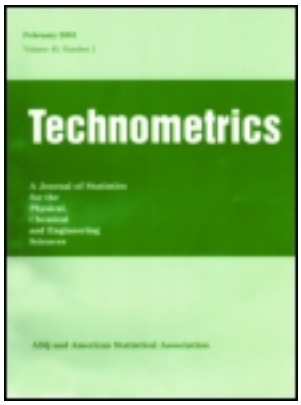

Technometrics
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/utch20

The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares

H. O. Hartley [a]
[a] Statistical Laboratory, Iowa State University, Ames, Iowa
Published online: 30 Apr 2012.

The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares*

H. O. HARTLEY

Statistical Laboratory
Iowa State University
Ames, Iowa

1. INTRODUCTION

The experimental scientist is frequently faced with the task of determining the functional relation between a 'response', y and a number of 'inputs', $x_1 , x_2 , \cdots , x_k$, with the help of empirical data. Often the mathematical form of the functional relation is assumed to be known and is written in the form of a 'regression function',

$$Y = f(x_1 , x_2 , \cdots , x_k ; \theta_1 , \theta_2 , \cdots , \theta_m), \tag{1}$$

where the $(k + m)$ variable function f is given mathematically but the unknown parameters, $\theta_1 , \cdots , \theta_m$, have to be estimated from a set of observed responses, y, and associated inputs $x_1 , \cdots , x_k$. As is well known, the method of estimation most frequently employed for the estimation of the θ_i in (1) is the method of Least Squares:—With this method the differences, $y - f$, between the observed responses, y, and the responses computed from the associated inputs $x_1 , \cdots , x_k$, through the regression function (1) are formed using a trial set of parameters $\theta_1 , \cdots , \theta_m$. The sum of squares of these differences, $Q = \sum (y - f)^2$, is then an m-variable function of the trial parameters $\theta_1 , \cdots , \theta_m$ and minimized as a function of these parameters.

When the regression function f is linear in the parameters the m equations resulting from setting $\partial Q/\partial \theta_i = 0$ for $i = 1, 2, \cdots , m$ are linear in the θ_i and are known as the 'normal equations' in multilinear regression estimation, a special case of the well developed theory of 'linear estimation'.

When the regression function is non-linear in the parameters, both the theory and the practice of the estimation procedure is considerably more difficult.

Most of this paper is concerned with the methodology of non-linear regression problems, that is, with the numerical technique of computing Least Squares estimates as solutions of the system of non-linear equations. After posing the problem (Section 2) we develop a modification of the well known 'Gauss-Newton' method of iterative solution. This new method, whilst sharing the advantageous features of the Gauss-Newton method, has the additional merits of a guaranteed convergence (Section 3).

* This work was sponsored by the U. S. Army Ordnance Corps under Contract No. DA-11-022-ORD-2732 Project No. 401-04-17-95-2633.

Whilst the solution of the normal equations in linear regression theory is usually unique and yields the *absolute* minimum of Q both these questions require special examination with non-linear problems and some of these questions are discussed in Section 4 where a theorem on uniqueness is also proved. The details of the numerical procedure are then spelt out in more detail and illustrated with a numerical example of fitting the exponential regression with asymptote involving three parameters (Section 5). Certain short cut methods are also discussed in this section. In a later section, (Section 6), certain questions concerning the optimality of the Least Squares estimation are invoked from the properties of maximum likelihood estimation. Comparisons with the literature are made in Section 7. The reader who is interested in the numerical procedure only and not with the mathematical proofs is advised to turn to the worked example of Section 5 and check the formulas by referring to the description of the general procedure given in Sections 2 and 3.

2. The Formulation of the Non-linear Regression Problem

Consider the following non-linear least squares problem:—Given n sets of 'observed' $(k+1)$−tuples $y_h\ ; x_{1h}\ , x_{2h}\ , \cdots\ , x_{kh} (h = 1, 2, \cdots, n)$,

$$f(x;\ \theta) \equiv f(x_1\ , \cdots\ , x_k\ ;\ \theta_1\ , \cdots\ , \theta_m); \tag{2}$$

it is required to determine a set of θ_i for which the sum of squares

$$Q(\theta) = \sum_{h=1}^{n} (y_h - f(x_h\ ;\ \theta))^2 = \text{Min.} \tag{3}$$

Here the symbol x_h stands for the k-vector with elements $x_{1h}\ , \cdots\ , x_{kh}$ and the symbol θ for the m-vector with elements $\theta_1\ , \cdots\ , \theta_m$. The function $f(x,\ \theta)$ is assumed to satisfy the following conditions:—

a. Introducing the first and second derivatives of $f(x,\ \theta)$ with regard to the θ_i as

$$\frac{\partial f}{\partial \theta_i} \equiv f_i(x,\ \theta); \qquad \frac{\partial^2 f}{\partial \theta_i\ \partial \theta_j} = f_{ij}(x,\ \theta), \tag{4}$$

the f_i and f_{ij} are assumed continuous functions of the θ for all k − tuples $x_h (h = 1, 2, \cdots, n)$.

The above assumption permits the following definitions to be made:—

$$\begin{aligned} \frac{\partial Q}{\partial \theta_i} &= Q_i(x;\ \theta) = -2 \sum_h (y_h - f(x_h\ ;\ \theta)) f_i(x_h\ ;\ \theta) \\ \frac{\partial^2 Q}{\partial \theta_i\ \partial \theta_j} &= Q_{ij} = -2 \sum_h (y_i - f(x_h\ ,\ \theta)) f_{ij}(x_h\ ,\ \theta) + 2 \sum_h f_i(x_h\ ,\ \theta) f_j(x_h\ ,\ \theta) \end{aligned} \tag{5}$$

b. The following assumption is equivalent to the well known assumption of non-degeneracy of rank in a linear least square problem:—We assume that for any (non-trivial) set of $u_i (i = 1, 2, \cdots, m)$ with $\sum u_i^2 > 0$,

$$\sum_{h=1}^{n} \left(\sum_{i=1}^{m} u_i f_i(x_h\ ,\ \theta) \right)^2 > 0 \tag{6}$$

for the 'observed' vectors x_h and for all θ in a bounded convex set S of the

parameter space $\theta_1, \cdots, \theta_m$. This assumption will usually be satisfied in practical situations.

c. The next assumption is a convenient one to ensure the convergence of the iterative process of solution to be described in Section 3 below:—Denote by

$$\mathbf{Q} = \lim_{\bar{S}} \inf Q(x, \theta), \tag{7}$$

where $\bar{S}$ is the complement to S. Then it is assumed that it is possible to find a vector ${}_0\theta$ in the interior of S such that

$$Q(x; {}_0\theta) < \mathbf{Q}. \tag{8}$$

The formal 'Least Squares equations' corresponding to the Least Squares problem (3) are given by

$$Q_i(x, \theta) = 0, \qquad i = 1, 2, \cdots, m \tag{9}$$

using the notation (4) and (5). In the next section we shall describe an iterative process which will converge to a solution of the Least Squares equations (9) under the assumptions a to c. A sharpening of assumption b is then necessary to ensure that (i) the solution reached gives the absolute minimum of the Least Squares form (3), and (ii) that the solution is unique.

3. The Modified Gauss-Newton Iteration

We give first a formal description of the process. Later, in Section 5, we discuss computational short cuts which will be seen to considerably simplify the numerical execution of the procedure:—

We start with the vector ${}_0\theta$ stipulated in assumption c. The first step is to compute 'corrections' to the elements ${}_0\theta_i$ of the starting vector ${}_0\theta$. These corrections will be proportional to the solutions D_i of the Gauss-Newton equations corresponding to (3). The latter equations are obtained in the familiar manner by substituting a multiple 1st order Taylor expansion of $f(x, \theta)$ at $\theta = {}_0\theta$ in (3) and writing the resulting linear least squares equations viz:—

$$2 \sum_{j=1}^{m} \{ \sum_{h} f_i(x_h ; {}_0\theta) f_j(x_h ; {}_0\theta) \} D_j = -Q_i(x; {}_0\theta), \tag{10}$$

where $Q_i(x; {}_0\theta)$ is given by (5). Because of assumption b the determinant of linear equations (10) has rank m and thus can always be solved, yielding the elements D_i of the vector D as solutions. Consider now the function

$$Q(v) = Q(x, {}_0\theta + v D), \quad \text{for} \quad 0 \leq v \leq 1, \tag{11}$$

and denote by v' the value of v for which $Q(v)$ is a minimum on the interval $0 \leq v \leq 1$. Defining the vector

$${}_1\theta = {}_0\theta + v' D \tag{12}$$

with elements ${}_1\theta_i = {}_0\theta_i + v' D_i$ we have obviously

$$Q(x, {}_1\theta) \leq Q(x, {}_0\theta) < \mathbf{Q}, \tag{13}$$

so that ${}_1\theta$ clearly lies in the interior of S. The above computation is now re-

peated at ${}_1\theta$ and so on. There results a sequence of vectors ${}_t\theta$, $t = 1, 2, \cdots$, all within the bounded convex set S, with

$$\lim_{t\to\infty} Q(x, {}_t\theta) = Q^* \quad \text{(say)}. \tag{14}$$

Consider a point of accumulation θ^* of this bounded sequence and subsequence ${}_\tau\theta$ with

$$\lim_{\tau\to\infty} {}_\tau\theta \to \theta^* \tag{15}$$

Now since clearly

$$\lim_{\tau\to\infty} Q(x, {}_\tau\theta) = Q(x, \theta^*) \leq Q(x, {}_0\theta) < \mathbf{Q}, \tag{16}$$

it follows from assumption c that θ^* must be an interior point of S. We now show that at this limit point θ^* the first partials $\mathbf{Q}_i$ must all be zero:—Introduce the D_i^* as the solution of the equations (corresponding to the equations (10))

$$2\sum_{j=1}^{m}\{\sum_h f_i(x_h, \theta^*)f_j(x_h, \theta^*)\} D_j^* = -Q_i(x, \theta^*). \tag{17}$$

Because of the continuity assumptions and (6)

$$\lim_{\tau\to\infty} {}_\tau D_i = D_i^*. \tag{18}$$

Further from (17) and (6)

$$\sum_{i=1}^{m} Q_i(x, \theta^*) D_i^* = -2\sum_h\{\sum_i f_i(x_h - \theta^*) D_i^*\}^2 < 0, \tag{19}$$

provided

$$\sum_{i=1}^{m} D_i^{*2} > 0. \tag{20}$$

But equation (19) implies that the differential of Q in the direction proportional to the D_i^* is negative. Therefore, it follows from (18) that for *all* ${}_\tau\theta$ in a small neighborhood of θ^* the differential of Q in the direction proportional to the ${}_\tau D_i$ would be smaller than a fixed quantity $-e$ (say). Since the second differential of Q in these directions# is bounded over a unit distance by a bound of say B, it follows that the minimum of Q in the direction of the ${}_\tau D_i$ must be below $Q(x, {}_t\theta)$ by at least the amount $ev - \frac{1}{2}Bv^2$ where v is the fractional distance moved in the direction proportional to the ${}_\tau D_i$ from the points ${}_\tau\theta$. Choosing $v^* = \min(1, |e|/B)$ we find that the minima of Q in the directions proportional to the ${}_\tau D_i$ would all be below $Q(x, {}_\tau\theta)$ by at least the amount $\frac{1}{2}ev^*$. This contradicts (14) which states that the $Q(x, {}_t\theta)$ of the original sequence t converge to Q^* which would also be the limit of the subsequence $Q(x, {}_\tau\theta)$. Thus, we reach a contradiction unless $\sum_{i=1}^{m} D_i^{*2} = 0$ which implies (because of the full rank of equations (17)) that

$$\sum_{i=1}^{m} Q_i^2(x, \theta^*) = 0. \tag{21}$$

This differential may be defined as the differential with regard to the variable v as defined by equation (11) with ${}_\tau\theta$ replacing ${}_0\theta$.

We have shown therefore that a subsequence ${}_r\theta$ of the sequence of vectors ${}_i\theta$ converges to a solution θ^* of the Least Squares equations

$$Q_i(x, \theta^*) = 0, \qquad i = 1, \cdots, m. \tag{22}$$

It is clear that for practically all problems of non-linear regression the original sequence of ${}_i\theta$ will converge to θ^*:—For if there were an infinite subsequence of ${}_{r'}\theta$ not converging to θ^* then a subsequence of these ${}_{r''}\theta$ would tend to a limit $\theta^{**} \neq \theta^*$. From the above argument we would now conclude that

$$Q(x, \theta^*) = Q(x, \theta^{**}) \tag{23}$$

and that θ^{**} must also be a stationary point, i.e.

$$Q_i(x, \theta^{**}) = 0 \qquad i = 1, 2, \cdots, m. \tag{24}$$

It is highly improbable that there be a regression surface and a set of observed x_h and y_h such that Q has two stationary points yielding *precisely* the same value of Q.

4. The Absolute Minimum of Q and the Uniqueness of the Least Squares Solution

It is clear that the above procedure will yield a solution of the Least Squares equations $Q_i = 0$. It is also clear that there must be an absolute minimum of Q inside S. If θ^+ is a vector at which this occurs, clearly $Q_i(x, \theta^+) = 0$. However, the vector θ^* to which our process has converged may not be equal to θ^+. It is therefore necessary to discuss conditions under which it does yield a θ^+ vector. Let us assume that θ^+ is the *only* vector yielding the *absolute* minimum of Q. We then have as a necessary consequence that the quadratic form

$$\sum_{ij} Q_{ij}(x, \theta^+)u_iu_j > 0 \tag{25}$$

is positive definite and hence will remain positive definite in a convex 'neighborhood region' of θ^+, say S^+. We now prove the following uniqueness theorem:—

Theorem 1.

In any convex region S^+ of the parameter space θ in which the quadratic form

$$\sum_{ij} Q_{ij}(x, \theta)u_iu_j > 0 \tag{26}$$

is positive definite there can not be more than one stationary point of $Q(x, \theta)$.

The proof is almost obvious. Suppose there were two stationary points θ' and θ'' in S^+. Consider the function

$$F(v) = \sum_i \{Q_i(x, \theta'v + \theta''(1 - v))(\theta_i' - \theta_i'')\}. \tag{27}$$

Clearly $F(0) = F(1) = 0$ and hence by Rolle's Theorem $df(v)/dv = 0$ for some value $\bar{v}$ with $0 \leq \bar{v} \leq 1$, so that

$$\sum_{ij} Q_{ij}(x, \theta'\bar{v} + \theta''(1 - \bar{v}))(\theta_i' - \theta_i'')(\theta_i' - \theta_i'') = 0, \tag{28}$$

which would contradict (26).

From Theorem 1 it is clear that *if* the region S involved in conditions b and c

TABLE 1

Fitting of Exponential Regression by Modified Gauss-Newton Iteration Cycle 1
$L_0 = 580$, $B_0 = -180$, $K_0 = -0.16$

1	2	3	4	5	6
x	y	K_0x	$\exp(K_0x)$	$x \exp(K_0x)$	$y - f$
−5	127	.800 000	2.22 554	−11.12 771	− 52. 403
−3	151	.480 000	1.61 607	− 4.84 822	−138. 107
−1	379	.160 000	1.17 351	− 1.17 351	10. 232
1	421	−.160 000	.85 214	.85 214	− 5. 615
3	460	−.480 000	.61 878	1.85 635	− 8. 620
5	426	−.800 000	.44 933	2.24 665	− 73. 121
			6.93 537	−12.19 430	−267. 634

of Section 2 can be chosen to be the region S^+ containing the absolute minimum θ^+ and comprising only vectors θ for which the quadratic form (26) is positive definite, *then* our iterative process will converge to the vector θ^+ yielding the absolute minimum of Q. In practice, the difficulty about the above proposition will be to find a starting vector ${}_0\theta$ which is known to be in the regionS^+. This is particularly difficult when the surface represented by $Q(x, \theta)$ may have numerous local minima and/or maxima and/or saddle points. Under such conditions it is necessary to 'search' the parameter space θ at a wide grid in an attempt to locate a point in the region S^+. If a starting vector ${}_0\theta$ can be located in the region S^+ our method can be used to converge upon θ^+, i.e. the absolute minimum.

If there is a problem in which the absolute minimum is not unique for all large samples the Least Squares principle ceases to be an appropriate method for estimation since it will be incapable of distinguishing between the two solutions. Finally, it should be stressed that only the (unique) vector θ^+ which yields the *absolute* minimum of Q is of interest in statistical estimation theory (see Section 6). If the Least Squares equations $Q_i = 0$ permit solutions θ other than θ^+ these do not necessarily share the properties which make θ^+ a desirable estimator (see Section 6).

5. Illustration by a Numerical Example and Computational Short Cuts

We now illustrate the computational procedure of our method with the help of a numerical example. In the course of this illustration certain computational short cuts will be introduced.

The data in columns 1 and 2 of Table 1 below have been taken from a 'Fertilizer Experiment' in which the $n = 6$ responses $y_h (h = 1, 2, \cdots, 6)$ represent the yields of wheat corresponding to six rates of application of fertilizer, x_h, which, on a coded scale, are given the values $x_h = -5, -3, -1, 1, 3, 5$. It is intended to fit to these data the exponential law of 'diminishing returns',

$$Y = f(x; L, B, K) = L + Be^{Kx} \tag{29}$$

(sometimes called Mitcherlisch's Law of diminishing returns) in which there

TABLE 2

Formulas for equations for D_1, D_2 and D_3

D_1	D_2	D_3	
n	$\sum e^{K_0 x}$	$B_0 \sum x e^{K_0 x}$	$= \sum (y - f)$
$\sum e^{K_0 x}$	$\sum e^{2K_0 x}$	$B_0 \sum x e^{2K_0 x}$	$= \sum (y - f) e^{K_0 x}$
$B_0 \sum x e^{K_0 x}$	$B_0 \sum x e^{2K_0 x}$	$B_0^2 \sum x^2 e^{2K_0 x}$	$= B_0 \sum (y - f) x e^{K_0 x}$

is only one input variable, x, but $m = 3$ parameters namely $\theta_1 = L$, the asymptotic yield for large rates of fertilizer application, $\theta_3 = K$ (usually negative), the exponential rate of response decrease and $\theta_2 = B$ (usually negative), defining the mid-point response (at $x = 0$) by $L + B$.

Clearly in accordance with (4)

$$f_1 = (\partial f / \partial L) = 1; \qquad f_2 = e^{Kx}; \qquad f_3 = Bxe^{Kx} \tag{30}$$

and the equations corresponding to (10) can be spelt out as shown in Table 2 below where, in usual fashion, the unknowns $D_1 D_2 D_3$ are written as column headings.

Here $(y - f)$ is the residual with f evaluated at the starting vector with elements L_0, B_0 and K_0. In our example these trial values were taken as

$$L_0 = 580 \qquad B_0 = -180 \qquad K_0 = -.160. \tag{31}$$

In Table 1 we show the stages of computations required to evaluate the elements of the matrix coefficients in Table 2:—Columns (1) and (2) show the original data x and y, column (3) the products $K_0 x$, column (4) exp $(K_0 x)$ from a suitable table of the exponential function (the decimal accuracy is more than is required for the initial cycles), column (5) the products x exp $(K_0 x)$ and column (6) the residuals

$$y - f \equiv y - 180 \exp \{K_0 x\} \equiv (2) - 180\,(4).$$

It is clearly convenient to divide the last row and column in the matrix of Table 2 by B_0 and thereby obtain D_1, D_2 and $D_3 B_0$ as solutions. All elements in Table 2 are then seen to be either the totals of columns (4), (5) and (6) in Table 1 or the sum of squares and cross products of these columns. They are set out in Table 3 below.

The solution of these equations yields

$$D_1 = -89.6838, \qquad D_2 = 58.8877, \qquad D_3 = -.063\ 117.$$

TABLE 3

Cycle 1, Equations for D_1, D_2 and D_3/B_0

D_1	D_2	$B_0 D_3$	
6.	6.935 37	−12.194 30	= −267.634
6.935 37	10.252 765	−31.093 050	= −370.782 306
−12.194 30	−31.093 050	157.927 907	= 1055.627 355

TABLE 4

Cycle 1, Computation of Starting Values for Cycle 2

		$v = 1$			$v = \frac{1}{2}$		
x	y	Kx	exp kx	$y - f$	Kx	exp Kx	$y - f$
−5	127	1.115 585	3.05 135	6.140	.957 795	2.605 94	− 15.868
−3	151	.669 351	1.95 297	−102.888	.574 667	1.776 56	−116.736
−1	379	.223 117	1.24 997	39.971	.191 559	1.211 14	26.136
1	421	− .223 117	.80 002	27.476	−.191 559	.825 67	10.102
3	460	− .669 351	.51 204	31.598	−.574 677	.562 89	9.538
5	426	−1.115 585	.32 772	− 24.725	−.957 795	.383 74	− 51.434

The next stage, in accordance with equation (11) would be to find the minimum value of Q as a function of v by substituting

$$L = L_0 + v\,D_1\,, \qquad B = B_0 + v\,D_2\,, \qquad K = K_0 + v\,D_3 \tag{32}$$

in f and hence Q, which then becomes a function of v. In order to find the minimum of Q we proceed by an approximate method:—We evaluate Q for $v = 0$, $v = \frac{1}{2}$ and $v = 1$ and determine the level of v for which the parabola through $Q(0)$, $Q(\frac{1}{2})$, $Q(1)$ attains its minimum from*

$$v_{\min} = \tfrac{1}{2} + \tfrac{1}{4}(Q(0) - Q(1))/(Q(1) - 2Q(\tfrac{1}{2}) + Q(0)). \tag{33}$$

In this formula $Q(0) = \sum (y - f)^2$ can be evaluated directly from the $(y - f)$ values given in column (6) of Table 1. We obtain $Q(0) = 27376.825$. The computation of $Q(\frac{1}{2})$ and $Q(1)$ requires the evaluation of the exponential regression for

$$L_0 + \tfrac{1}{2}D_1 = 535.2081, \quad B_0 + \tfrac{1}{2}D_2 = -150.55615, \quad K_0 + \tfrac{1}{2}D_3 = -.191559$$

and

$$L_0 + D_1 = 490.4162, \quad B_0 + D_2 = -121.11230, \quad K_0 + D_3 = -.223117.$$

This work is shown in Table 4 below. We obtain $Q(\frac{1}{2}) = 17400.6578$ and $Q(1) = 14586.01079$ and accordingly from (32) the fraction $v_{\min} = .99651$. Thus the minimum of Q on the line determined by (32) occurs at

$$L_1 = 495.207, \qquad B_1 = -124.2621, \qquad K_1 = -.219741.$$

The process is now repeated with L_1, B_1 and K_1 taking the place of the starting trial values L_0, B_0 and K_0. In Table 5 we give the four cycles required to converge upon the absolute minimum of Q. The final least squares estimates are

$$L = 523.3, \qquad B = -157.0, \qquad K = -.1994. \tag{33}$$

* An attractive alternative suggested to me by Dr. K. Ruedenberg is to find the minimum of the parabola passing through $Q(0)$ and $Q(1)$ with the negative slope S at $v = 0$ given by (19). This minimum occurs at $v_{\min} = -S/(Q(1) - Q(0) - S)\,2$. In any case the computer program should always test whether the new minimum $Q(v_{\min})$ is actually smaller then $Q(0)$ and if this check fails, should redo the computation on a segment of half-length.

TABLE 5

Convergence of Least Squares Estimates in 4 Cycles of Iteration with Starting Values (31)

	L	B	K	D_1	D_2	D_3	$Q(0)$
Cycle 1	580.000	−180.000	−.160 000	−89.5838	58.8877	−.063 117	27376.82
2	495.208	−124.262	−.219 741	33.0685	−38.9138	.032 337	14590.58
3	524.960	−159.273	−.190 648	− 5.7328	7.0241	−.013 351	13639.10
4	519.422	−152.488	−.203 545	5.2393	− 6.0729	.005 572	13394.35
Final	523.3	−157.0	−.1994				

In order to illustrate how in this example the convergence depends on the choice of the starting values, we have set out in Table 6 the analogous information to Table 5 for the starting values:

$$L_0 = 500, \qquad B_0 = -140, \qquad K_0 = -0.18. \tag{34}$$

These starting values, used merely for purposes of illustration, are actually 'closer' to the final estimates than the previous ones. For work with desk computers it usually pays to compute fairly good starting values from short cut estimation procedures. For work on high speed computers it is sometimes more convenient to avoid a special program for the computation of good starting values and consequently perform a larger number of cycles in the iteration process. It will be seen that in the present example with the starting values (34) after four cycles virtually the same final answers as in (33) are reached.

It is of interest to check whether the condition (6) is, in fact, satisfied which would mathematically ensure the convergence of the procedures just found to converge to the same limit in our numerical illustrations. Without loss of generality condition (6) can be written in the form

$$S = \sum_{h=1}^{n} (u_1 + u_2 e^{-Kx_h} + u_3 x_h e^{-Kx_h})^2 > 0 \tag{35}$$

for any set of u_1 , u_2 , u_3 with $u_1^2 + u_2^2 + u_3^2 = 1$.

Now $S = 0$ is clearly only possible if each of the quantities inside the $(\ )^2$ are zero. We now show that this is not possible if there are at least three different

TABLE 6

Convergence of Least Squares Estimates in 4 Cycles of Iteration with Starting Values (34)

	L	B	K	D_1	D_2	D_3	$Q(0)$
Cycle 1	500.000	−140.000	−.180 00	12.325	− 4.389	−.035 83	18428.00
2	511.156	−143.972	−.212 50	15.866	−17.516	.018 56	13433.10
3	521.040	−154.884	−.200 94	2.734	− 2.618	.001 826	13392.94
4	523.805	−157.532	−.199 09	− 0.717	.83976	−.000 817	13390.23
Final	523.3	−156.9	−.1997				

values of x_h among the n observed x_h-values. In fact, the equation

$$u_1 + u_2 e^{+Kx} + u_3 x e^{+Kx} = 0 \tag{36}$$

for any set of given K, u_1 , u_2 , u_3 with $u_1^2 + u_2^2 + u_3^2 = 1$ has at most two roots x. To prove this write (36) first in the form

$$g(x) \equiv (u_2 + u_3 x) + u_1 e^{-Kx} = 0 \tag{37}$$

and observe that $g(x) = 0$ can not have more than two distinct roots x. For if it had three or more distinct roots there would be at least one root of $d^2g(x)/dx^2 = 0$ which is not possible since $d^2g/dx^2 = u_1 K^2 e^{-Kx}$. (Note that if $u_1 = 0$, $g(x)$, the equation of a line, has at most one root if $u_1^2 + u_2^2 = 1$).

It should be pointed out that following standard regression analysis the three equations in Table 2 can be immediately reduced to two equations by elimination of D_1 with the help of the first equation thereby reading sums of squares and products of *deviations* of $y - f$, $e^{K_0 x}$ and $xe^{K_0 x}$ as coefficients for D_2 and D_3 .

Computations, not here presented, would show that in the present example the standard (unmodified) Gauss-Newton method would have worked quite well so that the benefits derived from our modification are not here apparent. They will, however, be considerable in situations in which the Gauss-Newton method does not converge. In the present example the differentials (30) of f with regard to the parameters are comparatively simple functions. However, these differentials frequently are analytically involved expressions whilst the function $f(x_1 , \cdots , x_k ; \theta_1 , \cdots , \theta_m)$ itself is easy to evaluate. In such situations one may prefer to evaluate $m + 1$ sets of f-functions, namely those for the 'change of one parameter at a time' pattern;

$${}_\tau\theta_t = {}_0\theta_t + \delta_{t\tau}\,\Delta; \qquad t, \tau = 1, 2, \cdots , n, \tag{38}$$

where $\delta_{t\tau}$ is the Kronecker δ and Δ is a conveniently chosen small increment. The differentials of f with regard to θ_t are then computed from the difference ratios. This method is particularly convenient on high speed computers as it only requires a program for the evaluation of the regression function f itself.

6. Statistical Properties of the Least Squares Estimators

We confine ourselves here to certain results which follow directly from the properties of maximum likelihood estimation. Let us assume then in this section that the observed responses y_h result from the model

$$y_h = f(x_{1h} , \cdots , x_{kh} ; \theta_1 , \cdots , \theta_m) + e_h ,$$

where the 'error residuals' are independent normal variates with mean zero and variance σ^2. Under these assumptions the Least Squares estimators are identical with the maximum likelihood estimators. Certain results by A. Wald (1949) can now be applied to the present situation. Wald proved under very general continuity conditions that the parameter vector θ which yields the absolute maximum of the likelihood is a consistent estimator of the true parameter vector θ_0 . This result, taken in conjunction with Cramer's (1946) results on consistent solutions of the maximum likelihood equations implies that the θ which yields the absolute maximum of the likelihood is fully efficient. Moreover,

since a consistent solution of the likelihood equations is unique it is very unlikely that there will be situations in which any solutions of the likelihood equations (i.e. the Least Squares equations $Q_i = 0$) other than θ will be fully efficient or, indeed, have any desirable properties for serving as estimator of θ_0 . Since we have merely shown that under the very general conditions a, b and c of Section 2 our process will converge to a solution of the Least Squares equations $Q_i = 0$ it is clear that it is necessary either to examine the uniqueness of this solution or, in case there is not unique solution, to take such precautions as to ensure that the absolute minimum of Q is attained by the process.

7. Some Comparisons with Alternative Least Squares Techniques

Only a few points of a general nature can be made at the present time. As experience in the use of the method is accumulated in various fields of applications its advantages and disadvantages will become better substantiated.

Compared with the well known method of steepest descent each step of the present procedure is computationally more elaborate. With the former the direction of movement in the parameter space is made proportional to the Q_i . That is, only the first derivatives of Q enter into the trial and error approximation of the surface Q. By contrast our procedure (as does the Gauss-Newton method) employs a first order approximation for the residuals whose squares comprise Q, and is therefore based on a second order approximation to Q. It should be noted that this approximation is not identical with the second order Taylor expansion of Q except in the special case when all second derivatives $\partial^2 f/\partial\theta_i\partial\theta_\tau$ are zero. In general the second order approximations will differ slightly and there has been some discussion in the literature (S. L. Piotrowsky (1948) and H. Wolf (1951)) as to their relative merits. It is, however, generally agreed that the number of iterative steps with a second order approximation to Q should be smaller and the unmodified Gauss-Newton method, if it does converge, will usually converge faster than the 'steepest descent' method.

It should be also pointed out that other modifications of the Gauss-Newton method have been attempted in the past notably the one by Levenberg (1944). However, all these methods require to a varying degree judgements which depend essentially on an advance knowledge of the higher differentials of Q in the near solution region. Sometimes such knowledge can be accumulated as the parameter space is explored in the course of the iteration, but such procedures do not lend themselves well to programming on modern computers.

Finally we should briefly mention the relation of our method of Least Squares solution to modern procedures of finding maxima (or minima) of statistical response surfaces subject to observational errors. Although there are certain features common to the two problems, there are essential differences in the economical aspects of different operations in the two problems. For example, with Least Squares work the evaluation of Q for a new point in the parameter space is a comparatively 'cheap' operation whilst the analogous determination of the response surface Q requires the often costly evaluation of the response for a new point in the 'factor space'. It is not surprising therefore that the most economical 'strategy' in the two problems will differ. A strategy suitable for

response surface work and based on quadratic programming has recently been sketched by E. M. L. Beale (1958). Earlier G. E. P. Box and K. B. Wilson (1951) had introduced the method of steepest ascent into this area and initiated a series of papers examining the capabilities and limitations of this method in industrial response problem.

It was due to the fact that I had a preview of the above report by E. M. L. Beale (1958) prepared by the S.T.R.G. at Princeton that I failed to make inquires from members of that group concerning other memoranda on the topic. It was not until after submission of this manuscript that my attention was drawn to the existence of an IBM Share program written in February 1959 by G. W. Booth and T. I. Peterson under the guidance of G. E. P. Box and M. E. Muller as a cooperative efford of the Statistical Techniques Research Group (Princeton), Department of Mathematics (Princeton University) and the Mathematics and Applications Department, Data Processing Division, IBM. This program, although it differs from the present procedure in detail, does in fact employ the same basic idea of moving in the direction of the Gauss-Newton solution of equations (10). Moreover my attention has been drawn to a mention of this idea earlier in a paper by G. E. P. Box (1958) p. 220 lines 13–18. Finally reference should be made to a published abstract of a paper by M. B. Wilk (1958) in which a different modification of the Gauss-Newton method is proposed. The above references have now been incorporated in the list which follows.

References

Beale, E. M. L. (1958), On an iterative method for finding a local minimum of a function of more than one variable. Technical Report No. 25, Stat. Tech. Research Group, Princeton University.

Booth, G. W. and Peterson, T. I. (1958), I.B.M. Share Program Pa. No. 687 WL NL1. 'Non-Linear Estimation'.

Box, G. E. P. and Wilson, K. B. (1951), On the experimental attainment of optimum conditions. *J. R. Statist. Soc.* (B) 13, 1–45.

Box, G. E. P. (1958), 'Use of statistical methods in the elucidation of basic mechanisms'. *Bull Inst. Intern. De Statistique*, 36, 215.

Cramer, H. (1946), *Mathematical Methods of Statistics*. Princeton University Press.

Levenberg, K. (1944), A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Math.* 2, 164–8.

Piotrowsky, S. L. (1948), Some remarks on the weight of unknowns as determined by the method of differential corrections. *Proc. Nat. Acad. Sci.*, U. S. A. 34, 23.

Wald, A. (1949), A note on the consistency of the maximum likelihood estimate. *Ann. Math. Stat.* 20, 595.

Wilk, M. B. (1958), 'An identity of use in non-linear Least Squares' Abstract 24, *Ann. Math. Statistics*. June 1958.

Wolf, H. (1951), Ist eine revision der gewichtskoefficienten-berechnung in der methode der kleinsten quadrate erforderlich? *Zeitsch. F. Angew. Math. & Mech.* 31, 153.