
DEPARTMENT OF STATISTICS

University of Wisconsin
Madison, Wisconsin 53706

TECHNICAL REPORT NO. 375

June, 1974

A QUICK AND DIRTY METHOD FOR SOLVING THE NON-LINEAR IMPLICIT REGRESSION PROBLEM

by

Grace Wahba* and Svante Wold**
University of Wisconsin
Madison, Wisconsin

Typist: Mary E. Arthur

** Statistician in Residence 1973-74.

Permanent address: Research Group for Chemometrics, Institute of Chemistry, Umeå University, S 901 87 Umeå, Sweden.

* Research supported by the United States Air Force under Grant No. AFOSR 72-2363-B.

A QUICK AND DIRTY METHOD FOR SOLVING THE NON-LINEAR IMPLICIT REGRESSION PROBLEM

by

Grace Wahba and Svante Wold

ABSTRACT

An approximate method is given to handle the implicit nonlinear regression problem with the model being defined implicitly by $F(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) = 0$; $i = 1,2,\ldots,N$. The observations Y_i are assumed to satisfy $Y_i = y_i + \varepsilon_i$ where ε_i is the measurement error.

The parameters $\underset{\sim}{\theta}$ and their standard errors are estimated by approximately minimizing $\sum (Y_i - y_i)^2 / w_i^2$. The method is based on approximating $Y_i - y_i$ by terms from the Taylor expansion of $F(y, \underset{\sim}{x}, \underset{\sim}{\theta})$. The method is tested on simulated data and applied to drug-protein binding data.

Key words. Estimation
Nonlinear
Least Squares
Implicit Models

1. Introduction.

In relating observed data to a theoretical model, one sometimes encounters the problem that the observed response (Y_i) cannot be related to the predictor variables (vector $\underset{\sim}{x}_i$) in the explicit way needed for the application of non-linear regression:

$$Y_i = f(\underset{\sim}{x}_i, \underset{\sim}{\theta}) + \varepsilon_i, \quad i = 1,2,\ldots,N. \tag{1}$$

Often, however, the theoretical relation can be implicitly formulated by means of the theoretical error-free response (y_i) and the predictor variables ($\underset{\sim}{x}_i$):[1]

$$F(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) = 0 \tag{2}$$

$$Y_i = y_i + \varepsilon_i \tag{3}$$

As usual, we assume that the vector of predictor variables ($\underset{\sim}{x}_i$) is free from errors and that the deviations (ε_i) are independent, random and having zero mean and variance $w_i^2\sigma^2$ where w_i is known. Furthermore, we assume that the parameter vector $\underset{\sim}{\theta}$ and the predictor variables $\underset{\sim}{x}_i$ define a unique y_i value by means of eq. (2).

Given observed data ($Y_i, \underset{\sim}{x}_i$; $i = 1,2,\ldots,N$) and the form of eq. (2), least squares estimates of the parameters are obtained by the minimization of

---

[1] The quantities Y_i, F, f and ε_i in eq.s (1) to (3) could be vectors, but for notational simplicity we discuss the problem in terms of the one-dimensional case.

$$Q_N(\underset{\sim}{\theta}) = \frac{1}{N} \sum_{i=1}^{N} (Y_i - y_i)^2 / w_i^2 . \qquad (4)$$

It might be tempting to estimate the parameters by instead minimizing

$$q_N(\underset{\sim}{\theta}) = \frac{1}{N} \sum_{i=1}^{N} F(Y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta})^2 \qquad (5)$$

but, as shown by Mezaki, Draper and Johnson (1973), this is an incorrect procedure which does not correspond to the minimization of the residual sum of squares. The result of using the incorrect form of eq. (5) can lead to a large spread in the parameter estimates as well as an appreciable bias.

Hence, it is of importance to calculate the estimates of $\underset{\sim}{\theta}$ by minimizing $Q_N(\underset{\sim}{\theta})$ in eq. (4) when the model is implicitly defined by eq.s (2) and (3). This problem has been treated by Deming (1943), Britt and Luecke (1973) and others. These treatments involve the minimization of eq. (4) under the constraints that eq. (2) is fulfilled for all points (i = 1,2,...,N) and require substantial programming efforts for the analysis according to a particular model.

In the present work, we derive an approximation of the deviation ε_i, by means of which eq. (4) can be minimized directly in a simple way, utilizing readily available standard programs. This approximation is good as long as the deviations (ε_i) are fairly small. Therefore, a limited number of simulations have been carried out to study the applicability of the methods. The methods have also been applied to drug-protein binding data,

providing estimates of the binding constants and their standard errors.

2. Approximate least squares estimates for the implicit model.

By simple Taylor expansions of the function $F_i(\eta) = F(\eta, \underset{\sim}{x}_i, \underset{\sim}{\theta})$ around the point $\eta = y_i$, we will derive an approximate expression for $\varepsilon_i = Y_i - y_i$. This derivation will rest on the assumptions that derivatives of F with respect to η and the elements of $\underset{\sim}{\theta}$ exist to the first and second order and that the second derivatives are small in a suitable region around $\eta = y_i$. Thus, eq. (2) gives:

$$0 = F_i(y_i) = F_i(Y_i) - (Y_i - y_i)\left.\frac{\partial F_i(\eta)}{\partial \eta}\right|_{\eta=y_i} - (Y_i - y_i)^2/2 \left.\frac{\partial^2 F_i(\eta)}{\partial \eta^2}\right|_{\eta=y_i} - \cdots \tag{6}$$

If quadratic and higher terms can be neglected, eq. (6) gives

$$\varepsilon_i \approx \varepsilon_i^* = F_i(Y_i) \Big/ \left.\frac{\partial F_i(\eta)}{\partial \eta}\right|_{\eta=Y_i} \tag{7a}$$

Noting that

$$\left.\frac{\partial F_i}{\partial \eta}\right|_{\eta=Y_i} \cong \left.\frac{\partial F_i}{\partial \eta}\right|_{\eta=y_i} + (Y_i - y_i) \left.\frac{\partial^2 F_i}{\partial \eta^2}\right|_{\eta=y_i}$$

and approximating $Y_i - y_i$ by ε_i^*, we also get (from eq. 6) the slightly more complicated expression

$$\varepsilon_i \approx \varepsilon_i^* + \frac{1}{2}(\varepsilon_i^*)^2 \left\{\frac{\partial^2 F_i}{\partial \eta^2} \Big/ \frac{\partial F_i}{\partial \eta}\right\}_{\eta=Y_i}$$

$$\equiv \varepsilon_i^{**} = \left\{F_i / \frac{\partial F_i}{\partial \eta} + \frac{1}{2} F_i^2 \frac{\partial^2 F_i}{\partial \eta^2} \Big/ \left(\frac{\partial F_i}{\partial \eta}\right)^3\right\}_{\eta=Y_i} \qquad (7b)$$

Hence, the least squares estimate can now be found by minimizing the approximate sum of residual squares. Substitution of eq. (7a) in eq. (4) gives:

$$\tilde{Q}_N(\underset{\sim}{\theta}) \cong \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{w_i} F(Y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) \left.\frac{\partial F(\eta, \underset{\sim}{x}_i, \underset{\sim}{\theta})}{\partial \eta}\right|_{\eta=Y_i}\right]^2 \qquad (8a)$$

Analogously, eq. (7b) leads to (in shorter notation):

$$\tilde{Q}_N(\underset{\sim}{\theta}) \approx \frac{1}{N} \sum_{i=1}^{N} \left\{\frac{1}{w_i} \left[F / \frac{\partial F}{\partial \eta} + \frac{1}{2} F^2 \frac{\partial^2 F}{\partial \eta^2} \Big/ \left(\frac{\partial F}{\partial \eta}\right)^3\right]\right\}^2_{\eta=Y_i} \qquad (8b)$$

The minimization of eq.s (8a) or (8b) with respect to $\underset{\sim}{\theta}$ can be carried out by standard methods, e.g. the Marquardt (1963) method or similar methods, which are available as standard packages at most computer centers. The user needs only to provide a small subroutine defining the function $\tilde{Q}_N$ in eq. (8) and the derivatives of this function with respect to all elements in the parameter vector $\underset{\sim}{\theta}$.

It should be noted that the goodness of the approximation (7) depends on the size of the second derivatives,

$$\left.\partial^2 F(\eta, \underset{\sim}{x}_i, \underset{\sim}{\theta}) / \partial \eta^2\right|_{\eta=Y_i} \qquad (9)$$

the approximation being exact when these quantities and higher partial derivatives are zero. Generally, there exist several choices of the function F in eq. (2). Specifically, if $H(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta})$ is a function strictly positive for all values of y_i, $\underset{\sim}{x}_i$ and $\underset{\sim}{\theta}$ in the region of interest, then the following expression is also an implicit definition of y_i:[2]

$$G(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) = H(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) \cdot F(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) = 0 \tag{10}$$

The function H can be chosen so as to minimize the second derivative of G with respect to η, thus making the approximation (7) as good as possible.

3. The approximate covariance matrix of the parameters.

Let $\hat{\underset{\sim}{\theta}}_N$ be the estimated parameter vector, which minimizes $\tilde{Q}_N(\underset{\sim}{\theta})$ according to eq. (8a). An approximate covariance matrix of $\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N$ can be found by using the definitions above. Hence

$$\frac{\partial}{\partial \theta_\nu} \tilde{Q}_N(\underset{\sim}{\theta}) \Big|_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N} = 0 = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{w_i^2} \left\{ \left[\frac{F_i}{\partial F_i / \partial \eta} \right] \frac{\left(\frac{\partial F_i}{\partial \eta} \frac{\partial F_i}{\partial \theta_\nu} - F_i \frac{\partial^2 F_i}{\partial \eta \partial \theta_\nu} \right)}{\left(\frac{\partial F_i}{\partial \eta} \right)^2} \right\}, \tag{11}$$

where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$, $\hat{\underset{\sim}{\theta}}_N = (\hat{\theta}_{1N}, \hat{\theta}_{2N}, \ldots, \hat{\theta}_{pN})$.

In addition, the following is true:

$$\begin{aligned} F(Y_i, x_i, \hat{\underset{\sim}{\theta}}_N) - F(y_i, \underset{\sim}{x}_i, \underset{\sim}{\theta}) = {} & F(Y_i, \underset{\sim}{x}_i, \hat{\underset{\sim}{\theta}}_N) - F(y_i, \underset{\sim}{x}_i, \hat{\underset{\sim}{\theta}}_N) \\ & + F(y_i, x_i, \hat{\underset{\sim}{\theta}}_N) - F(y_i, x_i, \underset{\sim}{\theta}). \end{aligned} \tag{12}$$

[2] We thank Professor H. Wold for this observation.

Using eq. (2) together with eq. (7a), neglecting second and higher order terms in $(\theta_\nu - \hat{\theta}_{\nu N})$, $\nu = 1,2,\ldots,p$; eq. (12) gives:

$$F(Y_i, x_i, \hat{\underset{\sim}{\theta}}_N) \approx \left(\frac{\partial F_i}{\partial \eta}\bigg|_{\eta=Y_i}\right) \varepsilon_i - \sum_{\nu=1}^{p} \left(\frac{\partial F_i}{\partial \theta_\nu}\bigg|_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}}\right) (\theta_\nu - \hat{\theta}_{\nu N}). \qquad (13)$$

Substitution of eq. (13) in eq. (11), again neglecting all higher order terms in both ε_i, $(\theta_\nu - \hat{\theta}_{\nu N})$ and $F(Y_i, \underset{\sim}{x}_i, \hat{\underset{\sim}{\theta}}_N)^2$ gives:

$$0 \approx \frac{1}{N} \sum_{i=1}^{N} \frac{1}{w_i^2} \left\{ \left[\varepsilon_i - \frac{1}{\frac{\partial F_i}{\partial \eta}} \sum_{\nu=1}^{p} \frac{\partial F_i}{\partial \theta_\nu} (\theta_\nu - \hat{\theta}_{\nu N}) \right] \left(\frac{\partial F_i}{\partial \theta_\nu} \Big/ \frac{\partial F_i}{\partial \eta} \right) \right\}_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}} \qquad (14)$$

Solving this system of p simultaneous equations for $\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N$ gives:

$$\sqrt{N}(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N) \approx A_N^{-1}(\hat{\underset{\sim}{\theta}}_N) \begin{pmatrix} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\varepsilon_i}{w_i^2} \left(\frac{\frac{\partial F_i}{\partial \theta_1}}{\frac{\partial F_i}{\partial \eta}} \right)_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}} \\ \vdots \\ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\varepsilon_i}{w_i^2} \left(\frac{\frac{\partial F_i}{\partial \theta_p}}{\frac{\partial F_i}{\partial \eta}} \right)_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}} \end{pmatrix},$$

where $A_N(\underset{\sim}{\theta})$ is the $p \times p$ matrix with $\mu\nu$ the entry $a_{N\mu\nu}(\underset{\sim}{\theta})$ given by

$$a_{N\mu\nu}(\underset{\sim}{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{w_i^2} \left[\left(\frac{\partial F_i}{\partial \theta_\mu} \right) \left(\frac{\partial F_i}{\partial \theta_\nu} \right) \right]_{\eta=Y_i} ,$$

and we are assuming that $A_N(\hat{\underset{\sim}{\theta}}_N)$ is non-singular.

Thus, the covariance matrix of $\sqrt{N}(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N)$ is approximately given by

$$E[\sqrt{N}(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N)' \sqrt{N}(\underset{\sim}{\theta} - \hat{\underset{\sim}{\theta}}_N)] \approx \sigma^2 \, A_N^{-1}(\hat{\theta}_N) \qquad (15)$$

which agrees with Britt and Luecke's eq. (32).

To the extent that the approximation leading to the expression (7) is valid, σ^2 may be estimated by $\hat{\sigma}_N^2$ given by

$$\hat{\sigma}_N^2 = \tilde{Q}_N(\hat{\underset{\sim}{\theta}}_N) \quad .$$

Finally, we remark that if $F_i(\eta, \underset{\sim}{\theta}) = H(\eta, y_i, \underset{\sim}{\theta}) \; G_i(\eta, \underset{\sim}{\theta})$ with $H > 0$, then to first order

$$\left[\left(\frac{\partial F_i}{\partial \theta_\mu} \right) \Big/ \left(\frac{\partial F_i}{\partial \eta} \right) \right]_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}} \approx \left[\left(\frac{\partial G_i}{\partial \theta_\mu} \right) \Big/ \left(\frac{\partial G_i}{\partial \eta} \right) \right]_{\substack{\eta=Y_i \\ \underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_N}} ,$$

since

$$\frac{\partial F_i}{\partial \theta_\mu} = H \frac{\partial G_i}{\partial \theta_\mu} + \frac{\partial H}{\partial \theta_\mu} G_i \; , \quad \frac{\partial F_i}{\partial \eta} = H \frac{\partial G_i}{\partial \eta} + \frac{\partial H}{\partial \eta} G_i$$

and $G_i(Y_i, \hat{\underset{\sim}{\theta}}_N) \approx 0$, if ε_i and $\theta_\mu - \hat{\theta}_{\mu N}$ are small.

4. Application to drug-protein binding data.

The absorption, distribution and elimination of a drug in the body can be greatly influenced by the extent of binding of the drug to macromolecules such as albumin (Martin 1965). The modeling of the binding of drugs to albumin is thus of considerable importance from the clinical point of view.

Assuming that the drug and protein interact via a number of absorption stops, and denoting the concentration of free drug as A_i and bound drug (to albumin) as R_i for a total drug concentration of C_i and a total protein concentration of P_i, the theoretical (error free) model for the relation between R_i (y) and C_i (x) corresponding to eq. (2) is (Edsall and Wyman 1958, Klotz and Hemston 1971, Perrin, Vallner and Wold 1974), assuming a non-cooperative process:

$$0 = F_i = R_i - P_i\{\sum_{k=1}^{M} n_j K_j (C_i - R_i)/[1 + K_j(C_i - R_i)]\} \quad (16)$$

$$C_i = R_i + A_i \quad (16a)$$

Here, the parameters M_j are interpreted as the number of drug molecules bound in the j:th class, M is the number of classes and K_j are the corresponding binding constants.

The estimation of the binding constants K_j, $j = 1,2,\ldots,M$, has commonly been made by minimizing

$$q = \sum_{i=1}^{M} [R_i - P_i\{\sum_{j=1}^{M} n_j K_j (C_i - R_i)/[1 + K_j(C_i - R_i)]\}]^2, \quad (17)$$

which corresponds to the adopting erroneous assumption that $C_i - R_i$ and R_i can be treated as x_i and y_i in non linear regression. This is obviously not the case since these quantities have the same magnitude of errors and, in addition, these errors are strongly correlated.

However, since C_i is accurately measured in comparison with R_i, eq. (16) corresponds to eq. (2) with $C_i = x_i$, $R_i = y_i$ and $K_j = \theta_j$. For the usual case that M is assumed to be 2, we further get by multiplication by the denominators

$$G_i = R_i[1 + K_1(C_i-R_i)][1 + K_2(C_i-R_i)] - P_i[n_1K_1(C_i-R_i)(1 + K_1[C_i-R_i]) + n_2K_2(C_i-R_i)(1 + K_2[C_i-R_i])] = 0 \tag{18}$$

and $R_{i,obs} = R_i + \varepsilon_i$

The second derivative of G with respect to R in eq. (18) is much smaller than that of F in eq. (16). (Here R plays the role of η.) The form G gives exact and approximate (ε^* or ε^{**}) residuals agreeing within 20%, the latter F only within 50%. (The "exact" residual is $Y_i - y_i(\hat{\theta}_N)$, where $y_i(\hat{\theta}_N)$ is the solution η of $F(\eta, x_i, \hat{\theta}_N) = 0$).

Consequently, the approximation of ε_i of eq. (7) was defined in terms of G in eq. (18). The parameters n_1, K_1, n_2 and K_2 were then estimated by minimizing either of

$$\tilde{Q}_N = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{w_i^2} [G/\frac{\partial G}{\partial R}]^2_{R=R_{i,obs}} \tag{19a}$$

$$\tilde{Q}_N = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{w_i^2} \left[G / \frac{\partial G}{\partial R} + \frac{1}{2} G^2 \frac{\partial^2 G}{\partial R^2} / \left(\frac{\partial G}{\partial R} \right)^3 \right]^2_{R=R_{i,obs}} \qquad (19b)$$

It is interesting to compare parameter values estimated by minimizing eq. (19) and (17) respectively; for dicoumarol-albumin data (N = 66) with the accuracy of measurement of R being about 1%, they differ by a factor of two or more (see table 1), showing the danger of using incorrect procedures.

Finally, a limited number of simulations was performed with 50 data points generated according to model (16) with M = 2 and $K_1 = 3 \times 10^6$, $K_2 = 3 \times 10^5$, $n_1 = n_2 = 1$ and subsequent addition of normally distributed random errors with variance $\sigma^2 \cdot R_i$. The results are shown in table 2. It is seen that for an error of measurement up to about 1%, eq. (19b) leads to fairly unbiased parameter estimates and standard errors. Equation (19a), however, breaks down for a relative error of measurement of between 0.1 and 1%.

Conclusions.

The nonlinear implicit regression problem can be handled in several ways. The approximate approach described in the present paper has the advantage of being very easy to program and use in practice.

The approximations involved are not serious as long as the relative accuracy of measurement is small; in the present example one percent or less. Then, the exact and approximate residuals

do not differ by more than approximately 10% and the methods work well in the examples tried.

Compared to earlier methods used for the estimation of binding constants for drug-protein interactions, the present method gives much less biased results, allows estimation of confidence intervals of the parameters, allows for weighting of the individual observations, and allows for the simple use of standard computer programs.

Table 1.

Estimated parameter values for dicoumarol-albumin data [Perrin and Vallner 1974]. Rows 1 and 2, minimization of eq. (19a); rows 3 and 4, eq. (19b); and row 5, eq. (17). Confidence intervals (95%) are based on eq. (15). Asterisks (*) denote fixed (non-estimated) values. S^2 denote resulting residual sum of squares, ψ the standard deviation of the relative difference between the exact and approximate residuals, $\psi = \{\sum\{(e_{approx} - e_{exact})/e_{exact}\}^2/N\}^{\frac{1}{2}}$.

	$K_1 \times 10^{-6}$	$K_2 \times 10^{-6}$	K_1	K_2	$S^2 \times 10^{12}$	ψ
1.	2.9 ± 15%	0.18 ± 8%	1.0*	1.0*	3.0	.08
2.	2.7 ± 25%	0.14 ± 19%	1.08 ± 10%	.97 ± 6%	2.8	.10
3.	2.7 ± 14%	0.18 ± 8%	1.0*	1.0*	2.9	.03
4.	3.9 ± 26%	0.22 ± 18%	0.82 ± 10%	1.19 ± 6%	3.0	.11
5.	14	0.33	0.53	1.45	—	—

Table II.

Results of simulations with data generated by eq. (16) with M = 2, $K_1 = 3\times10^6$, $K_2 = 0.3\times10^6$, $n_1 = n_2 = 1$, $C=1.0\times10^{-5}$ to 1.0×10^{-4} equally spaced (N=50). Random, normally distributed errors with mean 0 and standard deviation (SD) equal to R × σ were then added to the generated R-values. ψ is defined in the same way as in table I. Each row is the average of 25 runs with newly generated errors. The sample SD is the SD of the 25 estimates of K_i. The average estimated SD is based on eq. (15).

	Average K_1	Sample SD K_1	Average Estim. SD K_1	Average K_2	Sample SD K_2	Average Estim. SD K_2	ψ
Eq. (19a) σ = .001	3.03×10^6	.072	.063	$.300\times10^6$	.0019	.0018	. 012
Eq. (19b) σ = .001	3.01	.072	.063	.300	.0019	.0018	.0007
Eq. (19a) σ = . 01	3.92	1.0	. 60	.301	. 011	. 014	. 097
Eq. (19b) σ = . 01	2.79	. 57	. 41	.306	. 016	. 015	. 073

References.

[1]. Deming, W.E. (1943). Statistical adjustment of data. Wiley, New York.

[2] Edsall, J.T. and Wyman, J. (1958). Biophysical Chemistry, p. 591, Academic Press, New York.

[3] Klotz, I.M. and Hunston, D.L. (1971). Properties of graphical representations of multiple classes of binding sites. Biochemistry 10, 3065.

[4] Krenser, J. (1971). GPM/GPMNLC, extended gradient projection method nonlinear programming subroutines; reference manual, Madison Academic Computer Center, University of Wisconsin-Madison.

[5] Luecke, R.H. and Britt, H.I. (1973). The estimation of parameters in nonlinear, implicit models. Technometrics 15, 233.

[6] Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math. 11, 431.

[7] Martin, B.K. (1965). The potential effect of the plasma proteins on drug distribution. Nature 207, 274.

[8] Mezaki, R., Draper, N.R. and Johnson, R.A. (1973). On the violation of assumptions in nonlinear least squares by interchange of response and predictor variables. I&EC Fundamentals 12, 251.

[9] Perrin, J.H., Vallner, J.J. and Wold, S. (1974). An unbiased method for estimating binding parameters in a non-cooperative binding process. Biochim. Biophys. Acta (in press).