Least-squares fitting by visualization of the sum of squares space

Citation for published version (APA):

Herk, van, A. M. (1995). Least-squares fitting by visualization of the sum of squares space. *Journal of Chemical Education*, 72, 138-140. <https://doi.org/10.1021/ed072p138>

DOI:

[10.1021/ed072p138](https://doi.org/10.1021/ed072p138)

Document status and date:

Published: 01/01/1995

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 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.

Table 2. Calculated Average Lengths of Isotactic (from eqs 1, 3, and 5) and Syndiotactic (from eqs 2, 4, and 6) Sequences of a Polymer Chain for Various Values of P_m

P_m	$\langle n_i \rangle$			$\langle n_s \rangle$		
	eq 1	eq 3	eq 5	eq 2	eq 4	eq 6
0.1	1.112	1.112	1.111	9.956	10.046	10.033
0.2	1.245	1.246	1.248	4.996	5.017	5.028
0.4	1.670	1.671	1.668	2.497	2.500	2.496
0.5	2.000	2.002	2.002	1.997	1.999	1.998
0.6	2.498	2.500	2.502	1.663	1.664	1.666
0.8	4.961	4.981	4.977	1.252	1.253	1.252
0.9	9.907	9.998	9.980	1.113	1.113	1.111

tween theory (solid lines) and simulation (points) is obtained, as shown in Figures 3 and 4 for triad and tetrad probabilities, respectively (12).

(B) Using eqs 1, 3, and 5 and eqs 2, 4, and 6 calculate the average lengths of isotactic, $\langle n_i \rangle$, and syndiotactic, $\langle n_s \rangle$, sequences, respectively, as shown in Table 2. Good agreement is observed, usually to the third decimal place.

(C) Construct plots of number and weight fraction distributions of isotactic and syndiotactic sequences as function of n_i , as shown in Figures 6 and 7, for P_m values of 0.1, 0.5, and 0.9. Solid lines represent theoretical curves as described by eqs 7 and 8 and the points are simulation values that were converted to number fractions by

$$N(n_i) = N_{n_i} / v_{\text{tot, dyads}} \quad (9)$$

and weight fractions by

$$W(n_i) = N_{n_i} n_i / v_{\text{tot, dyads}} \quad (10)$$

- For the above exercises, the values of the parameters DP and number of molecules were chosen so as to avoid the effects of sample size and chain-end effects. The effect of sample size can be conveniently eliminated by keeping the product of DP and number of molecules to a value of 1,000,000, thus allowing chain-end effects to be analyzed. This effect can be demonstrated by constructing number and weight fraction distributions, as described in exercise 1C, for P_m values of 0.5 and 0.9 and DP values of 10, 20, 50, 100, 200, 300, 400, and 600. The conclusion that can be drawn from this exercise is that the minimum degree of polymerization depends on the P_m value, i.e. for $P_m = 0.5$, DP must be greater than 50; whereas, a $DP > 400$ is required for $P_m = 0.9$.

Literature Cited

- Bovey, F. A.; Tiers, G. V. D. *J. Polym. Sci.* **1960**, *44*, 173–182.
- Bovey, F. A.; Tiers, G. V. D. *Fortschr. Hochpolymer. Forsch.* **1963**, *3*, 139–195.
- Bovey, F. A. *Pure Appl. Chem.* **1966**, *12*, 525–544.
- Bovey, F. A. *Polym. Eng. Sci.* **1966**, *26*, 1419–1428.
- Tonelli, A. E. *NMR Spectroscopy and Polymer Microstructure: The Conformational Connection*; VCH Publishers: New York, 1989.
- Koenig, J. L. *Spectroscopy of Polymers*; American Chemical Society: Washington, 1992.
- Bovey, F. A. *Polymer Conformation and Configuration*; Academic Press: New York, 1969.
- Randall, J. C. *Polymer Sequence Determination: Carbon-13 NMR Method*; Academic Press: New York, 1977.
- Randall, J. C.; Hsieh, E. T. In *NMR and Macromolecules: Sequence, Dynamic and Domain Structure*; Randall, J. C., Ed.; American Chemical Society: Washington, 1984; pp 131–151.
- Ferguson, G. A. *Statistical Analysis in Psychology and Education*, 5th ed.; McGraw-Hill: New York, 1981.
- Lowry, G. G., Ed. *Markov Chains and Monte Carlo Calculations in Polymer Science*; Marcel Dekker: New York, 1970.
- Bovey, F. A. *Polymer Conformation and Configuration*; Academic Press: New York, 1969; p 16.
- Hiemenz, P. C. *Polymer Chemistry: The Basic Concepts*; Marcel Dekker: New York, 1984; pp 486–488.

Least-Squares Fitting by Visualization of the Sum of Squares Space

Alex M. van Herk

Eindhoven University of Technology
Department of Polymer Chemistry
P.O. Box 513
5600 MB Eindhoven, The Netherlands

It is evident that curve fitting of experimental data in the practice of chemistry is very important. One of the barriers for students to use for example the least-squares fitting method is the complexity of the software.

Recent articles in this *Journal* have introduced simple methods that can solve curve-fitting problems on personal computers (1–3). For the strictly linear function (1) and for the general nonlinear functions (2) iterative procedures are available. O'Neill et al. (3) introduced a noniterative procedure that is applicable to multilinear equations. So far, for nonlinear equations one still had to resort to iterative methods.

In this article a simple method is introduced that can be used for all types of functions involving up to three parameters to be fitted. The method gives the optimum parameters with their uncertainties in the form of unbiased joint confidence intervals and can weigh the experimental data correctly. Furthermore, in contrast to iterative methods, insight into the occurrence of local minima is gained.

This method could be used for teaching students the least-squares method in a way that makes it understandable. Of course modern iterative matrix methods can solve the same problems that can be solved by this method when care is taken to avoid zeroing in to local minima.

The Least-Squares Method

If we assume only errors in the dependent variables y_i and we have a set of p parameters Θ in the function $y_i = f(x_i, \Theta)$, the least-squares estimates of the parameters Θ are those that satisfy the minimum value of the sum of squares, $ss(\Theta)$, of the weighted residuals (further referred to as the sum of squares), calculated with i data points.

$$ss(\Theta) = \sum_i w_i (y_i - f(x_i, \Theta))^2 \quad (1)$$

w_i are the weighting factors, which are

$$w_i = \frac{1}{\sigma_i^2} \quad (2)$$

where σ_i is the uncertainty in y_i .

In many cases no statistical information about the uncertainty of y_i is available, and in that case one should make an assumption about the error structure in the y_i values. Most often the relative error in y_i is supposed to be constant and in that case w_i should be taken as $1/y_i^2$; this is equivalent to minimizing the relative residuals:

$$ss(\Theta) = \sum_i \left(\frac{y_i - f(x_i, \Theta)}{y_i} \right)^2 \quad (3)$$

When the absolute errors in y_i can be assumed constant, w_i in eq 1 can be taken as 1 for all datapoints.

Most statistical software packages then look for the minimum in the sum of squares space by iterative procedures. In the case of (multi)linear functions this can be done by evaluating the analytical expressions one obtains when solving the differential equations of the sum of squares function with respect to the parameters. For nonlinear equations, however, the advent of fast personal com-

s1	0.35	0.34	0.33	0.32	0.31	0.30	0.29	0.28	0.27	0.26	0.25
s2	0.38	0.103	0.084	0.071	0.064	0.064	0.071	0.086	0.110	0.142	0.235
0.39	0.109	0.088	0.074	0.065	0.063	0.069	0.083	0.105	0.135	0.176	0.227
0.40	0.116	0.093	0.077	0.066	0.063	0.067	0.079	0.100	0.130	0.169	0.219
0.41	0.123	0.098	0.080	0.068	0.063	0.066	0.077	0.096	0.124	0.163	0.212
0.42	0.130	0.104	0.084	0.070	0.064	0.065	0.074	0.092	0.119	0.157	0.205
0.43	0.137	0.109	0.087	0.072	0.064	0.064	0.072	0.089	0.115	0.151	0.198
0.44	0.145	0.115	0.091	0.075	0.065	0.064	0.070	0.086	0.111	0.146	0.192
0.45	0.152	0.121	0.095	0.077	0.066	0.063	0.069	0.083	0.107	0.141	0.186
0.46	0.159	0.126	0.100	0.080	0.068	0.063	0.068	0.081	0.104	0.137	0.181
0.47	0.167	0.132	0.104	0.083	0.069	0.064	0.067	0.079	0.100	0.133	0.176
0.48	0.174	0.138	0.108	0.086	0.071	0.064	0.066	0.077	0.097	0.129	0.171
0.49	0.182	0.144	0.113	0.089	0.073	0.065	0.065	0.075	0.095	0.125	0.167
0.50	0.189	0.150	0.118	0.092	0.075	0.065	0.065	0.073	0.092	0.122	0.162
0.51	0.197	0.156	0.122	0.095	0.077	0.066	0.064	0.072	0.090	0.118	0.158
0.52	0.204	0.162	0.127	0.099	0.079	0.067	0.064	0.071	0.088	0.115	0.155
0.53	0.212	0.168	0.131	0.102	0.081	0.068	0.064	0.070	0.086	0.113	0.151
0.54	0.219	0.174	0.136	0.106	0.083	0.069	0.064	0.069	0.084	0.110	0.148
0.55	0.226	0.180	0.141	0.109	0.085	0.070	0.064	0.068	0.082	0.108	0.145
0.56	0.233	0.186	0.145	0.112	0.088	0.072	0.065	0.068	0.081	0.105	0.142

Figure 1. Printout of the numerical values for the sum of squares in eq 6.

puters means it is no longer necessary to use iterative procedures.

One could simply calculate all of the sums of squares in a region for the parameters where we can be sure that realistic solutions for the parameters can be found. This can be done in a spreadsheet program like Excel or be programmed in, for example, Basic. By just keeping track of the lowest sum of squares value during the calculation of the sum of squares array one can obtain the least-squares sum $ss(\hat{\Theta})$ at the best set of parameters. When one takes a wide region for the parameters, there is no risk in finding only a local minimum. The latter aspect is one of the drawbacks of using the iterative procedures, where sometimes the result depends on the choice of the starting values of the parameters in the iterative process.

Furthermore we can visualize (in the case of two parameters) the sum of squares space by just printing out the numbers as is done in Figure 1 or plotting it as a surface as done in Figure 2. For three parameter several visualizations as a function of the third parameter can be plotted.

The minimum obtained with this procedure will of course coincide with values obtained by any other (iterative) least squares procedure when the same weighting of the y -values is used. It is also possible to first take a large overview on the sum of squares space and then zoom in onto the region of the minimum. This procedure was used previously to obtain best values for up to four parameters (4).

Parameter Uncertainties

The unbiased joint confidence regions for a level of confidence, z , around the best values found for the parameters can be constructed using the inequality:

$$ss(\Theta) \leq ss(\hat{\Theta}) \left(1 + \frac{p}{i-p} F_z(p, i-p) \right) \quad (4)$$

where $F_z(p, i-p)$ represents a value from the F -distribution at level z and at p and $i-p$ degrees of freedom.

In fact this inequality, which gives the uncertainty in the parameters, represents a contour line in the sum of squares space (Fig. 2).

In this approach no assumptions about linearity of the model functions $f(x_i, \Theta)$ are made with respect to the joint confidence regions, which in some statistical programs are taken as linearized confidence ellipsoids. In some cases it is proven that, where the function is nonlinear, approximate (linearized) confidence ellipsoids are inappropriate (5), whereas the unbiased joint confidence regions from this method are always correct and can for example be obtained by a contour plotting routine (Fig. 2) or by manually drawing the contour line in the sum of squares data in Figure 1.

It is useful to check the deviations between experimental and calculated points. If trends in the residuals are observed, the model may not be adequate for the data. When the actual residuals are in many cases larger than the estimated uncertainties in y_i , the model is not adequate either.

To check whether an alternative model B gives a significant improvement of the fit over model A, despite the fact that a different number of parameters p is used, the following statistical test can be made:

$$\frac{ss_B(\hat{\Theta}) - ss_A(\hat{\Theta}) / (p_A - p_B)}{ss_A(\hat{\Theta}) / (i - p_A)} \geq F_z(p_A - p_B, i - p_A) \quad (5)$$

If this inequality holds, model A provides the better fit at a confidence level z .

An Example in Copolymerization Modeling

One of the fields where using the right statistical procedures has been and still is an important issue is in the determination of reactivity ratios in copolymerization (6). The use of statistically invalid procedures on otherwise good data resulted in tremendous diversity of tabulated reactivity ratios.

A spectacular finding was the fact that the kinetics of copolymerization in most cases could not be described by

Figure 2. Three-dimensional visualization of the sum of squares plane. Contour plot at the 95% confidence limit.

Figure 3. Propagation rate constant as a function of the fraction styrene in the monomer feed ($s_1 = 0.31$, $s_2 = 0.4$).

the so-called ultimate model, where only the last monomeric unit determines the reactivity of a growing polymeric radical, but that a penultimate effect had to be introduced (7).

Fukuda studied the system styrene and methyl methacrylate at 40 °C with the rotating sector method (7). The classic work of Fukuda serves here as an example for the statistical method introduced before.

The propagation rate constant, k_p , can be described as a function of the monomer feed ratio f with the equations

$$\bar{k}_p = \frac{\bar{r}_1^2 + 2f_1f_2 + \bar{r}_2^2}{(\bar{r}_1f_1/\bar{k}_{11}) + (\bar{r}_2f_2/\bar{k}_{22})} \quad (6)$$

where

$$\bar{r}_1 = \frac{r_1'(f_1r_1 + f_2)}{f_1r_1' + f_2}$$

$$\bar{r}_2 = \frac{r_2'(f_2r_2 + f_1)}{f_2r_2' + f_1}$$

and

$$\bar{k}_{11} = \frac{k_{111}(f_1r_1 + f_2)}{f_1r_1 + f_2/s_1}$$

$$\bar{k}_{22} = \frac{k_{222}(f_2r_2 + f_1)}{f_2r_2 + f_1/s_2}$$

with

$$r_i = \frac{k_{iii}}{k_{ijj}} \quad (i \neq j)$$

$$r_i' = \frac{k_{jii}}{k_{jij}} \quad (i \neq j)$$

$$s_i = \frac{k_{jii}}{k_{iii}} \quad (i \neq j)$$

where k_{pmn} is the rate constant for the radical with penultimate unit p , a terminal unit m to add a monomer n ($p, m, n = 1$ or 2). In the so called restricted penultimate model it is assumed that $r_1 = r_1'$ and $r_2 = r_2'$ where the subscript 1 refers to styrene and the subscript 2 refers to methyl methacrylate.

The parameters $r_1 = r_1' = 0.523$, $r_2 = r_2' = 0.460$, $k_{111} = 120 \text{ L mol}^{-1} \text{ s}^{-1}$ and $k_{222} = 377 \text{ L mol}^{-1} \text{ s}^{-1}$ were fixed, leaving s_1 and s_2 as the only unknown parameters. The accepted procedure to obtain the two unknown parameters is the nonlinear least-squares curve-fitting method, which yielded in Fukuda's work $s_1 = 0.30$ and $s_2 = 0.53$ as optimum values.

Because Fukuda states that he has a constant relative error of $\pm 10\%$ in his values we used relative error minimization according to eq 3. For these data, we plotted the sum of squares space numerically in Figure 1 and as a surface in Figure 2. The sum of squares surface is plotted as $ss(\Theta)$ subtracted from a constant C to visualize the fitting optimum as a maximum ($C = 13.2$). As a minimum we obtained $s_1 = 0.31$ and $s_2 = 0.40$ with a relative sum of squares of 0.063. Also we found a local minimum at the same values Fukuda reported with a sum of squares at a somewhat higher value of 0.064.

This fact shows the strength of the method; there is less risk of obtaining only a local minimum. We also plotted the 95% unbiased joint confidence interval for s_1 and s_2 in Figure 2, which shows that the model is very insensitive to variations in s_2 , explaining the occurrence of several local minima. In a reanalysis of the original data of Fukuda, Davis et al. (8) calculated the linearized joint confidence region, which also showed the insensitivity of the model for s_2 .

Examination of the residuals showed that there is no systematic variation and that the actual deviations of the data with respect to the fitted k_p values are not bigger than the expected errors. From this example we can conclude that the penultimate model does fit the k_p data well as shown in Figure 3, but that the model is not sensitive for variations in s_2 . This is also shown by the fact that the original fit of Fukuda ($s_1 = 0.30$, $s_2 = 0.53$) gives a quite similar plot as that shown in Figure 3. Furthermore it is shown that local minima can be found in this sum of squares space. It is important to take into account the error structure in the y -values because this determines the choice of the weighting factors. In this example, when we use eq 1 with no weighting, we obtain $s_1 = 0.29$ and $s_2 = 0.59$.

The importance of the use of unbiased (delineated) joint confidence regions in model evaluation in copolymerization is also recognized by Hill et al. (9).

To calculate a sum of squares space in this example with a simple Basic program (listing available on request from the author), it takes only a few seconds on a personal computer to produce the matrix shown in Figure 1.

This statistical method has been applied with success in a second-year practical polymer course at the Eindhoven University of Technology.

Conclusion

A visualization of the sum of squares space together with obtaining the least-squares sum by the described technique is a general statistically valid method that can be applied to linear, multilinear, and nonlinear equations and also gives an unbiased joint confidence region.

The method is transparent to students, easy to explain, applicable to almost any problem, and easily realized in compact computer programs.

Acknowledgement

The author would like to thank H. Linssen for critically evaluating the manuscript.

Literature Cited

- Christian, S. D.; Tucker, E. E. *J. Chem. Educ.* **1984**, *61*, 788.
- Trindle, T. J. *Chem. Educ.* **1983**, *60*, 566–567.
- O'Neill, R. T.; Flaspohler, D. C. *J. Chem. Educ.* **1990**, *67*, 40–42.
- Van Herk, A. M.; Tullemans, A. H. J.; van Welzen, J.; German, A. L. *J. Mol. Catal.* **1988**, *44*, 269–277.
- Linssen, H. N. *Statistica Neerlandica* **1975**, *29*, 93–99.
- O'Driscoll, K. F.; Reilly, P. M. *Makromol. Chem. Macromol. Symp.* **1987**, *10/11*, 355–374.
- Fukuda, T.; Ma, Y.-D.; Inagaki, H. *Makromolecules* **1985**, *18*, 17–26.
- Davis, T. P.; O'Driscoll, K. F.; Piton, M. C.; Winnik, M. A. *J. Polym. Sci. Polymer Letters Ed.* **1989**, *27*, 181–185.
- Hill, D. J. T.; O'Donnell, J. H. *Makromol. Chem. Macromol. Symp.* **1984**, *10/11*, 375–394.