R. J. LeVeque

ARO Report 79-3
PROCEEDINGS OF THE 1979 ARMY NUMERICAL ANALYSIS AND COMPUTERS CONFERENCE

Approved for public release; distribution unlimited.
The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

SPONSORED BY
THE ARMY MATHEMATICS STEERING COMMITTEE ON BEHALF OF

THE OFFICE OF
THE CHIEF OF RESEARCH, DEVELOPMENT AND
ACQUISITION

Extensions and Uses of the Variable Projection Algorithm for Solving Nonlinear Least Squares Problems

Gene H. Golub*
Randall J. LeVeque*

Abstract. The variable projection algorithm for solving separable nonlinear least squares problems with a single data vector is well known[1]. We review that theory and present a modification of the algorithm for solving problems in which it is desired to fit more than one data vector with the same nonlinear parameters (though possibly different linear parameters) for each right hand side. We give an example from chemical kinetics and also show how such problems arise from an inverse differential equations problem as in compartmental analysis. A further modification is presented for dealing with total least squares problems: problems in which the independent variables as well as the observartions may have errors.

---

*Department of Computer Science, Stanford University, Stanford, CA 94305.

This work was supported in part by Army contract No. DAAG29-78-G-0179, Department of Energy contract No. EY-76-S-03-0326 PA#30 and by a National Science Foundation graduate fellowship. The paper was produced using TEX, a computer typesetting system created by Donald Knuth at Stanford.

1. Introduction.

The variable projection algorithm [1] has gained much popularity in recent years as a method for solving separable nonlinear least squares problems. A separable problem is one in which the model can be written in the form

$$\eta(t) \approx \sum_{j=1}^{n} \beta_j \phi_j(\alpha; t); \qquad \alpha \in \mathbb{R}^k, \tag{1.1}$$

where the ϕ_j are given functions of α and t and $\approx$ indicates approximation in the least squares sense. Given data $y = (\eta_1, \eta_2, \ldots, \eta_m)^T$ observed at times $t_1, t_2, \ldots, t_m$ respectively, the problem consists of finding the optimal parameters $\hat{b} = (\hat{\beta}_1, \ldots, \hat{\beta}_n)^T$, $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_k)^T$ which minimize the sum of the squares of the residuals, given by

$$\sum_{i=1}^{m} \left(\eta_i - \sum_{j=1}^{n} \beta_j \phi_j(\alpha; t_i) \right)^2.$$

If we let the matrix $\Phi(\alpha)$ consist of the components $\phi_j(\alpha; t_i)$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$, then the problem can be restated as

$$\text{minimize} \quad \|y - \Phi(\alpha) b\|^2 \qquad \text{over } b \in \mathbb{R}^n, \ \alpha \in \mathbb{R}^k.$$

Here we have used the 2-norm, $\|x\|^2 = \sum \xi_i^2$. The theory presented here can be easily extended to the case of weighted norms, $\|x\|_W^2 = \sum w_i \xi_i^2$.

The separable least squares problem which is probably most often encountered in practice is that of exponential fitting. In this case we want to fit

$$\eta_i \approx \sum_{j=1}^{n} \beta_j e^{\alpha_j t_i}, \qquad i = 1, 2, \ldots, m.$$

This leads to a matrix $\Phi(\alpha)$ of the form

$$\Phi(\alpha) = \begin{pmatrix} e^{\alpha_1 t_1} & \cdots & e^{\alpha_n t_1} \\ \vdots & & \vdots \\ e^{\alpha_1 t_m} & \cdots & e^{\alpha_n t_m} \end{pmatrix}.$$

In the next section we review the variable projection algorithm as modified by Kaufman[2]. We then present an extension of the algorithm to handle multiple right hand sides and discuss some uses of the algorithm in solving inverse differential equations and total least squares problems.

2. The variable projection algorithm.

For any fixed a, the problem of finding the optimal b corresponding to that a is a linear least squares problem. A solution is given by

$$b = \Phi^+(a)y$$

where $\Phi^+(a)$ is the pseudo-inverse of $\Phi(a)$. The idea behind the variable projection algorithm is to define the functional $r(a)$ by

$$\begin{aligned} r(a) &= \min_b \|y - \Phi(a)b\|^2 \\ &= \|y - \Phi(a)\Phi^+(a)y\|^2. \end{aligned} \tag{2.1}$$

This functional is then minimized with respect to a by means of the Levenberg-Marquardt algorithm (described below) to give the optimal $\hat{a}$. The linear parameters $\hat{b}$ are then recovered from $\hat{b} = \Phi^+(\hat{a})y$. It can be shown[1] that this method of solution yields the correct result provided that $\Phi(a)$ has constant rank in a neighborhood of the solution $\hat{a}$. The advantage of this technique is that the nonlinear functional $r(a)$ which must be minimized is now a function only of a. The parameters β_i do not appear explicitly and so the size of our nonlinear problem has been reduced.

This is not true in exponential fitting if some of the a's should be equal!

For any given a, there is an orthogonal matrix

$$Q(a) = \begin{pmatrix} Q_1(a) \\ Q_2(a) \end{pmatrix}$$

with

$$\begin{aligned} &Q^T(a)Q(a) = I, \\ &Q_1(a) \in \mathbb{R}^{n \times m}, \\ &Q_2(a) \in \mathbb{R}^{(m-n) \times m}, \end{aligned}$$

such that $Q(a)$ triangularizes $\Phi(a)$, that is,

$$\begin{pmatrix} Q_1(a) \\ Q_2(a) \end{pmatrix} \Phi(a) = \begin{pmatrix} U_1(a) \\ 0 \end{pmatrix}, \tag{2.2}$$

with $U_1(a)$ right triangular. From the invariance of the 2-norm under orthogonal transformations, we have that

$$\begin{aligned} \|y - \Phi(a)b\|^2 &= \|Q(a)(y - \Phi(a)b)\|^2 \\ &= \left\| \begin{pmatrix} Q_1(a)y \\ Q_2(a)y \end{pmatrix} - \begin{pmatrix} U_1(a) \\ 0 \end{pmatrix} b \right\|^2 \\ &= \|Q_1(a)y - U_1(a)b\|^2 + \|Q_2(a)y\|^2. \end{aligned}$$

Since the optimal b for any a is $b = \Phi^+(a)y = U_1^{-1}(a)Q_1(a)y$, the residual $r(a)$ is simply

$$r(a) = \|Q_2(a)y\|^2.$$

It is this form of r to which we apply the minimization algorithm.

The Levenberg-Marquardt algorithm is a general iterative procedure for minimizing $\|f(a)\|^2$ for a nonlinear m-vector valued function f. At the j^{th} stage we have an approximation $a^{(j)}$ to the solution and we compute the Jacobian $J(a^{(j)})$ of f,

$$J(a^{(j)}) = \left(\frac{\partial f(a^{(j)})}{\partial a_1}, \ldots, \frac{\partial f(a^{(j)})}{\partial a_k}\right),$$

where

$$\frac{\partial f}{\partial a_j} = \left(\frac{\partial f_1}{\partial a_j}, \ldots, \frac{\partial f_m}{\partial a_j}\right)^T.$$

We also choose a value for the "Marquardt parameter" ν_j and then solve the system

$$\begin{pmatrix} J(a^{(j)}) \\ \nu_j I \end{pmatrix} \delta^{(j)} \approx \begin{pmatrix} f(a^{(j)}) \\ 0 \end{pmatrix}$$

in the least squares sense for the correction $\delta^{(j)}$. The next iterate $a^{(j+1)}$ is then obtained as $a^{(j+1)} = a^{(j)} - \delta^{(j)}$. The parameter ν_j is used to control the length of the correction vector $\delta^{(j)}$. The correction $\delta^{(j)}$ minimizes

$$\|J(a^{(j)})\delta^{(j)} - f\|^2 + \|\nu_j \delta^{(j)}\|^2.$$

If $\nu_j = 0$, this becomes simply the Gauss-Newton algorithm. For further discussion of this minimization technique, see [5].

In the present context, $f(a) = Q_2(a)y$, so in order to apply this algorithm we must be able to compute the columns of the Jacobian matrix,

$$\frac{\partial (Q_2(a)y)}{\partial a_j} = \frac{\partial Q_2(a)}{\partial a_j} y.$$

Recall from (2.2) that $Q_2(a)\Phi(a) = 0$, so by differentiating we get

$$\frac{\partial Q_2(a)}{\partial a_j}\Phi(a) = -Q_2(a)\frac{\partial \Phi(a)}{\partial a_j}.$$

Golub and Pereyra[1] give an exact expression for $\partial Q_2(a)/\partial a_j$, but Kaufman[2] has suggested the simplification

$$\frac{\partial Q_2(a)}{\partial a_j} \approx -Q_2(a)\frac{\partial \Phi(a)}{\partial a_j}\Phi^+(a) \tag{2.3}$$

which works well in practice.

3. Extensions to multiple right hand sides.

Occasionally one wishes to solve a problem in which there are a number of vectors $y_1, y_2, \ldots, y_s$ of data each of which is to be fit by a model of the form (1.1). We allow the linear parameters b to be different for each data vector. If the nonlinear parameters a are also allowed to be different for each y_i, then there is no problem. We are simply faced with s distinct problems of the type already discussed. If however, a is constrained to be the same for each data vector, then we have a new problem. In this case, we wish to fit the matrix $Y = (y_1, y_2, \ldots, y_s) \in \mathbb{R}^{m\times s}$ by a matrix of the form $\Phi(a)B$ where $\Phi(a)$ is as before and $B \in \mathbb{R}^{n\times s}$. The minimization problem is now

$$\text{minimize} \quad \|Y - \Phi(a)B\|_F^2 \qquad \text{over } a \in \mathbb{R}^k,\ B \in \mathbb{R}^{n\times s}. \tag{3.1}$$

Here we use the Frobenius norm of the matrix, $\|X\|_F^2 = \sum_{i,j}\xi_{ij}^2$.

One example of this type of problem occurs in the analysis of data from the biological substance bacteriorhodopsin. The data vectors y_i consist of measurements of the amount of light absorbed by the substance at m different times during the course of a chemical reaction. We have s different data vectors, one for each of several wavelengths of light used. In this case the kinetic theory dictates that for each wavelength the absorbtion curve should be a sum of exponentials, and that the rate constants a should be the same for all wavelengths. Hence the data Y must be fit by a model of the form $\Phi(a)B$ where $\phi_j(a;t_i) = e^{a_j t_i}$.

One approach to solving this problem would be to write

$$\bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} \in \mathbb{R}^{ms}, \qquad \bar{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \in \mathbb{R}^{ns},$$

$$\bar{\Phi}(a) = \begin{pmatrix} \Phi(a) & & & \\ & \Phi(a) & & \\ & & \cdots & \\ & & & \Phi(a) \end{pmatrix} \in \mathbb{R}^{ms\times ns} \tag{3.2}$$

and then minimize $\|\bar{y} - \bar{\Phi}\bar{b}\|^2$ by means of the algorithm as previously discussed. However, the matrix $\bar{\Phi}(\alpha)$ is excessively large. The problem (3.1) has a special structure which is not sufficiently exploited by this approach.

Instead, for any α we proceed as before by finding $Q(\alpha)$ such that (2.2) holds. Since the Frobenius norm, like the 2-norm, is invariant under orthogonal transformations, we are led as before to the problem of simply minimizing $\|Q_2(\alpha)Y\|_F^2$. But $\|Q_2(\alpha)Y\|_F^2 = \|z(\alpha)\|^2$ where the vector $z(\alpha)$ is defined by

$$z(\alpha) = \begin{pmatrix} Q_2(\alpha)y_1 \\ \vdots \\ Q_2(\alpha)y_s \end{pmatrix}. \tag{3.3}$$

The Levenberg-Marquardt algorithm previously described can be applied directly to this problem, provided we can compute $\partial z(\alpha)/\partial \alpha_j$ for $j = 1, 2, \ldots, k$. But from (3.3) it is clear that

$$\frac{\partial z(\alpha)}{\partial \alpha_j} = \begin{pmatrix} \frac{\partial Q_2(\alpha)}{\partial \alpha_j} y_1 \\ \vdots \\ \frac{\partial Q_2(\alpha)}{\partial \alpha_j} y_s \end{pmatrix},$$

where $\partial Q_2(\alpha)/\partial \alpha_j$ is computed exactly as before.

So, in summary, at the j^{th} step of the minimization procedure we solve for the correction term $\delta^{(j)}$ from the linear system

$$\left(\begin{array}{ccc} \frac{\partial Q_2(a^{(j)})}{\partial a_1} y_1 & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial a_k} y_1 \\ \vdots & & \vdots \\ \frac{\partial Q_2(a^{(j)})}{\partial a_1} y_s & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial a_k} y_s \\ \hline \multicolumn{3}{c}{\nu_j I} \end{array}\right) \delta^{(j)} = \left(\begin{array}{c} Q_2(a^{(j)})y_1 \\ \vdots \\ Q_2(a^{(j)})y_s \\ \hline 0 \end{array}\right).$$

Note that this system is the same size as that which would result from using the alternative formulation (3.2). The advantage of the latter approach lies in the calculation of $Q_2(\alpha)$ and its derivatives. Here $Q_2(\alpha) \in \mathbb{R}^{(m-n)\times m}$ whereas (3.2) would lead to an $(m-n)s$ by ms matrix.

This algorithm has been applied to the bacteriorhodopsin problem described earlier with satisfactory results. For that problem, we had $m \approx 175$, $s = 5$. The correct number of exponential terms was not known *a priori*, so fits were computed with 3, 4, 5 and 6 terms. Figure 1 shows the residuals for one such calculation. For more information on this particular application, see [4].

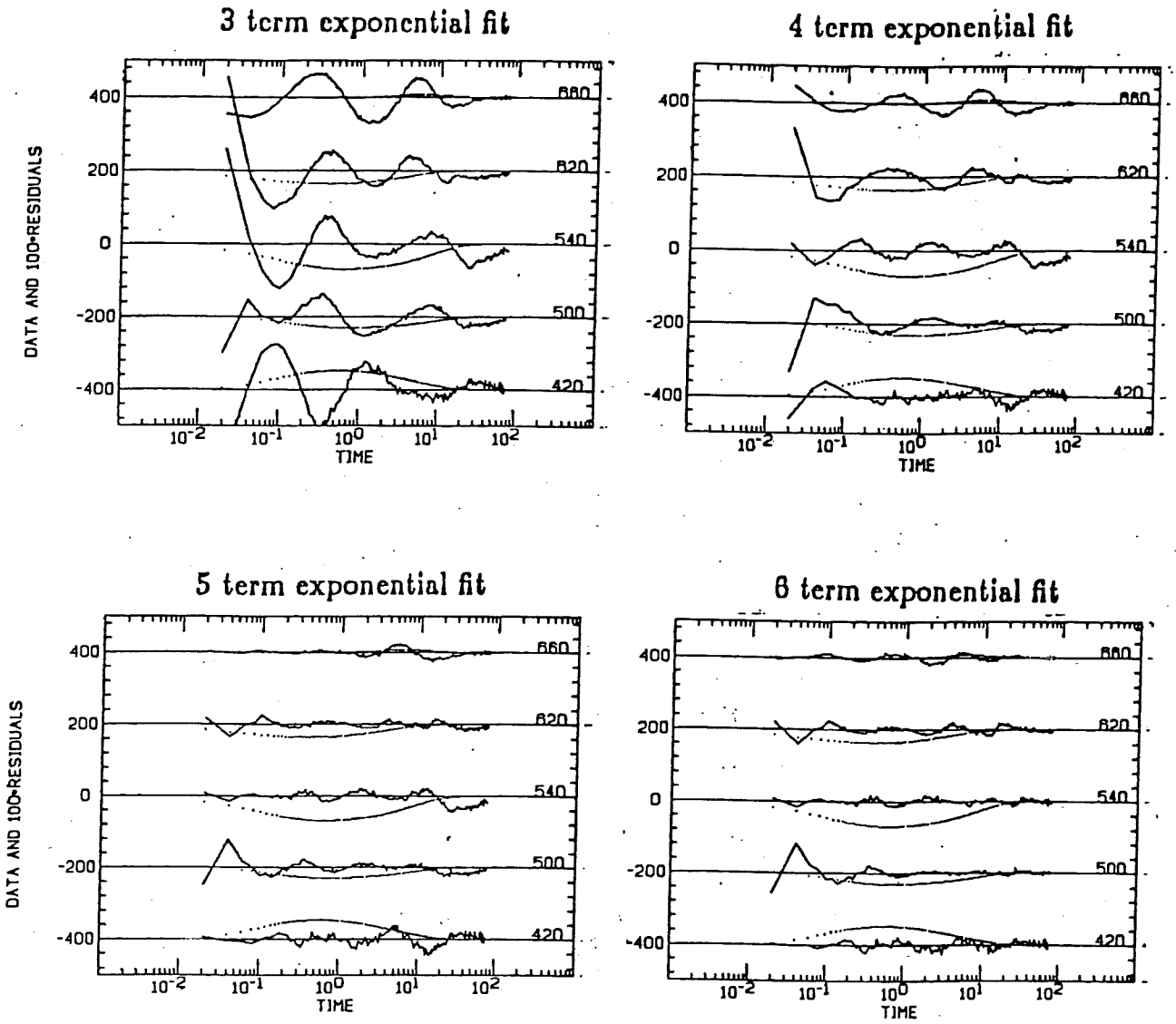

Figure 1
Bacteriorhodopsin data and magnified residuals

The data at each wavelength $\lambda = 420$ through $\lambda = 660$ is shown as points. The solid line is the residual from the fit obtained, multiplied by a factor of 100 to make it visible.

4. Inverse differential equations.

A special case of the exponential fitting problem of the previous section occurs when $s = k$, the number of right hand sides equals the number of exponential terms. In this case we may view the problem as a problem in inverse differential equations. We have vector-observations $y(t_j) \in \mathbb{R}^k$, $j = 1, \ldots, m$ which are assumed to be the values of the solution to some linear differential equation $y' = Ay$ at times t_j. The problem is then to approximate the eigenvalues of the unknown matrix A. The solution is given by the optimal nonlinear parameters obtained by solving $Y = \Phi(\alpha)B$ with

$$\Phi(\alpha) = \begin{pmatrix} e^{\alpha_1 t_1} & \cdots & e^{\alpha_k t_1} \\ \vdots & & \vdots \\ e^{\alpha_1 t_m} & \cdots & e^{\alpha_k t_m} \end{pmatrix}$$

and $B \in \mathbb{R}^{k \times k}$. This follows from the fact that the solution to $y' = Ay$ is of the form

$$y(t) = Ce^{At}y(0)$$

where $C \in \mathbb{R}^{k \times k}$. So, assuming A is diagonalizable,

$$\begin{aligned} y(t) &= CX\Lambda(t)X^{-1}y(0) \\ &\equiv \tilde{X}\Lambda(t)z, \end{aligned}$$

where

$$\Lambda(t) = \operatorname{diag}(e^{\alpha_1 t}, \ldots, e^{\alpha_k t}).$$

So in fact

$$\eta_i(t) = \sum_{j=1}^{k} \tilde{\xi}_{ij} e^{\alpha_j t} z_j = \sum_{j=1}^{k} e^{\alpha_j t} \beta_{ji}$$

where

$$\beta_{ji} = \tilde{\xi}_{ij} z_j.$$

5. Total least squares.

Up to this point we have assumed that the values of the independent variable t were known exactly. In that case our task was to minimize $\|Y - \Phi(\alpha; t)B\|_F^2$. Note that we have written $\Phi(\alpha; t)$ to show the dependence of Φ on t. For now we will consider the problem in which t as well as Y may have errors. In this case we

wish to solve a problem of the form

$$\begin{aligned} &\text{minimize} \quad \left(\|Y - \Phi(a;\tau)B\|_F^2 + \lambda^2\|\tau - t\|^2\right) \\ &\text{over } a \in \mathbb{R}^k,\ B \in \mathbb{R}^{k\times s},\ \tau \in \mathbb{R}^m. \end{aligned} \tag{5.1}$$

We can consider $a = \begin{pmatrix} a \\ \tau \end{pmatrix}$ as the vector of unknown nonlinear parameters. In order for the estimate of a to be consistent, the parameter λ^2 must be chosen in proportion to the ratio of variances of the errors. For more generality, the norms in (6.1) can be weighted norms without introducing any difficulties.

Again there are two possible approaches to solving this problem. The first approach uses the variable projection algorithm as presented in section 2 on the problem

$$\text{minimize} \quad \|\bar{\Phi}(a;\tau)\bar{B} - \bar{Y}\|_F^2 \qquad \text{over } a, \bar{B}$$

where

$$\bar{\Phi}(a;\tau) = \begin{pmatrix} \Phi(a;\tau) & 0 \\ 0 & \frac{\lambda}{s}\tau \end{pmatrix} \begin{matrix} \}m \\ \}m \end{matrix}$$

$$\bar{B} = \left(\begin{array}{c|c} B & \\ \hline 1 \ \cdots \ 1 & \end{array}\right), \quad \bar{Y} = \left(\begin{array}{c|c} Y & \\ \hline \frac{\lambda}{s}t \ \cdots \ \frac{\lambda}{s}t & \end{array}\right).$$

This is straightforward but requires that we be able to fix some of the linear parameters at 1. The variable projection algorithm can be easily modified to impose this constraint. Such a modification was first proposed by Krogh[3].

The second method which can be used for solving the total least squares problem involves a further modification of the functional to be minimized. Let $Q(a) = \begin{pmatrix} Q_1(a) \\ Q_2(a) \end{pmatrix}$ be the orthogonal matrix which triangularizes $\Phi(a;\tau)$. Then we wish to find a to minimize

$$r(a) = \|Q_2(a)Y\|_F^2 + \lambda\|\tau - t\|^2.$$

This residual functional can be rewritten in partitioned from as

$$r(a) = \|[Q_2(a)Y \mid \lambda(\tau - t)]\|_F^2.$$

The Levenberg-Marquardt algorithm can be used for this minimization. Recall from section 2 that we require the derivatives of the columns of $[Q_2(a)Y \mid \lambda(\tau - t)]$

with respect to each nonlinear parameter. The derivatives of $Q_2(a)y_j$ with respect to any α_i or τ_i can be computed exactly as before. For the final column we have

$$\frac{\partial(\tau - t)}{\partial \alpha_i} = 0,$$

$$\frac{\partial(\tau - t)}{\partial \tau_i} = e_i, \quad \text{the } i^{\text{th}} \text{ unit vector.}$$

So we see that at each stage of the minimization process we must solve for the corrections

$$\delta^{(j)} = \begin{pmatrix} \delta_\alpha^{(j)} \\ \delta_\tau^{(j)} \end{pmatrix}$$

from the system

$$\left(\begin{array}{ccc|ccc} \frac{\partial Q_2(a^{(j)})}{\partial a_1^{(j)}} y_1 & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial a_n^{(j)}} y_1 & \frac{\partial Q_2(a^{(j)})}{\partial \tau_1^{(j)}} y_1 & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial \tau_m^{(j)}} y_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial Q_2(a^{(j)})}{\partial a_1^{(j)}} y_s & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial a_n^{(j)}} y_s & \frac{\partial Q_2(a^{(j)})}{\partial \tau_1^{(j)}} y_s & \cdots & \frac{\partial Q_2(a^{(j)})}{\partial \tau_m^{(j)}} y_s \\ \hline 0 & & & I & & \\ \hline \multicolumn{6}{c}{\nu_j I} \end{array}\right) \begin{pmatrix} \delta_\alpha^{(j)} \\ \delta_\tau^{(j)} \end{pmatrix} = \left(\begin{array}{c} Q_2(a^{(j)})y_1 \\ \vdots \\ Q_2(a^{(j)})y_s \\ \hline \tau^{(j)} - t \\ \hline 0 \end{array}\right).$$

Both of the approaches described here were used on an example from Powell and Macdonald[7]. The set of data was originally given by Pearson[6] and the weights by York[9]. This data is given in Table 1. In this example both the times t_i and the observations y_i are weighted, by v_i and w_i respectively. This data was fitted by a linear polynomial $y \approx \beta_1 + \beta_2 x$, as was done in [7]. The results, $\beta_1 = 5.4799$, $\beta_2 = -0.48053$, agree with those reported there. The t_i were used as initial approximations to the nonlinear parameters and convergence to the accuracy shown occured in 5 iterations. The resulting fit is shown in figure 2. The given data points (t_i, y_i) are shown by X's, the points (τ_i, y_i) by boxes. Note the effect of the weights.

Table 1

Pearson's data and York's weights				
i	t_i	v_i	y_i	w_i
1	0.0	1000.0	5.9	1.0
2	0.9	1000.0	5.4	1.8
3	1.8	500.0	4.4	4.0
4	2.8	800.0	4.6	8.0
5	3.3	200.0	3.5	20.0
6	4.4	80.0	3.7	20.0
7	5.2	60.0	2.8	70.0
8	6.1	20.0	2.8	70.0
9	6.5	1.8	2.4	100.0
10	7.4	1.0	1.5	500.0

Total Least Squares fit to Pearson's Data with York's Weights

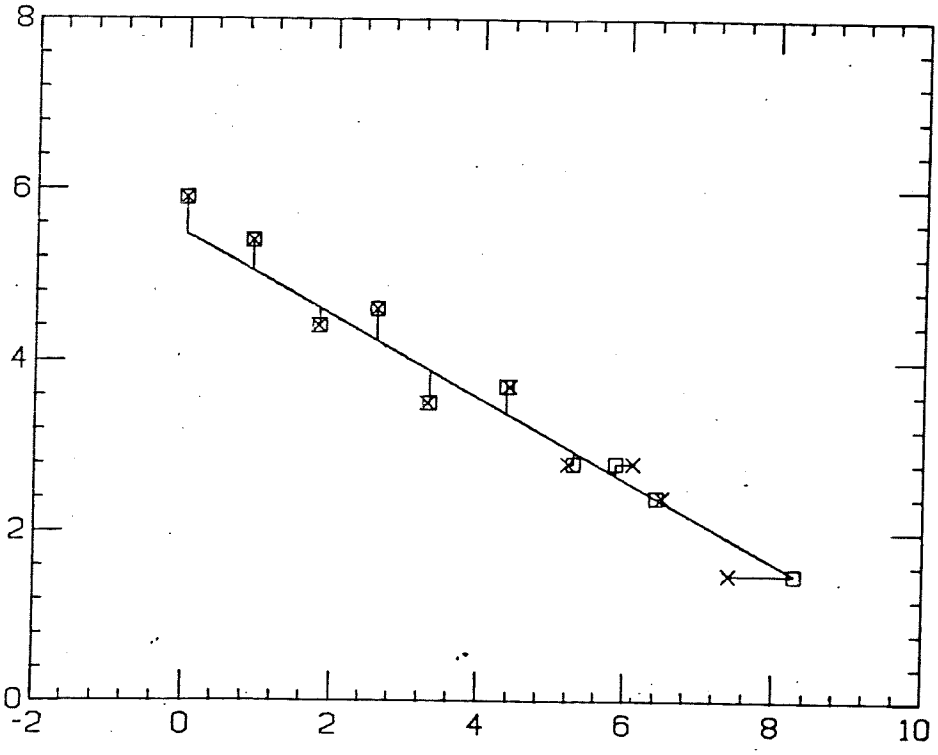

Figure 2

6. Acknowledgements.

The authors are indebted to Richard Lozier of the University of California at San Francisco for his support and expertise in the bacteriorhodopsin problem. Computer time for this research was provided by the Stanford Linear Accelerator Center of the U.S. Department of Energy.

7. References.

[1] G.H. Golub and V. Pereyra, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate, SINUM 10(1973), pp.413-432.

[2] L. Kaufman, A variable projection method for solving separable nonlinear least squares problems, BIT 15(1975), pp. 49-57.

[3] F.T. Krogh, Efficient implementation of a variable projection algorithm for nonlinear least squares problems, CACM 17(1974), pp. 167-169.

[4] R. Lozier, R. LeVeque, G. Golub, Simultaneous analysis of kinetic data obtained at several monitoring wavelengths: applications to the bacteriorhodopsin photocycle, to appear.

[5] M.R. Osborne, Some aspects of nonlinear least squares calculations, *Numerical Methods for Nonlinear Optimization*, Lootsma, et al., Academic Press, London, 1972

[6] K. Pearson, On lines and planes of closest fit to systems of points in space, *Phil. Mag.* 2(1901), pp. 559-572.

[7] D.R. Powell and J.R. Macdonald, A rapidly convergent iterative method for the solution of the generalised nonlinear least squares problem, *The Computer Journal* 15(1972) pp. 148-155.

[8] A. Ruhe and P.A. Wedin, Algorithms for separable nonlinear least squares problems, STAN-CS-74-434, Stanford Computer Science Department, 1974.

[9] D. York, Least-square fitting of a straight line, *Can. J. of Phys.* 44(1966), pp. 1079-1086.