

Modified Damped Least Squares: An Algorithm for Non-linear Estimation

ROBERT R. MEYER AND PHILIP M. ROTH*

*Shell Development Company,
A Division of Shell Oil Company,
Emeryville, California*

[Received 11 June 1971]

A new method for non-linear regression is described and compared to three other well-known methods for the non-linear least squares problem. Convergence theorems and computational results are given which indicate that the new method is more efficient and reliable than previously published methods.

1. Introduction

WE PRESENT in this paper a modified and formalized version of a non-linear estimation technique first explicated by Levenberg (1944). In a variety of applications during the past two years this algorithm has proved superior in both efficiency and reliability to all algorithms with which it has been compared.

One mode of classifying non-linear estimation techniques is by the nature of their use:

- (1) Limited: Application to a specific model or set of models; e.g., such functionals as the exponential, logarithm, and sum of exponentials.
- (2) General: Application to any model.

As consultants within a research laboratory, our need was for a reliable non-linear estimation algorithm of *general* applicability, one that displayed a high frequency of convergence for an unstructured collection of models and applications. The "standard available" general algorithm, which we call Damped Least Squares (DLS), is due to Marquardt (1963) and is the one upon which we have relied in the past. This technique is similar to Levenberg's, although Marquardt was apparently unaware of the former's work. Unfortunately, DLS displayed satisfactory convergence in our applications far less frequently than we would have liked. It was this situation that motivated our work in developing "still another non-linear estimation technique".

In Section 2 of this paper we provide in abbreviated form background material needed for the presentation. In Section 3 we present variants of the Gauss-Newton method, including one that we have introduced. These techniques are then compared with regard to their convergence properties. The associated mathematical analysis also sheds light on the reason for the superiority of the estimation method we propose.

In Section 4, we elaborate on certain features that supplement the basic technique of damped least squares and enhance its convergence characteristics, and present a

* Now with Systems Applications, Inc., Beverly Hills, California, 90212.

description of the logic for the entire algorithm. Computational results are compared with those of other algorithms in Section 5.

2. Background

A great many techniques have been proposed for the estimation of parameters in nonlinear models. With few exceptions, all methods can be assigned to one of three categories.

- (1) The Gauss-Newton method, or a variant of this method.
- (2) Gradient (or search) techniques, a search over the sum of squares surface with the parameters as a variable. A convenient subdivision is:
 - (a) Gradient techniques (first derivatives evaluated)
 - (i) Steepest descent
 - (ii) Second-order methods
 - (b) Search techniques (derivatives not evaluated)
- (3) A combination of iterant and gradient (or search) techniques.

In this section we outline the first two techniques, discuss their weaknesses, and indicate how a synthesis of the two procedures capitalizes on the strengths of the method while alleviating their shortcomings.

2.1. Linear Estimation

In linear estimation, we wish to find estimates, \mathbf{b} , of the parameters β , in a model of the form

$$y_k = \eta_k + \varepsilon_k \quad (1)$$

where

$$\eta_k = \sum_{i=1}^p \beta_i x_{ki} = f(\beta, \mathbf{x}_k) \quad (2)$$

where

β_j = expected value of b_j = $E(b_j)$

$\beta = (\beta_1, \dots, \beta_p)$

\mathbf{b} = vector of parameters = $(b_1 \dots b_p)$

y_k = k th observation

x_{kj} = the setting of the j th independent variable for the k th experiment (or any function of the independent variables, such as sin, cos, log, etc.)

$\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kp})$

ε_k = error associated with the k th observation, y_k . It is assumed to have an expected value of zero, or $E(\varepsilon_k) = 0$.

Least squares estimates of the parameters are defined as those estimates that minimize

$$S(\mathbf{b}) = \sum_{k=1}^N (y_k - f(\mathbf{b}, \mathbf{x}_k))^2 \quad (3)$$

where N = number of experiments. The vector of parameters that satisfies this criterion is

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (4)$$

$$\text{where } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} x_{11}, x_{12}, \dots, x_{1p} \\ x_{21}, x_{22}, \dots, x_{2p} \\ \vdots \\ \vdots \\ x_{N1}, x_{N2}, \dots, x_{Np} \end{bmatrix}.$$

The numerical solution for \mathbf{b} is direct. See Draper & Smith (1966) for further details.

2.2 Non-linear Estimation

In non-linear estimation we wish to estimate the parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ for models of the form

$$\eta_k = f(\boldsymbol{\theta}, \mathbf{x}_k) \quad (5)$$

where at least one θ_j is *not* linearly related to η_k . Some examples of this type of model are

$$\eta_k = \theta_1(1 - e^{-\theta_2 x_k})$$

and

$$\eta_k = \frac{\theta_1 x_{1k}}{1 + \theta_2 x_{2k}}.$$

The parameters are to be chosen to minimize

$$s(\boldsymbol{\theta}) = \sum_{k=1}^n (y_k - f_k(\boldsymbol{\theta}))^2, \quad (6)$$

where

$$f_k(\boldsymbol{\theta}) = f(\boldsymbol{\theta}, \mathbf{x}_k).$$

2.3. Gauss-Newton Method

Direct solution for the parameters in non-linear models is not possible. Many years ago, however, Gauss suggested an iterative solution for this estimation problem by analogy to the linear formulation. First, given a set of parameter values $\boldsymbol{\theta}^i$, expand $f_k(\boldsymbol{\theta})$ in a Taylor series, retaining only the linear terms;

$$f_k(\boldsymbol{\theta}) \approx f_k(\boldsymbol{\theta}^i) + f_{k1}(\theta_1 - \theta_1^i) + \dots + f_{kp}(\theta_p - \theta_p^i), \quad (7)$$

where

$$f_{kj} = \frac{\partial f(\boldsymbol{\theta}^i, \mathbf{x}_k)}{\partial \theta_j} \quad \text{for } j = 1, \dots, p.$$

Then define a linearized error function

$$\hat{\mathbf{e}}(\Delta) = \begin{pmatrix} \hat{e}_1(\Delta) \\ \vdots \\ \hat{e}_N(\Delta) \end{pmatrix} = \mathbf{F}\Delta - \mathbf{r}, \quad (8)$$

where

$$\mathbf{F} = \begin{bmatrix} f_{11} & \dots & f_{1p} \\ \vdots & & \vdots \\ f_{N1} & & f_{Np} \end{bmatrix},$$

$$\Delta = \begin{pmatrix} \theta_1 - \theta_1^i \\ \vdots \\ \theta_p - \theta_p^i \end{pmatrix},$$

$$\mathbf{r} = \begin{pmatrix} y_1 - f_1(\theta^i) \\ \vdots \\ y_N - f_N(\theta^i) \end{pmatrix}.$$

and

(The signs of the RHS of (8) are the opposite of what might be expected, but they lead to results that are more easily compared to the linear case.) The problem of minimizing

$$\sum_{j=1}^n \hat{e}_j(\Delta)^2$$

is a linear least squares problem whose solution is given by

$$\Delta^* = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{r}, \quad (9)$$

in analogy with equation (4).

The *Gauss-Newton method* can be succinctly stated in terms of the notation just developed.

- (A) Guess parameters θ^0 ,
- (B) Iterate in the following manner ($i = 0, 1, 2, \dots$),
 - (i) Calculate \mathbf{F} and \mathbf{r} at θ^i ,
 - (ii) Determine Δ^* ,
 - (iii) Set $\theta^{i+1} = \theta^i + \Delta^*$, and return to (i) with i increased by 1.[†]

This method converges rapidly if the initial parameters θ^0 are "good"; otherwise it may fail to converge.

2.4. Gradient and Search Methods

The second class of techniques, gradient and search methods, all rely on determination of the characteristics of the sum of squares surface (the contour of s in the space

[†] Of course, in the implementation of this algorithm (and the ones to follow) on a computer, only a finite number of iterations are performed. Termination criteria will be discussed in Section 4.

of the parameters) in the vicinity of current or "best" guesses for the parameters. The following is a brief outline of the gradient procedure:

- (1) Guess parameters, θ^0 ,
- (2) Determine the local direction of steepest descent on the sum of squares surface,
- (3) For some step size, determine new guesses for the parameters along the path of steepest descent or some path derived from it,
- (4) Return to step 2.

Box & Wilson (1951) have described a useful modification of this basic approach, and Steinmetz (1966) has presented a lucid description of the method and its diversity of application. A variety of quasi-Newton methods are discussed by Broyden (1967). These techniques require only first derivatives, but near the solution they display (at least in theory) the convergence characteristics of methods based on the use of second derivatives.

Search procedures are loosely classed as techniques that, while relying on examination of the sum of squares surface, do not require determination of first derivatives. Two examples are direct or pattern search, developed by Hooke & Jeeves (1961) and simplex search, described by Spendley, Hext & Hinsworth (1962).

The major problem encountered in using gradient or search techniques is not one of convergence; these methods normally converge. Rather it is the time required to achieve convergence. The basic gradient procedure is notoriously slow, although some modifications achieve convergence fairly rapidly. The point to be made here is that iterative techniques are generally rapid, but unstable with respect to convergence, whereas gradient or search techniques exhibit the opposite characteristics, good convergence and longer computation time.

Experience in working with these methods normally leads one to seek a compromise-gradient technique early in the calculation to assure convergence and iterative procedures toward the end of the calculation, when first derivatives are small and convergence is slow. Levenberg proposed such a compromise in 1944, before the advent of the computer. Some years later Marquardt (1963), unaware of Levenberg's work, developed a very similar technique at du Pont and has since made it available through SHARE. This program is commonly used in general applications, although, as we have mentioned, its convergence properties are less than satisfactory.

With this background to call upon, we undertook the development of a more reliable non-linear estimation technique. Our work led to modification of the original Levenberg algorithm that we now describe.

3. Variants of the Gauss-Newton Method

3.1. Damped Least Squares

In contrast to the Gaussian iterant and most other formulations, an iterative step of damped least squares is based on the minimization of

$$\tilde{s}(\Delta) = \sum_{k=1}^N e_k^2(\Delta) + \lambda \sum_{j=1}^P v_j \Delta_j^2 \quad (10)$$

where

λ = a weighting (or damping) factor

v_j = weighting factor associated with corrections to the parameter estimate, Δ_j .

The term,

$$\lambda \sum_{j=1}^P v_j \Delta_j^2$$

by its presence, prevents unwanted excursions in the iterative procedure (i.e., prevention of divergence early in the calculation). Divergence tends to occur for highly non-linear models and for poor initial guesses. The factor λ , acts as a control by determining the degree of contribution of the damping term. If λ is large, only small changes in the parameters can take place, and oscillations are thus damped. If λ is small, damping is reduced and freer oscillation is permitted, a desirable situation near the minimum of the sum of squares surface.

Casting equation (10) into matrix notation, we have

$$\tilde{s}(\Delta) = \hat{e}(\Delta)^T \hat{e}(\Delta) + \lambda \Delta^T V \Delta \quad (11)$$

where

$$V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}.$$

The minimization of \tilde{s} is also a linear least squares problem whose normal equations are

$$(F^T F + \lambda V) \Delta = F^T r. \quad (12)$$

The solution of this system of equations is given by

$$\Delta^*(\lambda, V) = (F^T F + \lambda V)^{-1} F^T r. \quad (13)$$

Comparison of (13) with (9) shows that the only difference is the addition of the term λV to the matrix which is inverted. Levenberg proved that for all sufficiently large λ

$$s(\theta^i + \Delta^*(\lambda, V)) < s(\theta^i). \quad (14)$$

An iterative procedure based upon this observation is the following:

- (A) Guess parameters θ^0 ,
- (B) Given θ^i ($i = 0, 1, 2, \dots$) determine θ^{i+1} by performing the iterative steps:
 - (i) Solve (12) for increasing values of λ until a value λ^* is found for which (14) is satisfied,
 - (ii) Set $\theta^{i+1} = \theta^i + \Delta^*(\lambda^*, V)$,
 - (iii) Increase the index i by 1, and return to (i).

3.2. Hartley's Variant

The most obvious drawback to DLS is that several systems of the form (12) may have to be solved at each iteration. In 1961 Hartley (1961) developed the properties of a modified Gauss-Newton (MGN) method that did not have this drawback. The iterative procedure of the MGN method is as follows:

- (A) Guess parameters θ^0 ,
- (i) Calculate F and r at θ^0 ,
- (ii) Determine Δ^* ,

- (iii) Determine a step-length factor γ_i^* such that $s(\theta^i + \gamma_i^* \Delta^*)$ is an "approximate minimum" of $s(\theta^i + \gamma \Delta^*)$ for $\gamma > 0$,
- (iv) Set $\theta^{i+1} = \theta^i + \gamma_i^* \Delta^*$ and return to (i) with i increased by 1.

It is easily shown that if the gradient of the sum of squares function does not vanish at θ^i , then for all sufficiently small step-length factors γ we have

$$s(\theta^i + \gamma \Delta^*) < s(\theta^i). \quad (15)$$

Thus, with the MGN method we can also (in theory) achieve a reduction of the sum of squares function at each iteration, but with the solution of only one least squares problem at each iteration.

3.3. Modified Damped Least Squares

The Modified Damped Least Squares (MDLS) method to be described below extends the DLS method in a manner analogous to Hartley's extension of the Gauss-Newton method. In MDLS the matrix of weights V for the i th iteration is taken to be the diagonal matrix D whose elements are those of the diagonal of the matrix $F^T F$. These diagonal entries are all positive unless one or more columns of F are identically zero, hence the underlying assumption in MDLS that no such column vanishes. (This is one of two suggestions for V advanced by Levenberg. The other called for setting $V = I$, the identity matrix. Note that adding the matrix λD to $F^T F$ is equivalent to multiplying the diagonal elements of $F^T F$ by the factor $(1 + \lambda)$.) We now state the basic algorithm of MDLS:

- (A) Guess parameters θ^0 ,
- (B) Iterate in the following manner ($i = 0, 1, 2, \dots$),
 - (i) Choose a damping factor λ_i ,
 - (ii) Determine $\Delta^*(\lambda_i, D)$,
 - (iii) Determine a step length factor $\bar{\gamma}_i$ such that $s(\theta^i + \bar{\gamma}_i \Delta^*(\lambda_i, D))$ is an "approximate minimum" of $s(\theta + \gamma \Delta^*(\lambda_i, D))$ for $\gamma > 0$,
 - (iv) Let $\theta^{i+1} = \theta^i + \bar{\gamma}_i \Delta^*(\lambda_i, D)$ and return to (i) with i increased by 1.

Again it is easily shown that if the gradient of $s(\theta)$ is non-vanishing at θ^i , then for sufficiently small values of γ we have

$$s(\theta^i + \gamma \Delta^*(\lambda_i, D)) < s(\theta^i).$$

It should also be noted that if $F^T F$ is singular (i.e., does not have an inverse), then Δ^* is not uniquely determined, whereas $\Delta^*(\lambda_i, D)$ is unique if $\lambda_i D$ is positive definite. Moreover, even when $F^T F$ is non-singular, the use of a positive damping factor improves the conditioning of the system of equations that must be solved. Proofs of these assertions are given by Meyer (1970).

3.4. Convergence Theory

In order to state convergence theorems, it is convenient to collect expressions for the first and second derivatives of the non-linear sum of squares $s(\theta)$. First define the error function

$$\mathbf{e}(\theta) = \begin{pmatrix} e_1(\theta) \\ \vdots \\ e_N(\theta) \end{pmatrix} = \begin{pmatrix} f_1(\theta) - y_1 \\ \vdots \\ f_N(\theta) - y_N \end{pmatrix} \quad (16)$$

so that

$$s(\theta) = \mathbf{e}^T(\theta)\mathbf{e}(\theta). \quad (17)$$

It is easily verified from the rules for differentiation of vectors and products that the gradient of $s(\theta)$ (a row vector with p components) is given by

$$\mathbf{s}'(\theta) = 2\mathbf{e}^T(\theta)\mathbf{e}'(\theta), \quad (18)$$

where

$$\mathbf{e}'(\theta) = \mathbf{F}(\theta) = \begin{pmatrix} \partial f_1(\theta)/\partial \theta_1 & \dots & \partial f_1(\theta)/\partial \theta_p \\ \vdots & \ddots & \vdots \\ \partial f_N(\theta)/\partial \theta_1 & \dots & \partial f_N(\theta)/\partial \theta_p \end{pmatrix}, \quad (19)$$

so that (18) may also be written as

$$\mathbf{s}'(\theta) = 2\mathbf{e}^T(\theta)\mathbf{F}(\theta). \quad (20)$$

By differentiating again, we can obtain the second derivative or Hessian matrix (of dimensions $p \times p$) of $s(\theta)$:

$$\mathbf{s}''(\theta) = 2\mathbf{F}(\theta)^T\mathbf{F}(\theta) + 2\mathbf{e}^T(\theta)\mathbf{F}'(\theta), \quad (21)$$

where

$$\mathbf{e}^T(\theta)\mathbf{F}'(\theta) \equiv \sum_{j=1}^n e_j(\theta)f_j''(\theta), \quad (22)$$

with

$$f_j''(\theta) \equiv (\partial^2 f_j(\theta)/\partial \theta_k \partial \theta_l), \quad (k, l = 1, \dots, p). \quad (23)$$

It is also necessary to rigorously define what is meant by an "approximate minimum" of the one-dimensional search performed at each iteration. One approach to the definition is to require a certain fixed percentage β , where $0 < \beta < 1$, of the decrease predicted from linearization. That is $\bar{\gamma}_i$ is taken to be the first integer power of $\frac{1}{2}$ for which the following inequality is satisfied:

$$s(\theta^i + \bar{\gamma}_i \Delta^*(\lambda_i, \mathbf{D})) \leq s(\theta^i) + \beta \cdot \bar{\gamma}_i s'(\theta^i) \Delta^*(\lambda_i, \mathbf{D}). \quad (24)$$

It is easily shown that such a $\bar{\gamma}_i$ does indeed exist (see Meyer, 1970). (Note that when β is chosen very close to zero, the inequality (24) is almost equivalent to

$$s(\theta^i + \bar{\gamma}_i \Delta^*(\lambda_i, \mathbf{D})) < s(\theta^i). \quad (25)$$

In fact, from a computational standpoint, since only a finite number of iterations can be performed, the satisfaction of (25) for the iterates actually obtained implies the existence of a range of values of β for which the iterates obtained would also satisfy (24). However, the inequality (25) by itself is not strong enough to guarantee the desired theoretical convergence properties.)

We will now state three convergence theorems for MDLS. The proofs of these three theorems are to be found in Meyer (1970).

THEOREM 1. Let $\{\theta^i\}$ be a sequence of iterates generated by MDLS using the rule given above for the selection of the step-length factor at the i th iteration. If there exists a subsequence $\{\theta^{n_i}\}$ converging to some set of parameters θ^* and if the corresponding subsequence $\{\mathbf{F}(\theta^{n_i})^T\mathbf{F}(\theta^{n_i}) + \lambda_{n_i} \mathbf{D}_{n_i}\}$ converges to some positive definite matrix Δ (\mathbf{D}_{n_i} denotes the diagonal matrix whose diagonal coincides with that of $\mathbf{F}(\theta^{n_i})^T\mathbf{F}(\theta^{n_i})$), then $\mathbf{s}'(\theta^*) = 0$.

There are a number of ways of guaranteeing that there will be a convergent subsequence of iterates. One of the simplest is to assume that for any $\bar{\theta} \in E^p$ the set $L(\bar{\theta}) \equiv \{\theta | s(\theta) < s(\bar{\theta})\}$ is compact. This hypothesis implies that for an arbitrary starting point θ^0 the sequence generated by MDLS will lie in a compact set. The next theorem states that the addition of further hypotheses leads to convergence of the *whole* sequence to a stationary point.

THEOREM 2. *If (a) $L(\bar{\theta})$ is compact for each $\bar{\theta} \in E^p$, (b) $s(\theta)$ has at most a finite number of stationary points having any given function value, (c) $F(\theta)^T F(\theta)$ is positive definite for all θ , and (d) there is an upper bound on the damping factor to be used at any iteration; then, given an arbitrary starting point θ^0 the iterates $\{\theta^i\}$ generated by MDLS will converge to a stationary point of $s(\theta)$.*

Note that Theorem 2 is a *global* convergence result that can be applied to MGN (by setting the upper bound on the damping factor to 0) as well as MDLS. A similar convergence result could also be obtained for DLS if the rule for the selection of the quasi-optimal damping factor at each iteration was analogous to the rule given for $\bar{\gamma}_i$. Recall that the Gauss-Newton method does not have this type of global convergence property.

It is clear from the expression (2.1) for the second derivative $s''(\theta)$ that the sum of squared errors function $s(\theta)$ will in general be non-convex. Hence there is no guarantee that a stationary point located by MDLS (or any other algorithm considered in this paper) will be the global solution of the problem of minimizing $s(\theta)$. The following theorem deals with the rate of convergence to stationary points that will at least be local minima. Since the theorem below assumes that $\lambda_i \rightarrow 0$ (as is usually the case in MDLS), the same convergence rate estimate applies to *all* of the methods even though it is stated for MDLS.

THEOREM 3. *Let $\{\theta^i\}$ be a sequence generated by MDLS which converges to a point θ^* . Let m be the minimum eigenvalue of $F(\theta^*)^T F(\theta^*)$ and M be the maximum of the absolute values of the eigenvalues of $e^T(\theta^*) F'(\theta^*)$. If $r \equiv M/m < 1$, $\beta < (1-r)/2$, and $\lambda_i \rightarrow 0$, then $\bar{\gamma}_i = 1$ for all sufficiently large i and $\limsup \|\theta^{i+1} - \theta^*\|/\|\theta^i - \theta^*\| \leq r$. Moreover, $s(\theta^*)$ will be an isolated local minimum of the function $s(\theta)$.*

Both Pereyra (1967) and Ben-Israel (1967) have also derived convergence results for the Gauss-Newton method. These, however, are sufficient conditions for convergence rather than global convergence results such as Theorem 2. In addition, the estimated rates of convergence given by Pereyra and Ben-Israel involve upper bounds on first and second derivatives in a sphere about the initial guess, θ^0 . The convergence rate estimate of Theorem 3 is more closely related to a result of Daniel (1968) for Newton-like methods.

From a qualitative standpoint, Theorem 3 indicates that the efficiency of the algorithms of Section 2 is related to the magnitude of the error, $e(\theta^*)$, at the stationary point to which the iterates converge. In particular if $e(\theta^*) = 0$, then $M = 0$ and therefore $r = 0$, so that the rate of convergence is superlinear. This is not surprising, since $e(\theta^*) = 0$ implies (in the context of Theorem 3) that $F(\theta^i)^T F(\theta^i) \rightarrow s''(\theta^*)$, so that the search directions $\Delta^*(\lambda_i, D)$ are a good approximation to the corrections $s''(\theta^{i-1})^{-1} s'(\theta^i)^T$ of Newton's method for determining a zero of $s'(\theta)$. In this instance, the limiting behaviour is similar in principle to the well-known Davidson-Fletcher-Powell method

and other quasi-Newton methods (Broyden, 1967) that generate matrices that converge to the Hessian matrix (or its inverse) in the limit. In the quasi-Newton methods the approximation to the Hessian is constructed using information from previous iterations, whereas algorithms derived from the Gauss-Newton method take advantage of the special form of the problem to construct an approximation to the Hessian based only on first derivatives at the most recent set of parameter values. Since the reliance of the quasi-Newton methods upon information from previous iterations has been observed to lead to numerical difficulties in some cases (McCormick & Pearson, 1968) the Gauss-Newton approach appears to be preferable in non-linear regression problems in which "a good fit" of the data is expected. Results of numerical experiments by Pitha & Jones (1966) and Bard (1970) agree with this conclusion.

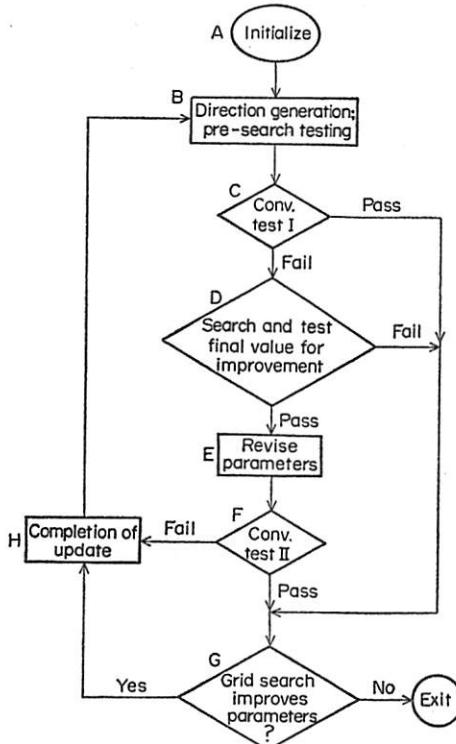


FIG. 1. Program structure.

4. A Computer Program

The flow charts below present the principal elements of the logic of the program written to implement MDLS. The main purpose of this section is to describe certain options that have been added to the basic algorithm. These options have been found to significantly increase efficiency and reliability in a variety of difficult problems.

Program Section B, whose logic is shown in Fig. 2, first attempts to solve the damped normal equations by the UNIVAC 1108 MATHPAK subroutine GJR, which uses the Gauss-Jordan algorithm. (It has been pointed out that greater accuracy may be attained in the solution of linear least squares problems by other algorithms, such as

those based upon Householder transformations (Businger & Golub, 1965; Golub, 1965) or Gram-Schmidt orthonormalization (Walsh, 1962). It would be interesting to explore the value of such algorithms in the iterative solution of non-linear least squares problems.) In the event that GJR fails to generate a solution to the damped normal equations for the current iteration (usually because of pivot tolerances), an alternate procedure is used to obtain a search direction. This procedure generates an appropriately scaled unit vector in the direction of the coordinate corresponding to the component of the gradient with greatest absolute value. This coordinate direction is also used if a direction is generated by GJR and this direction fails to meet a number of pre-search tests, including tests on the sizes of the components of the direction and on the sign of the inner product with the gradient. The basic idea of the pre-search tests is to reject directions that appear "unreasonable" or inaccurately determined.

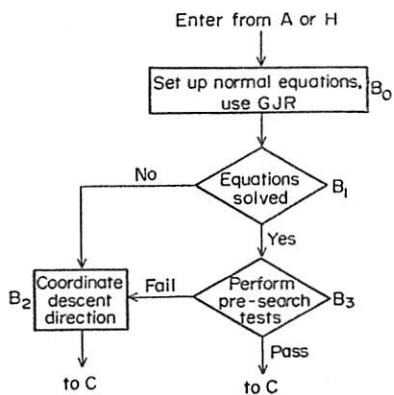


FIG. 2. Program section B.

In order to improve efficiency, the method of determining the step-length factor γ_i is not the bisection procedure cited above, but rather a modified iterative quadratic approximation scheme (program Section D). A number of similar schemes are described by Bard (1970). If, because of the numerical limitations of computer arithmetic, a step-length factor yielding a reduced sum of squares is not found, then a branch is made to program Section G. There, evaluations of the sum of squares function are performed with each parameter successively increased and decreased by 10% and 1% until a set of parameter values decreasing the sum of squares is found or until all $4p$ such evaluations have been performed. In the latter case, the program terminates, and the values of the sum of squares function at the perturbations provide a sensitivity analysis of the "optimal" parameter values.

The method of generating damping factors is in some respects similar to that proposed by Marquardt. The initial damping factor λ_0 is set to some nominal value such as 1/100 or 1/1000. At the i th iteration, λ_i is set to $4\lambda_{i-1}$ if θ^i was derived from θ^{i-1} by changing only one component; otherwise λ_i is set to $\lambda_{i-1}/4$.

A set of parameter values passes the convergence test built into the program when the relative change in *each* parameter from the previous iteration is less than a factor specified by the user. The value customarily used is 10^{-5} .

5. Evaluation of Computational Results

The results obtained in the eight examples present in detail below are summarized in Table 1. In all eight examples MDLS yielded the optimal solution, whereas DLS was successful in only the first three, for which it was slower than MDLS. For DLS the number of times a set of normal equations is solved is equal to the number of function evaluations, whereas in MDLS it is the same as the number of iterations. (An iteration is to be understood to mean the search performed in order to compute a set of corrections that yields a smaller sum of squared errors.) This consideration was not significant in the eight examples below because of the small number of parameters involved, but could result in a significant saving in computation time for problems with large (> 10) numbers of parameters.

The DLS program used to obtain the results below is one coded at Shell Development in 1967. Computational results with this version of DLS were better than those obtained using the SHARE version of DLS then available.

TABLE 1
Comparison of DLS and MDLS

Example	Iterations/Function evaluation	
	DLS	MDLS
1	19/29	4/4
2	38/61	17/32
3	29/46	16/29
4	Failure	10/25
5	Failure	14/46
6	Failure	24/40
7	Failure	22/35
8	Failure	7/12

Example 1

Model.

$$\frac{\theta_1 \theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2}.$$

Data.

x_1	x_2	y
1.0	1.0	0.126
2.0	1.0	0.219
1.0	2.0	0.076
2.0	2.0	0.126
0.1	0.0	0.186

Initial values. $\theta^0 = (10.39, 48.83, 0.74)$, $s(\theta^0) = 0.0365$.

Optimal values. $\theta^* = (3.13, 15.16, 0.78)$, $s(\theta^*) = 0.4 \times 10^{-4}$.

Comments. This is a problem originally presented by Box & Hunter (1965).

Examples 2 and 3

Model. $f_1(\theta) = 10(\theta_2 - \theta_1)$, $f_2(\theta) = (1 - \theta_1)$,
 $s(\theta) = f_1^2(\theta) + f_2^2(\theta)$.

Initial values for problem 2. $\theta^\circ = (-1.2, 1.0)$, $s(\theta^\circ) = 24.2$.

Initial values for problem 3. $\theta^\circ = (-0.86, 1.14)$, $s(\theta^\circ) = 19.5$.

Optimal values. $\theta^* = (1.0, 1.0)$, $s(\theta^*) = 0$.

Comments. This is Rosenbrock's parabolic valley problem (Rosenbrook, 1963) which appears frequently in the mathematical programming literature. The contours of the sum of squares surface form a steep-sided valley following the curve $\theta_2 = \theta_1^2$. Although both DLS and MDLS reach the solution by moving along the parabolic valley, MDLS "tracks" along the valley in a much more efficient manner than DLS. The iterates generated in the course of solving problem 2 are shown in the following table:

Iteration	$s(\theta)$	θ_1	θ_2
1	0.198558+02	-0.101019+01	0.622796-00
2	0.182580+02	-0.865121-00	0.363995-00
3	0.169675+02	-0.735965-00	0.168095-00
4	0.157172+02	-0.618375-00	0.204758-01
5	0.144868+02	-0.509703-00	-0.895959-01
6	0.132855+02	-0.408079-00	-0.169669-00
7	0.121234+02	-0.312397-00	-0.224915-00
8	0.108976+02	-0.200220-00	-0.267436-00
9	0.960473+01	-0.555588-01	-0.288298-00
10	0.824338+01	0.987233-01	-0.263001-00
11	0.679077+01	0.261945-00	-0.181306-00
12	0.521214+01	0.440688-00	-0.271380-01
13	0.344921+01	0.639693-00	0.227016-00
14	0.155894+01	0.819847-00	0.548597-00
15	0.218406-00	0.961242-00	0.877413-00
16	0.225656-03	0.100000+01	0.998498-00
17	0.224265-13	0.100000+01	0.100000+01

It should be noted that these problems illustrate the possibility of using MDLS in the solution of systems of non-linear equations. Here we are determining a root of the system of two equations: $f_1(\theta) = 0$, $f_2(\theta) = 0$. If the damping factor is fixed at 0 and a step-length factor of 1 is used at each iteration, the method is equivalent to Newton's method for solving systems of non-linear equations. However, since the step-length factor is allowed to vary, and an alternative method for generating search directions is provided, MDLS can be expected to be more reliable than Newton's method.

Examples 4 and 5

Model. $\theta_3(e^{-\theta_1 x_1} + e^{-\theta_2 x_2})$.

x_1	x_2	Example 4	Example 5
		y	y
0	0	40.2	40.0
0.6	0.4	11.0349	10.0
0.6	1.0	4.48869	5.0
1.4	1.4	2.46137	2.5
2.6	1.4	2.45137	2.5
3.2	1.6	1.82343	2.0
0.8	2.0	1.00094	1.0
1.6	2.2	0.741352	0.7
2.6	2.2	0.741352	0.8
4.0	2.2	0.741352	0.7
1.2	2.6	0.406863	0.4
2.0	2.6	0.406862	0.4
4.6	2.8	0.301411	0.3
3.2	3.0	0.223291	0.22
1.6	3.2	0.165418	0.2
4.2	3.4	0.122545	0.1
2.0	3.8	0.067254	0.05
3.2	3.8	0.067254	0.07
2.8	4.2	0.036910	0.03
4.2	4.2	0.036910	0.03
5.4	4.4	0.027343	0.03
5.6	4.8	0.015006	0.02
3.2	5.0	0.011117	0.01

$$\theta^\circ = (12.0, 1.0, 25.0)$$

Initial sum of squares, Example 4. $s(\theta^\circ) = 187.0$.

Initial sum of squares, Example 5. $s(\theta^\circ) = 226.9$.

Optimal values, Example 4. $\theta^* = (14.3, 1.5, 20.1)$, $s(\theta^*) = 10^{-12}$.

Optimal values, Example 5. $\theta^* = (31.5, 1.51, 19.9)$, $s(\theta^*) = 1.25$.

Comments. This double-decay type model tends to lead to difficulties in parameter estimation. The data for problem 4 were generated by evaluating the model using the parameter values (14.3, 1.5, 20.1), insuring the existence of a set of parameter values that would give 0 error. To obtain the data for problem 5, error was introduced by reducing each data value to 1 or 2 significant figures. In both problems, DLS quit after failing to solve damped normal equations corresponding to overly large values of $\theta_1 (> 100)$. In RDLS, the change in θ_1 at each iteration was limited in magnitude to its initial value (12.0), and no such difficulties arose.

Box (1958) has observed that high correlations among parameters are a common occurrence for models in which the number of parameters p exceed the number of independent variables v by two or more. This correlation arises from the structure of the model and not from a poor choice of experiments. The models selected for Examples 6, 7, 8, all have the property that $p - v = 2$. Estimation of parameters for these models constitutes a stringent test of the capability of a general purpose algorithm to deal with built-in difficulties.

*Examples 6 and 7**Model.* $\theta_1 + \theta_2 e^{\theta_3 x}$.*Data.*

<i>x</i>	Example 6 <i>y</i>	Example 7 <i>y</i>
1	16.7242	16.7
5	16.8262	16.8
10	16.9657	16.9
15	17.1198	17.1
20	17.2902	17.2
25	17.4785	17.4
30	17.6865	17.6
35	17.9165	17.9
40	18.1706	18.1
50	18.7619	18.7

Initial values. $\theta^0 = (20, 2, 0.5)$, $s(\theta^0) \approx 2 \times 10^{22}$.*Optimal values for problem 6.* $\theta^* = (15.5, 1.2, 0.02)$, $s(\theta^*) = 10^{-12}$.*Optimal values for problem 7.* $\theta^* = (15.67, 0.999, 0.022)$, $s(\theta^*) = 0.006$.

Comments. The set of data for Example 6 was generated using the parameter values $(15.5, 1.2, 0.02)$, and the data for Example 7 were obtained by retaining only the first three decimals of each function value. Depending upon the initial damping factor, DLS failed either because θ_3 became large and negative (-10^5) leading to failure in evaluating the exponential or because θ_2 became small in absolute value (10^{-5}) leading to the failure of GJR. Convergence to good parameter values was obtained with MDLS.

*Example 8**Model.* $\theta_1 + e^{\theta_2/(x+\theta_3)}$.*Data.*

<i>x</i>	<i>y</i>
50	34,780
55	28,610
60	23,650
65	19,630
70	16,370
75	13,720
80	11,540
85	9,744
90	8,261
95	7,030
100	6,005
105	5,147
110	4,427
115	3,820
120	3,307
125	2,872

Initial values. $\theta^0 = (0.02, 4,000, 250)$, $s(\theta^0) = 1.7 \times 10^9$.

Optimal values. $\theta^* = (0.0056, 6,181.4, 345.2)$, $s(\theta^*) = 88$.

Comments. The data for this example give the resistance of a thermistor as a function of temperature. Attempts at solution with DLS had failed, but rapid convergence to the solution was obtained with MDLS with $\lambda^0 = 0$.

The results summarized in Table 1 are typical of the relative performance of DLS and MDLS. In all problems that have been tested, MDLS has either been more efficient than DLS in obtaining optimal values for the parameters or had succeeded where DLS has failed to obtain the optimal values. In addition to the test problems described above, MDLS has been used successfully in the solution of a number of practical problems, involving reaction kinetics of propylene, reaction rate equations for the peroxidation of ethyl-benzene, viscosity of concentrated resin solutions, and the analysis of spectral data (see Meyer, 1970).

We are indebted to Susan E. Post for her efforts in writing the computer code, in developing a number of examples, and in patiently running and rerunning a large number of test problems.

REFERENCES

- BARD, Y. 1970 *SIAM J. Num. Analysis* 7(1), 3031–3050.
 BEN-ISRAEL, A. 1967 *Israeli J. Math.*, 27–36.
 BOX, G. E. P. 1958 *Bull. Inst. int. Statist.* 36, 215–225.
 BOX, G. E. P. & HUNTER, W. G. 1965 *Technometrics* 7, 23.
 BOX, G. E. P. & WILSON, K. B. 1951 *Jl R. statist. Soc.* XIII(B), 1–45.
 BROYDEN, C. G. 1967 *Maths Comput.* 21, 368–381.
 BUSINGER, P. & GOLUB, G. 1965 *Num. Math.* 7, 269–276.
 DANIEL, J. W. 1968 On Newton-like methods. *The University of Wisconsin Computer Sciences Dept. Tech. Rep.* No. 21.
 DRAPER, N. R. & SMITH, H. 1966 *Applied regression analysis*. New York: John Wiley.
 GAUSS, C. F. "Theory of least squares". (English translation by H. F. Trotter). *Princeton Univ. Stat. Tech. Res. Group, Tech. Rep.* No. 5.
 GOLUB, G. 1965 *Num. Math.* 7, 206–216.
 HARTLEY, H. O. 1961 *Technometrics* 3.
 HOOKE, R. & JEEVES, T. A. 1961 *J. Ass. comput. Mach.* 8, 212–229.
 LEVENBERG, K. 1944 *Q. appl. Math.* 2, 164–168.
 McCORMICK, G. P. & PEARSON, J. D. 1968 *Variable metric methods and unconstrained optimization*. Paper presented at the Joint Conference on Optimization, University of Keele, March 25–28.
 MARQUARDT, D. W. 1963 *J. Soc. ind. appl. Math.* 11, 431–441.
 MEYER, R. R. 1970 Theoretical and computational aspects of non-linear regression. *Tech. Pap. P-1819, Shell Development Co., Emeryville, California*.
 PEREYRA, V. 1967 *SIAM J. Num. Analysis* 4, 27–36.
 PITHA, J. & JONES, R. N. 1966 *Can. J. Chem.* 44, 3031–3050.
 ROSENBROCK, H. H. 1963 *Comput. J.* 3, 163.
 SPENDLEY, W., HEXT, G. R. & HIMSWORTH, F. R. 1962 *Technometrics* 4, 441–461.
 STEINMETZ, H. L. 1966 *I & EC*, 58(1), 33–39.
 WALSH, P. J. 1962 *Communs Ass. comput. Mach.* 5, 511–513.
 WILDE, D. J. 1964 *Optimum seeking methods*, pp. 145–157. New York: Prentice Hall.