

# Analytical and numerical computation of error propagation of model parameters

Gyula Tasi · Dóra Barna

Received: 14 March 2011 / Accepted: 21 April 2011 / Published online: 7 May 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** Linear and nonlinear regression analyses based on the least-squares method play a fundamental role in the evaluation of scientific data. A large number of valuable papers have dealt with various applications of the least-squares method in the chemical literature. They, however, usually contain tremendous formulas for computing the error estimates of the estimated parameters. This paper presents a simple numerical solution based on the well-known simplex method to this problem. Elaborate enzyme kinetic data published earlier have been chosen to test the simplex method extended with error estimation. The capability of the numerical method is demonstrated by the revision of the originally calculated error propagation. Our program might prove useful in handling either chemical or biological data.

**Keywords** Least-squares method · Parameter estimation · Simplex method · Error estimation · Variance-covariance matrix

## 1 Introduction

Linear and nonlinear regression analyses based on the least-squares (LS) method play a fundamental role in the evaluation of scientific data [1]. These methods have been successfully utilized in the evolution of current active databases in the field of thermochemistry (ATcT [2] and NEAT [3]) and molecular spectroscopy (MARVEL [4]).

The various applications of the LS method are well-documented, a large number of valuable papers have been devoted to the topic in the chemical and biological literature too, e.g., in the field of structure-property/activity relationships [5–8] or of

---

G. Tasi (✉) · D. Barna  
Department of Applied and Environmental Chemistry, University of Szeged, Rerrich B. tér 1,  
6720 Szeged, Hungary  
e-mail: tasi@chem.u-szeged.hu

fitting helices to data points [9–12]. However, some of them contain slightly cumbersome derivations and tremendous formulas for computing the error estimates of the estimated parameters. In this paper, a simple numerical solution to the problem is presented which can be utilized in handling of either chemical or biological data.

## 2 Methods

### 2.1 The least-squares method

The least-squares method was invented more or less independently by three famous scientists in the nineteenth century [13]: Adrien Marie Legendre (1805), Robert Adrain (1808), and Carl Friedrich Gauss (1809).

Let us suppose that we have experimental data points  $\{(X_i, Y_i), i = 1, \dots, m\}$  containing uncorrelated random errors both in the  $X_i$  and  $Y_i$  values, and a mathematical model  $f(\mathbf{x}; \mathbf{p})$  with parameters  $\{p_i, i = 1, \dots, n\}$  collected in the parameter vector  $\mathbf{p}$ , which gives the assumed quantitative relationship between the variables:

$$\begin{aligned} X_i &= x_i + \varepsilon_{xi} \\ Y_i &= y_i + \varepsilon_{yi} \\ y_i &= f(x_i; \mathbf{p}) \end{aligned} \quad (1)$$

where  $x_i$  and  $y_i$  are the precise, error-free quantities. For the sake of simplicity, we consider only one dependent and one independent variable in this paper. It is worth noting, however, that for a correct model the roles of the variables, from statistical point of view, do not matter: they are interchangeable. If our model function is linear in all parameters, we are talking about linear LS (LLS) and linear regression analysis. Otherwise, we are talking about nonlinear LS (NLS) and nonlinear regression analysis.

Let us assume that our random errors are normally distributed with zero expectation values, nonzero variances, and they are independent of each other:

$$\begin{aligned} E(\varepsilon_{yi}) &= 0 \\ D^2(\varepsilon_{yi}) &= \sigma_{yi}^2 \\ \text{Cov}(\varepsilon_{yi}, \varepsilon_{yj}) &= 0 \quad (i \neq j) \end{aligned} \quad (2)$$

$$\begin{aligned} E(\varepsilon_{xi}) &= 0 \\ D^2(\varepsilon_{xi}) &= \sigma_{xi}^2 \\ \text{Cov}(\varepsilon_{xi}, \varepsilon_{xj}) &= 0 \quad (i \neq j) \end{aligned} \quad (3)$$

We also assume implicitly here that there are no correlations between the errors of the independent and the dependent variables.

According to the LS philosophy, our main goal is to determine the optimal  $\hat{\mathbf{p}}$  vector, which minimizes the objective function  $S(\mathbf{p})$  in Eq. 4, that is, the weighted sums of the residuals  $R_{xi} = X_i - \hat{x}_i$  and  $R_{yi} = Y_i - \hat{y}_i$  in the general case [14–16] with weights  $w_{yi} = \sigma_{yi}^{-2}$  and  $w_{xi} = \sigma_{xi}^{-2}$ .

$$S(\mathbf{p}) = \sum_{i=1}^m w_{yi} R_{yi}^2 + \sum_{i=1}^m w_{xi} R_{xi}^2 \quad (4)$$

Minimization of the objective function thus minimizes the sum of the weighted perpendicular distances between the experimental and the computed points. If we use only the first term of Eq. 4, i.e.,  $\forall \sigma_{xi} = 0$ , and  $\forall \sigma_{yi} = \sigma$  (marked as zero  $\delta x$  and uniform  $\delta y$  in ref. [17]), we are talking about ordinary least-squares (OLLS or ONLS) method. For the weighted least-squares (WLLS or WNLS) method, the errors in the  $y$  values are not all the same (marked as zero  $\delta x$  and nonzero  $\delta y$  in ref. [17]). In the general case, the errors both in the  $x$  and  $y$  values are nonzero (marked as nonzero  $\delta x$  and nonzero  $\delta y$  in ref. [17]): general least-squares (GLLS or GNLS) method. Besides  $\delta x$  and  $\delta y$ , we shall also use the  $SE(x)$  and  $SE(y)$  notations for the standard errors of the independent and dependent variables, respectively.

It can be shown [16] that, under some circumstances, Eq. 4 can be reduced to Eq. 5, to a much simpler form, introducing appropriate weights.

$$S(\mathbf{p}) = \sum_{i=1}^m w_i R_{yi}^2 \quad (5)$$

For instance, if the noise in  $x$  is not so large then Eq. 6 gives the weights [16] in Eq. 5.

$$w_i = \frac{1}{\sigma_{yi}^2 + \left(\frac{\partial f}{\partial x} \sigma_{xi}\right)^2} \quad (6)$$

Using a simple straight line with two parameters,  $a$  and  $b$ , as a model function:

$$y_i = f(x_i; a, b) = ax_i + b \quad (7)$$

Equations 5 and 6 are exact, and Eq. 6 reads as follows:

$$w_i = \frac{1}{\sigma_{yi}^2 + a^2 \sigma_{xi}^2} \quad (8)$$

It is worth mentioning that Irvin and Quickenden have presented an approximate, iterative solution [18] based on Eqs. 5 and 8 to the GLLS problem. Their algorithm can also be used in the OLLS and WLLS cases.

## 2.2 The simplex method

The sequential simplex method in its variable size form was developed by Nelder and Mead [19] in 1965 to minimize functions of several variables. However, the procedure can also be used, with little tricks, to locate not only local minima but also local maxima and saddle points. The simplex procedure is a robust, direct-search method: it needs

smooth function's values only. Its usefulness in chemistry has been demonstrated in several papers [20–22].

In general, it is easy to find good approximations to the optimal parameter values, from which the initial simplex can be generated. Since the simplex method uses function's values only, it is important to scale the values of the parameters to keep them approximately in the same range. In ambiguous cases or to check the quality of the point obtained by the simplex procedure, it is advisable to start the whole search anew from the last point. It can turn out that the procedure terminated abnormally and the new start would result in a better point with lower function's value.

Unfortunately, the simplex method did not possess originally error analysis. This imperfection, however, can easily be remedied [19, 23]. Only the variance-covariance matrix  $\mathbf{V}$  should be computed at the minimum as follows:

$$\begin{aligned} \mathbf{V} &= 2s_{r2}\mathbf{H}^{-1} \\ s_{r2} &= \frac{S(\hat{\mathbf{p}})}{m-n} \\ H_{ij} &= \left. \frac{\partial^2 S}{\partial p_i \partial p_j} \right|_{\mathbf{p}=\hat{\mathbf{p}}} \end{aligned} \quad (9)$$

where  $s_{r2}$  is the residual variance and  $\mathbf{H}^{-1}$  is the inverse of the computed Hessian  $\mathbf{H}$  built up from the second-order derivatives of the objective function with respect to the parameters. The Hessian of the objective function can be calculated analytically or numerically. If we do not want to spend frequently too much time to derive the analytical form of the Hessian, the numerical computation should then be our choice.

Numerical computation of the first- and second-order derivatives of the objective function with respect to the parameters is straightforward using various finite-difference formulas [24]. The gradient vector of the objective function consists of the first-order derivatives as components. At a stationary (critical) point, the gradient vector vanishes. The Hessian is a real and symmetric matrix with real eigenvalues and orthogonal eigenvectors. Through the analysis of the eigenvalues, it is possible to check the character of the stationary point obtained. For example, if the gradient vector is null vector, and all the eigenvalues are positive, then our stationary point is a local minimum. The standard eigenvalue equation of the Hessian can be solved and its inverse matrix can be generated via numerical methods detailed e.g., in ref. [25].

The estimated standard errors of the optimal parameter values are obtained from the diagonal elements of the variance-covariance matrix, while the correlation coefficients representing the pair correlations of the optimal parameter values are derived from the off-diagonal covariances and diagonal variances:

$$SE(\hat{p}_i) = \sqrt{\text{Var}(\hat{p}_i)} = \sqrt{V_{ii}} \quad (10)$$

$$r_{\hat{p}_i \hat{p}_j} = \frac{\text{Cov}(\hat{p}_i, \hat{p}_j)}{SE(\hat{p}_i) SE(\hat{p}_j)} = \frac{V_{ij}}{\sqrt{V_{ii}} \sqrt{V_{jj}}} \quad (11)$$

**Table 1** The experimental enzyme kinetic data<sup>a</sup>

[S] <sup>b</sup>	SE([S])	<i>r</i> <sup>c</sup>	SE( <i>r</i> )
0.138	0.011	0.148	0.015
0.220	0.015	0.171	0.015
0.291	0.017	0.234	0.019
0.560	0.028	0.324	0.023
0.766	0.031	0.390	0.023
1.460	0.044	0.493	0.025

<sup>a</sup> The experimental data are from ref. [17]

<sup>b</sup> The concentration of nicotinamide mononucleotide per millimole

<sup>c</sup>  $\mu$  moles of nicotinamide-adenine dinucleotide formed in 1 min/mg of enzyme protein

### 3 Results and discussion

For testing the capability of the numerical method in linear and nonlinear regression analyses, an instructive example concerning the Michaelis-Menten enzyme kinetics presented elsewhere [17] has been chosen. The simplex method extended with error estimation provided a useful tool for the revision of the originally calculated error propagation. The Michaelis-Menten equation, Eq. 12, is a nonlinear equation, which is frequently used as a first approximation to evaluate the experimental data of enzyme-catalyzed reactions.

$$r = \frac{V_{\max}[S]}{K_M + [S]} \quad (12)$$

In Eq. 12, *r* is the initial rate;  $V_{\max}$  is the maximum rate; [S] is the substrate concentration; and  $K_M$  is the Michaelis constant. Equation 13 is the reverse form of Eq. 12.

$$[S] = \frac{K_M r}{V_{\max} - r} \quad (13)$$

Of the linearized forms of Eqs. 12 and 13, only the Woolf-Lineweaver-Burk form [26], Eq. 14, and its reverse form, Eq. 15, are considered in this paper.

$$\frac{1}{r} = \frac{K_M}{V_{\max}} \frac{1}{[S]} + \frac{1}{V_{\max}} \quad (14)$$

$$\frac{1}{[S]} = \frac{V_{\max}}{K_M} \frac{1}{r} - \frac{1}{K_M} \quad (15)$$

The experimental data set, listed in Table 1, was published by Chong [17].

Equations 14 and 15 need the experimental data in Table 1 to be transformed. Salter's paper [27] gives the general formula for error propagation. The proper use of the formula results in Eqs. 16 and 17 for the errors of the reciprocal quantities.

**Table 2** The transformed experimental data of Table 1

$1/[S]$	$SE(1/[S])$	$1/r$	$SE(1/r)$
7.2464	0.5776	6.7568	0.6848
4.5455	0.3099	5.8480	0.5130
3.4364	0.2008	4.2735	0.3470
1.7857	0.0893	3.0864	0.2191
1.3055	0.0528	2.5641	0.1512
0.6849	0.0206	2.0284	0.1029

**Table 3** Results of the LS fits on the data in Tables 1 and 2<sup>a</sup>

Eq	Method	This work		Literature data <sup>b</sup>	
		$K_{Mc}$	$V_{max}^d$	$K_{Mc}$	$V_{max}^d$
14	OLLS	0.4406(1185)	0.5853(1039)	0.4406(906)	0.5853(1039)
	WLLS	0.5552(739)	0.6715(455)	0.5552(556)	0.6715(455)
	GLLS	0.5721(742)	0.6797(436)	0.5721(562)	0.6797(439)
15	OLLS	0.4891(1404)	0.6228(1219)	0.4891(1404)	0.6228(1219)
	WLLS	0.6190(557)	0.7016(246)	0.6190(558)	0.7016(247)
	GLLS	0.5721(742)	0.6797(436)	0.5721(750)	0.6797(439)
12	ONLS	0.5965(692)	0.6904(373)	0.5965(692)	0.6904(373)
	WNLS	0.5869(821)	0.6849(495)	0.5869(821)	0.6849(495)
	GNLS	0.5860(771)	0.6854(452)	0.5860(708)	0.6854(429)
13	ONLS	0.6244(394)	0.7039(149)	–	–
	WNLS	0.5906(526)	0.6912(232)	–	–
	GNLS	0.5860(779)	0.6854(458)	–	–

<sup>a</sup> Values in parentheses represent one standard deviation in the final digits of the quoted values

<sup>b</sup> The literature data are from ref. [17]

<sup>c</sup> Michaelis constant in millimole/L

<sup>d</sup> Maximum initial rate in  $\mu$  moles of nicotinamide-adenine dinucleotide formed in 1 min/mg of enzyme protein

$$SE(1/[S]) = SE([S])/[S]^2 \tag{16}$$

$$SE(1/r) = SE(r)/r^2 \tag{17}$$

Table 2 shows the properly transformed experimental data of Table 1.

The results of the LS fits of the data in Tables 1 and 2 are listed in Table 3. It can be seen that the estimated values of the parameters obtained by the simplex method are the same as those presented in ref. [17]. However, the estimated errors differ in numerous cases. For instance, let us see the first row of Table 3. The same optimal  $K_M$  value has different error estimates in the two cases: 0.1185 and 0.0906, respectively. To shed light on the cause of this difference, let us calculate the error estimate analytically with the error propagation formula given by Salter [27].

Using Eq. 14 in the form of Eq. 7, we get

$$a = \frac{K_M}{V_{\max}} \quad (18)$$

$$b = \frac{1}{V_{\max}} \quad (19)$$

$$c = f(a, b) = \frac{a}{b} = K_M \quad (20)$$

Standard OLLS fit gives the following results

$$\hat{a} = 0.75280 \pm 0.07815 \quad (21)$$

$$\hat{b} = 1.70845 \pm 0.30326 \quad (22)$$

$$\text{Cov}(\hat{a}, \hat{b}) = -0.01935 \quad (23)$$

Applying the error propagation formula given in ref. [27], we can compute the variance of  $c$  as follows:

$$\begin{aligned} \text{Var}(\hat{c}) &= [\text{grad } f]^T \cdot \mathbf{V} \cdot \text{grad } f \\ &= (1/\hat{b} - \hat{a}/\hat{b}^2) \begin{pmatrix} \text{Var}(\hat{a}) & \text{Cov}(\hat{a}, \hat{b}) \\ \text{Cov}(\hat{a}, \hat{b}) & \text{Var}(\hat{b}) \end{pmatrix} \begin{pmatrix} 1/\hat{b} \\ -\hat{a}/\hat{b}^2 \end{pmatrix} \\ &= \frac{\text{Var}(\hat{a})}{\hat{b}^2} + \frac{\hat{a}^2 \text{Var}(\hat{b})}{\hat{b}^4} - 2 \frac{\hat{a} \text{Cov}(\hat{a}, \hat{b})}{\hat{b}^3} \\ &= 0.00209 + 0.00612 + 0.00584 \end{aligned} \quad (24)$$

If we consider only the first two terms in calculating the standard error of  $c$ , that is, we neglect the covariance of  $a$  and  $b$ , then Eq. 25 is obtained.

$$\text{SE}(\hat{c}) = \sqrt{\text{Var}(\hat{c})} = \sqrt{0.00209 + 0.00612} = 0.0906 \quad (25)$$

However, if we take the correlation of the two parameters into account, that is, we consider all the three terms in Eq. 24, we get Eq. 26.

$$\text{SE}(\hat{c}) = \sqrt{\text{Var}(\hat{c})} = \sqrt{0.00209 + 0.00612 + 0.00584} = 0.1185 \quad (26)$$

It can be seen that the simplex method with error estimation gives the very same number. Similar derivations reveal that all the error estimates of this work are correct. Furthermore, for a correct model the roles of the variables do not matter: we should get the same results after interchanging the roles of the variables. In this work, both the GLLS and GNLS methods provide the same results for the reverse forms of the fitted functions (see the lines 3 and 6 as well as the lines 9 and 12 in Table 3). However, this is not so for the literature data: they differ from each other.

The simplex method with error estimation can also be used for global least-squares fit. For example, it is a very good test exercise to determine the absolute zero temperature via global least-squares fit [28] using the simplex method. We can get the same results as presented in ref. [28] without doing lengthy math derivations concerning the propagation of the errors of the parameters.

## 4 Conclusion

The simplex procedure with error estimation can be used for linear and nonlinear regression analyses of experimental data according to the least-squares philosophy. Only the numerical or analytical computation of the Hessian is needed to obtain the variance-covariance matrix from which all pieces of information can be extracted about the variances and the covariances of the estimated parameters. The numerical method presented can cope with a large number of parameter estimation problems occurring in the field of chemistry and biology.

## References

1. E.B. Wilson Jr., *An Introduction to Scientific Research* (Dover Publications, New York, 1990)
2. B. Ruscic, R.E. Pinzon, M.L. Morton, G.V. Laszevski, S.J. Bittner, S.G. Nijssure, K.A. Amin, M. Minkoff, A.F. Wagner, *J. Phys. Chem. A* **108**, 9979 (2004)
3. A.G. Császár, T. Furtenbacher, *Chem. Eur. J.* **16**, 4826 (2010)
4. T. Furtenbacher, A.G. Császár, J. Tennyson, *J. Mol. Spectrosc.* **245**, 115 (2007)
5. J. Chen, X. Xue, K.W. Schramm, X. Quan, X. Yang, A. Kettrup, *Comput. Biol. Chem.* **27**, 165 (2003)
6. J.W. Chen, T. Harner, K.W. Schramm, X. Quan, X.Y. Xue, A. Kettrup, *Comput. Biol. Chem.* **27**, 405 (2003)
7. K. Hasegawa, M. Arakawa, K. Funatsu, *Comput. Biol. Chem.* **27**, 211 (2003)
8. K. Hasegawa, S. Matsuoka, M. Arakawa, K. Funatsu, *Comput. Biol. Chem.* **27**, 381 (2003)
9. J. Åquist, *Comput. Chem.* **10**, 97 (1986)
10. J.A. Christopher, R. Swanson, T.O. Balwin, *Comput. Chem.* **20**, 339 (1996)
11. P. Enkhbayar, S. Damdinsuren, M. Osaki, N. Matsushima, *Comput. Biol. Chem.* **32**, 307 (2008)
12. P.C. Kahn, *Comput. Chem.* **13**, 185 (1989)
13. S.M. Stigler, *Annals of Statistics* **9**, 465 (1981)
14. W.E. Deming, *Statistical Adjustment of Data* (John Wiley & Sons, New York, 1943)
15. D.R. Powell, J.R. Macdonald, *Comput. J.* **15**, 148 (1972)
16. J. Wolberg, *Prediction Analysis* (Van Nostrand, Princeton, 1967)
17. D.P. Chong, *J. Chem. Educ.* **71**, 489 (1994)
18. J.A. Irvin, T.I. Quickenden, *J. Chem. Educ.* **60**, 711 (1983)
19. J.A. Nelder, R. Mead, *Comput. J.* **7**, 308 (1965)
20. D.S. Amenta, C.E. Lamb, J.J. Leary, *J. Chem. Educ.* **56**, 557 (1979)
21. S.G. Lieb, *J. Chem. Educ.* **74**, 1008 (1997)
22. C.L. Shavers, M.L. Parsons, S.N. Deming, *J. Chem. Educ.* **56**, 307 (1979)
23. G.R. Phillips, E.M. Eyring, *Anal. Chem.* **60**, 738 (1988)
24. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, New York, 1972)
25. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1994)
26. H. Lineweaver, D. Burk, *J. Am. Chem. Soc.* **56**, 658 (1934)
27. C. Salter, *J. Chem. Educ.* **77**, 1239 (2000)
28. C. Salter, *J. Chem. Educ.* **80**, 1033 (2003)