on} \ k+1 \ \text{observations}.
\end{align*}
\]

Some remarks, closely parallel to those made earlier, include:
- Note that the "gain factor" \( \Gamma(k+1) \) does not depend on the observation and can be determined beforehand for the whole estimation procedure.
- For a one dimensional case and \( C = 1 \)
\[
\begin{align*}
P_{k+1} &= P_{k+1} = \frac{Q_{k+1} N}{Q_{k+1} + N} \\
\Gamma(k+1) &= \Gamma(k+1) = \frac{P_{k+1}}{N} = \frac{Q_{k+1}}{Q_{k+1} + N}
\end{align*}
\]
- Note also that a time dependence of e.g. matrix \( \Delta \) does not change the essence of the problem, at least if this time dependence is known completely.
- A comparison of these recursive equations for state estimation, representing the K.-B. filter, with those derived in section 7.1 for the recursive technique for parameter estimation will indicate a close correspondence. A study of the relation between those techniques can be found in Ho (1962).

- The estimate is linear in the observations. This is due to the assumption of Gaussian-distributed variables. For those cases this filter is optimal in the least squares sense. For the non-Gaussian situation it provides the optimal-linear filters.

In stead of the derivation along the lines of Bayes' estimation as given before, it is also possible to obtain those filter equations along different lines. Consider the simplest case. Assume the observations are given by

\[ y(k) = x + n(k) \]

An estimator, linear in the observations \( y(1), \ldots, y(k) \) has to be found:

\[ \hat{x} = q'y = q'_y y(1) + \ldots + q'_y y(k) \]

such that \( E = \mathbb{E}[\epsilon^2] = \mathbb{E}[(x - \hat{x})^2] \) is minimized. By differentiation of \( E \) with respect to \( q'_y \), one obtains the necessary conditions:

\[
\frac{\partial E}{\partial q'_y} = -2 \mathbb{E}[(x - \hat{x}) y(i)] = 0 \quad i = 1, \ldots, k \quad (12.31)
\]

These equations formulate the so called orthogonality conditions indicating that \( (x - \hat{x}) \), the error in the estimate, has to be orthogonal to the observations. From the set of equations (12.31) one can determine the coefficients \( q'_1, \ldots, q'_k \), which constitute the filter. In a straightforward way these orthogonality conditions can be derived for more general and complex situations. Based on these expressions equations for the different filtering jobs can be derived.

Still another way to look at the filter equations is by recognizing the close parallel between the process and the filter (model) structure. This parallel had already been recognized with respect to Wiener filtering (e.g. cf. Smith, 1958). It has been developed into a theory on "observers" (Luenberger, 1966).
In the previous sections state estimation procedures have been considered for simple cases: linear process, Gaussian noise. For a non-linear process the techniques remain valid through linearizing around a known approximate solution of the process equations. This may be indicated by equations (12.3) and (12.4), where the latter expression can be used in stead of the non-linear one if the matrices $A$, $B$, $C$ and $D$ consist of the proper partial derivatives taken with the approximate solution in mind.

Another approach is called extended Kalman filter. In this method the linearizing is being done around the best estimate $\hat{x}$ of the state vector (and the measurable input vector $u$ if present). If the process is described by e.g.:

$$
\begin{align*}
\dot{x}(k+1) &= f[x(k), k] + v(k) \\
y(k) &= g[x(k), k] + u(k)
\end{align*}
$$

then the filter equations can be:

$$
\hat{x}(k+1|k) = \hat{x}(k+1|k) + \Sigma(k+1) \left[ y(k+1) - g[\hat{x}(k+1|k), k] \right]
$$

where, as before,

$$
\hat{x}(k+1|k) = f[\hat{x}(k|k), k]
$$

is the best estimate of $x(k+1)$ based on $k$ observations.

The close relationship between recursive parameter- and state estimation has been noted before. For a process that is linear in the parameters the correspondence is easily illustrated. Assume the process is described by:

$$
\begin{align*}
y(k) + a_1 y(k-1) + \ldots + a_n y(k-n) &= \\
&= b_1 u(k-1) + \ldots + b_n u(k-n) + n(k)
\end{align*}
$$

Consequently the process and observation equations are:

$$
\begin{align*}
\dot{x}(k+1) &= x(k) \\
y(k) &= c' x(k) + n(k)
\end{align*}
$$
with:

\[
\begin{bmatrix}
    a_1 \\
    \vdots \\
    a_n \\
    b_1 \\
    \vdots \\
    b_n
\end{bmatrix}
\]

\[
\begin{bmatrix}
    -y(k-1) \\
    \vdots \\
    -y(k-n) \\
    u(k-1) \\
    \vdots \\
    u(k-n)
\end{bmatrix}
\]

In this way the parameter estimation problem can be solved along the lines of the filter algorithms given in this chapter.

12.3 Further aspects, applications and examples

There remain many aspects that should be discussed, e.g.,
- the possibilities of divergence of the estimation results, i.e.
  occurrence of a covariance matrix which is much smaller than the actual errors in the estimate and its possible causes;
- the investigation of the residuals

\[
y(k+1) - \mathcal{L} \hat{x}(k)
\]

which should have the statistical characteristics of a white-noise sequence and which may be used to verify the a priori knowledge, e.g. about the level of the observation noise;
- the case of non-white observation noise which can be taken care off by including the noise colouring filter into the process model;
- the case of correlated signal and noise which also can be treated by extending the process model;
- the case of more than one sampling rate for the different input and output signals and/or for the estimation sequence;
- the effect of inaccurate or incomplete a priori knowledge; the possibilities
of constructing adaptive filters for those cases;
the case of processes with continuous-time signals, which can be described by
the following equations:

process:
\[
\begin{align*}
\dot{x} &= A x + B u \\
y &= C x + n \\
E[u] &= 0 \\
E[n] &= 0 \\
E[u(t) u'(\theta)] &= \Pi \delta(t-\theta) \\
E[n(t) n'(\theta)] &= \mathcal{N} \delta(t-\theta) \\
E[u(t) n'(\theta)] &= 0
\end{align*}
\]

filter:
\[
\begin{align*}
\dot{\hat{x}} &= A \hat{x} + \Gamma [y - C \hat{x}] \\
\Gamma &= P C' N^{-1} \\
\dot{P} &= A P + PA' - P C' N^{-1} C P + B U B' \\
\dot{\hat{x}}(0) &= \hat{x}_0 \\
\dot{P}(0) &= P_0
\end{align*}
\]

In view of the limited goal of this chapter, to be an introductory review for
the next one, the reader is referred to the references given, for more details
about these filter procedures.

Some papers on applications of the K.B.-filter in navigation can be found in
Richman and Thau (1966), Schmidt et al. (1968), Toda et al. (1969), Leondes

Applications in communication are given in Irwin and O'Neal (1968), Snijder
(1969). References to applications in industrial control and with respect to
biological objects can be found in chapter 14.
As illustrations of these filter algorithms a few elementary examples are given in the figures 12.7 - 12.9.

These figures are taken from a simulation on a PDP-8 computer by Mr. G.J. Schoenmakers at the Technical University, Eindhoven, Netherlands.

Figure 12.7 represents a static case, i.e. the estimation $\hat{x}(k)$ of a constant signal from observed samples that are contaminated with "white" noise. The pictures illustrate the influence of (assumed) a priori knowledge about the quantity to be estimated. In a) and b) it is assumed that $\hat{x}(0) = 0$, which is not correct, but with a large value of $P(0)$, expressing a large degree of uncertainty with respect to that value of $\hat{x}(0)$. In a) seven runs are given for different white noise sequences, in b) the development of the standard deviation around the expected value of $\hat{x}(k)$ is given.

In c) and d) also $\hat{x}(0) = 0$ is chosen, but now with a small value of $P(0)$, i.e. an unjustified certainty about the $x(0)$ chosen. Again c) shows seven runs and in d) the development of the standard deviation around the expected value of $\hat{x}(k)$ is given. Note that the estimate approaches its true value slowly due to the wrong values that were assumed for $\hat{x}(0)$ and $P(0)$.

In e) and f) the value of $\hat{x}(0)$ was chosen as the right value with a small $P(0)$; e) gives seven runs, f) shows the standard deviation around the expected value.

In g) the pictures a) and b) are combined.

Figures 12.8 and 12.9 show a one dimensional dynamic case; the estimation of a low frequency signal from observations that are also contaminated with white noise.
Figure 12.8  a) shows the process noise u(k) and the resulting signal x(k), that has to be estimated.
   b) shows both x(k) and its estimate $\hat{x}(k)$.
   c) shows the error $\hat{x}(k)-x(k)$ and x(k), the signal to be estimated.
   d) shows the error $\hat{x}(k)-x(k)$ and the additive noise that was used in this example.

Figure 12.9 shows the influence of the (assumed) a priori (un)certainty about x(0).

Figure 12.10 represents again a static one, but now for non-sampled signals. The upper diagram represents the filtering by an ordinary simple (first order, constant coefficient) low pass filter, the lower diagram gives the output of a (first order) K.B. filter. Both filters have the same input. The K.B. filter has a time-varying gain which is proportional to $t^{-1}$. Note that in this case the variance of the output diminishes whereas for the constant-coefficient filter the variance tends to a constant value.


Ho, V.C. (1962). The methods of least squares and optimal filtering. Memo 3329-PR, RAND Corp

Irwin, J.D. and J.B. O'Neal (1968) The design of optimum dpcm encoding systems via the Kalman predictor. Preprints JACC, Ann Arbor, Mich; IEEE, 130-136


Schmidt, S.F. et al. (1968). Case study of Kalman filtering in the C-5 aircraft navigation system. Case Studies in System Control workshop JACC, IEEE group on autom. control


Toda, N.F. et al (1969). Region of Kalman filter convergence for several autonomous navigation modes. AIEEE Journ., 7, nr. 4

Additional literature


Fig. 12.1

Fig. 12.2

Fig. 12.3

Fig. 12.4
Fig. 12.5

Fig. 12.6
$u(k)$

$x(k)$, signal to be estimated

$x(k)$

$\hat{x}(k)$

$\hat{x}(k) - x(k)$; error

$x(k)$

$\hat{x}(k) - x(k)$

$n(k)$

Many aspects of modern control theory are centered around the notion of state. For applications of these theories one has to have available, explicitly or implicitly, a measurement or estimate of the state variable. This can be found along the lines given in the previous chapter if both the structure and the parameters of the process are known. If the parameters are not known then in many schemes a combination of parameter and state estimation is needed.

A straightforward approach advocated in a number of papers is the extension of the state vector with the unknown parameters. As before the process is described by

\[
\begin{align*}
\dot{x} &= f(x, u, a, \nu, t) \\
y &= g(x, u, c, n, t)
\end{align*}
\]

(cf. eq. (4.64). The process parameter vector \(a\), if it is known or expected to be constant, is treated as a function of time by the equation

\[
\dot{a} = 0
\]  

(13.2)

This leads to a set of equations of the following form:

\[
\begin{bmatrix}
\dot{x} \\
\dot{a}
\end{bmatrix} = \begin{bmatrix}
f(x, u, a, \nu, t) \\
a
\end{bmatrix}  
\]

(13.3)

It will be clear that, even if the dynamics of the process is linear, this equation can not properly be separated into a form

\[
\begin{bmatrix}
\dot{x} \\
\dot{a}
\end{bmatrix} = M \begin{bmatrix}
x \\
a
\end{bmatrix}
\]

(13.4)

where \(M\) does not depend on either \(x\) or \(a\). Consequently even for a process with linear dynamics and which is linear in the parameters the problem of combined parameter and state estimation is a nonlinear one.
This implies that all approaches to this problem will have the characteristics of the model-adjustment technique, i.e. they will be of an iterative type.

Equation (13.1) provides a general representation of the behaviour of (nonlinear) continuous processes. For sampled processes a corresponding representation is:

\[
\begin{align*}
    x_{t(k+1)} &= f(x_{t(k)}, u_{t(k)}, \theta_{t(k)}, \eta_{t(k)}, k) \\
    y_{t(k)} &= g(x_{t(k)}, u_{t(k)}, \xi_{t(k)}, \eta_{t(k)}, k)
\end{align*}
\]  \tag{13.5}

Based on a priori knowledge about the process and about the way it is operated or controlled the representation may be simplified. Anyhow in the pertinent literature such simplifications are introduced frequently in order to make the estimation problem more tractable.

This may result in many process descriptions, some of which are indicated

\[\text{In this chapter we will emphasize the sampled type of representations. This is motivated by the predominant use of digital computers for estimation purposes. In many cases the equivalent continuous type of representation can be derived by careful limiting procedures, taking into account the mathematical complications in describing the effect of white noise sources (Ito / Stratonovich).}\]

in the diagram below and are illustrated in fig. 13.1.
Among the other alterations that may hold are:

\[
\begin{align*}
\mathbf{u}(k) &= 0  \\
\mathbf{a}(k) &= \mathbf{a}  \\
\mathbf{c}(k) &= \mathbf{c}
\end{align*}
\]
The condition that the noise is additive is a very convenient one. From a simple example:

\[ y(k) = g(x(k)) + n(k) \]

it follows that

\[ p(y(k) | x(k)) = p(y(k) - g(x(k))) = p(n(k)) \]

(13.6)

Consequently that conditional probability function is equal to the (assumed or known) probability density function of the noise. Particularly if this function is Gaussian then this way of approach leads to acceptable expressions as has already been demonstrated in the previous chapter and as will be noticed again in the sequel.

13.2. Derivation of an "error function".

For the moment we consider a "state/parameter vector" \( x \) which may consist of:

- state variables and parameters together or
- state variables (for a nonlinear process) or
- parameters for a process that is nonlinear-in-the-parameters.

In this way these different types of estimation problems may be discussed together.

Again the problem can be approached from different points of view:
- Bayes' estimation
- maximum likelihood estimation
- least squares estimation

Again one may discuss the relations between the different estimates for special cases (even cost-function; Gaussian probability density functions; white noise) as it has been done in chapter 5.

In all cases the methods lead to a minimization of an "error function" \( E \) with respect to the state/parameter vector \( x(i) \) over an interval \( i = 0, \ldots, k \).

First we summarize the derivation of such an "error function".
The process description chosen is (cf.: Cox, 1964):

\[ x(k+1) = f(x(k), k) + v(k) \]  
\[ y(k) = g(x(k), k) + n(k) \]  

where \( v(k) \) and \( n(k) \) are Gaussian random vectors, with zero means and covariance matrices:

\[
\begin{align*}
\mathbb{E}[v(k) v'(j)] &= \delta_{kj} v(k) \\
\mathbb{E}[n(k) n'(j)] &= \delta_{kj} n(k) \\
\mathbb{E}[v(k) n'(j)] &= 0
\end{align*}
\]

i.e. the noise sequences are assumed to be "white".

Bayes' estimation. As before the following notation is used:

\[ X_k = [x(0), \ldots, x(k)] \]
\[ Y_k = [y(0), \ldots, y(k)] \]

Again the conditional probability density function \( P(X_k | Y_k) \) of \( X_k \), given the observations \( Y_k \), has to be found. From Bayes' rule:

\[ P(X_k | Y_k) = \frac{P(Y_k | X_k) P(X_k)}{P(Y_k)} \]  

(13.10)

Just like in chapter 5 a cost or loss function can be used to reduce that a posteriori probability density function to simple numerical values, i.e. a minimum cost or minimum loss estimate. On the other hand one might be satisfied by determining those numerical values of the estimate for which the conditional probability function eq. (13.10) obtains its maximum value. For Gaussian probability density functions and even cost or loss functions both methods will result in the same estimate.
Under the conditions given by (13.7) to (13.9) let us therefore consider the maximisation of eq. (13.10) and discuss each of the terms of the right hand side of that equation.

\( \rho(Y_k) \) is the most simple part. As it contains only the numerical values of the observations \( Y_k \) it just acts as a normalizing constant in order to make the integral

\[
\int \rho(X_k | Y_k) \, dX_k = \text{equal to one}
\]

For maximization purposes it does not play any role and thus can be neglected.

\( \rho(Y_k | X_k) \) also results in a simple expression due to the analogy with eq. (13.6) and the fact that the sequence \( \{ n(k) \} \) is non correlated in time ("white"):

\[
\rho(Y_k | X_k) = \prod_{i=0}^{K} \rho \left( Y_i - g(x(i), i) \right) = c_i \prod_{i=0}^{K} \exp \left\{ -\frac{t}{2} \left[ Y_i - g(x(i), i) \right] N^{-1}(i) \left[ Y_i - g(x(i), i) \right] \right\}
\]

where again normalizing factors are taken care of by the constant \( c_i \). Using the shorthand notation \( \Xi, A, Z = \| \Xi \|_A \) it follows:

\[
\rho(Y_k | X_k) = c_i \exp \left\{ -\frac{1}{2} \sum_{i=0}^{K} \left\| Y_i - g(x(i), i) \right\|_N^{-1} \right\}
\]

can be rewritten by recognizing that

\[
\rho(X_i) = \rho(x(i) | \Xi_{i-1}) = \rho(x(i) | X_{i-1}) \rho(X_{i-1})
\]

Due to eq. (13.7) and the assumption that the sequence \( \{ y(i) \} \) is non correlated in time (white) it follows that

\[
\rho(x(i) | \Xi_{i-1}) = \rho(x(i) | x(i-1), \ldots, x(0)) = \rho(x(i) | x(i-1))
\]

Consequently

\[
\rho(X_k) = \rho(x(0)) \prod_{i=1}^{K} \rho(x(i) | x(i-1))
\]
Also due to eq. (13.7) 
\[ p(x(t+1)|x(t)) \]

is Gaussian with
mean
\[ f(x(t), i) \]
covariance matrix
\[ G(i) Y(i) G'(i) \]

The a priori knowledge 
\[ p(x(0)) \]
is also assumed to be Gaussian with:
mean
\[ \mu_x(0) \]
covariance matrix
\[ P(0) \]

Consequently the term under discussion can be written as
\[
p(x_k) = c_2 \exp \left\{ -\frac{1}{2} \left[ x(0) - \mu_x(0) \right]^T P(0)^{-1} \left[ x(0) - \mu_x(0) \right] \right\} \\
- \frac{1}{T} \exp \left\{ -\frac{1}{2} \left[ x(t+1) - f(x(t), i) \right]^T G(i) Y(i) G'(i) \right\} \left[ x(t+1) - f(x(t), i) \right] \\
+ \frac{1}{2} \sum_{i=0}^{k} \left\{ x(t+i) - f(x(t+i), i) \right\}^T G(i) Y(i) G'(i) \left[ G(i) Y(i) G'(i) \right]^{-1} \]

provided the matrix \( G(i) Y(i) G'(i) \) can be inverted. Again \( c_2 \) is a normalizing constant. In the shorthand notation
\[
p(x_k) = c_2 \exp \left\{ -\frac{1}{2} \| x(0) - \mu_x(0) \|_{P(0)}^{-1} \right\} \\
- \frac{1}{T} \exp \left\{ -\frac{1}{2} \| x(t+1) - f(x(t), i) \|_{G(i) Y(i) G'(i)}^{-1} \right\} \\
- \frac{1}{2} \sum_{i=0}^{k} \| x(t+i) - f(x(t+i), i) \|_{G(i) Y(i) G'(i)}^{-1} \}
\]

If we combine these terms there results:
\[
p(x_k | y_k) = c_3 \exp \left\{ -E_k \right\} \quad (13.11)
\]

where
\[
E_k = \frac{1}{2} \left\| x(0) - \mu x(0) \right\| P^{-1}(0) + \\
+ \frac{1}{2} \sum_{i=0}^{k} \left\| y(i) - G(x(i), i) \right\| N^{-1}(i) + \\
+ \frac{1}{2} \sum_{i=0}^{k-1} \left\| x(i+1) - f(x(i), i) \right\| \left[ G V G' \right]^{-1} \]
\]

(13.12)

Maximizing eq. (13.11) with respect to \( x_k = [x(0), \ldots, x(k)] \) implies minimization of \( E_k \). The last term of (13.12) can be replaced by

\[
\frac{1}{2} \sum_{i=0}^{k-1} \left\| y(i) \right\| \frac{1}{Y(i)}
\]

and the constraint expressed by eq. (13.7). Consequently one may also minimize:

\[
E'_k = \frac{1}{2} \left\| x(0) - \mu x(0) \right\| P^{-1}(0) + \\
+ \frac{1}{2} \sum_{i=0}^{k} \left\| y(i) - G(x(i), i) \right\| N^{-1}(i) + \\
+ \frac{1}{2} \sum_{i=0}^{k-1} \left\| y(i) \right\| \frac{1}{Y(i)} + \\
\lambda'(i) \left[ x(i+1) - f(x(i), i) - G(i) y(i) \right] \]

(13.13)

where \( \lambda'(i) \) is a vector of Lagrange multipliers. In this formulation \( G V G' \) does not have to be invertible.

Note that \( E'_k \) consists of three components, which can be recognized as being related to:
- the assumption on the state/parameter vector at \( t = 0 \)
- the additive "observation noise" \( \eta \)
- the "process noise" \( \nu \) (together with the constraints given by the process description).

The weighting matrices \( P^{-1}(0) \), \( N^{-1}(i) \) and \( V^{-1}(i) \) provide a measure for the relative weight of these contributions to the "error" \( E'_k \).

Note also that \( y(i) \) for \( i = 0, \ldots, k \) are the numerical values of the observations on the process.
If no a priori knowledge on $x(0)$ is available, then $P^{-1}(0) \rightarrow 0$ and the first term of $E_k$ drops out.

Note that in the process description (13.7) it has been assumed that there is no measurable input sequence $\{u(i)\} \ i = 0, 1, \ldots, k$. If such an input is present then that description and the subsequent derivation can easily be adapted to incorporate that extra information.

13.3. Minimization of the "error function"

The error function (13.13) was derived under the following conditions:
- "observation noise" $n$: additive, Gaussian, "white";
- "process noise" $v$: additive, Gaussian, resulting in a "coloured" contribution due to matrix $G$.

For non-Gaussian noises the expression (13.13) may be assumed as an acceptable error function on intuitive grounds. In that case its minimization leads to a least squares estimate.

By minimizing (13.13) one finds the estimate $\hat{x}(j|k)$, i.e. the estimation of the state/parameter vector at time $j$ based on the observations $0, 1, \ldots, k$.

Again one may recognize the cases:

- $0 \leq j < k$ interpolation (smoothing)
- $j = k$ estimation of the present state (filtering)
- $j > k$ extrapolation (if a simple extension of the expression for $E_k$ is used).

Several approaches to this minimization problem are available e.g.:
- the dynamic programming formulation leading to a sequential procedure with the possibility of using new observations that becomes available, but also having its inherent problems with respect to the memory capacity needed in the computer (Bellman: curse of dimensionality)
- the discrete maximum principle or
- the discrete Euler–Lagrange equations;

these two methods result in a nonlinear two-point boundary value problem (TPBVP)

- the same, but linearizing the process equations about the point \( x^*(k|k) \)

This leads to approximated estimates, but may result in recursive schemes in which new observations can be processed as soon as they become available.

A straightforward derivation of the necessary conditions for a minimum of \( E_k \) can be found by differentiation:

\[
\begin{align*}
\frac{\partial E'_k}{\partial x(i)} &= -\left[ y(i) - \mathbf{g}(x(i), i) \right]' \mathbf{N}^{-1}(i) \frac{\partial \mathbf{g}(x(i), i)}{\partial x(i)} + \\
&\quad + \lambda'(i-1) - \lambda'(i) \frac{\partial \mathbf{f}(x(i), i)}{\partial x(i)} \quad (13.14) \\
\frac{\partial E'_k}{\partial \lambda(i)} &= x(i+1) - \mathbf{f}(x(i), i) - \mathbf{G}(i) \nu(i) \quad (13.15) \\
\frac{\partial E'_k}{\partial \nu(i)} &= \mathbf{V}(i) \mathbf{V}^{-1}(i) + \lambda'(i) \mathbf{G}(i) \quad (13.16)
\end{align*}
\]
By elimination of $y(i)$ and noting that these are the conditions for the optimal estimates $\hat{x}(i|k)$ and $\hat{x}(i)$ there results:

\[
\hat{x}(i+1|k) = \mathcal{P}(\hat{x}(i|k), i) - \mathcal{Q}(i) \mathcal{V}(i) \mathcal{G}(i) \mathcal{A}(i) \tag{13.17}
\]

\[
\mathcal{A}'(i-1) = \mathcal{A}'(i) \frac{\partial \mathcal{F}(x(i), i)}{\partial x(i)} \bigg|_{\hat{x}(i) = \hat{x}(i|k)} + \left[ y(i) - g(\hat{x}(i|k), i) \right]' \mathcal{N}(i) \left. \frac{\partial z(x(i), i)}{\partial x(i)} \right|_{\hat{x}(i) = \hat{x}(i|k)} \tag{13.18}
\]

The expressions (13.17) and (13.18) are called the canonic equations.

In the same way follow the boundary conditions for $i = 0$ and $i = k$:

\[
\hat{x}(0|k) = \mu_\lambda(0) + \left[ y(0) - g(\hat{x}(0|k), 0) \right]' \mathcal{N}(0) \left. \frac{\partial z(x(0), i)}{\partial x(0)} \right|_{\hat{x}(0) = \hat{x}(0|k)} + \mathcal{A}'(0) \left. \frac{\partial \mathcal{F}(x(0), i)}{\partial x(0)} \right|_{\hat{x}(0) = \hat{x}(0|k)} \mathcal{P}(0) \tag{13.19}
\]

\[
\mathcal{A}(k) = 0 \tag{13.20}
\]

The expressions (13.17) to (13.20) define a type of the celebrated TPBVP.
Of the host of literature on optimal control problems many publications are relevant to our problem. It is far beyond the scope of this chapter to try to give an extensive discussion of the many methods discussed or proposed, e.g. (cf. Cuenod and Sage, 1967):
- gradient method
- quasilinearization
- invariant imbedding

We just give some notes on Quasilinearization. This is an interesting and powerful technique. Its development can be traced from the Newton-Raphson method for determining the root of an algebraic equation, through the generalization of this method by Kantorovich for functional equations. Under the name quasilinearization it was developed by Bellman and Kalaba (1965).

For this exposition we use the continuous type of process description given before, but the output $y$ is assumed to be linearly dependent on the state $x$:

$$
\left[ \begin{array}{c} \dot{x} \\ \dot{a} \end{array} \right] = \left[ \begin{array}{c} f(x, u, a, y, t) \\ 0 \end{array} \right]$$  \hspace{1cm} (13.21)

$$
y(t) = C \cdot x(t)$$

$x$ and $a$ have the dimension $n$ and $m$. Over the interval of process observation there are available

$$
y(j) = y_j \hspace{1cm} j = 1, 2, \ldots, m+n$$  \hspace{1cm} (13.22)

i.e. numerical values for the measurements at $(m+n)$ instances of time.

The expressions (13.21) provide $(m+n)$ equations that have to satisfy the boundary conditions given by (13.22). This boundary value problem can in the following way be tackled by means of the method of quasilinearization. The $(i+1)$-th approximation of the time functions $x(t)$ and $a(t)$ over the time interval of observation are
\[
\dot{x}^{(i+1)} = f(x^{(i)}, u, a^{(i)}) + \frac{F_1^{(i)}}{x^{(i+1)}} [x^{(i+1)} - x^{(i)}] + \frac{F_2^{(i)}}{a^{(i)} - a^{(i)}} \]
\]
\[
\dot{a}^{(i+1)} = 0
\]

with the conditions:
\[
C x^{(i+1)} = y^{(i+1)} = y_j
\]
(13.24)

The matrices \( F_1^{(i)} \) and \( F_2^{(i)} \) are the Jacobians:
\[
F_1^{(i)} = \left[ \frac{\partial f}{\partial x} \right]_{x^{(i)} a^{(i)}} \quad \text{and} \quad F_2^{(i)} = \left[ \frac{\partial f}{\partial a} \right]_{x^{(i)} a^{(i)}}
\]
(13.25)

For the \((i+1)\)th iteration the problem statement is:
\[
\begin{cases}
\dot{x}^{(i+1)} = F_1^{(i)} x^{(i+1)} + F_2^{(i)} a^{(i+1)} + w^{(i)} \\
y^{(i)} = C x^{(i+1)}
\end{cases}
\]
(13.26)

where \( w^{(i)} \) is available from the previous iteration:
\[
w^{(i)} = f(x^{(i)}, u, a^{(i)}) - F_1^{(i)} x^{(i)} - F_2^{(i)} a^{(i)}
\]
(13.27)

For the first iteration one starts with an initial guess \( x^{(0)} \) and \( a^{(0)} \).

The iteration process can be terminated when the difference between the solutions of two successive iterations, i.e. \( x^{(i+1)} - x^{(i)}, a^{(i+1)} - a^{(i)} \), fulfill pre-assigned conditions.

By Bellman and Kalaba (1965) it is shown that the sequence of functions \( x^{(0)}(t), \ldots, x^{(i+1)}(t) \) converges to the true solution in a quadratic way, if \( f(x, u, a) \) is a convex function and if the elements of the Jacobian above and below the main diagonal are positive.
Ohap and Stubberud (1965) have shown that also in cases where these conditions are not fulfilled convergence may be obtained.

So far it has been assumed that the number of measurements equals the number of equation \((n+m)\). In view of disturbances on the observations in practice, the number of measurements \(k\) will be chosen \(k > (n+m)\). In that case one may minimize the squared error between the measurements \(y_j\) and the "model" output \(y(i)\).

**Simplifications.** In the preceding discussions the process was assumed to have nonlinear dynamics. In section 13.1, it was indicated already that, in spite of linear dynamics, the combined parameter and state estimation problem is of a nonlinear nature. Yet one may expect that a simpler type of process will result in a simpler estimation problem. Only partial insight into this relationship between complexity of the process dynamics and the complexity of the estimation problem seems to be available.

As an interesting case the following example is cited (Anderson et al, 1969). Assume a process with linear dynamics, that is constant in time, to be described by eq. (13.28)

\[
\begin{align*}
\dot{x}(k+1) &= A \dot{x}(k) + \nu(k) \\
y(k) &= x(k) + \eta(k)
\end{align*}
\]

(13.28)

\(\nu(k)\) and \(\eta(k)\) are assumed to be "white". Note that the entire state is observed. The authors show that the following estimates for \(A\)

\[
\hat{A}_n = \left[ \sum_{i=3}^n y(i) y'(i-2) \right] \left[ \sum_{i=3}^n y(i-1) y'(i-2) \right]^{-1}
\]

converge with probability one to \(A\) and that a Kalman/filter updated by these \(\hat{A}_n\) has an estimate which converges with probability one to that of the correct Kalman filter. Consequently both parameters and states are properly estimated.
13.4. Concluding remarks

As it is virtually impossible to do justice to the many interesting publica­tions on relevant methods for minimizing the "error functions" and solving TPBVVP the reader is referred to:

- the monograph by Sage and Melsa (1971) on this topic;
- the papers mentioned under "additional literature".


Additional literature


fig. 13-1
Appendix A

LIST OF SYMBOLS USED

Only the most important symbols are given. General rules (with a few exceptions):

- bold face lower-case symbols indicate vectors
- bold face capitals indicate matrices
- where applicable capitals denote the $F$, $L$- or $Z$-transform of lower-case variables.

A
B - transfer function; dynamic operator
C
D - differentiation operator
E - even function(al) of the error signal; error or loss function
F - function(al), $F[ ]$
G - transfer function; dynamic operator, $F( )$
H
I
J - criterion function
K
L - likelihood function
M - model
N - normal distribution function; $N(\mu, \sigma^2)$, mean $\mu$, variance $\sigma^2$
N - noise power, e.g. at correlator output or process output
O
P - process
P - probability, $P[ ]$
Q
R - Lyapunov function
S - System
S - signal power, e.g. at correlator output or process output
T - time, integration interval
U
V
W - transform of the corresponding time function
X
Y
Z
- system matrix
- distribution matrix
- output or measurement matrix
- lower triangular matrix; input-output matrix
- matrix of partial derivatives of vector function \( f \)
- identity matrix
- information matrix
- matrix
- covariance of the additive noise
- "covariance matrix" of an estimate
- estimation matrix
- positive definite (weighting) matrix
- (similarity) transformation matrix
- matrix of sample values of signals \([u_0, \ldots, u_m]\)
- matrix of sample values of output signals
- coefficients of differential equation, difference equation, transfer function of the process; (varying) parameters of the process.

- coefficients in (orthogonal) developments

- (generalized) error signal
- basis vectors
- frequency

- function; weighting function; impulse response; sample points of the impulse response; unit-sample-response function

- integer numbers

- additive noise (vector)

- probability density function

- weighting factor
- eigen vector
- Laplace transform variable
- time
- measurable input (vector) of the process (and model); voltage
- output of the linear operators in the (generalized) model; parameter sensitivity functions
- unmeasurable noise (vectors) of the process
\( w \) - output of the model; weighting function
\( x \) - process state (vector)
\( y \) - output (vector) of the process, either or not corrupted with noise
\( z \) - shift operator
\{ \begin{align*}
A & \quad \text{estimates of } A, B, C, \quad A, B \\
B & \\
C & \\
P & \quad \text{filter;}
\end{align*} \}

\begin{align*}
E & \quad \text{expectation; ensemble average; } E[ ] \\
F & \quad \text{Fourier transform; inverse Fourier transform; } F[ ] \\
F^+ & \quad \text{Fourier transform acting on positive time components only} \\
F^- & \quad \text{Fourier transform acting on negative time components only}
\end{align*}

\begin{align*}
L & \quad \text{one-sided Laplace transform; inverse Laplace transform; } L[ ] \\
L^2 & \quad \text{two-sided } " " " " " "
\end{align*}

\begin{align*}
M & \quad \text{dynamic operator; } M[ ] \\
O & \quad \text{policy operator; } O[ ] \\
P & \quad \text{quantizing operator; } O[ ] \\
R & \quad \text{mapping;}
\end{align*}

\begin{align*}
S & \quad \text{class of systems;}
T & \quad \text{class of input signals;}
U & \\
W & \\
X & \\
Y & \\
Z & \quad \text{z-transform; inverse z-transform; } Z[ ]
\end{align*}
\( \Gamma \) (Gamma) - gain factor
\( \Gamma \) (Gamma) - matrix of (gain) factors
\( \Delta \) (Delta) - variation
\( \Theta \) (Theta)
\( \Lambda \) (Lambda)
\( \Xi \) (Xi)
\( \Pi \) (Pi) - continued product
\( \Pi_i \) - orthogonal polynomials
\( \Sigma \) (Sigma) - continued sum
\( \Upsilon \) (Upsilon)
\( \Phi \) (Phi) - transition or principal matrix; \( \Phi(t) \), \( \Phi(k) \)
\( \nabla \Phi \) - power spectral density function; \( \Phi(\omega) \)
\( \Psi \) (Psi) - correlation function: \( \Psi(\tau) \)
\( \Omega \) (Omega) - matrix of process input and output samples;
\( \nabla \) - gradient; \( \nabla_\beta \) gradient with respect to \( \beta \)
\( \alpha \) \( \beta \) (alpha) \( \beta \) (beta) \{ \text{estimates of } a, b; \text{coefficients of differential equation, difference equation, transfer function of the model; adjustable parameters of the model; components of the coefficient vector} \}

\( \gamma \) \( \delta \) (delta) \{ \text{gain factor; coefficient of orthogonal development} \}

\( \delta_m \) \{ \text{Kronecker delta} \}

\( \varepsilon \) (epsilon) \{ \text{base of Naperian logarithms} \}

\( \varepsilon \) \{ \text{small real constant} > 0 \}

\( \zeta \) (zeta) \{ \text{components of the gradient vector} \}

\( \eta \) (eta) \{ \text{time delay interval, sample interval, integration variable} \}

\( \theta \) (theta) \{ \text{parameter vector with components } \begin{bmatrix} \beta_j, \alpha_i \end{bmatrix} \}

\( i \) (iota) \{ \text{scalar quantities} \}

\( \kappa \) \( \lambda \) (lambda) \{ \text{parameter error vector with components } \begin{bmatrix} \alpha_i - a_i, \beta_j - b_j \end{bmatrix} \}

\( \mu \) (mu) \{ \text{first (general) moment of a random variable; e.g.} \}

\( \nu \) (nu) \{ \text{dummy integration variables} \}

\( \xi \) (xi) \{ \text{model state (vector)} \}

\( \pi \) (pi) \{ \text{polynomials} \}

\( \pi \) \{ \text{3.14159} \}

\( \rho \) (rho) \{ \text{correlation coefficient} \}

\( \sigma \) (sigma) \{ \text{standard deviation} \}

\( \sigma^2 \) \{ \text{second central moment of random variable; variance e.g.} \}

\( \sigma_x^2 = \mathbb{E} \left[ \left( x - \mu \right)^2 \right] \)

\( \tau \) (tau) \{ \text{time (age) variable; time constant} \}
\( \nu \) (upsilon)
\( \varphi \) (phi) - characteristic function
\( \phi \) (phi) - phase angle
\( \chi \) (chi)
\( \psi \) (psi) - reference vector
\( \omega \) (omega) - radian frequency variable
Subscripts and superscripts

\( (n) \)  
\( u; \dot{u} \)  
- n-th; first derivative

\( u(k) \)  
- sample value at \( t = k \theta \)

\( u_i \)  
- i-th component of the vector \( u \)

\( u_i \)  
- vector during the i-th measurement interval

\( u(i) \)  
- i-th approximation

\( u(k+1) \)  
- new observation on \( u \) at time \( k+1 \)

\( \beta(i) \)  
- vector after the i-th adjustment

\( E(i) \)  
- value after the i-th measurement

\( \ast \)  
- e.g. \( u^\ast \) = sampled version of time-function \( u(t) \)

\( \ast \)  
- e.g. \( G^\ast \) = pulse transfer function

\( \text{def.} \)  
- e.g. \( G(\omega) \) = complex conjugate of \( G(\omega) \)

\( e.g. \overline{x} = \text{ensemble average over infinite ensemble} \)

\( e.g. \overline{x}_N = \text{ensemble average over finite ensemble} = \frac{1}{N} \sum_{i=1}^{N} x_i \)

\( e.g. \langle x \rangle = \text{time average over infinite time} = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} x(t) dt \)

\( e.g. \langle x \rangle_T = \text{time average over finite time} = \frac{1}{T} \int_{0}^{T} x(t) dt \)

\( e.g. \langle x, y \rangle = \text{some type of (nonlinear) signal processing operation:} \)

\( e.g. \langle x, y \rangle = \int_{0}^{T} x(t) y(t) dt \)

\( \ast \)  
- e.g. \( x \ast h = \text{convolution} \)

\( \ast \)  
- e.g. \( H(\omega) = \text{phase angle contribution} \)

\( | | \)  
- absolute value; determinant e.g. Jacobian

\( \ast \)  
- identity

\( || || \)  
- norm

\( \| \|_2^R \)  
- positive definite form with weighting matrix \( R \)

\( \text{def.} \)  
- e.g. \( x^{\text{def.}} \) by definition
\[\text{sgn}\[x\] \quad \text{signum, e.g. } \text{sgn } e = \frac{e}{|e|}\]

\[\text{var}\[x\] \quad \text{variance}\]

\[\text{cov}\[x\] \quad \text{covariance}\]

\[(-1) \quad \text{e.g. } A(-1): \text{generalized inverse}\]

\[\mid \quad \text{e.g. } p(z|x): \text{conditional}\]

\[^{\hat{}} \quad \text{e.g. } \hat{\beta}: \text{optimum estimate}\]

\[^{\sim} \quad \text{e.g. } \tilde{\varphi}: \text{approximation of correlation function}\]

\[\text{Re}\[x\] \quad \text{real part of } x\]

\[\text{Im}\[x\] \quad \text{imaginary part of } x\]

\[\exp[x] \quad e \text{ to the power } x\]

\[\quad \text{e.g. } U': \text{transpose}\]

\[\int_{k}^{\ldots} \text{d}y \quad k-\text{fold integral}\]

\[M \quad \text{model, e.g. } y_M\]
Appendix B  SUMMARY OF NOTIONS FROM PROBABILITY THEORY AND STOCHASTIC SIGNALS.

In writing the text it is supposed that the reader has some knowledge of probability theory and stochastic signals. This summary is designed only for easy reference of properties and notation. The contents of this appendix will be used freely in this text without further reference.

Sets. The relations between "events" are governed by the rules of set theory: Notions: set, subset, element, empty set.

Operations; cf. fig. B.1:

- $\bar{A}$ complement
- $A \cup B$ union
- $A \cap B$ intersection

Properties:

<table>
<thead>
<tr>
<th>commutative</th>
<th>associative</th>
<th>distributive</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A \cup B = B \cup A$</td>
<td>$(A \cup B) \cup C = A \cup (B \cup C)$</td>
<td>$A \cap (B \cap C) = (A \cap B) \cap (A \cap C)$</td>
</tr>
<tr>
<td>$A \cap B = B \cap A$</td>
<td>$(A \cap B) \cap C = A \cap (B \cap C)$</td>
<td>$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability. To each "event" $A$ of a class of possible events in a simple experiment a number $P[A]$ is assigned. This number is called probability.

Axiom I $P[A] \geq 0$

Axiom II $P[\Omega] = 1$ if $\Omega$ is a certain event

$P[\emptyset] = 0$ if $\emptyset$ is an impossible event

Axiom III $P[A \cup B] = P[A] + P[B]$ if $A \cap B = \emptyset$; i.e. if the events are mutually exclusive.

For cases where there is an infinite number of events:
axiom IV \[ P[A_1 \cup A_2 \cup A_3 \cup \ldots] = \sum_{i=1}^{\infty} P[A_i] \]

if \( A_i \cap A_j = \emptyset \) for each \( i \neq j \).

Combined experiments. The outcomes of two simple experiments, viz. \( A_i \) and \( B_j \), are considered as a (combined) event \([A_i, B_j]\).

\[ 0 \leq P[A_i, B_j] \leq 1 \]

\[ \sum_{i,j} P[A_i, B_j] = 1 \]

\[ P[A_i] = \sum_j P[A_i, B_j] \quad \text{if} \quad B_k \cap B_j = \emptyset \quad \text{for each} \quad k \neq j \]

\[ P[B_j] = \sum_i P[A_i, B_j] \quad \text{if} \quad A_k \cap A_i = \emptyset \quad \text{for each} \quad k \neq i \]

The symbols \( \sum_i \) and \( \sum_{i,j} \) mean summation over all \( i \), viz. all combinations \( i,j \).

Conditional probability. Definition:

\[ P[A_i|B_j] = \frac{P[A_i, B_j]}{P[B_j]} \quad \text{if} \quad P[B_j] \neq 0 \]

conssequently


For independent events:

\[ P[A_i|B_j] = P[A_i] \]

\[ P[A_i, B_j] = P[A_i] P[B_j] \]

In most cases the event is (or can be interpreted by means of) a numerical value or a set of numerical values.

A (real) random variable \( X \) is a quantity that can have different values in

In this appendix a random variable is denoted by an upper case type. This
The random variable can be **discrete**, if it can have a finite or denumerable set of distinct values, or it can be **continuous**. Based on these ideas the following notions can be recognized.

<table>
<thead>
<tr>
<th>random variable $X$.</th>
<th>discrete</th>
<th>continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>one-dimensional case</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>distribution function</strong></td>
<td>$F(x) = P[X \leq x]$</td>
<td>$F(x) = P[X \leq x]$</td>
</tr>
<tr>
<td>(cf. fig. B.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>probability(density)function</strong></td>
<td>$p_i = P[X = x_i]$</td>
<td>$p(x) = \frac{dF(x)}{dx}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>probability</strong></td>
<td>$\sum_{i} p_i \leq 1$</td>
<td>$\sum_{i} p_i = 1$</td>
</tr>
<tr>
<td></td>
<td>$0 \leq p(x)$</td>
<td>$0 \leq p(x)$</td>
</tr>
<tr>
<td></td>
<td>$\int_{-\infty}^{\infty} p(x)dx = 1$</td>
<td>$\int_{a}^{b} p(x)dx = 1$</td>
</tr>
<tr>
<td><strong>properties</strong></td>
<td>$F(x) = \sum_{i, x_i \leq x} p_i$</td>
<td>$F(x) = \int_{-\infty}^{x} p(x)dx$</td>
</tr>
<tr>
<td></td>
<td>$F(-\infty) = 0$</td>
<td>$F(\infty) = 1$</td>
</tr>
</tbody>
</table>
Examples of density functions

Two-dimensional case

(joint) distribution function

\( F(x,y) = P[X \leq x, Y \leq y] \)

(joint) probability (density) function

\( p(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} \)

\( p(x,y) = \begin{cases} \lambda^x e^{-\lambda} & x = 0, 1, 2, \ldots x! \\ 0 & -\infty < x < 0 \end{cases} \)

Binomial:

\( P[X=x] = \binom{n}{x} p^x (1-p)^{n-x} \)

\( x = 0, 1, \ldots n \)

Poisson:

\( P[X=x] = \frac{\lambda^x e^{-\lambda}}{x!} \)

\( x = 0, 1, 2, \ldots \)

Normal (c.f. fig. B-3):

\( p(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right] \)

\( -\infty \leq x \leq \infty \)

Exponential:

\( p(x) = \begin{cases} \lambda e^{-\lambda x} & 0 \leq x < \infty \\ 0 & -\infty < x < 0 \end{cases} \)

Properties

\( 0 \leq p_{ij} \leq 1 \)

\( \sum_{i} \sum_{j} p_{ij} = 1 \)

\( F(x,y) = \sum_{\text{all } x_i \leq x} \sum_{\text{all } y_i \leq y} p_{ij} \)

\( \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x,\eta) \, dx \, d\eta = 1 \)

\( F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(x,\eta) \, dx \, d\eta \)

\( F(-\infty, y) = 0 \)

\( F(x, -\infty) = 0 \)

\( F(\infty, \infty) = 1 \)
marginal distribution function

\[ F(x, \infty) = P\left( X \leq x, Y \leq \infty \right) = P\left( X \leq x \right) = P\left( X \leq x_1 \right) \]

\[ F(\infty, y) = P\left( X \leq \infty, Y \leq y \right) = P\left( Y \leq y \right) = P\left( Y \leq y_j \right) \]

marginal probability (density function)

\[ p_i = \sum_j p_{ij} \]

\[ p_j = \sum_i p_{ij} \]

p(x) = \frac{dF(x, \infty)}{dx} = \int_{-\infty}^{\infty} p(x, y) \, dy

p(y) = \frac{dF(\infty, y)}{dy} = \int_{-\infty}^{\infty} p(x, y) \, dx

\% From here on the assumption will be used that a change of argument indicates a different function.

conditional distribution function

\[ F(x \mid y) = P\left( X \leq x \mid Y = y \right) \]

\[ F(y \mid x) = P\left( Y \leq y \mid X = x \right) \]

conditional probability (density) function

\[ P(X = x_1 \mid Y = y_j) = \frac{p_{ij}}{p_j} \]

\[ P(Y = y_j \mid X = x_1) = \frac{p_{ij}}{p_i} \]

independence of \( X, Y \)

\[ F(x, y) = F(x)F(y) \]

\[ P_{ij} = p_ip_j \]

most important example of a bivariate density function:

\[ p(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left[-\frac{1}{2} Q \right] \]

\[ Q = 1 - r^2 \left\{ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2r \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right\} \]

bivariate normal (cf. fig. B.4):

\[ p(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left[-\frac{1}{2} Q \right] \]
more-dimensional case

(joint) distribution function

\[ F(x, y, z, \ldots) = \mathbb{P} \{X \leq x, Y \leq y, Z \leq z, \ldots\} = \mathbb{P} \{X \leq x, Y \leq y, Z \leq z, \ldots\} \]

etc., analogous to the previous case.

In many cases it is a quite elaborate procedure to work with the expressions for these (probability) functions. Fortunately quite often it is permissible (and more convenient) to work with parameters in stead of such functions. This leads to the following notions.

<table>
<thead>
<tr>
<th>random variable ( X )</th>
</tr>
</thead>
<tbody>
<tr>
<td>discrete</td>
</tr>
<tr>
<td>continuous</td>
</tr>
</tbody>
</table>

one-dimensional case

expectation

\[ \mathbb{E} \{ \ldots \} = \sum_i \ldots \pi_i \quad \mathbb{E} \{ f(X) \} = \sum_i f(x_i) \pi_i \]

property: linearity

\[ \mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y] \]

\( n \)th moment

\[ \mathbb{E}[X^n] = \sum_i x_i^n \pi_i \quad \mathbb{E}[X^n] = \int_{-\infty}^{+\infty} x^n \pi(x) dx \]

first moment; mean; expectation

\[ \mu = \mathbb{E}[X] \quad \mu = \int_{-\infty}^{+\infty} x \pi(x) dx \]

\( n \)th central moment

\[ \mathbb{E}[(X - \mu)^n] \]

first central moment

\[ \mathbb{E}[(X - \mu)] = \sigma \]
second central moment; variance
spread; standard deviation

two-dimensional case

expectation

\[ \mathbb{E}[\cdots] = \sum_{i,j} p_{ij} \]
\[ \mathbb{E}[f(x,y)] = \sum_{i,j} f(x_i, y_j) p_{ij} \]

mean

\[ \mu_x = \mathbb{E}[X] \]
\[ \mu_x = \sum_{i,j} x_i p_{ij} \]

variance

\[ \sigma_x^2 = \mathbb{E}[(X-\mu_x)^2] \]
\[ \sigma_x^2 = \sum_{i,j} x_i^2 p_{ij} - \mu_x^2 \]

covariance

\[ \sigma_{xy}^2 = \text{cov}[X,Y] = \mathbb{E}[(X-\mu_x)(Y-\mu_y)] = \mathbb{E}[XY] - \mu_x \mu_y \]

\[ \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \]

correlation coefficient

conditional expectation

\[ \mathbb{E}[X|Y=y_j] = \frac{\sum_{i} x_i p_{ij}}{p_j} \]
\[ \mathbb{E}[X|Y] = \frac{\int x p(x,y) dx}{p(y)} \]

property

\[ \mathbb{E} [\mathbb{E}[X|Y]] = \mathbb{E}[X] \]

independence of \( X, Y \)

\[ \mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y] \]
\[ \rho_{xy} = 0 \]

\[ \sum_{i,j} x_i y_j p_{ij} = \sum_{i} x_i p_i \{ \sum_{j} y_j p_j \} = \mu_x \mu_y \]
\[ = \frac{\sum_{i} x_i p_i \{ \sum_{j} y_j p_j \} - \mu_x \mu_y}{\sum_{i} x_i p_i \{ \sum_{j} y_j p_j \} } = \mu_x \mu_y \]
A **stochastic process** is a family of random variables \( X(t) \), where \( t \) is a parameter running over a set \( T \).

A realization or sample function of a stochastic process is an assignment of a possible value to \( X(t) \) for each \( t \) of \( T \).

In this text \( t = \) time and \( T = \) the time interval of observation. In that case we may choose as:

- a sample function \( \sim \) a (stochastic) signal
- a stochastic process \( \sim \) an ensemble of stoch. signals.

Conceptually such an ensemble might consist of a (large) number of physical systems of (macroscopically) identical structure and under identical environmental conditions, c.f. fig. B.5. If each system produces a stochastic signal then we may assume that the statistical properties of all these signals are the same.

In that situation one can define the probability of the amplitude \( X(t_1) \) at time \( t_1 \):

\[
F_1(x_1, t_1) = P[X(t_1) \leq x_1] \quad \text{and}
\]

\[
p_1(x_1, t_1) = \frac{d F_1(x_1, t_1)}{dx_1}
\]

or

\[
p_1(x_1, t_1) \Delta x_1 \sim P[x_1 < X(t_1) \leq x_1 + \Delta x_1]
\]

(if \( F_1 \) can be differentiated). Fig. B.6 shows some signals (ensemble elements) that pass through the interval \( (x_1, x_1 + \Delta x_1] \) at time \( t_1 \).

A more-detailed description of the random process can be given by studying the ensemble at two instants of time:

\[
F_2(x_1, t_1; x_2, t_2) = P[X(t_1) \leq x_1; X(t_2) \leq x_2] \quad \text{and}
\]

\[
p_2(x_1, t_1; x_2, t_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_2(......) \quad \text{or}
\]

\[
p_2(x_1, t_1; x_2, t_2) \Delta x_1 \Delta x_2 \sim P[x_1 < X(t_1) \leq x_1 + \Delta x_1; x_2 < X(t_2) \leq x_2 + \Delta x_2]
\]
In the same way one may find probability density functions $p_3$, $p_4$, ... relating the amplitude probabilities at 3, 4, ... instances of time, that give a more detailed description of the stochastic process.

As done before for random variables one may look for parameters that can be used, in stead of probability density functions, for characterizing a stochastic process. These are called ensemble moments. Some examples:

$$
\mu_X(t) = \frac{\mathbb{E}[X(t)]}{\phi} = \int_{-\infty}^{\infty} x p(x, t) \, dx,
$$

$$
\mu^2_X(t) = \frac{\mathbb{E}[X^2(t)]}{\phi} = \int_{-\infty}^{\infty} x^2 p(x, t) \, dx,
$$

$$
\psi_{XX}(t_1, t_2) = \frac{\mathbb{E}[X(t_1)X(t_2)]}{\phi} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, t_1; x_2, t_2) \, dx_1 \, dx_2.
$$

The latter expression, the autocorrelation function, is of great importance for describing the properties of a stochastic process. Analogously one may describe the relation between two stochastic processes $X$ and $Y$ by the crosscorrelation function $\psi_{XY}(t_1, t_2)$.

These definitions hold for general stochastic processes, which may be nonstationary. A very important subclass is that of the stationary processes, i.e. for which holds that the statistical properties are invariant with respect to shifting along the time axis (cf. fig. B.5):

$$
F_n(x_1, t_1 + \theta; \ldots; x_n, t_n + \theta) = F_n(x_1, t_1; \ldots; x_n, t_n) \quad \text{each } \theta
$$

One consequence is, that for such stationary processes all parameters are invariant under a timeshift:

$$
\mu_X(t_1) = \mu_X(t) = \mu_X,
$$

$$
\psi_{XX}(t_1, t_2) = \psi_{XX}(t, t + \tau) = \psi_{XX}(\tau) \quad \text{with } \tau = t_2 - t_1.
$$

For such processes it makes sense to use the notion of time average or time moments. If the corresponding time and ensemble moments are equal:

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T X(t) \, dt = \frac{1}{\phi} \int_{-\infty}^{\infty} x \, p(x) \, dx,
$$

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T X^2(t) \, dt = \frac{1}{\phi} \int_{-\infty}^{\infty} x^2 \, p(x) \, dx.
$$
\[ \psi_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T X(t) Y(t+\tau) \, dt = \frac{X(t) Y(t+\tau)}{\text{time}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \, y \, p(x, y, \tau) \, dx \, dy \]

e tc.

then the process is called ergodic.

Some properties of autocorrelation functions for stationary processes:

\[ \psi_{xx}(\tau) \text{ exists if } \mathbb{E}[X^2(t)] \text{ exists} \]

\[ \psi_{xx}(0) = \text{average "power" } > 0 \text{ if } \mathbb{E}[X^2(t)] \neq 0 \]

\[ \psi_{xx}(\tau) = \psi_{xx}(-\tau) \]

\[ |\psi_{xx}(\tau)| \leq \psi_{xx}(0) \]

if \( X(t) \) has a periodic component, so has \( \psi_{xx}(\tau) \)

if \( X(t) \) has no periodic component and no d.c. component

then \( \psi_{xx}(\tau) \to 0 \) for \( \tau \to \infty \)

For a presentation of a stochastic process in the frequency domain one may use the auto- and cross spectral density functions (power spectrum), \( \phi_{xx}(j\omega) \) and \( \phi_{xy}(j\omega) \) which can be found through Fourier or two-sided Laplace transform of the respective correlation functions (Wiener-Khintchine relations):

\[ \left\{ \begin{array}{l}
\phi_{xy}(j\omega) = \mathcal{F} \left[ \psi_{xy}(\tau) \right] = \int_{-\infty}^{\infty} \psi_{xy}(\tau) \, e^{-j\omega \tau} \, d\tau \\
\psi_{xy}(\tau) = \mathcal{F}^{-1} \left[ \phi_{xy}(j\omega) \right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{xy}(j\omega) \, e^{j\omega \tau} \, d\omega
\end{array} \right. \]

or

\[ \left\{ \begin{array}{l}
\phi_{xy}(s) = \mathcal{L}_2 \left[ \psi_{xy}(\tau) \right] = \int_{-\infty}^{\infty} \psi_{xy}(\tau) \, e^{-s \tau} \, d\tau \quad \alpha < \text{Re}[s] < \beta \\
\psi_{xy}(\tau) = \mathcal{L}_2^{-1} \left[ \phi_{xy}(s) \right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{xy}(s) \, e^{s \tau} \, ds \quad \alpha < c < \beta
\end{array} \right. \]

Fig. B.8 shows some transform pairs. Note the use of the \( j \)-operator in the argument of the \( \mathcal{F} \) transformed variable; this supports the parallelism between \( \mathcal{F} \)- and \( \mathcal{L}_2 \)-transformation through \( j\omega \leftrightarrow s \).

One aspect of the \( \mathcal{L}_2 \) transformation needs special attention. For the transformation of correlation functions one may chose \( c = 0 \), i.e. the convergence
strip contains the \( j\omega \) axis. This axis divides the \( s \)-plane:
- the left hand part, \( \text{Re}[s] < 0 \), corresponds to \( \tau > 0 \)
- the right hand part, \( \text{Re}[s] > 0 \), corresponds to \( \tau < 0 \)
cf. fig. B.9.

As an application of signal characterization in time and frequency domain consider fig. B.10. We want to determine the crosscorrelation function for \( Y \) and \( Z \) as well as the cross power spectral density (ergodic case).

\[
\begin{align*}
\psi_{yz}(\tau) &= \frac{Y(t) Z(t+\tau)}{\mathcal{L}_2[H(t)]} \\
&= \left\{ \int_{-\infty}^{+\infty} g(\theta) X(t-\theta) \, d\theta \right\} \left\{ \int_{-\infty}^{+\infty} h(\xi) X(t+\tau-\xi) \, d\xi \right\} \\
&= \int_{-\infty}^{+\infty} g(\theta) \int_{-\infty}^{+\infty} h(\xi) \, X(t-\theta) X(t+\tau-\xi) \, d\theta \, d\xi \\
&= \int_{-\infty}^{+\infty} g(\theta) \int_{-\infty}^{+\infty} h(\xi) \, \psi_{xx}(\tau-\xi+\theta) \, d\theta \, d\xi
\end{align*}
\]

This represents two convolutions, one after the other. By two-sided Laplace transforms one obtains the cross power spectral density function:

\[
\phi_{yz}(s) = \mathcal{L}_2[\psi_{yz}(\tau)] = \int_{-\infty}^{+\infty} \psi_{yz}(\tau) e^{-st} \, d\tau = \\
= \int_{-\infty}^{+\infty} g(\theta) e^{-s\theta} \int_{-\infty}^{+\infty} h(\xi) e^{-s\xi} \int_{-\infty}^{+\infty} \psi(\tau-\xi+\theta) e^{-s(\tau-\xi+\theta)} \, d\tau \, d\xi \\
= \mathcal{L}_2[g(t)] \mathcal{L}_2[h(t)] \mathcal{L}_2[\psi_{xx}(\tau)]
\]

with

\[
\begin{align*}
\mathcal{L}_2[g(t)] &= \mathcal{L}_2[h(t)] \\
\phi_{xx}(s) &= \mathcal{L}_2[\psi_{xx}(\tau)] \\
g(t) &= h(t) = 0 \text{ for } t < 0 \\
(\text{physical realizability})
\end{align*}
\]

From this derivation the formalism is clear:

\[
\phi_{yz}(s) = \mathcal{L}_2[g(t)] \mathcal{L}_2[h(t)] \phi_{xx}(s)
\]
The transfer function with $-s$ changes the first subscript, the transfer function with $+s$ changes the second subscript of the power spectrum. Fig. B.11 illustrates this procedure for two very simple examples. A corresponding notation $G(s) \rightarrow G(j\omega)$ and $G(-s) \rightarrow G^*(j\omega)$ can be used for the respective Fourier transforms.
fig. B.1

F(x)

discrete

continuous

fig. B.2

p(x)

fig. B.3
Fig. B.4

P(x,y) = 0

P(x,y) = 0.5

P(x,y) = 0.8

Fig. B.5

Stationary process

Non-stationary process
Fig. B.9

Fig. B.10
### Frequency Domain

<table>
<thead>
<tr>
<th>$H(s)$</th>
<th>$s$-plane</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_{xx}(s)$</td>
<td></td>
</tr>
<tr>
<td>$\phi_{xy}(s)$</td>
<td></td>
</tr>
<tr>
<td>$\phi_{yx}(s)$</td>
<td></td>
</tr>
<tr>
<td>$\phi_{yy}(s)$</td>
<td></td>
</tr>
</tbody>
</table>

### Time Domain

<table>
<thead>
<tr>
<th>$h(\tau)$</th>
<th>( h )</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{xx}(\tau)$</td>
<td>&quot;white&quot; noise</td>
</tr>
<tr>
<td>$\psi_{xy}(\tau)$</td>
<td>&quot;colored&quot; noise</td>
</tr>
<tr>
<td>$\psi_{yx}(\tau)$</td>
<td></td>
</tr>
<tr>
<td>$\psi_{yx}(\tau)$</td>
<td></td>
</tr>
<tr>
<td>$\psi_{yy}(\tau)$</td>
<td></td>
</tr>
</tbody>
</table>

**Fig B.11**
Appendix C  SUMMARY OF MATRIX OPERATIONS.

**notions**

vector: \[ \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \]

matrix: \[ A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \]

\( k \times 1 \) matrix \hspace{1cm} m \times n \) matrix

notions:
dimension, elements, row, column, cofactor, trace
diagonal matrix, unit matrix,

**matrix algebra**

**transpose:**

\[ x' = \begin{bmatrix} x_1 & \cdots & x_k \end{bmatrix} \]

\[ A' = \begin{bmatrix} a_{ij} \end{bmatrix} \]

\[ [A']' = A \]

\[ [A + B]' = A' + B' \]

\[ [AB]' = C' B' A' \]

square matrix \( A \):

if \( A' = A \) symmetric;

if \( A' = -A \) skewsymmetric;

if \( A' A = A A' = I \) orthogonal;

if \( A' A = A \) idempotent

**determinant:**

\[ |A| = \sum_{j=1}^{n} a_{ij} \text{ cofactor } (a_{ij}) \hspace{1cm} \text{if } A \text{ } n \times n \text{ matrix} \]

\[ |A| = 0 \Rightarrow \text{matrix singular} \]

\[ |AB| = |A||B| \]

**rank:**

order of its largest non-singular square submatrix;

number of linearly independent row (column) vectors.

\[ \text{rank } C = \text{rank } A \times C \times B \hspace{1cm} \text{if } A \text{ and } B \text{ nonsingular; } \]

\( C \) and \( A \times C \times B \) equivalent matrices.
adjoint:
\[ \tilde{A}^* = \left[ \begin{array}{c} \text{matrix of} \\ \text{cofactors} \end{array} \right] \]
\[ \tilde{A} \tilde{A} = A \tilde{A} = |A| I \]

addition:
associative
commutative

multiplication:
associative \[ A \begin{bmatrix} B & C \end{bmatrix} = \begin{bmatrix} A B & A C \end{bmatrix} \]
distributive \[ A \begin{bmatrix} B + C \end{bmatrix} = A B + A C \]
not commutative, except \[ C A = A C \quad \text{if} \quad C = c I \]
scalar multiplication \[ c A B = A c B = A B c \]
vector dot (inner) product \[ \begin{bmatrix} x' \end{bmatrix} \begin{bmatrix} y \end{bmatrix} \rightarrow \text{scalar} \]
dyadic (outer) product \[ \begin{bmatrix} x \end{bmatrix} \begin{bmatrix} y' \end{bmatrix} \rightarrow \text{matrix} \]

matrix multiplication:
an expression ending with column vector is col. vect. \[ A B x = y \]
an expression beginning with row vector is row vect. \[ \begin{bmatrix} x' \end{bmatrix} B C = d' \]
an expression beginning with row vector and ending with column vector is scalar \[ x' A y = c \]

quadratic forms:
\[ \begin{bmatrix} x_1, x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha x_1^2 + 2bx_1x_2 + Cx_2^2 \]
\[ x' A_1 x \quad \text{quadratic form} \quad \begin{cases} > 0 & \text{posit. definite} \\ < 0 & \text{negat. definite} \end{cases} \]
\[ x' A_2 y \quad \text{bilinear form} \quad \begin{cases} \geq 0 & \text{nonnegat. definite} \\ \leq 0 & \text{nonposit. definite} \end{cases} \]
for each \( x \neq 0 \)

\( A_1 \) can always be made symmetric
i.e. \( A_1 = A_1^* \); if \( Q \) is posit. definite then
\[ \begin{bmatrix} x - x_0 \end{bmatrix}^* Q^{-1} \begin{bmatrix} x - x_0 \end{bmatrix} = c \text{ ellipsoid, center } x_0 \]
\[ |\text{semi-axes}| = \sqrt{\text{charact. roots of } Q} \]
singularity:

\[
A \text{ singular if } Ax = 0 \quad \text{or } A'x = 0
\]

for any \(x \neq 0\)

inverse:

\[
A^{-1}A = I = A A^{-1}
\]

\[
A^{-1} = |A|^{-1} \tilde{A}
\]

\[
[A^{-1}]' = [A']^{-1}
\]

\[
[A^{-1}]^{-1} = A
\]

\[
[ABC]^{-1} = C^{-1}B^{-1}A^{-1}
\]

if \(A' = A^{-1}\) orthogonal;

for \(A' = A, C' = C\):

\[
[A + BCB']^{-1} = A^{-1} - A^{-1}B [C^{-1} + B'A^{-1}B]^{-1} B'A^{-1} = \mathbb{I}
\]

\[
\mathbb{I} = [A + BCB'] [A^{-1} - A^{-1}B [C^{-1} + B'A^{-1}B]^{-1} B'A^{-1}] =
\]

\[
= I - B [C^{-1} + B'A^{-1}B]^{-1} B'A^{-1} + 
+ B C B'A^{-1}
- BCB'A^{-1}B [C^{-1} + B'A^{-1}B]^{-1} B'A^{-1} =
\]

\[
= I + B \left[ C - [C^{-1} + B'A^{-1}B]^{-1} [I + CB'A^{-1}B] \right] B'A^{-1}
= I + BC \left[ I - I \right] B'A^{-1} = I
\]

\[
[CB'] [A + BCB']^{-1} = [C^{-1} + B'A^{-1}B]^{-1} B'A^{-1}
\]

\[
[A + BCB']^{-1} BC = A^{-1} B [C^{-1} + B'A^{-1}B]^{-1}
\]
possible modifications $C \rightarrow -C$

$$A \rightarrow -A$$
$$A \rightarrow -A^{-1}$$
$$C \rightarrow C^{-1}$$
$$B \rightarrow B'$$

characteristic roots:

$$A - \lambda I$$

characteristic matrix; $A$ square

$$|A - \lambda I| = f(\lambda)$$

characteristic function

$$|A - \lambda I| = f(\lambda) = 0$$

characteristic equation

$$\lambda_1, \ldots, \lambda_n$$

characteristic roots; eigen values

$$[A - \lambda_i I]x_i = 0 \Rightarrow x_1, \ldots, x_n$$

characteristic vectors; eigen vectors

$$x_i = \frac{x_i}{|x_i|}$$

$$R = [x_1, x_2, \ldots, x_n]$$

$A$ real symmetric then there is at least one $R$ such that:

$$R'AR = D$$

$$A = RD R'$$

$D$ - diagonal

if $A$ real symmetric $\Rightarrow \lambda_i$ real

if $A$ real symmetric $\Rightarrow \lambda_i > 0 \Rightarrow$ positive definite

if $A$ real symmetric $\Rightarrow \lambda_i \geq 0 \Rightarrow$ non negative definite

similarity transformation:

$$B = T^{-1} AT$$

$A$ and $B$ similar matrices

$$T T^{-1} = T^{-1} T = I$$

$$A = TB T^{-1}$$

$$A^{-1} = TB^{-1} T^{-1}$$

congruence transformation:

$$B = T' AT$$

$A$ and $B$ symmetric

$A$ and $B$ congruent matrices

$B$ can be made diagonal with elements $-1, 0, 1$;

number of $1's$ is rank (invariant)
Orthogonal transformation:

\[ B = T^{-1} A T \]

\[ T^{-1} = T' \implies T \text{ orthogonal matrix.} \]

Symmetric idempotent matrix:

\[ A^2 = A \]

\[ \begin{align*}
    A^2 x_i &= A x_i = \lambda_i x_i \\
    A^2 x_i &= A \lambda_i x_i = \lambda_i^2 x_i \\
\end{align*} \]

\[ \lambda_i = 0 \text{ or } 1 \]

\[ [I - A]^2 = I^2 - 2A + A^2 = I - A \text{ also idempotent.} \]

Generalized inverse:

\[ A^{(-1)} \]

\[ A A^{(-1)} A = A \]

\[ A^{(-1)} = A^{-1} \text{ if } A \text{ square, non singular} \]

\[ A^{(-1)} A = I \implies \text{inverse from the left} \]

\[ A A^{(-1)} = I \implies \text{inverse to the right} \]
matrix calculus

differentiation with respect to scalar:

\[
\frac{dA}{dt} = \left[ \frac{da_{ij}}{dt} \right]
\]
\[
\frac{dx}{dt} = \left[ \frac{dx_1}{dt} \quad \frac{dx_2}{dt} \quad \cdots \quad \frac{dx_k}{dt} \right]
\]
\[
\frac{d}{dt} (A + B) = \frac{dA}{dt} + \frac{dB}{dt}
\]
\[
\frac{d}{dt} (ABC) = \frac{dA}{dt} BC + A \frac{dB}{dt} C + AB \frac{dC}{dt}
\]
\[
\frac{d}{dt} A^n = \frac{dA}{dt} A^{n-1} + A \frac{dA}{dt} A^{n-2} + \cdots + A^n \frac{dA}{dt}
\]
\[
\frac{d}{dt} A^{-1} = -A^{-1} \frac{dA}{dt} A^{-1}
\]
\[
\frac{d}{dt} |A| = \text{trace} \left[ \frac{dA}{dt} \tilde{A} \right]
\]

differentiation with respect to vector:

\[
\frac{\partial}{\partial x} = \nabla_x \overset{\text{def}}{=} \left[ \frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \cdots \quad \frac{\partial}{\partial x_k} \right]
\]

gradient operator

\[
\frac{\partial}{\partial x'} = \nabla_{x'} \overset{\text{def}}{=} \left[ \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k} \right]
\]

scalar c

\[
\nabla_x c = \left[ \frac{\partial c}{\partial x_1} \quad \frac{\partial c}{\partial x_2} \quad \cdots \quad \frac{\partial c}{\partial x_k} \right]
\]

vector; \quad \nabla_{x'} c = \left[ \frac{\partial c}{\partial x_1, \partial x_1} \quad \frac{\partial c}{\partial x_1, \partial x_k} \quad \cdots \quad \frac{\partial c}{\partial x_k, \partial x_1} \quad \frac{\partial c}{\partial x_k, \partial x_k} \right]
vector $\mathbf{y}$

$$\nabla_x \mathbf{y'} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_k} & \cdots & \frac{\partial y_m}{\partial x_k} \end{bmatrix} \quad \quad \nabla_x \mathbf{y} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_k} \end{bmatrix}$$

$$\nabla_x \mathbf{x'} = I$$

$$\nabla_x \begin{bmatrix} \mathbf{y'} \mathbf{x'} \end{bmatrix} = \mathbf{y} = \nabla_x \begin{bmatrix} \mathbf{x'} \mathbf{y} \end{bmatrix}$$

$$\nabla_x \begin{bmatrix} \mathbf{x'} \mathbf{x} \end{bmatrix} = 2\mathbf{x}$$

$$\nabla_x \begin{bmatrix} \mathbf{x'} \mathbf{A} \mathbf{x} \end{bmatrix} = \mathbf{A} \mathbf{x} + \mathbf{A}' \mathbf{x}$$

$$\nabla_x \begin{bmatrix} \mathbf{x'} \mathbf{A} \end{bmatrix} = \mathbf{A}$$

both $\mathbf{z}$ and $\mathbf{y}$ a function of $\mathbf{x}$

$$\nabla_x \begin{bmatrix} \mathbf{y'} \mathbf{z} \end{bmatrix} = \left[ \nabla_x \mathbf{y} \right] \mathbf{z} + \left[ \nabla_x \mathbf{z} \right] \mathbf{y}$$

$\mathbf{A}$ symmetric, non negative definite, not a function of $\mathbf{x}$:

$$\nabla_x \begin{bmatrix} \mathbf{y'} \mathbf{A} \mathbf{y} \end{bmatrix} = \nabla_x \left[ \mathbf{y'} \mathbf{A}' \mathbf{A} \mathbf{y} \right] = 2 \left[ \nabla_x \begin{bmatrix} \mathbf{y'} \mathbf{A}' \mathbf{y} \end{bmatrix} \right] \mathbf{y} \mathbf{y'} = 2 \left[ \nabla_x \mathbf{y'} \right] \mathbf{A} \mathbf{y}$$

where

$$\mathbf{A} = \mathbf{R}' \mathbf{D} \mathbf{R} = \left[ \mathbf{R}' \mathbf{D}' \mathbf{R} \right] \mathbf{D}$$

$$\left| \nabla_x \mathbf{y} \right| \quad \text{Jacobian or functional determinant}$$

stationary value:

scalar $c$

$$\left. \frac{\partial c}{\partial \mathbf{x}} \right|_{\mathbf{x} = \hat{\mathbf{x}}} = 0 \quad \min. \text{ if } \left. \frac{\partial}{\partial \mathbf{x}'} \frac{\partial c}{\partial \mathbf{x}} \right|_{\mathbf{x} = \hat{\mathbf{x}}} \quad \text{positive definite}$$
scalar \( e' e \) with \( e = z - U \beta \)

\[
\frac{\partial e' e}{\partial \beta} = -2 U' e = -2 U' [z - U \beta]
\]

\[
\frac{\partial}{\partial \beta'} \frac{\partial e' e}{\partial \beta} = 2 U' U
\]

matrix power series and exponential:

\[
f(A) \overset{\text{def.}}{=} \sum_{i=0}^{\infty} \alpha_i A^i
\]

\[
e^{At} \overset{\text{def.}}{=} I + At + \frac{[Ae]^2}{2!} + \frac{[Ae]^3}{3!} + \ldots
\]

Cayley-Hamilton \( \Rightarrow \) each matrix polynomial and matrix power series

that converges can be expressed in a linear combination of \( A^n, \ldots, A^0 I \), if \( A \) is an \( n \times n \) matrix. Consequently

\[
e^{At} = \alpha_0 I + \alpha_1 A + \alpha_2 A^2 \quad \text{if} \ A \text{ is a } 3 \times 3 \text{ matrix}
\]

Schwarz's inequality (outline of proof):

\[
\begin{bmatrix} A x + B y \end{bmatrix}' \begin{bmatrix} A x + B y \end{bmatrix} \geq 0
\]

\[
\begin{bmatrix} x' A' A x \quad x' A' B y \quad y' B' A x \quad y' B' B y \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} C & D' \\
C' & D' \end{bmatrix} \begin{bmatrix} x' A' A x + C' D y \end{bmatrix} \begin{bmatrix} x' A' A x + C' D y \end{bmatrix} \geq 0
\]

\[
= 0 \quad \text{if one chooses} \quad x = -C' D y
\]

Consequently

\[
y' \left[ E - D' C^{-1} D \right] y \geq 0
\]

or

\[
[B'B] \geq [B'A] [A'A]^{-1} [A'B]
\]
References


Mann, F. ( ). Introduction to Determinants and Matrices.


Smith, J.S. (1964). Matrix functions and applications, IEEE Spectrum, 1,

I: Matrix operations and generalized inverses. March, 208-220

II: Functions of a matrix

April, 102-108

III: Application of matrices to systems analysis, May, 100-109

IV: Matrix functions and constituent matrices, June, 123-131

V: Similarity reductions by rational or orthogonal matrices, July, 103-109


Appendix $E$. TRANSIENTS IN STATISTICAL AVERAGE AND VARIANCE OF CORRELATION MEASUREMENTS;

Product of two random signals

The system of fig. $E.1$ is considered. This system is slightly more general than is necessary for our purpose. Both $x$ and $y$ are stationary random signals, elements of ergodic ensembles, which are multiplied together. The potentiometer "modulates" the product with a time function $f(t)$ which may, for example, be a step function due to the closing of a switch. (If one or both of the time functions $x$ and $y$ are "modulated" in power, then the time function $f(t)$ can represent that modulation). This modulated product is fed to a block $G$ with an impulse response $g(t)$.

Consequently assuming $f(t) = 0$ for $t < 0$ and $G$ having initial conditions zero:

$$r(t) = \int_0^t f(t - \theta) x(t - \theta) y(t - \theta) q(\theta) d\theta \quad (E.1)$$

If $f(t)$ is a unit step function at $t = 0$ then:

$$r(t) = \int_0^t x(t - \theta) y(t - \theta) q(\theta) d\theta \quad (E.2)$$

If the block denoted by $G$ is a pure integrator then $g(t)$ is a step function as well and by change of variable one finds:

$$r(t) = \int_0^t x(\tau) y(\tau) d\tau \quad (E.3)$$

So both the low pass filter and the pure-integrator case can be represented by the same formula, equation (E.2).

One wants to obtain an expression for the statistical average and for the variance of the output $r$ as a function of time.

a. Statistical average

The statistical average of $r$ is:

$$\mu_r(t) = E [r(t)] = E \left[ \int_0^t f(t - \theta) x(t - \theta) y(t - \theta) q(\theta) d\theta \right] =$$

$$= \int_0^t f(t - \theta) E \left[ x(t - \theta) y(t - \theta) \right] q(\theta) d\theta =$$

$$= \psi_{xy}(\theta) \int_0^t f(t - \theta) q(\theta) d\theta \quad (E.4)$$
If \( f(t) \) is a step function:

\[
f(t) = \begin{cases} 
1 & \text{for } t > 0 \\
0 & \text{for } t < 0
\end{cases}
\]
then \( \mu_r(t) = \psi_{xy}(\theta) \int_{0}^{t} q(\theta) d\theta \) (E.5)

If an integrator is used:

\[
s(t) = \begin{cases} 
1 & \text{for } t > 0 \\
0 & \text{for } t < 0
\end{cases}
\]
then \( \mu_r(t) = \psi_{xy}(\theta) t \) (E.6)

b. Variance

The variance of \( r \) is:

\[
\sigma^2_r(t) = \mathbb{E} \left[ \left( r(t) - \mu_r(t) \right)^2 \right] = \mathbb{E} \left[ r^2(t) \right] - \mu_r^2(t) \quad \text{(E.7)}
\]

This can be written as follows:

\[
\sigma^2_r(t) = \mathbb{E} \left[ \int_{0}^{t} f(t-\theta) x(t-\theta) q(\theta) d\theta \right] + \mathbb{E} \left[ \int_{0}^{t} f(t-\theta) y(t-\theta) q(\theta) d\theta \right] + \mathbb{E} \left[ \int_{0}^{t} f(t-\theta) y(t-\theta) q(\theta) d\theta \right] - \left\{ \psi_{xy}(\theta) \int_{0}^{t} f(t-\theta) q(\theta) d\theta \right\}^2 =
\]

\[
= \int_{0}^{t} \int_{0}^{t} f(t-\theta)f(t-\xi) \mathbb{E} \left[ x(t-\theta) y(t-\xi) \right] q(\theta) q(\xi) d\theta d\xi + \mathbb{E} \left[ \int_{0}^{t} f(t-\theta) q(\theta) d\theta \right]^2 \quad \text{(E.8)}
\]

Now we have to determine a fourth-order moment. Consider the moment generating function \( P(s) \) defined by:

\[
P(s) = \mathbb{E} \left[ \exp s x \right] =
\]

\[
= \iiint_{-\infty}^{+\infty} \exp \left\{ s_1 x_1 + s_2 x_2 + s_3 x_3 + s_4 x_4 \right\} p(x) dx_1 dx_2 dx_3 dx_4 \quad \text{(E.9)}
\]

where:

\[
S = \begin{bmatrix} 
  s_1 \\
  s_2 \\
  s_3 \\
  s_4
\end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 
  x_1 \\
  x_2 \\
  x_3 \\
  x_4
\end{bmatrix}
\]
and \( p(x) \) is the four dimensional probability density function of the random process under consideration.

Differentiation of \( P(s) \) with respect to \( s_1, s_2, s_3 \) and \( s_4 \) results in the type of fourth-order moment that we need:

\[
\frac{\partial^4 P(s)}{\partial s_1 \partial s_2 \partial s_3 \partial s_4} = \mathbb{E}[x_1 x_2 x_3 x_4]
\]

Let us assume that the random variables have zero mean values and that joint probability density function is Gaussian:

\[
p(x) = \frac{1}{(2\pi)^2 |M|^{1/2}} \exp \left\{ -\frac{1}{2} x'M^{-1} x \right\}
\]

with

\[
M = \begin{bmatrix} \Psi_{11} & \cdots & \Psi_{1n} \\ \vdots & \ddots & \vdots \\ \Psi_{n1} & \cdots & \Psi_{nn} \end{bmatrix} \quad \text{and} \quad \Psi_{ij} = \mathbb{E}[x_i x_j] = \psi_{ji}
\]

Consequently its moment generating function can be found to be:

\[
P(s) = \frac{1}{(2\pi)^2 |M|^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} x'M^{-1} x + \frac{i}{2} x's + \frac{1}{2} s'x \right\} dx_1 dx_2 dx_3 dx_4
\]

By completing the square:

\[
\frac{1}{2} \left\{ x'M^{-1} x - x's - s'x \right\} =
\]

\[
= -\frac{1}{2} \left\{ x'M^{-1} x - s'M^{1/2} \right\} \left( M^{1/2} x - M^{1/2} s \right) - s'M s
\]

and taking into account the properties of integrals of probability density functions, the resulting expression is:

\[
P(s) = \exp \left\{ \frac{i}{2} s'M s \right\}
\]

Differentiation according to eq. (E.10) leads to:

\[
\mathbb{E}[x_1 x_2 x_3 x_4] = \psi_{12} \psi_{34} + \psi_{13} \psi_{24} + \psi_{14} \psi_{23}
\]
For the case under consideration one finds:

\[
\mathbb{E} \left[ x(t-\theta) y(t-\varepsilon) x(t-\xi) y(t-\zeta) \right] = \\
= \psi_{xy}^2(\varepsilon) + \psi^2_{xx}(\theta-\zeta) \psi_{yy}(\theta-\xi) + \psi_{xy}(\theta-\xi) \psi_{yx}(\theta-\zeta) = \\
= \psi_{xy}^2(\varepsilon) + R(\theta - \xi)
\]

where the term \( R(\theta - \xi) \) is substituted for the sake of convenience. As

\[
\psi_{xx}(-\nu) = \psi_{xx}(\nu) \quad \psi_{yy}(-\nu) = \psi_{yy}(\nu) \\
\psi_{xy}(-\nu) = \psi_{yx}(\nu)
\]

it follows that \( R(-\nu) = R(\nu) \). As the first term \( \psi_{xy}^2(\varepsilon) \), gives rise to

\[
\int_0^t \int_0^\varepsilon f(t-\theta) f(t-\xi) \psi_{xy}^2(\varepsilon) q(\theta) q(\xi) \, d\theta \, d\xi = \\
= \left\{ \psi_{xy}(\varepsilon) \int_0^\varepsilon f(t-\theta) q(\theta) \, d\theta \right\}^2
\]

this term cancels the corresponding one in equation (E.8). Thus a substitution of eq. (E.13) in eq. (E.8) gives:

\[
\sigma_r^2(t) = \int_0^t \int_0^{t-\xi} f(t-\theta) f(t-\xi) R(\theta - \xi) q(\theta) q(\xi) \, d\theta \, d\xi
\]

By putting \( \theta - \xi = \nu \) and by dividing the region of integration of fig. E.2 into two parts according to fig. E.3 one obtains

\[
|\text{Jacobian}| = 1
\]

For \( I_1 : 0 \leq \xi \leq \theta \leq t \) \quad \Rightarrow \quad 0 \leq \xi \leq (\xi + \nu) \leq t

For \( I_2 : 0 \leq \theta \leq \xi \leq t \) \quad \Rightarrow \quad 0 \leq \theta \leq (\theta - \nu) \leq t

Thus the variance as a function of time can be written as the sum of two integrals:
\[ \sigma_r^2(t) = \int_0^t d\nu R(\nu) \int_0^{t-\nu} f(t-\xi) g(\xi) \, d\xi + \int_0^t d\nu R(\nu) \int_0^{t+\nu} f(t-\theta) g(\theta) \, d\theta \]

As \( R(-\nu) = R(\nu) \), both integrals are equal and:

\[
\sigma_r^2(t) = 2 \int_0^t d\nu R(\nu) \int_0^{t-\nu} f(t-\nu-\xi) f(t-\xi) g(\nu+\xi) g(\xi) \, d\xi
\]

with

\[ R(\nu) = \psi_{pp}(\nu) \psi_{qq}(\nu) = \psi_{pq}(\nu) \psi_{qp}(\nu) \]

If \( f(t) \) is a step function:

\[
f(t) = \begin{cases} 
1 & \text{for } t > 0 \\
0 & \text{for } t < 0 
\end{cases}
\]

then

\[
\sigma_r^2(t) = 2 \int_0^t d\nu R(\nu) \int_0^{t-\nu} f(t-\nu-\xi) g(\nu+\xi) g(\xi) \, d\xi
\]

If an integrator is used then

\[
g(t) = \begin{cases} 
1 & \text{for } t > 0 \\
0 & \text{for } t < 0 
\end{cases}
\]

then

\[
\sigma_r^2(t) = 2 \int_0^t (t-\nu) R(\nu) \, d\nu
\]

c. Discussion of these results.

First it has to be noted that the results hold in general for stationary random processes for which one can determine \( \psi_{xy}(0) \) and the fourth order moment

\[
E \left[ x(t-\theta) y(t-\theta) x(t-\xi) y(t-\xi) \right].
\]

As an example a Gaussian process has been taken.

Secondly it has to be mentioned that although the time-delay parameter \( \tau = 0 \) has been chosen this does not restrict the results.

If one wishes to determine \( \mu_r(t) \) and \( \sigma_r^2(t) \) for another argument say \( \tau = \tau' \), then this can be obtained by introducing a virtual time delay element in the process-generating block-diagram; fig. E.4.
Finally it can be seen that the formulas that have been derived can also be used for the study of the transients in statistical average and variance for the circuit shown in fig. E.5, where a stationary signal is modulated by a time function $f(t)$ and then fed to a linear operator with $g(t)$ as impulse response.

**Product of a sinusoidal and a random signal**

Now the system under consideration is given by fig. E.6. Consequently we restrict ourselves to a discussion of the influence of the additive noise. As in chapter 10 the integration interval is restricted to $(0, T)$ with $T = k T_1$, $k$ integer and $T_1 = 2\pi/\omega_1$.

a. **Statistical average**

In chapter 10 it was mentioned already that the noise does not influence the value of the statistical average due to the independence of signal and noise.

b. **Variance**

In chapter 10 the expressions for the variance were found to be:

$$\sigma_{c}^2(T) = \frac{4}{u_0^2 r_1^2} \int_0^T \int_0^T \psi_{nn}(\theta - \tau) \left\{ \begin{array}{c} \sin \omega_1 \tau \\ \cos \omega_1 \tau \\ \sin \omega_1 \tau \end{array} \right\} \left\{ \begin{array}{c} \sin \omega_1 \theta \\ \cos \omega_1 \theta \\ \sin \omega_1 \theta \end{array} \right\} d\tau d\theta$$

These expressions can be reworked like equation (E.15). Again by substitution of $\theta - \tau = \nu$ and by dividing the region of integration of fig. E.2 into two parts according to fig. E.3 (with $t = T$) one obtains:

$$\sigma_{c}^2(T) = \frac{4}{u_0^2 r_1^2} \int_0^T \int_0^T \psi_{nn}(\nu) \left[ \int_0^T \sin \omega_1 \tau \sin \omega_1 (\nu + \tau) d\tau + \int_0^T \int_0^T \psi_{nn}(\nu) \sin \omega_1 (\nu - \tau) \sin \omega_1 \theta d\tau d\theta. \right.$$
As \( \psi_{mn}(\xi, \nu) = \psi_{mn}(\nu) \) both integrals are equal and

\[
\sigma_{\xi}^2(T) = \frac{\partial}{\partial T} \int_{0}^{T} \frac{1}{\nu^2} \psi_{mn}(\nu) \left[ \sin \omega \nu \cos \frac{\nu(T-\nu)}{2} \sin \omega (\xi + \nu) \right] d\nu
\]

where

\[
\sin \omega \nu \sin \omega (\xi + \nu) = \frac{1}{2} \cos \omega \nu - \frac{1}{2} \cos \omega (2\xi + \nu)
\]

which leads to

\[
\sigma_{\xi}^2(T) = \frac{4}{\nu^2} \int_{0}^{T} \frac{1}{\nu^2} \psi_{mn}(\nu) \left[ \left( T - \frac{1}{2} \nu \right) \cos \omega \nu + \frac{1}{\omega^2 \nu} \sin \omega \nu \right] d\nu
\]

as given in eq. (10.24). In the same way the expressions for \( \sigma_{\xi}^2(T) \) and \( \sigma_{cd}^2(T) \) can be derived.
As mentioned already in section 1.6 the parameter and state estimation techniques can be used:

- in different ways of application, e.g.
  - "diagnostic" applications
  - control application
  - automatic (industrial) measurements
  - automatic (industrial) decisions
  - automatic (industrial) adjustments
  - pattern recognition

- in different fields of human endeavour, e.g.
  - communication engineering
  - power engineering
  - mechanical engineering
  - aeronautical engineering
  - chemical engineering
  - physics
  - geology
  - economy
  - biology/medicine.

In applications directed to automatic control there is the problem how the estimation procedures fit into that control problem. This is briefly discussed in section 14.1. The remaining part of the chapter is devoted to particular fields of applications, illustrated by specific examples. These are taken from the fields of:

- physical/mechanical/chemical processes
- nuclear reactors
- power plants and power distribution systems
This list of "fields" does not constitute a unique division of types of applications, nor are the examples given necessarily the most important or the most spectacular ones. By use of the literature references the picture of applications can be adapted to the interest of the reader. With this chapter the list of literature is divided according to the topics of the sections.

14.1 Automatic control

In chapter 1 it was stressed already that the application of feedback has been recognized as an effective expedient for combatting uncertainties, but that there are limitations:

- very strict requirements for optimization, leading to optimal or self-optimizing control;

- very large parameter variations, leading to the use of adaptive control.

In such cases the need for very good or for permanently updated knowledge of the process under control is of crucial importance. In section 1.4 it was mentioned that Fel'dbaum's dual control principle implies two functions, performed simultaneously:

- to study or learn the characteristics of the process;

- to direct the process to the required state.

It is unfortunate that such a coherent treatment of the problems at hand leads to a theory that as yet results in exorbitant computational problems, even for simple cases; cf. Fel'dbaum (1960, 1961; 1965), Mendes, 1970; cf. also Snegur, Tomashov and Dubina (1967).

Consequently attempts have been made to simplify this general problem. This is being done by an a priori assumption that the design of a control scheme can be divided into two parts: identification and control. In analogy with the theory of stochastic control we refer to this assumption as the separation hypothesis. The approach is very natural, in particular if we consider
the multitude of techniques which have been developed for the design of systems with known process dynamics and known environments. However, it is seldom true that optimum solutions are obtained if a process is identified and the results of the identification are used in a design procedure, developed under the assumption that the process and its environment are known precisely. It can be necessary to modify the control strategy to take into account the fact that the identification is not precise.

It seems to be a worthwhile problem to investigate rigorously under what conditions the separation hypothesis is valid. Initial attempts in this direction have been made by Åström and Wittenmark (1969).

Apart from the obvious fact that it is desirable to choose a class of model \( \mathcal{J} \) for which there is a control theory available, there are also many other interesting questions in the area of identification and control, e.g.:

- Is it possible to obtain rational choices of model structures and criteria for the identification if we know that the results of identification will be used to design control strategies?
- What "accuracy" is required of the solution of an identification problem if the separation hypothesis should be valid at least with a specified error?

Partial answers to these questions are given by Åström and Wittenmark (1969) for a restricted class of problems.

Many of the methods proposed rely on estimation techniques. Also, however, it is possible to mechanize the self-optimizing and (self-)adaptive principles in a direct way. As a typical example we mention the on-line hunting for the optimum or the on-line tuning of PID controllers. In such cases the introduction of testsignals on the quantities of interest (setpoints, viz. coefficients) may be needed in order to determine the direction of the optimum.

One of the classes of adaptive systems is indicated by the name "model-reference" approach. One of the early contributions along this line was by Whitaker et al (1958). This method can be illustrated for an airplane control system indicated in fig. 14.1. At different altitudes, due to the difference in air-density, the effectiveness of the rudders change. This results in a change of the airplane dynamics and consequently in the command properties of the plane. On the pilot commands the (second order) model gives a response as it is desired for the airplane as a whole. By automatic adjustment of a gain factor $\kappa$ the loop gain of the control system is made as high as permissible, such that this closed loop forms a "stiff" system. The adjustment of the maximum permissible gain is done by detection of the small limit cycle. If no limit cycle is detected, then $\kappa$ is increased, if the amplitude of the limit cycle gets too large then $\kappa$ is diminished. Consequently no use of estimation techniques is needed.

Now we consider two simple examples of control systems in which use is made of estimation techniques for adaptive purposes. The first one is a simple servo system as described by Young (1965, 1969). The principle is indicated in fig. 14.2.
The purpose of the block "estimation scheme" is to determine the coefficients of the transfer function

\[ H(s) = \frac{1}{a_2 s^2 + a_1 s + a_0} \]

Based on these estimates the transfer function \( G(s) \) of an adaptive network is adjusted in such a way that, as well as possible, \( G(s) = H^{-1}(s) \). If that could be realized completely then the characteristics of the feedback system are determined by the model, which can be chosen properly. For the estimation scheme analog filters were used that enable one to derive a signal and its derivatives without using pure differentiators, cf. fig. 14.3. and section 9.2. By sampling and analog/digital conversion a set of simultaneous equations in the three unknown parameters is made, from which the estimate \( \alpha \) for \( a = [a_2, a_1, a_0] \) can be determined by explicit methods.

The second example of a control system in which use is made of estimation techniques is taken from Rumold and Speth (1968), cf. also Speth (1969, 1971). In the case described the process consists of an electric speed control system with the transfer functions indicated in fig. 14.4:

\[
\frac{1}{1 + s \tau_2} \quad \text{viz.} \quad \frac{1}{s \tau_1}
\]

where \( \tau_1 \) is fixed by the moment of inertia and \( \tau_2 \) is the time constant of a torque control unit. Both the parameters \( \alpha_1 = \tau_1 \) and \( \alpha_2 = \tau_1 \tau_2 \) are estimated by an implicit (model adjustment) method. In order to prevent differentiation again use in made of analog filters that provide a signal and its derivative; cf. fig. 14.5 where \( \alpha_1 \) and \( \alpha_2 \) are adjusted according to the gradient method based on an integral squared error criterion. (In order to eliminate the d.c. component in e a low pass filter is included). Based on these parameter values the controller can be adjusted by using explicit rules, based on a "generalized
symmetrical optimum", as referred to in the cited publications. It has been demonstrated that this type of scheme works satisfactorily; it has been implemented in practical applications. Cf. also Spiegel (1969).

In concluding this section it is worth mentioning the different structures of self-adaptive systems according to fig. 14.6 as given by Speth (1971). Among the differences and resulting properties of these schemes there are:

1) first column - estimation of the closed-loop-characteristics.
   The estimation results can be made independent of the disturbances, if those disturbances are not related to the system input function.

2) second column - estimation of the forward path.
   If the control loop to be adjusted is unstable then the estimation is fast due to the large signal amplitudes. Under ordinary conditions, however, one of the signals used for the estimation will be small and the estimation will be slow. Disturbance signals will influence the estimation results due to the "circulating noise".

a) first row - adaptation inside the part being estimated.
   As the controller $C$ belongs to the part being estimated it is not necessary to determine completely the unknown parameters; it is sufficient to test whether they are bigger or smaller than the desired value.

b) second row - adaptation outside the part being estimated.
   A fast change of the adaptive element does not interfere with the estimation procedure.
14.2 Physical/mechanical/chemical processes.

Much has been written on the choice, derivation, use and evaluation of models for different purposes. Introductory remarks on the use of industrial models.

- for stabilizing control and
- for optimizing control

together with some examples of dynamic and static models by Van der Grinten can be found in the survey paper Eykhoff, Van den Grinten, Kwakernaak and Veltman (1966).

There is an abundant literature on modelling of the great variety of industrial processes. Some general references: Stout (1961), Murrill et al. (1969) and IEE (1969). Again it is far beyond the scope of this section to try to provide a coherent and complete picture of the many applications. In addition to the references mentioned explicitly the reader is referred to the additional literature mentioned.

An interesting table listing experiments that were made using pseudo-random binary time functions as perturbation signals and using correlation techniques was given by Godfrey (1970). A detailed discussion of such techniques to an oil-refinery can be found in Godfrey (1969).

As specific applications can be referred to:

- distillation column: Rees (1966)


- water gas shift reactor: Wise, Price and Rippin (1967), Price and Rippen (1968),
As an example we consider the topic of the last-mentioned paper. Schematically a gas chromatograph can be represented by fig. 14.7. It consists of the following parts: 1) a cylinder filled with gas, e.g. hydrogen; 2) a control value which regulates a constant flow of gas; 3) an injection device by which the sample to be tested is brought into the gas stream; 4) the "column", a tube in which the separation of the sample components takes place. This column is enclosed in a thermostat; 5) the detector, measuring some physical property of the gas, e.g. the heat conductivity; 6) the recorder.

The very small sample, mostly liquid, evaporates in the carrier gas flow. Due to the physical properties of the column the sample molecules are delayed with respect to the carrier gas; this delay is specific for each kind of molecule. Consequently at the end of the column the components of the sample are detected as a sequence of peaks in the recorded chromatogram; cf. fig. 14.8. The time delay since the time of injection (retention time) $t_R$ is characteristic for a component, the area $A$ under the peak provides a measure for the quantity of that
component in the sample. Also the width (standard deviation $\sigma$) of the peak is of importance.

Although the attention is focussed on the properties of the input sample, these properties can also be attributed to the process (measuring device) which causes the separation. Consequently the problem can be formulated as that of estimating (for each peak) the parameter vector $\beta = [t_R, A, \sigma]^T$. In practice the chromatogram may be disturbed by noise, e.g. due to pollution in the carrier gas, changes of temperature, electrical noise of detector and amplifier. A technique described in the reference cited, assuming the noise to be stationary and additive, is: estimating the properties of the noise when there is no measurement signal; filtering the signal by a "noise-whitening filter"; estimation of $\beta$.

14.3 Nuclear reactors

Also in this field a variety of types of estimation schemes has been proposed and performed. As few specific examples are cited:

- the measurement of the under-critical (neutron) multiplication factor, e.g. Bastl (1966). This may be approached by introducing a neutron source which is modulated by a pseudo-random binary sequence and by correlation of that (delayed) signal with measurements of the neutron density level.

- the determination of kinetic parameters of the reactor from correlation measurements based on the statistical fluctuations of the neutron chain reaction; cf. Wilkie (1967); Liewers and Butler (1967) who use the polarity correlation technique; Roman, Hsu and Habegger (1970) who applied quasi-linearization.

- parameter estimation of coupled reactors, cf. Hendrikson and Murphy (1968), Schwalm (1971),
o state estimation, where the time-dependent reactivity is considered to be
one of the state variables, using Kalman Bucy filter theory; cf. Venerus
and Bullock (1970), or combined parameter and state estimation by means

14.4 Power plants and power distribution systems.

o In order to deal with frequent load changes in an efficient way much know-
ledge and even parametric control may be necessary in power plants. Some
examples of literature on estimation problems: Stanton (1965), Debelle,
Foureau and Vazquez (1966), Thompson (1967).

o For the load dispatching problem use can be made of state (and parameter)
estimation, as it is discussed in e.g. Habegger and Barley (1969), Debs and

o For the adjustment of load frequency controller parameters one needs to
know the transferfunction $G_f$ as indicated in fig. 14.9, where $u$ is the
control signal of the generating unit; $P_g$, $P_c$ and $P_e$ are the generated, the
consumed and the "exported" power; and where $f$ is the frequency. Cf.

14.5 Telecommunication.

It is quite evident that there is an intimate relationship between many
problems in the telecommunication area and the field of parameter and state
estimation, e.g. the dominating role of disturbances (noise), causing
fundamental limitations to the possible achievements. Such relationships are
clear in the problem of separating signal and noise, in estimating para-
eters of signals (e.g. distance and velocity of the object in radar observ-
Some special applications are indicated briefly.
The attenuation and phase characteristics of transmission lines change with length, type of cable, temperature, humidity, etc. Equalizers are being used to compensate as well as possible for undesirable characteristics and for changes that may occur in those properties. Variable equalizers have been discussed already by Bode (1938).

Automatic equalization has become a popular subject of research and application. Many papers have been written on adaptive filters that adjust automatically to the (changes in) characteristics of the transmission line in such a way that e.g. the intersymbol interference is minimized. The techniques used can be considered as model-adjustment procedures where the "model" (equalizer) is placed in series with the changing object, cf. fig. 14.10. A popular choice of equalizer is a tapped delay line with an adjustable attenuator at every tap. A survey of automatic equalization is given in Kanal (1969); many papers on this topic are mentioned as additional literature.

Adaptive filters used as frequency-selective limiters; cf. Orth (1968).


Adaptive antenna systems; cf. Widrow et al. (1967).

Estimation of the characteristics of multipath communication channels; cf. Goldman (1968).

Estimation or measurement of electronic components by self-adjusting bridges (cf. Mulders and Lückers, 1965) or by the model-adjustment technique (cf. fig. 1.7).
14.6 Air and space vehicles.

In section 14.1 control-type of applications have been indicated that have been used on high performance airplanes. Such adaptive systems may be directed towards compensation of:
- variations of vehicle dynamics, i.e. providing a closed loop control system with good stability and maneuvering properties in spite of drastic changes that the open loop characteristics of the vehicle may undergo;
- external disturbances due to atmospheric turbulence and outer loop command signals. A survey of such applications to aerospace vehicles written by R.K. Smyth and H.L. Ehlers can be found in the reference: AGARD 1970.
In the same reference a number of papers on specific adaptive control schemes can be found. Of the host of literature only a few are mentioned under additional references.

For determining unknown air and spacecraft parameters several techniques have been applied, e.g.:
- an iterative weighted least squares estimation of satellite dynamics (Farrell et al., 1969).
- estimation of aircraft parameters by quasilinearization techniques (Larson and Fleck, 1969) and by maximum likelihood methods (Mehra, 1970).

In line with the earliest application of least squares estimation techniques by Gauss for determining celestial state variables, these types of procedures have been used extensively for orbit calculations of space vehicles. Cf. the literature cited in chapter 12.
14. 7 **Biological objects**

Due to the complexity of living creatures, the difficulty of measuring crucial quantities in a direct way and the emphasis on more accurate diagnoses as well as treatment of diseases, the estimation techniques may be regarded as (potentially) valuable aids. This development is closely connected to the recognition of system/control aspects in biological objects, as is exemplified by a number of books, e.g. Bayliss (1966), Milsum (1966), Mesarovic (1968), Stark (1968).

**Human operator.** For many years a very popular object of study has been the human being as an element of a control system; cf. fig. 1.5. One of the early contributions is Tustin (1947). The interest in identification and estimation studies on such human operators stems primarily from assignments of controlling various kinds of land-, water-, air-, space- and lunar vehicles. Crucial issues in such studies are: the difficulty of the assignment, the variations in time of the dynamic properties due to learning and fatigue, the adaptive capability shown in reactions to sudden changes in the vehicle dynamics, e.g. due to failure, etc. Very much attention has been paid to these types of problems; early surveys are: Clymer and Ax (1960), Clymer (1960) and Cosgriff and Briggs (1960). An extensive bibliography has been published by Costello and Higgins (1966); many interesting contributions are published in the proceedings of the annual NASA-university conferences on manual control (NASA).

The wide variety of papers published is caused by the many aspects that can be discerned, e.g.

- the type of assignment: pursuit or compensatory, cf. fig. 14.11;
  - single or more axes (degrees of freedom).
- the type of information desired: impulse response;
  - transfer function coefficients, either constant or stochastically time varying;
  - nonlinearities; sampling frequency.
- the type of input characteristics: deterministic multi-frequency signal; stochastic signal with prescribed probability density and power spectrum curves.
- the process dynamics; unstable; nonlinear; sudden changes/failures.
- the type of signal preprocessing, e.g. quickening, prediction.
- influence of the environment on the human operator
- the type of estimation techniques used.

Only a very limited number of publications is mentioned under additional literature.

The fact that the human operator is embedded in closed loop, cf. fig. 14.12, results in the crucial problem of circulating noise, i.e. the signals \( n \) and \( e \) are correlated. Consequently the use of only the signals \( e \) and \( x \) may lead to biased estimates. One of the methods to deal with this problem is to have recourse to e.g. the technique where instrumental variables are derived from the signal \( u \).

Most NASA studies on the human operator have been directed to highly qualified and trained men in simulations of extraordinary missions. We will briefly summarize an interesting experiment directed to the man-in-the-street, viz. a bicycle rider. This description is based on private communication\(^\text{1})\) and Stassen (1969). The experimental set up is shown in fig. 14.13. The simulator consists of a bicycle frame at a fixed position on which forces can be applied by an hydraulic control system. The task of the test person is to "ride" this bicycle, i.e. to keep the inherently unstable system in balance. The hydraulic system takes care of the natural forces that occur due to the gyroscopic moment of the front wheel and the centrifugal moment of the bike when riding a curved path.

The sign conventions defined by fig. 14.14 and the following notations are used:

\[
\begin{align*}
\theta & \text{ angle between handlebar and frame} \\
\beta & \text{ angle between upper (part of the) body and frame} \\
\phi & \text{ angle between 0-z axis and frame}
\end{align*}
\]

\(^1\) Delft University of Technology, Dept. of Mechan. Engng, lab. for Measurement and Control Engineering; R.G. Boiten, H.G. Stassen, A. van Lunteren; Annual report 1969 of the Man-Machine Systems Group
\[ \text{moment of inertia} \] of bike and lower (part of the) body with respect to the 0-x axis
\[ M \] static moment
\[ J_\bot \] moment of inertia of the upper (part of the) body with respect to the 0-x axis
\[ \mu \] static moment
\[ z, \] distance between 0-x axis and 0-x axis
\[ v \] forward velocity
\[ A, B \] constants

With a few simplifying assumptions the moments around the 0-x axis can be equated as follows:

- gyroscopic moment due to the front wheel
  \[ A \ddot{v} \dot{y}(t) + B v^2 \ddot{z}_2(t) = \]

\[ = \left\{ \dot{y}(t) - M \ddot{y}(t) \right\} + \left\{ (J_\bot + z, \mu) \ddot{y}(t) - \mu \ddot{y}_2(t) \right\} \]

- centrifugal moment when riding a curved path
- moment due to a deviation of the frame and the lower part of the body from the equilibrium position
- moment due to a deviation of the upper part of the body with respect to the frame

The diagram of fig. 14.15 represents the experimental set up as a block diagram.

By means of a test-signal course the human operator is given an assignment; the course deviation \( E(t) \) acts as an error signal.

In this experimental situation the interest is focused on the dynamic properties of the human operator. As models for the human operator are chosen:

- for the relation of frame rotation \( \rightarrow \) handlebar rotation:
  \[ H(t) = (\alpha_{11} + \frac{\alpha_{12}}{s} + \alpha_{13} s) e^{-\tau_s} \]
Use is made of a model-adjustment technique by means of a small digital computer. Again one has to be aware of the bias problems due to circulating noise. One of the lines of research is the testing of the effects of drugs and alcohol on the different parameters.

**Blood circulation.** Numerous attempts have been made to build (electrical) models of (parts of) the (human) blood circulation. The cardiovascular system is a distributed system; the dynamic behaviour of which has to be described by partial differential equations. Different kinds of approximations by lumped circuits have been studied; Noordergraaf (1963), Beneken (1965), De Pater (1966), e.g. by simulations on an analog computer.

One of the objectives of such studies may be the wish to determine parameters of the blood circulation that are difficult to measure in a direct way, e.g. the compliance of the arteries or veins. According to fig. 14.16 the analog computer model may be adjusted by a digital computer in such a way that the measurements from the model compare as well as possible with those from the patients. Under the additional literature a few publications are cited.

**Electric heart action.** From the electrical phenomena of the heart muscle (depolarisation wave) the resulting electrocardiogram (e.c.g.) can be measured at the skin of the body. Given an adequate model of the heart and the properties of the surrounding parts of the body the e.c.g. can be simulated. Such a model may consist of say \( m \) current dipoles, assigned to the different segments of the heart muscle, where each dipole is characterized by the time function of its intensity. If such a time function can be characterized by \( n \) parameters, then the whole model contains \( m \times n \) parameters.

A crucial problem now is whether it is possible to determine those parameters from the measured e.c.g. If that were the case then the location and severity...
of heart defects could be determined by estimation, viz. model-adjustment
techniques. Relevant references are: Bellman et al. (1964), Damen and Piceni
(1971).

Of the other applications to biological objects we mention only the model building:
o with respect to muscles; Bahler (1968); Gottlieb, Agarwall and Stark (1969)
o with respect to the human lens accommodation; cf. O'Neill, Snathanan and

14.8 Economical systems
For many years already econometricians have been working on the development of
models of essential parts of national economies and other sectors of economic
activities. Again it is outside the scope of this text to discuss examples of
these projects. Only a few general references to pertinent literature may

Extensions of these techniques have been made to industrial dynamics, urban
dynamics and world dynamics (Forrester 1961, 1969, 1971)

14.9 Learning systems and pattern recognition
The reader is reminded of the close relationships that there are between the pa­
rameter estimation procedures and the problems that are discussed under the terms
learning systems and pattern recognition. Among those relationships there are the
choice of model structures and the ways in which weighting coefficients (para­
meters) may be adjusted. Again only a few publications are mentioned.

Due to the nature of this chapter it contains either too much or far too little,
depending on the interests of the reader. It is only meant to demonstrate the wide
variety of goals and means that have been proposed and used. The quite extensive
list of literature can be instrumental in providing more information. The reader is reminded of the possibilities offered by the Citation Index in finding more recent publications that refer to a particular paper.

Originated by definite needs in particular areas, stimulated by the potential possibilities and the demonstrated capabilities, and supported by the decreasing costs of information processing equipment the field of parameter and state estimation is developping rapidly. It can clearly be recognized as a means for improving human well-being in many respects; eg. by economic gains, by increased safety of airplanes through applying identification techniques for maintenance, by providing better models for socio-economic processes, by contributing to possibilities of mass health screening, etc.
Literature

Please note: as this chapter is divided into well-defined quite homogeneous sections so is the literature. For each section first the references and then the additional literature is given.

(14.1) Automatic control

References


Fel'dbaum, A.A. (1960, 1961; 1965) cf. references Chapter 1.

Gibson, J.E. (1960), Mechanising the adaptive principle. Control engng., 7, 109-


Kalman, R.E. (1958). Design of self-optimalizing control systems. Trans. ASME, 80,


(14.2) Physical/mechanical/chemical processes

References


Paiziev, E. (1968). Identification of complex object of control by means of partial correlation and dispersion functions of random functions. Automation and remote control,


Rocco, V. (1968). Building a dynamic model to simulate in real time the evolution of a rotating cement kiln. (in Italian) Automaz. e Strum., 16, 146-152.


(14.3) Nuclear reactors

References


Toyoda, J. et al. (1970). An application of state estimation to short-term load forecasting. *IEEE Trans. on power apparat. and systems*, PAS-89, 1678-


(14.6) Air and space vehicles

References

dynamic coefficients determined from ex-
perimential data. Journal of Spacecraft and
Rockets, 7, 1265-1267.

linearization techniques. Proceedings Nat.

Mehra, K. (1970). Maximum likelihood ident-
fication of aircraft parameters. Joint
Automatic Control Conference,
Atlanta, Ga. 442-444.

of adaptive control application to aero-
space vehicles. AGARD Advanced Control
System Concepts, 1-12.

Advanced Control System Concepts.

identification in linear continuous vibratory
systems. Joint Automatic Control Conference,
11th, Atlanta, Ga., 556-564.

(1970). The place of adaptive control techni-
quies in the future development of automatic
flight control systems. AGARD Advanced Control
System Concepts, 87-100.

Farrel, J.L., J.K. Newton, J.A. Miller and J.J.
Connelly. (1969). Optimal estimation of rotation-
coupled flexural oscillations. Joint Automatic
Control Conference, Boulder, Colo.
433-441.

regulation technique for self-adaptive control
of flexible aircraft. AGARD Advanced Control
System Concepts, 121-138.

adaptive manoeuvre demand control systems:
some practical design considerations. AGARD
Advanced Control System Concepts, 149-168.


(14.7) Biological objects

References


Pollack, G.H., R.V. Reddy and A. Noordegraaf


Watt, T.B., Jr. and R.M. Goldwyn (1966). Clinical estimation of human vascular properties via a mathematical model describing the contour of the arterial pressure pulse. 19th Annual Conf. on Engng. in Med. & Biol., San Francisco, 8, 182-

Westerhof, N. and A. Noordergraaf (1966). Wave travel and input impedance for various types of networks in an electrical model of the human systemic arterial tree, with and without elastic tapering. 19th Annual Conf. on Engng. in Med. & Biol., San Francisco, 8, 90-


estimation of the characteristics of the
Closed loop
forward path

adaptation

inside the part that is being estimated

outside the part that is being estimated
Fig. 14.7

Fig. 14.8

Fig. 14.9

Fig. 14.10
fig. 14.11.

fig. 14.12.

fig. 14.13.
Fig. 14.16