

Algoritme voor het schatten van de parameters in ARMA-modellen met meetfouten op de in- en uitvoer

Citation for published version (APA):

Vregelaar, ten, J. M. (1985). *Algoritme voor het schatten van de parameters in ARMA-modellen met meetfouten op de in- en uitvoer*. (Memorandum COSOR; Vol. 8510). Technische Hogeschool Eindhoven.

Document status and date:

Gepubliceerd: 01/01/1985

Document Version:

Uitgevers PDF, ook bekend als Version of Record

Please check the document version of this publication:

- A submitted manuscript is the 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](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

TECHNISCHE HOGESCHOOL EINDHOVEN

Onderafdeling der Wiskunde en Informatica

Memorandum-COSOR 85-10

Algoritme voor het schatten van de
parameters in Arma-modellen met
meetfouten op de in- en uitvoer

door

J.M. ten Vregelaar

Eindhoven, Nederland

Juni 1985

<u>0. Inhoud</u>	<u>pag.</u>
1. Inleiding	1
2. Probleemstelling	2
3. Het schatten van de parameters	5
4. Algoritme m.b.v. Q - R decompositie	9
5. Q - R decompositie gebaseerd op Householder transformaties	12
6. Bepaling van λ uit $R\lambda = u$	23
7. Bepaling van η , ξ	25
8. Bepaling van de gradiënt $\frac{\partial f}{\partial \gamma_k}$	27
9. Conclusies	30
10. Referenties	33

1. Inleiding

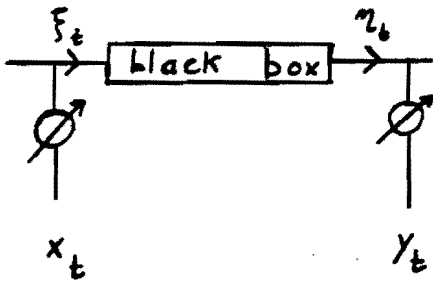
In dit memorandum wordt een algoritme beschreven voor het schatten van de parameters in Arma-modellen met meetfouten op in- en uitvoer. We beschouwen hier het MIMO (multiple input - multiple output) geval en nemen aan dat het totaal aantal waarnemingen van de in- en uitgangen (= N) groot is vergeleken met het totaal aantal parameters $ps^2 + (q + 1)sr$.

In [EIS] staat een algoritme beschreven voor het SISO-geval (single input - single output), dat gebruik maakt van zgn. Toeplitz-afstanden om een grootschalige matrix te inverteren. In de hier gehanteerde, meer inzichtelijke methode, komt een Q-R decompositie in de plaats van matrixinversie.

Na de probleemstelling in paragraaf 2 volgt in de paragrafen 3 en 4 de methode die aanleiding geeft tot het algoritme. In het algoritme onderscheiden we 4 stappen die achtereenvolgens in de paragrafen 5 t/m 8 aan de orde komen. De cruciale stap is de Q-R decompositie, waarvan aan het eind van paragraaf 5 een algoritme in pseudo-algol vorm is opgenomen.

Arma-modellen met ruis op de in- en uitgangen vinden o.a. toepassing in de econometrie.

2. Probleemstelling



ξ_t : afmeting r ,
 η_t : afmeting s .

We beschouwen een dynamisch lineair tijdinvariant MIMO (multiple input - multiple output) systeem, dat als volgt gemodelleerd wordt (zgn. ARMA-model):

$$\eta_t = \underbrace{\alpha_1 \eta_{t-1} + \dots + \alpha_p \eta_{t-p}}_{\text{Auto Regressief deel}} + \underbrace{\beta_0 \xi_t + \beta_1 \xi_{t-1} + \dots + \beta_q \xi_{t-q}}_{\text{Moving Average deel}}, \quad (1)$$

$t = 1, 2, \dots$. De dieptes p en q zijn bekend ($p, q \geq 0$).

De parameters α_i en β_i zijn $s \times s$ resp. $s \times r$ matrices.

Er zijn met fout gemeten waarnemingen beschikbaar:

$$y_t = \eta_t + e_t ; \quad x_t = \xi_t + e'_t \quad t = 1, 2, \dots, N . \quad (2)$$

We maken de volgende aannames voor de fouten e_t en e'_t :

$$\left. \begin{aligned} & \mathbb{E}e_t = 0 ; \quad \mathbb{E}e'_t = 0 \\ & e_1, \dots, e_N, e'_1, \dots, e'_N \text{ zijn onafhankelijk} \\ & \text{var } e_t = \sigma^2 I_s ; \quad \text{var } e'_t = \sigma^2 I_r \end{aligned} \right\} \quad (3)$$

Opmerkingen.

1. We veronderstellen dat voor grote N het beginneffect te verwaarlozen is, waardoor (4) gerechtvaardigd is. Eventueel kunnen ξ_t en η_t voor $t \leq 0$ geschat worden.
2. Aanname (5d) garandeert de identificeerbaarheid van de parameters, zie [EIS].

3. Het schatten van de parameters

We bepalen schatters voor α_i , β_j , η en ξ door de restkwadratensom te minimaliseren:

$$\min_{\alpha_i, \beta_j, \eta, \xi} \|x - \xi\|^2 + \|y - \eta\|^2 \quad \text{onder } A\eta + B\xi = 0 . \quad (6)$$

Als we aannemen dat de fouten e en e' normaal verdeeld zijn dan zijn de oplossingen van (6) de maximum likelihood schatters voor de parameters.

Na het invoeren van de notaties

$$\gamma = (\alpha_1, \dots, \alpha_p, \beta_0, \dots, \beta_q) , \quad \delta = \begin{bmatrix} \eta \\ \xi \end{bmatrix} , \quad z = \begin{bmatrix} y \\ x \end{bmatrix} \quad \text{en } D = (A \mid B)$$

gaat (6) over in

$$\min_{\gamma, \delta} \|z - \delta\|^2 \quad \text{onder } D\delta = 0 . \quad (7)$$

Omdat A regulier is, is D van volle rijenrang.

Voor vaste γ geeft toepassing van Lagrange

$$L(\delta, \lambda) = \frac{1}{2}(z - \delta)^T(z - \delta) + \lambda^T D\delta ,$$

we lossen δ en λ op uit $\nabla L = 0$, dus uit

$$\begin{cases} \delta + D^T \lambda = z & (8a) \\ D\delta = 0 . & (8b) \end{cases}$$

Uit (8) volgt $DD^T\lambda = Dz$, dus $\lambda = (D^+)^T z$ (9a)

met $D^+ := D^T(DD^T)^{-1}$.

Dan is

$$\delta = (8a) z - D^T\lambda = z - D^T(D^+)^T z = (I - D^+D)z \quad (9b)$$

want $D^T(D^+)^T = D^+D$.

Dus $z - \delta = D^+Dz$.

Probleem (7) is nu gereduceerd tot

$$\min_{\gamma} f(\gamma) \quad \text{met } f(\gamma) = \|D^+Dz\|^2. \quad (10)$$

Tot slot van deze paragraaf leiden we uitdrukkingen af voor de eerste en tweede afgeleiden van $f(\gamma)$, die we eventueel benutten bij de oplossing van probleem (10).

Noteer $f_k := \frac{\partial f}{\partial \gamma_k}$ met γ_k is een element van een parametermatrix, analoog

D_k, λ_k, δ_k .

Uit (8) volgt

$$\delta_k + D^T\lambda_k = -D^T\lambda \quad (11a)$$

$$D\delta_k = -D_k\delta \quad (11b)$$

Los δ_k en λ_k op uit (11):

$$D\delta_k + DD^T\lambda_k \stackrel{(11a)}{=} -DD^T\lambda \quad \text{dus } DD^T\lambda_k \stackrel{(11b)}{=} D_k\delta - DD^T\lambda,$$

waaruit volgt

$$\lambda_k = (DD^T)^{-1} D_k \delta - (D^+)^T D_k^T \lambda \quad (12a)$$

$$\delta_k = -D^+ D_k \delta + (D^+ D - I) D_k^T \lambda \quad (12b)$$

(11a)

Eerste afgeleide

Differentieer $f(\gamma) = \|z - \delta\|^2$:

$$f_k = -2(z - \delta)^T \delta_k \stackrel{(8a)}{=} -2\lambda^T D_k \delta \stackrel{(11b)}{=} 2\lambda^T D_k \delta \quad (13)$$

voor γ_k afkomstig uit een α_i -matrix is $f_k = 2\lambda^T A_k \eta$

en voor γ_k uit een β_j -matrix is $f_k = 2\lambda^T B_k \xi$.

Tweede afgeleide

Noteer

$$f_{k\ell} = \frac{\partial^2 f}{\partial \gamma_\ell \partial \gamma_k},$$

dan is

$$\begin{aligned} \frac{1}{2} f_{k\ell} &= \frac{\partial}{\partial \gamma_\ell} (\lambda^T D_k \delta) \\ &= \lambda_\ell^T D_k \delta + \lambda^T D_{k\ell} \delta + \lambda^T D_k \delta_\ell \end{aligned}$$

D_k constant

$$= \lambda_\ell^T D_k \delta + \lambda^T D_{k\ell} \delta_\ell$$

$$(12) \quad = [\delta^T D_\ell^T (DD^T)^{-1} - \lambda^T D_\ell^T D^+] D_k \delta + \lambda^T D_k [-D^+ D_\ell \delta + (D^+ D - I) D_\ell^T \lambda]$$

$$= \delta^T D_\ell^T (DD^T)^{-1} D_k \delta - \lambda^T (D_\ell^T D^+ D_k + D_k^T D^+ D_\ell) \delta - \lambda^T D_k (I - D^+ D) D_\ell^T \lambda. \quad (1)$$

Opmerking

1. In termen van A, B, y en x is $f(\gamma)$ te schrijven als (zie (10)),

$$f(\gamma) = (Ay + Bx)^T (AA^T + BB^T)^{-1} (Ay + Bx) . \quad (15)$$

2. Uit (8a) volgt dat δ naast (9b) ook geschreven kan worden als

$$\left. \begin{aligned} \delta &= z - D^T \lambda , & \text{dus } \eta &= y - A^T \lambda \\ & & \xi &= x - B^T \lambda \end{aligned} \right\} . \quad (16)$$

4. Algoritme met behulp van Q - R decompositie

We beschrijven in deze paragraaf een algoritme dat bij gegeven γ de waarde $f(\gamma)$ en de gradiënt $\frac{\partial f}{\partial \gamma}$ berekent. Bij geschikt gekozen startwaarden, zal een numeriek minimaliseringsalgoritme dat gebruik maakt van de gradiënt het gezochte minimum bepalen.

We hebben probleem (6) herschreven tot:

$$\min_{\gamma} f(\gamma) , \quad \text{met } f(\gamma) = \|D^+ Dz\|^2 \quad (17a)$$

en

$$\frac{\partial f}{\partial \gamma_k} = \begin{cases} 2\lambda^T A_k^T \eta & \text{als } \gamma_k \text{ afkomstig uit een } \alpha_i\text{-matrix} \\ 2\lambda^T B_k^T \xi & \text{als } \gamma_k \text{ afkomstig uit een } \beta_j\text{-matrix,} \end{cases} \quad (17b)$$

$$(17c)$$

verder is

$$\lambda = (D^+)^T z \quad (17d)$$

$$\eta = y - A^T \lambda \quad (17e)$$

$$\xi = x - B^T \lambda . \quad (17f)$$

In [EIS] wordt voor elke waarde van γ voor het SISO-geval een matrix geïnverteerd m.b.v. zgn. Toeplitz-afstanden. Een alternatief wordt gegeven door [GOL], p. 410 e.v.: Q - R decompositie. In de volgende paragraaf komt de uitvoering van de Q - R decompositie uitvoerig aan de orde. Hier gaan we in op de consequenties ervan voor de berekening van $f(\gamma)$ en $\frac{\partial f}{\partial \gamma_k}$.

Gelet op de nevenvoorwaarde $D\delta = 0$ (zie (7)) maken we een Q-R decompositie van de matrix D^T die van volle kolommenrang is:

$$D^T = Q \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad Q \text{ is orthogonaal, } R \text{ is rechtsboven, band en regulier.}$$

Als $Q = (Q_1 \mid Q_2)$ volgt

$$D^T = Q_1 R. \quad (18)$$

Dan is

$$D^+ = D^T (DD^T)^{-1} \stackrel{(18)}{=} Q_1 R^{-T} \quad (19a)$$

en

$$D^+ D = Q_1 Q_1^T. \quad (19b)$$

Uitgedrukt in z , Q , R volgt voor f en λ (zie (17 a,d))

$$f(\gamma) = \|Q_1^T z\|^2, \quad \lambda = R^{-1} Q_1^T z.$$

Definieer

$$u := Q_1^T z \quad (20a)$$

dan is

$$f(\gamma) = \|u\|^2 \quad (20b)$$

en

$$\lambda = R^{-1} u. \quad (20c)$$

Voor de berekening van $f(\gamma)$ en $\frac{\partial f}{\partial \gamma_k}$ hebben we nodig $u = Q_1^T z$, $R \stackrel{(18)}{=} Q_1^T D^T$ en $\lambda = R^{-1} u$, waarbij de waarnemingsvector z gegeven is.

Breid de matrix D^T uit met de kolomvector z dan geldt

$$Q^T (D^T \mid z) = \begin{bmatrix} R & u \\ \hline O & w \end{bmatrix}, \quad (21)$$

met $w = Q_2^T z$.

Aan de hand van bovenstaande afleiding komen we tot het volgende algoritme voor de berekening van $f(\gamma)$ en $\frac{\partial f}{\partial \gamma_k}(\gamma)$ als γ gegeven is.

1. Maak een Q-R decompositie van D^T ; via (21) volgen R, u en $f = \|u\|^2$.

2. Los λ op uit $R\lambda = u$, R is rechtsboven, band.

3. Bereken $\eta = y - A^T\lambda$

en $\xi = x - B^T\lambda$ (zie (17 e, f)).

4. Bereken

$$\frac{\partial f}{\partial \gamma_k} = 2\lambda^T D_k \begin{bmatrix} \eta \\ \xi \end{bmatrix} \quad (\text{zie (17 b, c)}).$$

In de volgende 4 paragrafen worden de 4 stappen in het algoritme gedetailleerd beschreven.

5. Q - R decompositie gebaseerd op Householdertransformaties

5.1. Algemeen

Zij A een $m \times n$ matrix ($m \geq n$), van volle kolommenrang.

Dan is er een orthogonale matrix Q (zie b.v. [GOL], p. 147 e.v.)

zodanig dat

$$Q^T A = \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad R \text{ is rechtsboven en regulier.}$$

Door de eis dat de diagonaalelementen van R positief zijn is deze $Q - R$ decompositie van A eenduidig bepaald.

Een methode ter berekening van Q en R wordt gegeven door gebruik te maken van Householder matrices (zie [GOL], p. 38-43).

Als $v \in \mathbb{R}^n$, $v \neq 0$ dan is $P := I - 2vv^T/v^T v$ een $n \times n$ Householder matrix.

Het is duidelijk dat P symmetrisch en orthogonaal is. De voor ons doel beoogde toepassing van Householder matrices is de volgende.

Stel $x \in \mathbb{R}^n$

definieer $v = x \pm \|x\|e_1$ dan is $(I - 2vv^T/v^T v)x = \mp \|x\|e_1$. (22)

Opmerking. Om cijferverlies tegen te gaan, nemen we $v = x + \text{sign}(x_1) \|x\|e_1$ (zie [STE], p. 233).

Onderstaand voorbeeld laat zien hoe we Q en R berekenen voor gegeven A.

Voorbeeld. Stel $m = 6$, $n = 5$ en Householdermatrices P_1 en P_2 zijn berekend zodanig dat

$$P_2 P_1 A = \begin{bmatrix} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{bmatrix}.$$

Bepaal nu een 4×4 Householder matrix \tilde{P}_3 zodanig dat

$$\tilde{P}_3 \begin{bmatrix} x \\ x \\ x \\ x \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{en neem } P_3 = \text{diag}(I_2, \tilde{P}_3)$$

dan is

$$P_3 P_2 P_1 A = \begin{bmatrix} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix}.$$

Uiteindelijk is $P_n P_{n-1} \dots P_1 A = \begin{bmatrix} R \\ 0 \end{bmatrix}$ met R rechtsboven.

P_i is symmetrisch en orthogonaal, $i = 1, \dots, n$.

Definieer $Q = P_1 \dots P_n$ dan is $A = Q \begin{bmatrix} R \\ 0 \end{bmatrix}$.

5.2. Speciale Q - R decompositie

We willen nu bovenstaand recept toepassen op de matrix D^T . Om de Toeplitz structuur in de matrix D^T niet kapot te maken, voeren we een speciaal op die structuur toegesneden Q - R decompositie uit.

Volgens (21) zijn we geïnteresseerd in de werking van Q^T op $(D^T|z)$.

We schrijven Q^T als produkt van Ns matrices, die alle orthogonaal en symmetrisch zijn:

$$Q^T = \underbrace{H_{N,s} \dots H_{N,1}}_{P_N} \dots \underbrace{H_{2,s} \dots H_{2,1}}_{P_2} \underbrace{H_{1,s} \dots H_{1,1}}_{P_1} \quad (23)$$

We zullen nu d.m.v. de werking van $H_{1,1}$ laten zien hoe de blok Toeplitz-structuur behouden blijft.

De werking van $H_{j,\ell}$ in het algemeen ($j = 1, \dots, N$; $\ell = 1, \dots, s$) komt daarna aan de orde.

Definieer

$$a_i := \alpha_i^T, \quad i = 1, \dots, p \quad \text{en} \quad a_0 := -I_s$$

$$b_i := \beta_i^T, \quad i = 0, \dots, q$$

$$m := \max(p, q),$$

dan kunnen we, door eventueel enkele blokken a_i of b_i nul te nemen, $(D^T|z)$ schrijven als

$$(D^T | z) = \left[\begin{array}{ccc|ccc} a_0 & \dots & a_m & & y_1 & \\ & \dots & & & \vdots & \\ & & & & & y_N \\ \hline b_0 & \dots & b_m & & x_1 & \\ & \dots & & & \vdots & \\ & & & & & x_N \end{array} \right]$$

Vermenigvuldiging met $H_{1,1}$

Beschouw de eerste kolom in $(D^T | z)$. Deze bevat slechts $(r + 1)$ elementen die ongelijk aan 0 zijn: $a_{0,1,1}, b_{0,1,1}, \dots, b_{0,r,1}$.

Laat \tilde{h} een $(r + 1) \times (r + 1)$ Householder matrix zijn, waarvoor geldt dat

$$\tilde{h} \begin{bmatrix} a_{0,1,1} \\ b_{0,1,1} \\ \vdots \\ b_{0,r,1} \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

We kunnen nu dezelfde \tilde{h} ook op de overige $(N - 1)$ onderelkaarstaande a_0 en b_0 blokken laten werken door $H_{1,1}$ als volgt te kiezen:

en het "resterende deel" van de matrix A^T blok Toeplitz blijven door de speciale keuze van de $H_{j,l}$ -matrices.

Vermenigvuldiging met $H_{j,l}$

We hoeven slechts de werking van een $(r + 1) \times (r + 1)$ Householder matrix \tilde{h} op onderstaande blokken en vectoren te berekenen om de "nieuwe" $(D^T | z)$ helemaal te kennen. Door het behoud van de Toeplitz structuur blijven de blokdiagonalen constant (in de matrix A^T alleen de j^e t/m N^e blokrij). De veranderde elementen behouden hun "naam":

$$\begin{bmatrix} a_{i,l,l} & \dots & a_{i,l,s} \\ b_{i,1,l} & \dots & b_{i,1,s} \\ \vdots & & \vdots \\ b_{i,r,l} & \dots & b_{i,r,s} \end{bmatrix} := \tilde{h} \begin{bmatrix} a_{i,l,l} & \dots & a_{i,l,s} \\ b_{i,1,l} & \dots & b_{i,1,s} \\ \vdots & & \vdots \\ b_{i,r,l} & \dots & b_{i,r,s} \end{bmatrix} \quad i=0, \dots, \min(m, N-j) \quad (24a)$$

$$a_{0,l,l} := -\text{sign}(a_{0,l,l}) \sqrt{a_{0,l,l}^2 + b_{0,1,l}^2 + \dots + b_{0,r,l}^2}$$

en $b_{0,1,l} := \dots = b_{0,r,l} := 0$ (zie (22,25))

$$\text{en } \begin{bmatrix} y_{k+j,l} \\ x_{k+1} \end{bmatrix} := \tilde{h} \begin{bmatrix} y_{k+j,l} \\ x_{k+1} \end{bmatrix} \quad k = 0, 1, \dots, N-j. \quad (24b)$$

Hierin is \tilde{h} Householder,

$$\tilde{h} = I - 2vv^T/v^T v, \quad (25)$$

met

$$v = (a_{0,l,l} + \text{sign}(a_{0,l,l}) \sqrt{a_{0,l,l}^2 + b_{0,1,l}^2 + \dots + b_{0,r,l}^2}, b_{0,1,l}, \dots, b_{0,r,l})^T$$

De volgende keuze voor de matrices $H_{j,\ell}$ in (23) zorgt voor handhaving van de Toeplitz structuur in de matrix D^T :

$$\underbrace{H_{j,\ell}}_{(s+r)N \times (s+r)N} = \begin{bmatrix} \begin{array}{c|c} I^{(j-1)} & \\ \hline h_{11} & h_{12} \\ & \ddots \\ & h_{11} & h_{12} \end{array} & \\ \hline \begin{array}{c|c} h_{12}^T & h_{22} \\ & \ddots \\ h_{12}^T & h_{12} \end{array} & \begin{array}{c} I \\ \ddots \\ I \end{array} \end{bmatrix} \quad j = 1, \dots, N; \ell = 1$$

waarbij

$$\underbrace{h_{11}}_{s \times s} := \text{diag}(1, \dots, 1, \underbrace{\tilde{h}_{1,1}}_{\uparrow \ell^e \text{ plaats}}, 1, \dots, 1) ; \quad \underbrace{h_{12}}_{s \times r} = \begin{bmatrix} \varnothing \\ \hline \tilde{h}_{1,2} \dots \tilde{h}_{1,r+1} \\ \hline \varnothing \end{bmatrix} \leftarrow \ell^e$$

$$\underbrace{h_{22}}_{r \times r} = \begin{bmatrix} \tilde{h}_{2,2} & \dots & \tilde{h}_{2,r+1} \\ \vdots & & \vdots \\ \tilde{h}_{2,r+1} & \dots & \tilde{h}_{r+1,r+1} \end{bmatrix}$$

De werking van $H_{j,\ell}$ bestaat uit het schoonvegen van de ℓ^e -kolom in alle b_0 blokken op de blokdagonaal $N+1, j; N+2, j+1; \dots; 2N-j+1, N$. Na elke $H_{j,\ell}$ is er een blokdagonaal met b_0 -blokken nul geworden.

Schema: Werking van Q^T op $(D^T|z)$

Het volgende schema beschrijft globaal de werking van Q^T op $(D^T|z)$; we definiëren daartoe $R := (r_{ij})$, r_{ij} is $(s \times s)$ -blok voor $i, j = 1, \dots, N$

en $u := \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}$ met u_i ($s \times 1$)-vector.

$$P_1(D^T | z) = \left[\begin{array}{c|c} r_{1,1} \dots r_{1,m+1} & u_1 \\ \hline \boxed{a_0} \dots a_m \dots a_m \dots a_0 & y_2 \\ & \vdots \\ & y_N \\ \hline 0 \quad \boxed{b_0} \dots b_{m-1} \dots b_{m-1} \dots b_0 & x_1 \\ & \vdots \\ & x_{N-1} \\ \hline & 0 \\ & w_N \end{array} \right] \dots$$

$$P_{N-m} \dots P_1(D^T | z) = \left[\begin{array}{c|c} r_{1,1} \dots r_{1,m+1} & u_1 \\ & \vdots \\ & u_{N-m} \\ \hline \boxed{a_0} \dots a_{m-1} \dots a_0 & y_{N-m+1} \\ & \vdots \\ & y_N \\ \hline \boxed{b_0} \dots b_{m-1} \dots b_0 & x_1 \\ & \vdots \\ & x_m \\ \hline & w_{m+1} \\ & \vdots \\ & w_N \end{array} \right] \dots$$

$$P_N \dots P_1 (D^T | z) = \left[\begin{array}{c|c} \begin{array}{ccc} r_{1,1} & \dots & r_{1,m+1} \\ & \ddots & \\ & & r_{N-m,N} \\ & & \vdots \\ & & r_{N,n} \end{array} & \begin{array}{c} u_1 \\ \vdots \\ u_N \end{array} \\ \hline \bigcirc \end{array} \right] \quad (26)$$

R en u zijn nu bepaald, (zie 21).

Toelichting bij het schema

1. We hernoemen de b_i -blokken na elke vermenigvuldiging met $H_{j,s}$. Verder is na elke vermenigvuldiging met $H_{j,s}$ de j^e blokrij van R en de j^e blokcomponent van u gevonden:

$$r_{j,j} := a_0, \dots, r_{j,j+\min(m, N-j)} := a_{\min(m, N-j)}$$

en

$$u_j := y_j.$$

Deze blijven onveranderd, evenals w_{N-j+1} .

2. Omdat voor $j = 2, \dots, N-m$ het aantal blokken a_i in de j^e blokrij één meer bedraagt dan het aantal blokken b_i in de $(N+1)^e$ blokrij van (de getransformeerde) D^T blijft het aantal blokken b_i constant ($= m$) voor $j = 2, \dots, N-m$: enerzijds verdwijnt het b_0 blok, anderzijds wordt het b_m -blok gevormd

dat na henummering de naam b_{m-1} krijgt; in (24 a) is weliswaar het b_m -blok in het rechterlid nul, maar het a_m -blok niet.

3. Omdat na elke stap a_0 rechtsboven blijft is slechts een $(r+1) \times (r+1)$ ondermatrix van de $(s+r) \times (s+r)$ matrix, die op $\begin{bmatrix} a_0 \\ b_0 \end{bmatrix}$ werkt van belang. De $(r+1) \times (r+1)$ Householder matrix \tilde{h} (in (24)) wordt aangevuld met $s-1$ termen 1 op de diagonaal.

In het bijzonder zijn alle $r_{i,1}$ ($i = 1, \dots, N$) in (26) rechtsboven.

4. Uit $\tilde{h} = I - 2vv^T/v^T v$ volgt

$$\tilde{h}_{i,j} = \delta_{i,j} - \frac{2}{\|v\|^2} v_i v_j \quad i, j = 1, \dots, r+1 .$$

Bovenstaande beschrijving geeft aanleiding tot het volgende algoritme ter bepaling van R, u en $f(\gamma)$.

Invoer $a_0, \dots, a_m, b_0, \dots, b_m, z = \begin{bmatrix} y \\ x \end{bmatrix}$

Uitvoer R, u, f .

Beschouw het geval $m \geq 1$. Het commentaar is vermeld tussen accolades. De waarden voor j en ℓ corresponderen met de vermenigvuldiging met $H_{j,\ell}$.


```
begin f := 0 ;  
  
  for j := 1(1) N do  
    begin if 2 ≤ j ≤ N - m then bm := 0 ; {zie toelichting punt 2}  
      for ℓ := 1(1) s do  
  
        begin v := (a0,ℓ,ℓ, b0,1,ℓ ... b0,r,ℓ)T ;  
          v1 := v1 + sign(v1) ||v|| ;           {zie (25)}  
                                                    {zie (24 a)}  
          for i := 0(1)min(m,N-j) do "verander ai,ℓ*, bi" ;  
            for k := 0(1)N-j do "verander yk+j,ℓ, xk+1" {zie (24b)}  
  
          end ;  
          for i := 0(1)min(m,N-j) - 1 do  
            begin bi := bi+1 {Hernummer de bi's}  
              rj,j+i := ai  
            end ;  
            rj,j+min(m,N-j) := amin(m,N-j) ; uj := yj ; {zie toelichting  
                                                                punt 1}  
            f := f + ||uj||2  
          end  
        end ;  
      end ;  
    end ;  
  end ;
```

6. Bepaling van λ uit $R\lambda = u$

Via (2.6) schrijven we $R\lambda = u$ als

$$\begin{bmatrix} r_{1,1} & \dots & r_{1,m+1} & & & \\ & \ddots & & & & \\ & & & & & \\ & & & & r_{N-m,N} & \\ & & & & & r_{N,N} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \vdots \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ \vdots \\ \vdots \\ u_N \end{bmatrix} \quad (27)$$

met $r_{i,j} : s \times s$, $\lambda_i, u_i : s \times 1$.

Noteer $\lambda_{i,j}$ en $u_{i,j}$ voor de j^e component van de vectoren λ_i resp. u_i ($i = 1, \dots, N; j = 1, \dots, s$), zie ook par. 7 en 8.

R en u zijn bekend (zie paragraaf 5), R is een rechtsboven en bandmatrix, ook alle $r_{i,i}$ zijn rechtsboven voor $i = 1, \dots, N$.

Schrijf (27) uit door de N blokkrijen te vermenigvuldigen met λ , te beginnen met de N^e :

$$r_{N,N} \lambda_N = u_N$$

$$r_{N-1,N-1} \lambda_{N-1} = u_{N-1} - r_{N-1,N} \lambda_N$$

\vdots

$$r_{N-m,N-m} \lambda_{N-m} = u_{N-m} - r_{N-m,N-m+1} \lambda_{N-m+1} - \dots - r_{N-m,N} \lambda_N$$

$$r_{N-m-1,N-m-1} \lambda_{N-m-1} = u_{N-m-1} - r_{N-m-1,N-m} \lambda_{N-m} - \dots - r_{N-m-1,N-1} \lambda_{N-1}$$

\vdots

$$r_{1,1} \lambda_1 = u_1 - r_{1,2} \lambda_2 - \dots - r_{1,m+1} \lambda_{m+1}$$

Los uit de eerste vergelijking λ_N op, uit de tweede λ_{N-1}, \dots , uit de N^e λ_1 .

Algoritme.

```

For i:=N(-1)1 do begin rechterlid := ui -  $\sum_{k=i+1}^{\min(N,i+m)} r_{i,k} \lambda_k$  ;
    "los  $\lambda_i$  op uit  $r_{i,i} \lambda_i = \text{rechterlid}$ "
end;

```

Het deelprobleem $\text{los } \lambda \text{ op uit } r\lambda = u$, is hetzelfde probleem met dit verschil dat r nu een $s \times s$ matrix is en λ en u $s \times 1$ vectoren, en verder is r wel rechtsboven, maar geen bandmatrix.

$$\begin{bmatrix} r_{1,1} & \dots & r_{1,s} \\ & \ddots & \vdots \\ & & r_{s,s} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_s \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_s \end{bmatrix}$$

uitgeschreven geeft

$$\begin{aligned}
 r_{s,s} \lambda_s &= u_s \\
 r_{s-1,s-1} \lambda_{s-1} &= u_{s-1} - r_{s-1,s} \lambda_s \\
 &\vdots \\
 r_{1,1} \lambda_1 &= u_1 - r_{1,2} \lambda_2 - \dots - r_{1,s} \lambda_s .
 \end{aligned}$$

Het bijbehorende algoritme luidt nu:

```

for i := s(-1)1 do begin rechterlid := ui -  $\sum_{k=i+1}^s r_{i,k} \lambda_k$  ;
     $\lambda_i := \text{rechterlid} / r_{i,i}$ 
end ;

```


Uit (28) volgt nu

$$\eta_{i,j} = y_{i,j} + \lambda_{i,j} - \sum_{k=1}^{\min(p,N-i)} (\alpha_k^T \lambda_{k+1})_j \quad i = 1, \dots, N; j = 1, \dots, s$$

$$\xi_{i,j} = x_{i,j} - \sum_{k=0}^{\min(q,N-i)} (\beta_k^T \lambda_{k+1})_j \quad i = 1, \dots, N; j = 1, \dots, r$$

Omdat

$$(\alpha_k^T \lambda_{k+1})_j = \sum_{\ell=1}^s (\alpha_k^T)_{j,\ell} \lambda_{k+1,\ell} = \sum_{\ell=1}^s (\alpha_k)_{\ell,j} \lambda_{k+1,\ell} \quad j = 1, \dots, s$$

en

$$(\beta_k^T \lambda_{k+1})_j = \sum_{\ell=1}^s (\beta_k^T)_{j,\ell} \lambda_{k+1,\ell} = \sum_{\ell=1}^s (\beta_k)_{\ell,j} \lambda_{k+1,\ell} \quad j = 1, \dots, r$$

vinden we voor η en ξ

(29a)

$$\eta_{i,j} = y_{i,j} + \lambda_{i,j} - \sum_{k=1}^{\min(p,N-i)} \sum_{\ell=1}^s (\alpha_k)_{\ell,j} \lambda_{k+1,\ell} \quad i = 1, \dots, N; j = 1, \dots, s$$

(29b)

$$\xi_{i,j} = x_{i,j} - \sum_{k=0}^{\min(q,N-i)} \sum_{\ell=1}^s (\beta_k)_{\ell,j} \lambda_{k+1,\ell} \quad i = 1, \dots, N; j = 1, \dots, r$$

8. Bepaling van de gradiënt $\frac{\partial f}{\partial \gamma_k}$

We berekenen de gradiënt $\frac{\partial f}{\partial \gamma_k}$ uit de formules (17 b,c)

$$\frac{\partial f}{\partial \gamma_k} = \begin{cases} 2\lambda^T A_k \eta & \text{als } \gamma_k \text{ afkomstig uit een } \alpha_i\text{-matrix} & (30a) \\ 2\lambda^T B_k \xi & \text{als } \gamma_k \text{ afkomstig uit een } \beta_j\text{-matrix} & (30b) \end{cases}$$

De vectoren λ , η en ξ zijn bekend (uit paragraaf 6 en 7), en

$$A_k = \frac{\partial A}{\partial \gamma_k}, \quad B_k = \frac{\partial B}{\partial \gamma_k}$$

We voeren eerst de vermenigvuldiging $\lambda^T A_k$ resp. $\lambda^T B_k$ uit, dit levert sN - resp. rN -rijvectoren.

Bedenk dat $\lambda^T = (\lambda_{1,1} \dots \lambda_{1,s} \lambda_{2,1} \dots \lambda_{2,s} \dots \lambda_{N,1} \dots \lambda_{N,s})$

Na 3 speciale γ_k 's volgt het algemene resultaat.

1) $\gamma_k = (\beta_0)_{i,j}$

Dan is

$$B_k = \begin{matrix} 0.s+i & \rightarrow & [& 1 & & &] \\ 1.s+i & \rightarrow & & & 1 & & \\ \vdots & & & & & \ddots & \\ (N-1)s+i & \rightarrow & & & & & 1 \end{matrix}$$

$$\begin{matrix} \uparrow & & \uparrow & & \uparrow \\ 0.r+j & & 1.r+j & \dots & (N-1)r+j \end{matrix}$$

B_k bevat N elementen 1, de overige zijn 0.

Gevolg

$$\lambda^T B_k = (\lambda_{1,i} \quad \lambda_{2,i} \quad \dots \quad \lambda_{N,i}) \quad i = 1, \dots, s$$

$$\text{plaats:} \quad \begin{matrix} \uparrow & \uparrow & \uparrow \\ j & r+j & (N-1)r+j \end{matrix} \quad j = 1, \dots, r$$

De overige elementen in deze rN -rijvector zijn 0.

2) $\gamma_k = (\alpha_1)_{i,j}$

$$A_k = \begin{matrix} s+i & \rightarrow & [& 1 & & & & &] \\ 2s+i & \rightarrow & & & 1 & \dots & & & \\ \vdots & & & & & & & & \\ (N-1)s+i & \rightarrow & & & & & & & 1 \\ & & \uparrow & & \uparrow & & & & \uparrow \\ & & j & & s+j & \dots & & & (N-2)s+j \end{matrix}$$

A_k bevat $(N-1)$ elementen 1. Nu is

$$\lambda^T A_k = (\lambda_{2,i} \quad \lambda_{3,i} \quad \dots \quad \lambda_{N,i} \quad 0) \quad i, j = 1, \dots, s$$

$$\text{plaats:} \quad \begin{matrix} \uparrow & \uparrow & \uparrow \\ j & s+j & (N-2)s+j \end{matrix}$$

3) $\gamma_k = (\beta_1)_{i,j}$ analoog aan 2).

$$\lambda^T B_k = (\lambda_{2,i} \quad \lambda_{3,i} \quad \dots \quad \lambda_{N,i} \quad 0) \quad i = 1, \dots, s$$

$$\text{plaats} \quad \begin{matrix} \uparrow & \uparrow & \uparrow \\ j & r+j & (N-2)r+j \end{matrix} \quad j = 1, \dots, r$$

Algemeen geldt:

$$\lambda^T \frac{\partial A}{\partial (\alpha_k)_{i,j}} = (\lambda_{k+1,i} \quad \lambda_{k+2,i} \quad \dots \quad \lambda_{N,i} \quad 0 \dots 0) \quad i, j = 1, \dots, s$$

↑
↑
↑
↑

plaats j
s+j
(N-k-1)s+j
k = 1, \dots, p

$$\lambda^T \frac{\partial B}{\partial (\beta_k)_{i,j}} = (\lambda_{k+1,i} \quad \lambda_{k+2,i} \quad \dots \quad \lambda_{N,i} \quad 0 \dots 0) \quad i = 1, \dots, s$$

↑
↑
↑
↑

j
r+j
(N-k-1)r+j
j = 1, \dots, r

k = 0, \dots, q

Berekening van de afgeleide uit (30) resulteert in

$$\frac{\partial f}{\partial (\alpha_k)_{i,j}} = 2 \sum_{n=k+1}^N \lambda_{n,i} \eta_{n-k,j}, \quad \begin{matrix} k = 1, \dots, p \\ i, j = 1, \dots, s \end{matrix} \quad (31a)$$

$$\frac{\partial f}{\partial (\beta_k)_{i,j}} = 2 \sum_{n=k+1}^N \lambda_{n,i} \xi_{n-k,j}, \quad \begin{matrix} k = 0, \dots, q \\ i = 1, \dots, s \\ j = 1, \dots, r \end{matrix} \quad (31b)$$

9. Conclusies

We geven een samenvatting van de hier beschreven methode die tot een algoritme leidt voor het schatten van de parameters in Arma-modellen met meetfouten op in- en uitvoer.

We bepalen de schatters voor de parameters γ door de residuenkwadraten-som te minimaliseren met de modelvergelijking als nevenvoorwaarde (z is de waarnemingsvector van de uitgangs- ingangskolom δ):

$$\min_{\gamma, \delta} \|z - \delta\|^2 \text{ onder } D\delta = 0, \quad D(\gamma) \text{ van volle rijenrang.}$$

Met behulp van de Lagrangiaan $L = \frac{1}{2}\|z - \delta\|^2 + \lambda^T D\delta$ vinden we het probleem zonder nevenvoorwaarde

$$\min_{\gamma} f(\gamma) \text{ met } f(\gamma) = \|D^+ Dz\|^2, \quad D^+ = D^T(DD^T)^{-1}.$$

Omdat we f niet expliciet als functie van γ kennen, lossen we dit probleem numeriek op. Bij geschikt gekozen beginschattingen zal een numeriek algoritme dat gebruikmaakt van de afgeleiden

$$f_i = 2\lambda^T D_i^T \delta \text{ met } \lambda = (D^+)^T z, \quad \delta = z - D^T \lambda$$

ons schattingen van de parameters leveren. Vanwege de vervelende termen $(DD^T)^{-1} = R^{-1}R^{-T}$ en $D^+D = Q_1Q_1^T$ in (14) benutten we de tweede afgeleide niet.

Als we een Q-R decompositie van D^T geconstrueerd hebben ($Q = (Q_1 Q_2)$) dan wordt

$$f = \|u\|^2, \text{ met } u = Q_1^T z$$

$$\text{en } \lambda = R^{-1} u.$$

Het algoritme voor de berekening van f en f_i als functie van γ ziet er globaal als volgt uit:

1) Bepaal R, u uit $Q^T(D^T | z) = \left[\begin{array}{c|c} R & u \\ \hline 0 & * \end{array} \right]$ en $f = \|u\|^2$.

Door een speciale Q - R decompositie gebaseerd op Householder matrices wordt de Toeplitz-structuur in de matrix D^T goed gebruikt.

2. Los λ op uit $R\lambda = u$. R is een rechtsboven en bandmatrix. Het is mogelijk de blokstructuur te handhaven bij de terugsubstitutie. Overigens verwachten we dat dit probleem goed geconditioneerd is als de parameterwaarden in D^T niet te groot zijn.

3) Bepaal $\delta = z - D^T \lambda$.

4) Bepaal $f_i = 2\lambda^T D_i^T \delta$ ($i = 1, \dots, ps^2 + (q + 1)sr$).

Door de bewerkingen op een geschikte wijze te organiseren komen bij de laatste twee stappen de formules (29) en (31) tevoorschijn.

Merk op dat nergens expliciet gebruik is gemaakt van vrije parametrisatie in de matrix $D = (A|B)$, het verhaal gaat dus ook op voor gestructureerde

parametrisaties, die zich in o.a. economische toepassingen voordoen. Alleen formule (31) voor de afgeleide behoeft enige aanpassing.

Of en zo ja, hoe snel, een computerprogramma gebaseerd op bovenstaand algoritme ons parameterschattingen levert heeft nu onze interesse. We zullen dan de resultaten vergelijken met die van het algoritme beschreven in [EIS]; eventueel verdient de keuze van de beginschatting nog wat aandacht.

Het toekomstig onderzoek richt zich verder op de asymptotische eigenschappen van de schatters en op toepassingen in de vorm van econometrische modellen.

In een later stadium zullen we het geval bekijken dat enkele van de ingangen exact te regelen zijn, dus niet met fout gemeten.

10. Referenties

- [EIS] Eising F., Linssen H.N., Rietbergen H.,
System identification from noisy measurements of inputs and
outputs, System and control letters 2, 348-353, 1983.
- [GOL] Golub G.H., van Loan C.F.,
Matrix Computations, North Oxford Academic, Oxford, 1983
- [STE] Steweart G.W.,
Introduction to matrix computations, Academic Press, New York,
1973.