--NOTES—

A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES*

BY KENNETH LEVENBERG[1] (*Frankford Arsenal*)

The standard method for solving least squares problems which lead to non-linear normal equations depends upon a reduction of the residuals to linear form by first order Taylor approximations taken about an initial or trial solution for the parameters.[2] If the usual least squares procedure, performed with these linear approximations, yields new values for the parameters which are not sufficiently close to the initial values, the neglect of second and higher order terms may invalidate the process, and may actually give rise to a larger value of the sum of the squares of the residuals than that corresponding to the initial solution. This failure of the standard method to improve the initial solution has received some notice in statistical applications of least squares[3] and has been encountered rather frequently in connection with certain engineering applications involving the approximate representation of one function by another. The purpose of this article is to show how the problem may be solved by an extension of the standard method which insures improvement of the initial solution.[4] The process can also be used for solving non-linear simultaneous equations, in which case it may be considered an extension of Newton's method.

Let the function to be approximated be $h(x, y, z, \cdots)$, and let the approximating function be $H(x, y, z, \cdots; \alpha, \beta, \gamma, \cdots)$, where $\alpha, \beta, \gamma, \cdots$ are the unknown parameters. Then the residuals at the points, $(x_i, y_i, z_i, \cdots)$, $i=1, 2, \cdots, n$, are

$$f_i(\alpha, \beta, \gamma, \cdots) = H(x_i, y_i, z_i, \cdots; \alpha, \beta, \gamma, \cdots) - h(x_i, y_i, z_i, \cdots), \tag{1}$$

and the least squares criterion requires the minimization of

$$s(\alpha, \beta, \gamma, \cdots) = \sum_1^n f_i^2. \tag{2}$$

(It is assumed that the weights of the residuals are unity. If not, consider the func-

* This paper was read before the Annual Meeting of the American Mathematical Society in Chicago, Ill., on Nov. 26, 1943. Manuscript received Feb. 2, 1944.

[1] The writer wishes to thank Dr. J. G. Tappert, under whose direction the method of damped least squares was developed, and Dr. H. B. Curry, for valuable suggestions and guidance.

[2] E. T. Whittaker and G. Robinson, *The calculus of observations*, Blackie and Son, London, 1937, p. 214.

[3] E. B. Wilson and R. R. Puffer, *Least squares and laws of population growth*, Proc. Amer. Acad. Arts and Sci. (Boston), **68**, 285–382 (1933).

[4] Another extension of the standard method, which requires the use of second partial derivatives, is given by Wilson and Puffer (l.c.).

A different kind of approach, not based upon the standard method, is given by Cauchy, *Méthode générale pour la résolution des systèmes d'équations simultanées*, C. R. Acad. Sci. Paris, **25**, 536–538 (1847). See also a paper by H. B. Curry, not yet published, (abstract in Bull. Amer. Math. Soc., **49**, 859 (1943), abstract No. 278).

tion f_i to be the product of the residual and the square root of the corresponding weight.) Choosing an initial solution, $p_0=(\alpha_0, \beta_0, \gamma_0, \cdots)$, at which it is assumed that s does not have a stationary value, the first order Taylor expansions of the residuals are taken about p_0, giving a set of linear approximations to the residuals,

$$f_i(\alpha, \beta, \gamma, \cdots) \cong F_i(\alpha, \beta, \gamma, \cdots) = f_i(p_0) + \frac{\partial f_i}{\partial \alpha}\Delta\alpha + \frac{\partial f_i}{\partial \beta}\Delta\beta + \frac{\partial f_i}{\partial \gamma}\Delta\gamma + \cdots, \tag{3}$$

where $\Delta\alpha=\alpha-\alpha_0$, $\Delta\beta=\beta-\beta_0$, $\cdots$, and the partial derivatives are evaluated at p_0. Now, the standard method consists of minimizing

$$S(\alpha, \beta, \gamma, \cdots) = \sum_1^n F_i^2 \tag{4}$$

by setting the partial derivatives of S with respect to the various parameters equal to zero, yielding the usual linear normal equations,

$$\begin{aligned} \frac{1}{2}\frac{\partial S}{\partial \alpha} &= [\alpha\alpha]\Delta\alpha + [\alpha\beta]\Delta\beta + [\alpha\gamma]\Delta\gamma + \cdots + [\alpha 0] = 0, \\ \frac{1}{2}\frac{\partial S}{\partial \beta} &= [\beta\alpha]\Delta\alpha + [\beta\beta]\Delta\beta + [\beta\gamma]\Delta\gamma + \cdots + [\beta 0] = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots, \end{aligned} \tag{5}$$

where the notation [ ] is a symbol of summation, so that, e.g.,

$$[\alpha\alpha] = \sum_1^n \left(\frac{\partial f_i}{\partial \alpha}\right)^2, \qquad [\alpha\beta] = \sum_1^n \left(\frac{\partial f_i}{\partial \alpha}\cdot\frac{\partial f_i}{\partial \beta}\right), \qquad [\alpha 0] = \sum_1^n \left(\frac{\partial f_i}{\partial \alpha}\cdot f_i\right), \quad \text{etc.}$$

However, as pointed out above, the values of the increments, $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$, $\cdots$, obtained by solving equations (5), may be so large in absolute value as to invalidate the approximations (3) so that the decrease in S may not correspond to a decrease in s.

In such cases, it would seem advisable to limit or "damp" the absolute values of the increments of the parameters in order to improve the first order Taylor approximations (3) and to minimize simultaneously the sum of the squares of the approximating residuals (4) under these damped conditions. In order to make both the increments and the residuals small in absolute value, the least squares idea can be employed. The sum of the squares of both the residuals and the increments may be minimized. More precisely, the expression to be minimized will be

$$\overline{S}(\alpha, \beta, \gamma, \cdots) = wS(\alpha, \beta, \gamma, \cdots) + a(\Delta\alpha)^2 + b(\Delta\beta)^2 + c(\Delta\gamma)^2 + \cdots, \tag{6}$$

where $a, b, c, \cdots$ are a system of positive constants or weighting factors expressing the relative importance of damping the different increments, and w is a positive quantity expressing the relative importance of the residuals and increments in this minimizing process. If we denote the point at which $\overline{S}$ takes its minimum, for any positive value of w, by $p_w=(\alpha_w, \beta_w, \gamma_w, \cdots)$, and set

$$Q(\alpha, \beta, \gamma, \cdots) = a(\Delta\alpha)^2 + b(\Delta\beta)^2 + c(\Delta\gamma)^2 + \cdots, \tag{7}$$

it is seen, under the assumption that s is not stationary at p_0, that

$$wS(p_w) < wS(p_w) + Q(p_w) = \bar{S}(p_w) < \bar{S}(p_0) = wS(p_0) + Q(p_0) = wS(p_0),$$

whence
$$S(p_w) < S(p_0). \tag{8}$$

Also, denoting the standard least squares solution by p_∞ (the reason for the notation is discussed later), we have

$$wS(p_w) + Q(p_w) = \bar{S}(p_w) < \bar{S}(p_\infty) = wS(p_\infty) + Q(p_\infty) < wS(p_w) + Q(p_\infty),$$

whence
$$Q(p_w) < Q(p_\infty). \tag{9}$$

Inequality (8) shows that the minimization of (6) will diminish the sum of the squares of the approximating residuals, S, and (9) shows that the increments given by the standard least squares solution will be improved in the sense that the weighted sum of their squares, Q, will be reduced. That the sum of the squares of the true residuals, s, can be diminished, will be proved shortly.

To minimize (6) and obtain p_w, the partial derivatives of $\bar{S}$ with respect to the various parameters are put equal to zero, and we get

$$\frac{\partial \bar{S}}{\partial \alpha} = w\frac{\partial S}{\partial \alpha} + 2a\Delta\alpha = 0, \quad \frac{\partial \bar{S}}{\partial \beta} = w\frac{\partial S}{\partial \beta} + 2b\Delta\beta = 0, \quad \cdots .$$

When we divide through by $2w$, and substitute the expressions for the partial derivatives of S from (5), the "damped normal equations" become

$$\begin{aligned} ([\alpha\alpha] + aw^{-1})\Delta\alpha + [\alpha\beta]\Delta\beta + [\alpha\gamma]\Delta\gamma + \cdots + [\alpha 0] &= 0, \\ [\beta\alpha]\Delta\alpha + ([\beta\beta] + bw^{-1})\Delta\beta + [\beta\gamma]\Delta\gamma + \cdots + [\beta 0] &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & . \end{aligned} \tag{10}$$

These equations are seen to be the same as the ordinary normal equations (5), except for the coefficients of the principal diagonal, which are increased by quantities proportional to the weighting factors $a, b, c, \cdots$, respectively. Since the symmetry of the matrix of the coefficients of equations (5) is preserved, simplified methods of solution of linear simultaneous equations, which take full advantage of such symmetry,[5] may be used to solve equations (10). It is to be noted that the standard method of least squares corresponds to $w \to \infty$, and is thus a special case of the method here given, which may be termed the method of "damped least squares."

If we denote the number of parameters by k, it is seen from the determinantal solution of equations (10) that, in the neighborhood of $w=0$,

$$\Delta\alpha = \alpha_w - \alpha_0 = \frac{-[\alpha 0]w^{1-k}bcd\cdots + \cdots}{w^{-k}abc\cdots + \cdots} = -[\alpha 0]a^{-1}w + \cdots,$$

whence
$$\left(\frac{d\alpha_w}{dw}\right)_{w=0} = -[\alpha 0]a^{-1}, \tag{11}$$

and similarly for the other parameters. Now

$$\frac{ds(p_w)}{dw} = \frac{\partial s}{\partial \alpha}\cdot\frac{d\alpha}{dw} + \frac{\partial s}{\partial \beta}\cdot\frac{d\beta}{dw} + \cdots, \tag{12}$$

[5] P. S. Dwyer, *The solution of simultaneous equations*, Psychometrika, **6**, 101–129 (1941).

and, from the definition of the summation symbols, we find that the partial derivatives of s at p_0 are given by

$$\frac{\partial s}{\partial \alpha} = 2[\alpha 0], \qquad \frac{\partial s}{\partial \beta} = 2[\beta 0], \cdots . \tag{13}$$

Hence the substitution of (11) and (13) in (12) yields

$$\left(\frac{ds}{dw}\right)_{w=0} = -2\{[\alpha 0]^2 a^{-1} + [\beta 0]^2 b^{-1} + \cdots\}. \tag{14}$$

This derivative is negative since the partial derivatives in (13) are not all zero, by the assumption that s does not have a stationary value at p_0. Therefore, $s(p_w)$ is decreasing at $w=0$, thus insuring that values of w can be found for which the sum of the squares of the true residuals (2) will be reduced.

The best value of w to use may theoretically be determined directly by solving

$$\frac{ds(p_w)}{dw} = 0; \tag{15}$$

however, this equation is generally complex in practice. By writing

$$s(p_w) \cong s(p_0) + w\left(\frac{ds}{dw}\right)_{w=0}, \tag{16}$$

and setting the left side of (16) equal to zero on the assumption that p_0 was chosen so that the decreased value $s(p_w)$ will be small, the approximate formula,

$$w \cong -\frac{s(p_0)}{ds/dw_{w=0}} = \frac{\frac{1}{2}s(p_0)}{[\alpha 0]^2 a^{-1} + [\beta 0]^2 b^{-1} + \cdots}, \tag{17}$$

is obtained.[6] If necessary, this value may be improved by calculating $s(p_w)$ for several different trial values of w, so that an approximate minimum may be located graphically. Experience with the method, especially in connection with fitting a particular function $H(x, y, z, \cdots; \alpha, \beta, \gamma, \cdots)$, enables one to get an idea of the general order of magnitude of the best value of w so that very few trial values of w should suffice. If so desired, the improved set of values of the parameters may be further improved (if the true minimum has not already been reached), by a repetition of the process, considering this improved set as a new initial solution.

So far, the weighting system $a, b, c, \cdots$ has been left arbitrary, the only restriction being that the weighting factors be positive. If we set the criterion that these factors be chosen so that the directional derivative of s, taken at $w=0$ along the curve $\alpha=\alpha_w$, $\beta=\beta_w$, $\cdots$, should have its minimum value, namely, the negative gradient, we have

$$\frac{ds}{dw}\left\{\left(\frac{d\alpha}{dw}\right)^2 + \left(\frac{d\beta}{dw}\right)^2 + \cdots\right\}^{-1/2} = -\left\{\left(\frac{\partial s}{\partial \alpha}\right)^2 + \left(\frac{\partial s}{\partial \beta}\right)^2 + \cdots\right\}^{1/2}, \tag{18}$$

where the derivatives are taken at $w=0$. Substitution of (14), (11), (13) in (18) gives us

[6] This type of approximation was used by Cauchy (l.c.).

$$\left\{[\alpha 0]^2 a^{-1} + [\beta 0]^2 b^{-1} + \cdots\right\}\left\{\alpha[0]^2 a^{-2} + [\beta 0]^2 b^{-2} + \cdots\right\}^{-1/2} = \left\{[\alpha 0]^2 + [\beta 0]^2 + \cdots\right\}^{1/2}, \quad (19)$$

and this is satisfied when the factors $a, b, c, \cdots$ are all equal. Without loss of generality, they may be taken equal to unity. For this weighting system, the formation of the damped normal equations (10) may be thought of as being accomplished simply by the addition of a positive constant, $1/w$, to the coefficients of the principal diagonal of the standard normal equations (5). Another weighting system which has been used successfully is, $a=[\alpha\alpha]$, $b=[\beta\beta]$, $\cdots$; in this case the damped normal equations are formed by multiplying the principal diagonal coefficients of the standard normal equations by a constant greater than unity, $1+1/w$.

The nature of the damping which we have imposed upon the parameter variables can be given a simple geometric interpretation. For instance, if the unity weighting system is considered, the "overshooting" of the solution is prevented by damping the distance (k dimensional) from the initial solution point, since Q is then the square of this distance. By this restriction of k dimensional distance (which would appear to be a natural way to prevent overshooting), we are not obliged to decide on an arbitrary preassigned procedure restricting the variables individually, as is done, for example, by the method of Cauchy (l.c.). The greater freedom given the individual variables by the method of damped least squares may account for the fact that it has solved, with a comparatively rapid rate of convergence, types of problems which are of much greater complexity than those to which the principle of least squares is ordinarily applied.

ON THE DEFLECTION OF A CANTILEVER BEAM*

By H. J. BARTEN (*Washington Navy Yard*)

In spring theory it is sometimes necessary to compute the deflection of a cantilever beam for which the squares of the first derivatives cannot be neglected as is done in classical beam theory. This problem is thus placed in the same category as the problem of the elastica.

The solution given in this note can be applied to a cantilever of any stiffness. The difference between the deflection as found by the classical beam theory and that found by the present method is, however, noticeable only in the case of beams of low stiffness.

The clamped end of the beam is taken as the origin of coordinates and downward deflections are considered as positive. A point on the beam may be identified by four quantities of which only one is independent. These four quantities are the two rectangular coordinates x and y, the arc length s measured from the origin of coordinates, and the deflection angle θ which is the angle between the tangent to the curve at the point under discussion and the horizontal. We may thus identify this point by the symbol (x, y, s, θ). The subscript L is used to identify the value of these quantities at the free end of the beam. Before deflection a vertical load P is applied at the point $(L, 0, L, 0)$. The beam has a uniform cross section of moment of inertia I and is com-

* Received Feb. 21, 1944.