SIAM J. OPTIMIZATION
Vol. 1, No. 1, pp. 1-17, February 1991

VARIABLE METRIC METHOD FOR MINIMIZATION*

WILLIAM C. DAVIDON†

Abstract. This is a method for determining numerically local minima of differentiable functions of several variables. In the process of locating each minimum, a matrix which characterizes the behavior of the function about the minimum is determined. For a region in which the function depends quadratically on the variables, no more than N iterations are required, where N is the number of variables. By suitable choice of starting values, and without modification of the procedure, linear constraints can be imposed upon the variables.

Key words. variable metric algorithms, quasi-Newton, optimization

AMS(MOS) subject classifications. primary, 65K10; secondary, 49D37, 65K05, 90C30

A belated preface for ANL 5990. Enrico Fermi and Nicholas Metropolis used one of the first digital computers, the Los Alamos Maniac, to determine which values of certain theoretical parameters (phase shifts) best fit experimental data (scattering cross sections) [8]. They varied one theoretical parameter at a time by steps of the same magnitude, and when no such increase or decrease in any one parameter further improved the fit to the experimental data, they halved the step size and repeated the process until the steps were deemed sufficiently small. Their simple procedure was slow but sure, and several of us used it on the Avidac computer at the Argonne National Laboratory for adjusting six theoretical parameters to fit the pion-proton scattering data we had gathered using the University of Chicago synchrocyclotron [9]. To see how accurately the six parameters were determined, I varied them from their optimum values, and used the resulting degradations in the fit to estimate a six-by-six error matrix. This matrix approximates the inverse of a Hessian matrix of second derivatives of the objective function f, and specifies a metric in the space of gradients ∇f. Conjugate displacements in the domain of a quadratic objective function change gradients by amounts which are orthogonal with respect to this metric. The key ideas that led me to the development of variable-metric algorithms were 1) to update a metric in the space of gradients during the search for an optimum, rather than waiting until the search was over, and 2) to accelerate convergence by using each updated metric to choose the next search direction. In those days, we needed faster convergence to get results in the few hours between expected failures of the Avidac's large roomful of a few thousand bytes of temperamental electrostatic memory.

Shortly after joining the theoretical physics group at Argonne National Laboratory in 1956, I programmed the first variable-metric algorithms for the Avidac and used them to analyze the scattering of pi mesons by protons [10]. In 1957, I submitted a brief article about these algorithms to the Journal of Mathematics and Physics. This article was rejected, partly because it lacked proofs of convergence. The referee also found my notation "a bit bizarre," since I used "+" rather than "$k+1$" to denote updated quantities, as in "$x_+ = x + \alpha s$" rather than "$x_{k+1} = x_k + \alpha_k s_k$." While I then turned to other research, another member of our theoretical physics group, Murray Peshkin, modified and adapted one of these programs for Argonne's IBM 650. An

* This belated preface was received by the editors June 4, 1990; accepted for publication August 10, 1990. The rest of this article was originally published as Argonne National Laboratory Research and Development Report 5990, May 1959 (revised November 1959). This work was supported in part by University of Chicago contract W-31-109-eng-38.

† Department of Mathematics, Haverford College, Haverford, Pennsylvania 19041-1392.

Argonne physicist who used the program, Gilbert Perlow, urged me to publish a description of the algorithm so that he and Andrew Stehney could refer to it in a paper they were writing about their analysis of the radioactive decay of certain fission products [7]. Their reference thirteen was the first one to the report I was preparing [ANL-5990].

While this report focused mainly on the particular variable-metric algorithm which seemed to work best, it divided all algorithms of this type into five parts:

1. Choose a step direction s by acting on the current gradient $\mathbf{g}$ with the current metric in gradient space. If in this metric, $\mathbf{g}$ has a sufficiently small magnitude, then go to step 5.

2. Estimate the location of an optimum in the direction $\mathbf{s}$; e.g., by making a cubic interpolation. Go to this location if a sufficient change in the objective function is expected, else choose a new direction.

3. Evaluate the objective function f and and its gradient ∇f at the location x chosen in step 2 and estimate the directional derivative $\partial \nabla f(x+\alpha s)/\partial \alpha|_{\alpha=0}$ of ∇f at x.

4. Update the metric in gradient space, so that it yields $\mathbf{s}$ when acting on the directional derivative estimated in step 3. Return to step 1.

5. Test the current metric and minimizer. If these seem adequate, then quit, else return to step 1.

The hunting metaphor used in the report to name these five parts was chosen with tongue in cheek, since I expected the report would be read mostly by friends who knew I opposed killing for sport. The report would have been clearer had it first presented just the basic algorithm, with only those features needed to optimize quadratic objective functions, without the various "bells and whistles," which were added to accelerate convergence for certain nonquadratic objective functions; e.g., Formula 6.1 for the components $g_{\mu s}(x) = \lambda \partial g_\mu(x+\alpha s)/\partial \alpha|_{\alpha=0}$ of the directional derivative of the gradient would simplify to $g_\mu^+ - g_\mu \to g_{\mu s}$. It would also have been clearer if the rank-two update to the metric presented in the body of the report (later known as the Davidon–Fletcher–Powell (DFP) update) had been compared with the symmetric rank-one update (which was relegated to the appendix because it had not worked as well on certain problems).

Optimization algorithms were not among my research interests for several years after writing ANL-5990, and I returned to them only after others had called my attention to Fletcher and Powell's pioneering work on the subject [11].

1. Introduction. The solution to many different types of physical and mathematical problems can be obtained by minimizing a function of a finite number of variables. Among these problems are least-squares fitting of experimental data, determination of scattering amplitudes and energy eigenvalues by variational methods, the solution of differential equations, etc. With the use of high-speed digital computers, numerical methods for finding the minima of functions have received increased attention. Some of the procedures which have been used are those of optimum gradient [1], conjugate gradients [2], the Newton–Raphson iteration (see, e.g., [3], [4]) and one by Garwin and Reich [5]. In many instances, however, all of these methods require a large number of iterations to achieve a given accuracy in locating the minimum. Also, for some behaviors of the function being minimized, the procedures do not converge.

The method presented in this paper has been developed to improve the speed and accuracy with which the minima of functions can be evaluated numerically. In addition, a matrix characterizing the behavior of the function in the neighborhood of the minimum is determined in the process. Linear constraints can be imposed upon the variables by suitable choice of initial conditions, without alteration of the procedure.

2. Notation. We will employ the summation convention:

$$a^{\mu}b_{\mu} \equiv \sum_{\mu=1}^{N} a^{\mu}b_{\mu}.$$

In describing the iterative procedure, we will use symbols for memory locations rather than successive values of a number; e.g., we would write $x+3 \to x$ instead of $x_i+3=x_{i+1}$. In this notation, the sequence of operations is generally relevant. The following symbols will be used.

x^{μ}: $\mu = 1, \cdots, N$: the set of N independent variables.

$f(\mathbf{x})$: the value of the function to be minimized evaluated at the point $\mathbf{x}$.

$g_{\mu}(\mathbf{x})$: the derivatives of $f(\mathbf{x})$ with respect to x^{μ} evaluated at x:

$$g_{\mu}(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x^{\mu}}.$$

$h^{\mu\nu}$: a nonnegative symmetric matrix which will be used as a metric in the space of the variables.

Δ: The determinant of $h^{\mu\nu}$.

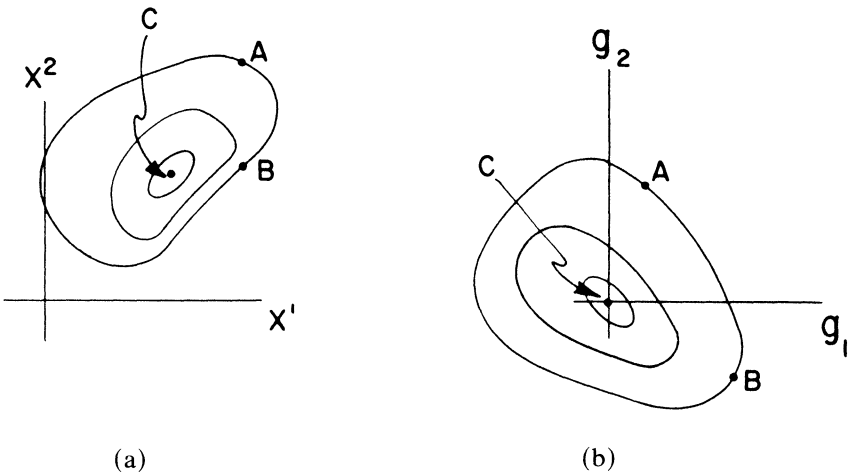
ε: 2 times fractional accuracy to which the function $f(\mathbf{x})$ is to be minimized.

d: a limiting value for what is to be considered as a "reasonable" minimum value of the function. For least-squares problems, d can be set equal to zero.

K: an integer which specifies the number of times the variables are to be changed in a random manner to test the reliability of the determination of the minimum.

3. Geometrical interpretation. It is convenient to use geometrical concepts to describe the minimization procedure. We do so by considering the variables x^{μ} to be the coordinates of a point in an N-dimensional linear space. As shown in Fig. 1(a), the set of $\mathbf{x}$ for which $f(\mathbf{x})$ is constant forms an $N-1$ dimensional surface in this space. One of this family of surfaces passes through each $\mathbf{x}$, and the surface about a point is characterized by the gradient of the function at that point:

$$g_{\mu}(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x^{\mu}}.$$



(a) (b)

FIG. 1. *Geometrical interpretation of* x^{μ} *and* $g_{\mu}(\mathbf{x})$.

These N components of the gradient can in turn be considered as the coordinates of a point in a different space, as shown in Fig. 1(b). As long as $f(\mathbf{x})$ is differentiable at all points, there is a unique point $\mathbf{g}$ in the gradient space associated with each point $\mathbf{x}$ in the position space, though there may be more than one $\mathbf{x}$ with the same $\mathbf{g}$.

In the neighborhood of any one point A the second derivatives of $f(\mathbf{x})$ specify a linear mapping of changes in position, $\mathbf{dx}$, onto changes in gradient $\mathbf{dg}$, in accordance with the equation

$$dg_\mu = \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \, dx^\nu. \tag{3.1}$$

The vectors $\mathbf{dx}$ and $\mathbf{dg}$ will be in the same direction only if $\mathbf{dx}$ is an eigenvector of the Hessian matrix:

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|.$$

If the ratios among the corresponding eigenvalues are large, then for most $\mathbf{dx}$ there will be considerable difference in the directions of these two vectors.

All iterative gradient methods, of which this is one, for locating the minima of functions consist of calculating $\mathbf{g}$ for various $\mathbf{x}$ in an effort to locate those values of $\mathbf{x}$ for which $\mathbf{g} = 0$, and for which the Hessian matrix is positive definite. If this matrix were constant and explicitly known, then the value of the gradient at one point would suffice to determine the minimum. In that case the change desired in $\mathbf{g}$ would be $-\mathbf{g}$, so we would have

$$-g_\mu = \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \, \Delta x^\nu, \tag{3.2}$$

from which we could obtain Δx^ν by multiplying (3.2) by the inverse of the matrix

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|.$$

However, in most situations of interest,

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|$$

is not constant, nor would explicit evaluation at points that might be far from a minimum represent the best expenditure of time.

Instead, an initial trial value is assumed for the matrix,

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|^{-1}.$$

This matrix, denoted by $h^{\mu\nu}$, specifies a linear mapping of all changes in the gradient onto changes in position. It is to be symmetric and nonnegative (positive definite if there are no constraints on the variables). After making a change in the variable $\mathbf{x}$, this trial value is improved on the basis of the actual relation between the changes in $\mathbf{g}$ and $\mathbf{x}$. If

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|$$

is constant, then, after N iterations, not only will the minimum of the function be determined, but also the final value of $h^{\mu\nu}$ will equal

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|^{-1}.$$

We shall subsequently discuss the significance of this matrix in specifying the accuracy to which the variables have been determined.

The matrix $h^{\mu\nu}$ can be used to associate a squared length to any gradient, defined by $h^{\mu\nu}g_\mu g_\nu$. If the Hessian matrix were constant and $h^{\mu\nu}$ were its inverse, then $\frac{1}{2}h^{\mu\nu}g_\mu g_\nu$ would be the amount by which $f(\mathbf{x})$ would exceed its minimum value. We therefore consider $h^{\mu\nu}$ as specifying a metric, and when we refer to the lengths of vectors, we will imply their lengths using $h^{\mu\nu}$ as the metric. We have called the method a "variable metric" method to reflect the fact that $h^{\mu\nu}$ is changed after each iteration.

We have divided the procedure into five parts which, to a large extent, are logically distinct. This not only facilitates the presentation and analysis of the method, but it is convenient in programming the method for machine computation.

4. Ready: Chart 1. The function of this section is to establish a direction along which to search for a relative minimum, and to box off an interval in this direction within which a relative minimum is located. In addition, the criterion for terminating the iterative procedure is evaluated.

Those operations which are only performed at the beginning of the calculation and not repeated on successive iterations have been included in Chart 1. These include the loading of input data, initial printouts, and the initial calculation of the function and its gradient. This latter calculation is treated as an independent subroutine, which may on its initial and final calculations include some operations not part of the usual iteration, such as loading operations, calculation of quantities for repeated use, special printouts, etc. A counter recording the number of iterations has been found to be a convenience, and is labeled I.

The iterative part of the computation begins with "READY 1." The direction of the first step is chosen by using the metric $h^{\mu\nu}$ in the relation

$$-h^{\mu\nu}g_\nu \to s^\mu. \tag{4.1}$$

The "component" of the gradient in this direction is evaluated through the relation

$$s^\mu g_\mu \to g_s. \tag{4.2}$$

From (4.1) and (4.2) we see that $-g_s$ is the squared length of $\mathbf{g}$, and hence the improvement to be expected in the function is $-\frac{1}{2}g_s$. The positive definiteness of $h^{\mu\nu}$ insures that g_s is negative, so that the step is in a direction which (at least initially) decreases the function. If its decrease is within the accuracy desired, i.e., if $g_s + \varepsilon > 0$, then the minimum has been determined. If not, we continue with the procedure.

In a first effort to box in the minimum, we take a step which is twice the size that would locate the minimum if the trial $h^{\mu\nu}$ were

$$\left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|^{-1}.$$

However, in order to prevent this step from being unreasonably large when the trial $h^{\mu\nu}$ is a poor estimate, we restrict the step to a length such that $(\lambda s^\mu)g_\mu$, the decrease in the function if it continued to decrease linearly, is not greater than some preassigned maximum, $2(f-d)$. We then change x^μ by

$$x^\mu + \lambda s^\mu \to x^{+\mu}, \tag{4.3}$$

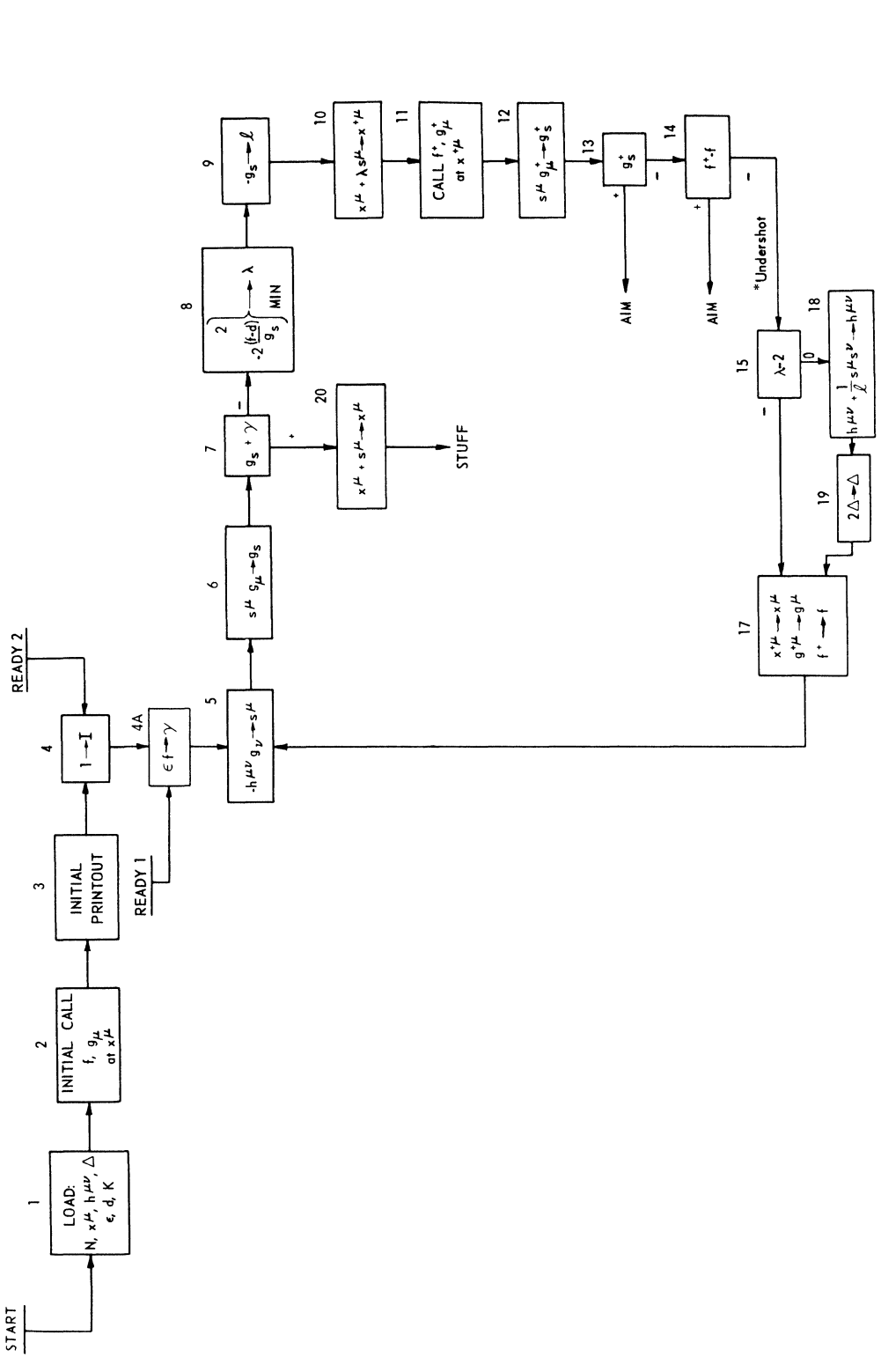


CHART 1: *Ready.*

* Optionally Printed Statements

and calculate the new value of the function and its gradient at $x^{+\mu}$. If the projection $s^\mu g_\mu^+ = g_s^+$ of the new gradient in the direction of the step is positive, or if the new value of the function f^+ is greater than the original f, then there is a relative minimum along the direction s between $\mathbf{x}$ and $\mathbf{x}^+$, and we proceed to "Aim" where we will interpolate its position. However, if neither of these conditions is fulfilled, the function has decreased and is decreasing at the point $\mathbf{x}^+$, and we infer that the step taken was too small. If the step had been limited by the preassigned change in the function d, we double d. If the step had been taken on the basis of $h^{\mu\nu}$, we modify $h^{\mu\nu}$ so as to double the squared length of s^μ, leaving the length of all perpendicular vectors unchanged. This is accomplished by

$$h^{\mu\nu} + \frac{1}{\ell} s^\mu s^\nu \to h^{\mu\nu}, \tag{4.4}$$

where ℓ is the squared length of s^μ. This doubles the determinant of $h^{\mu\nu}$. The process is then repeated, starting from the new position.

5. Aim: Chart 2. The function of this section is to estimate the location of the relative minimum within the interval selected by "Ready." Also, a comparison is made of the improvement expected by going to this minimum with that from a step perpendicular to this direction.

Inasmuch as the interpolation is along a one-dimensional interval, it is convenient to plot the function along this direction as a simple graph (see Fig. 2).

The values of f and f^+ of the function at points $\mathbf{x}$ and $\mathbf{x}^+$ are known, and so are its slopes, g_s and g_s^+, at these two points. We interpolate for the location of the minimum by choosing the "smoothest" curve satisfying the obundary conditions at $\mathbf{x}$ and $\mathbf{x}^+$, namely, the curve defined as the one which minimizes

$$\int_0^\lambda d\alpha \left(\frac{d^2 f}{d\alpha^2}\right)^2$$

over the curve. This is the curve formed by a flat spring fitted to the known ordinates and slopes at the end points, provided the slope is small. The resulting curve is a cubic, and its slope at any α $(0 \leqq \alpha \leqq \lambda)$ is given by

$$g_s(\alpha) = g_s - \frac{2\alpha}{\lambda}(g_s + z) + \frac{\alpha^2}{\lambda^2}(g_s + g_s^+ + 2z), \tag{5.1}$$

where

$$z = \frac{3(f - f^+)}{\lambda} + g_s + g_s^+.$$

The root of (5.1) that corresponds to a minimum lies between 0 and 1 by virtue of the fact that $g_s < 0$ and either $g_s^+ > 0$ or $z < g_s + g_s^+$. It can be expressed as

$$\alpha_{\min} = \lambda(1 - a),$$

where

$$a = \frac{g_s^+ + \mathcal{Q} - z}{g_s^+ - g_s + 2\mathcal{Q}} \tag{5.2}$$

and

$$\mathcal{Q} = (z^2 - g_s g_s^+)^{1/2}.$$

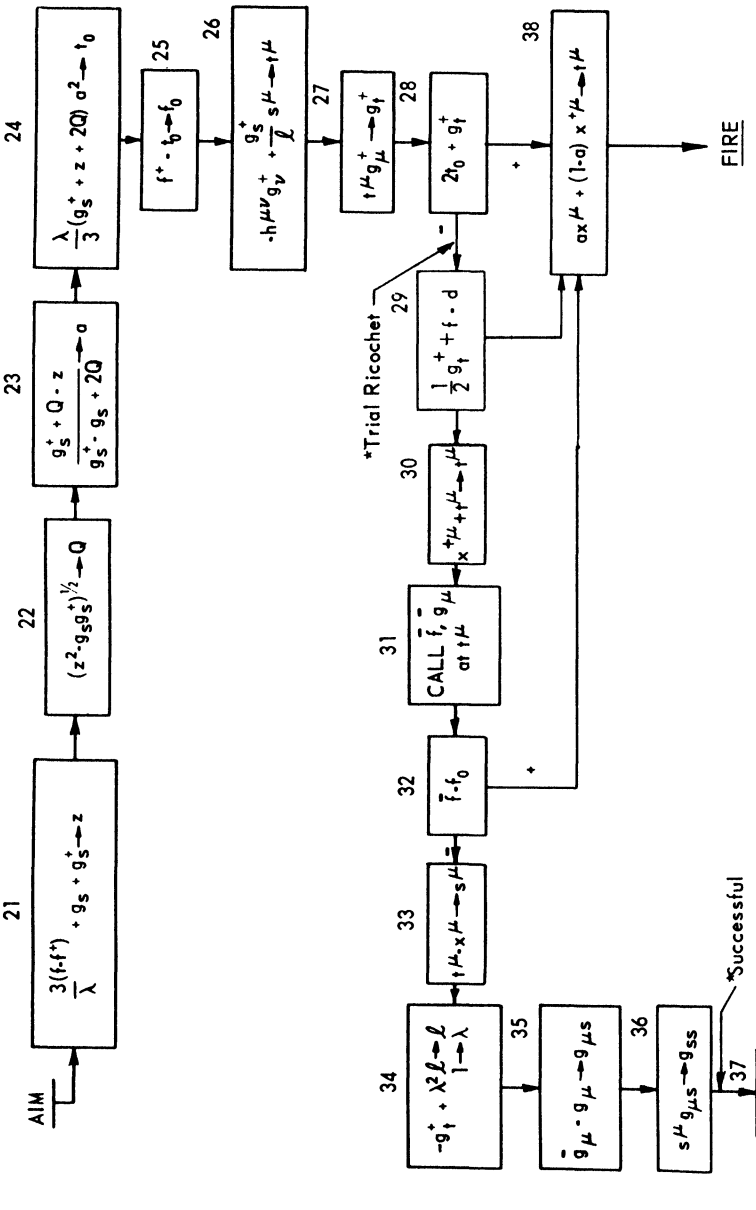


CHART 2: *Aim.*

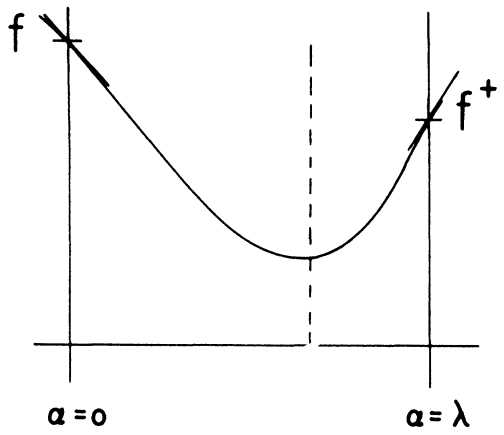


FIG. 2. *Plot of $f(\mathbf{x})$ along a one-dimensional interval.*

The particular form of (5.2) is chosen to obtain maximum accuracy, which might otherwise be lost in taking the difference of nearly equal quantities. The amount by which the minimum in f is expected to fall below f^+ is given by

$$\int_{(\lambda-\mathbf{a}\lambda)}^{\lambda} d\alpha\, g_s(\alpha) = \frac{1}{3}(g_s^+ + z + 2\mathcal{Q})\, a^2\lambda. \tag{5.3}$$

The anticipated change is now compared with what would be expected from a perpendicular step. On the basis of the metric $h^{\mu\nu}$, the step to the optimum point in the $(N-1)$-dimensional surface perpendicular to s^μ through $x^{+\mu}$ is given by

$$-h^{\mu\nu}g_\nu^+ + \frac{g_s^+}{\ell}\, s^\mu \to t^\mu. \tag{5.4}$$

The change in f to be expected from this step is $\frac{1}{2}t^\mu g_\mu^+$. Hence, the decision whether to interpolate for the minimum along **s** or to change **x** by use of (5.4) is made by comparing $g_t^+ = t^\mu g_\mu^+$ with expression (5.3).

The purpose of allowing for this option is to improve the speed of convergence when the function is not quadratic. Consider the situation of Fig. 3. The optimum point between **x** and $\mathbf{x}^+$ is point *A*. However, by going to point *B*, a greater improvement can be made in the function. When the behavior of the function is described by a curving valley, this option is of particular value, for it enables successive iterations to proceed around the curve without backtracking to the local minimum along each step. However, if evaluation of the function at this new position does not give a better value than that expected from the interpolation, then the interpolated position is used. Should the new position be better as expected, it is then desired to modify $h^{\mu\nu}$ to incorporate the new information obtained about the function. The full step taken is stored at s^μ, and its squared length is the sum of the squares of the step along **s** and the perpendicular step, i.e., $s^\mu = -g_t^+ + \lambda^2\ell$. The change in the gradient resulting from this step is stored at $g_{\mu s}$ and these quantities are used in § 7 in a manner to be described.

For the interpolated step, we set

$$ax^\mu + (1-a)x^{+\mu} \to t^\mu. \tag{5.5}$$

By direct use of the x^μ instead of the s^μ, greater accuracy is obtained in the event that a is small. After making this interpolation, we proceed to "Fire."

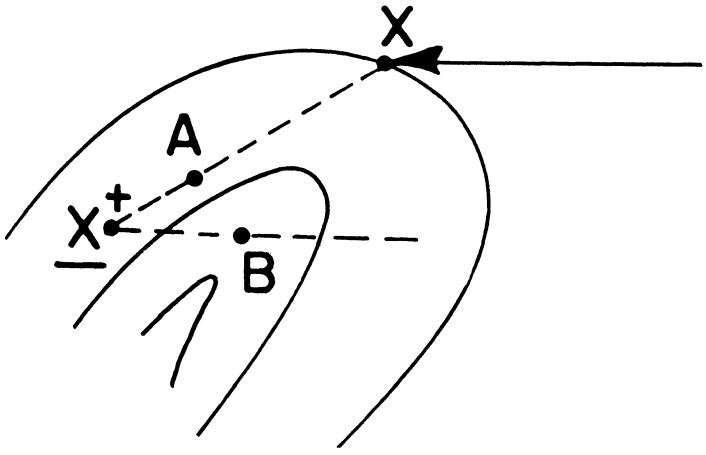


FIG. 3. *Illustration of procedure for nonquadratic functions. Point A is the optimum point along* (x, x^+); *point B is the location for the new trial.*

6. Fire: Chart 3. The purposes of this section are to evaluate the function and its gradient at the interpolated point and to determine if the local minimum has been sufficiently well located. If so, then the rate of change of gradient is evaluated (or, more accurately, λ times the rate of change) by interpolating from its values at $\mathbf{x}$, $\mathbf{x}^+$, and at the interpolated point.

If the function were cubic, then f at the interpolated point would be a minimum, the component of the gradient at this point along $\mathbf{s}$ would be zero, and the second derivative of the function at the minimum along the line would be $2\mathcal{Q}/\lambda$. However, as the function will generally be more complicated, none of these properties of f and its derivatives at the interpolated point will be exactly satisfied. We designate the actual value of f and its gradient at the interpolated point by $\bar{f}$ and $\bar{g}_\mu$. The component of $\bar{g}_\mu$ along s is $s^\mu \bar{g}_\mu = \bar{g}_s$. Should $\bar{f}$ be greater than f or f^+ by a significant amount $(>\varepsilon)$, the interpolation is not considered satisfactory and a new one is made within that part of the original interval for which f at the end point is smaller.

From the values of the gradient g_μ, $\bar{g}_\mu$, and g^+_μ at three points along a line, we can interpolate to obtain its rate of change at the interpolated point. With a quadratic interpolation for the gradient, we obtain

$$(\bar{g}_\mu - g_\mu)\frac{a}{1-a} + (g^+_\mu - \bar{g}_\mu)\frac{1-a}{a} \to g_{\mu s}, \tag{6.1}$$

where $g_{\mu s}/\lambda$ is the rate of change of the gradient at the interpolated point. The component of $g_{\mu s}$ in the direction of s, namely, $s^\mu g_{\mu s} = g_{ss}$, can be expressed as

$$\bar{g}_s\left(\frac{a}{1-a} - \frac{1-a}{a}\right) + 2\mathcal{Q} \to g_{ss}. \tag{6.2}$$

If the interpolated point were a minimum, then $\bar{g}_s = 0$ and $g_{ss} = 2\mathcal{Q}$.

An additional criterion imposed upon the interpolation is that the first term on the left of (6.2) be smaller in magnitude than $\mathcal{Q}$. Among other things, this insures that the interpolated value for the second derivative is positive. If this criterion is not fulfilled, no interpolation is made, and the matrix $h^{\mu\nu}$ is changed in a less sophisticated manner.

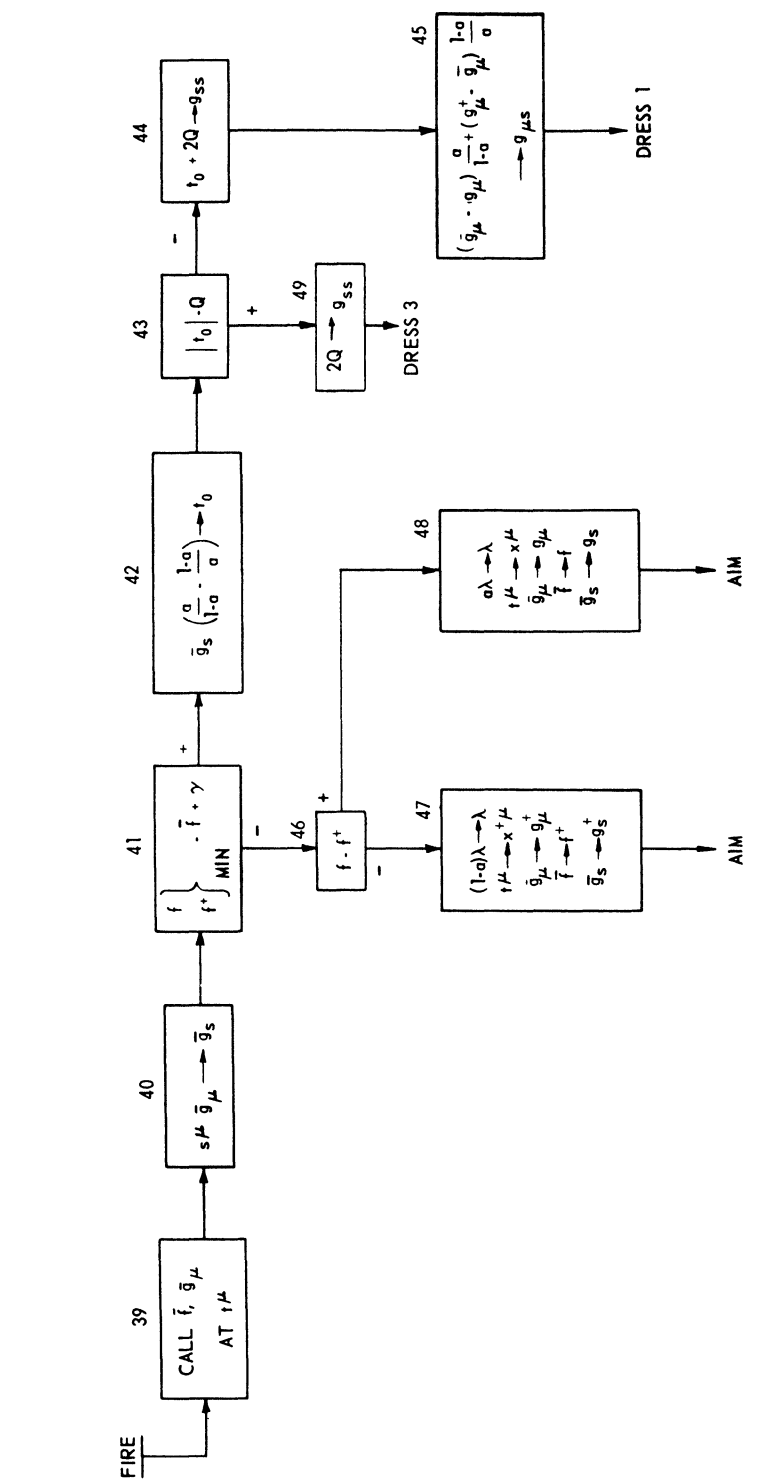


CHART 3: *Fire.*

7. Dress: Chart 4. The purpose of this section is to modify the metric $h^{\mu\nu}$ on the basis of information obtained about the function along the direction **s**. The new $h^{\mu\nu}$ is to have the property that $(h^{\mu\nu})'g_{\nu s} = \lambda s^{\mu}$, and must retain the information which the preceding iterations had given about the function.

If the vector $h^{\mu\nu}g_{\nu s} = t^{\mu}$ were in the direction of $\mathbf{s}^{\mu}$, then it would be sufficient to add to $h^{\mu\nu}$ a matrix proportional to $s^{\mu}s^{\nu}$. If t^{μ} is not in the direction of s^{μ}, the smallest squared length for the difference between s^{μ} and $(h^{\mu\nu} + \alpha s^{\mu}s^{\nu})g_{\nu s}$ is obtained when $\alpha = (\lambda / g_{ss}) - (1/\ell)$. For this value of α, the squared length of the difference is $t_0 - (g_{ss}/\ell)$ where t_0 is the square length of **d**, namely, $h^{\mu\nu}d_{\mu}d_{\nu}$. When this quantity is sufficiently small ($<\varepsilon$), the matrix $h^{\mu\nu}$ undergoes the change:

$$h^{\mu\nu} + \left(\frac{\lambda}{g_{ss}} - \frac{1}{\ell}\right) s^{\mu}s^{\nu} \to h^{\mu\nu}. \tag{7.1}$$

The corresponding change in the determinant of $h^{\mu\nu}$ is

$$\frac{\lambda \ell}{g_{ss}} \Delta \to \Delta. \tag{7.2}$$

When the vectors t^{μ} and s^{μ} are not sufficiently colinear, it is necessary to modify $h^{\mu\nu}$ by a matrix of rank two instead of one, i.e.,

$$h^{\mu\nu} - \frac{t^{\mu}t^{\nu}}{t_0} + \frac{\lambda}{g_{ss}} s^{\mu}s^{\nu} \to h^{\mu\nu}. \tag{7.3}$$

Then the change in the determinant of $h^{\mu\nu}$ is

$$\frac{\lambda g_{ss}}{t_0} \Delta \to \Delta. \tag{7.4}$$

After the matrix is changed, the iteration is complete; after printing out whatever information is desired about this part of the calculation, a new iteration is begun. This is repeated until the function is minimized to within the accuracy required.

8. Stuff: Chart 5. The purposes of this section are to test how well the function has been minimized and to test how well the matrix $h^{\mu\nu}$ approximates

$$\left\| \frac{\partial^2 f}{\partial x^{\mu}\, \partial x^{\nu}} \right\|$$

at the minimum. This is done by displacing point **x** from the location of the minimum in a random direction.

The displacement of point **x** is chosen to be a unit length in terms of $h^{\mu\nu}$ as the metric. When

$$h^{\mu\nu} = \left\| \frac{\partial^2 f}{\partial x^{\mu}\, \partial x^{\nu}} \right\|^{-1},$$

such a step will increase f by half the square of the length of the step.

If the direction were to be randomly distributed, then it would not be satisfactory to choose the range of each component of t_{μ} independently; rather, the range for the t_{μ} should be such that $h^{\mu\nu}t_{\mu}t_{\nu}$ is bounded by preassigned values. However, this refinement has not been incorporated into the charts nor the computer program. The length of the step has been chosen equal to one so that the function should increase by $\frac{1}{2}$ when each random step is taken.

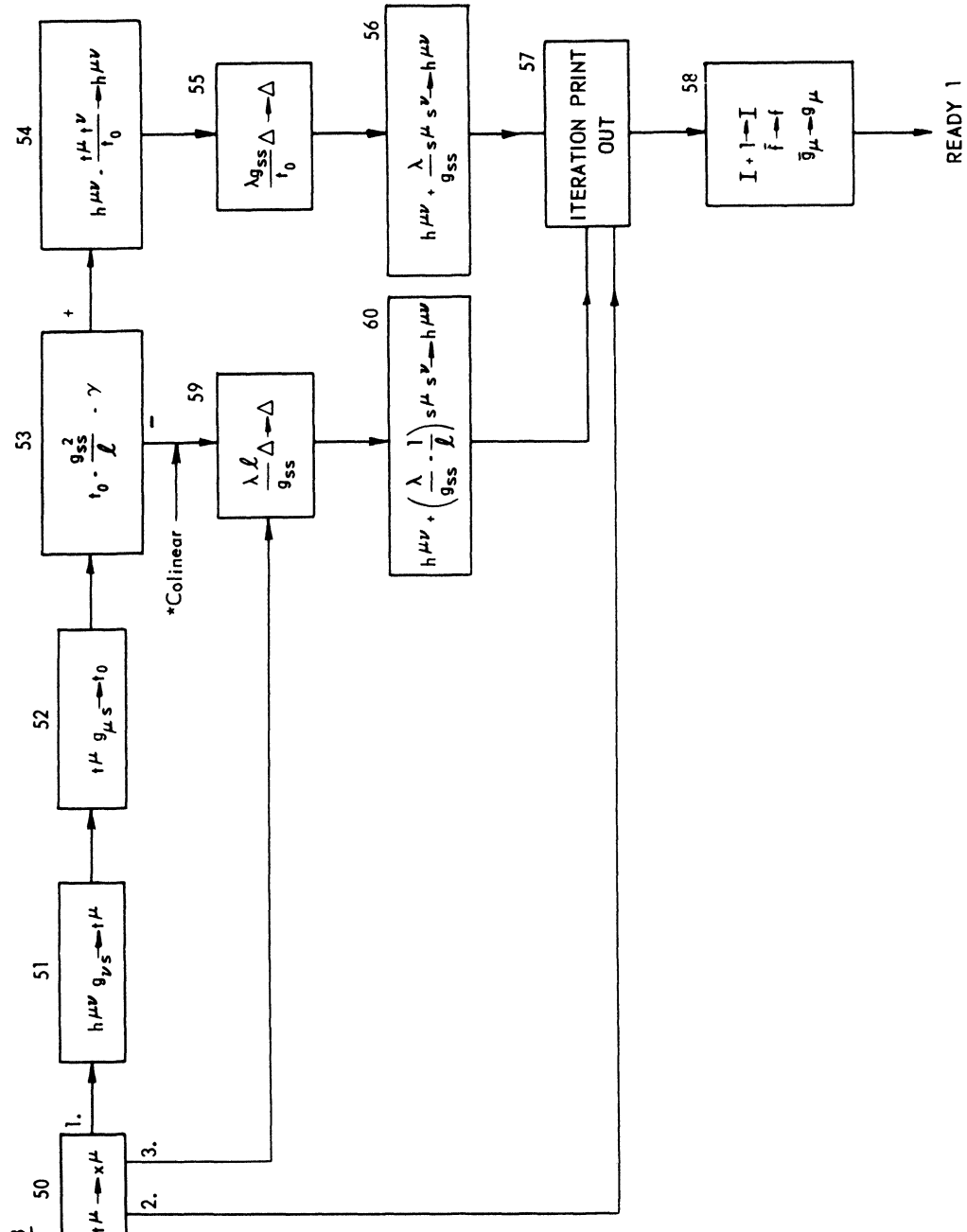


CHART 4: *Dress.*

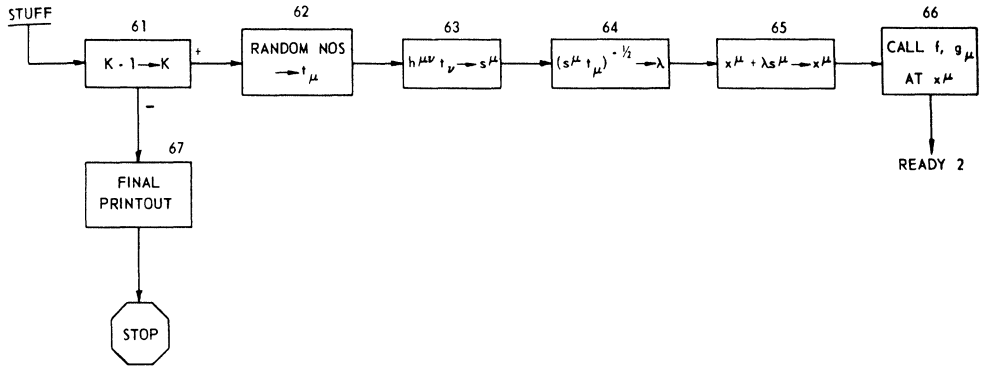


CHART 5: *Stuff.*

Significance of $h^{\mu\nu}$. We examine a least-squares analysis to illustrate how the initial trial value for $h^{\mu\nu}$ is chosen, and what its final value signifies. In this case, the function to be minimized will be chosen to be $\chi^2/2$, where χ^2 is the statistical measure of goodness of fit. The function $\chi^2/2$ is the natural logarithm of the relative probability for having obtained the observed set of data as a function of the variables χ^μ being determined.

The matrix

$$h^{\mu\nu} = \left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|^{-1}$$

then specifies the spreads and correlations among the variables by

$$\langle \Delta x^\mu \Delta x^\nu \rangle = \frac{\int d^N x (x^\mu - \langle x^\mu \rangle)(x^\nu - \langle x^\nu \rangle)\, e^{-\chi^2/2}}{\int d^N x \, e^{-\chi^2/2}} \approx h^{\mu\nu}. \tag{8.1}$$

The diagonal elements of $h^{\mu\nu}$ give the mean-square uncertainty for each of the variables, while the off-diagonal elements determine the correlations among them. The full significance of this matrix (the error matrix) is to be found in various works on statistics (see, for example, [6]). It enables us to determine the uncertainty in any linear function of the variables, for, if $u = a_\mu x^\mu$, then

$$\begin{aligned} \langle u \rangle &= a_\mu \langle x^\mu \rangle \\ \langle (\Delta u)^2 \rangle &= a_\mu a_\nu (\langle x^\mu x^\nu \rangle - \langle x^\mu \rangle \langle x^\nu \rangle) \\ &= a_\mu a_\nu h^{\mu\nu}. \end{aligned} \tag{8.2a}$$

If u is a more general function of $\mathbf{x}$, then if, in a Taylor expansion about the value of $\mathbf{x}$, derivatives higher than first can be ignored, we have

$$\begin{aligned} \langle u(\mathbf{x}) \rangle &= u(\langle \mathbf{x} \rangle) \\ \langle \Delta u(\mathbf{x}) \rangle^2 &= \frac{\partial u}{\partial x^\mu}(\langle \mathbf{x} \rangle) \frac{\partial u}{\partial x^\nu}(\langle \mathbf{x} \rangle) h^{\mu\nu}. \end{aligned} \tag{8.2b}$$

If it is possible to estimate the accuracy with which the variables are determined, the use of such estimates in the initial trial value of $h^{\mu\nu}$ will speed the convergence of the minimization procedure. Suppose, for example, that to fit some set of experi-

mental data, it is estimated that the variables x^μ have the values:

$$\begin{aligned} x^1 &= 3.0 \pm 0.1 \\ x^2 &= 28.0 \pm 2 \\ x^3 &= 10^4 \pm 10^2. \end{aligned} \tag{8.3}$$

Then, the initial values for x^μ and $h^{\mu\nu}$ would be

$$\begin{aligned} x^\mu &= (3.0 \quad 28.0 \quad 10^4) \\ h^{\mu\nu} &= \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 10^4 \end{pmatrix}. \end{aligned} \tag{8.4}$$

If this estimate is even correct to within a couple of orders of magnitude, the number of iterations required to locate the minimum may be substantially fewer than that for some more arbitrary choice, such as the unit matrix.

If it is desired to impose linear constraints on the variables, this can be readily done by starting with a matrix $h^{\mu\nu}$, which is no longer positive definite, but which has zero eigenvalues. For the constraints

$$\begin{aligned} a_\mu x^\mu &= \alpha \\ b_\mu x^\mu &= \beta, \end{aligned} \tag{8.5}$$

etc., the matrix $h^{\mu\nu}$ must be chosen so that

$$\begin{aligned} h^{\mu\nu} a_\nu &= 0 \\ h^{\mu\nu} b_\nu &= 0, \end{aligned} \tag{8.6}$$

and the starting value for x^μ must satisfy (8.5). For example, if x^3 is to be held constant, all elements of $h^{\mu\nu}$ in the third row and third column are set equal to zero and x^3 is set equal to the constant value.

When constraints are imposed instead of setting Δ equal to the determinant of $h^{\mu\nu}$ $(=0)$, it is set equal to the product of the nonzero eigenvalue of $h^{\mu\nu}$. Then, except for roundoff errors, not only will the conditions (8.6) be preserved in subsequent iterations, but also Δ will continue to equal the product of nonzero eigenvalues.

Though Δ is not used in the calculations, its value may be of interest in estimating how well the variables have been determined, since $\sum_\mu h^{\mu\mu}$ gives the sum of the eigenvalues of $h^{\mu\nu}$, while Δ gives their product. The square root of each of these eigenvalues is equal to one of the principal semiaxes of the ellipse formed by all $\mathbf{x}$ for which $f(\mathbf{x})$ exceeds its minimum value by $\frac{1}{2}$.

9. Conclusion. The minimization method described has been coded for the IBM-704 using Fortran. Experience is now being gathered on the operation of the method with diverse types of functions. Parts of the procedure, not incorporating all of the provisions described here, have been in use for some time in least-squares calculations for such computations as the analysis of $\pi - P$ scattering experiments [10], for the analysis of delayed neutron experiments [7], and similar computations. Though full mathematical analysis of its stability and convergence has not been made, general considerations and numerical experience with it indicate that minima of functions can be generally more quickly located than in alternate procedures. The ability of the

metric, $h^{\mu\nu}$, to accumulate information about the function and to compensate for ill-conditioned $g_{\mu\nu}$ is the primary reason for this advantage.

10. Acknowledgment. The author wishes to thank Dr. G. Perlow and Dr. M. Peshkin for valued discussions and suggestions, and Mr. K. Hillstrom for carrying out the computer programming and operation.

Appendix.[1] If we have the gradient of the function at a point in the neighborhood of a minimum together with $\mathbf{G}^{-1}$, where

$$\mathbf{G} = \left\| \frac{\partial^2 f}{\partial x^\mu \, \partial x^\nu} \right\|,$$

then, neglecting terms of higher order, the location of the minimum would be given in matrix notation by

$$\xi = x - \mathbf{G}^{-1}\nabla. \tag{1}$$

In the method to be described, a trial matrix is used for $\mathbf{G}^{-1}$ and a step determined by (1) is taken. From the change in the gradient resulting from this step, the trial value is improved and this procedure is repeated. The changes made in the trial value for $\mathbf{G}^{-1}$ are restricted to keep the hunting procedure "reasonable" regardless of the nature of the function. Let $\mathbf{H}$ be the trial value for $\mathbf{G}^{-1}$. Then the step taken will be to the point

$$x^+ = x - \mathbf{H}\nabla. \tag{2}$$

The gradient at x^+, ∇^+, is then evaluated. Let $D = \nabla^+ - \nabla$ be the change in the gradient as a result of the step $S = x^+ - x = -H\nabla$. We form the new trial matrix by

$$H^+_{\mu\nu} = H_{\mu\nu} + a(\mathbf{H}\nabla^+)_\mu (\mathbf{H}\nabla^+)_\nu. \tag{3}$$

The constant a is determined by the following two conditions:

1. The ratio of the determinant of $\mathbf{H}^+$ to that of $\mathbf{H}$ should be between R^{-1} and R, where R is a preassigned constant greater than 1. This is to prevent undue changes in the trial matrix and, in particular, if $\mathbf{H}$ is positive definite, $\mathbf{H}^+$ will be positive definite also.
2. The nonnegative quantity

$$\Delta = D\mathbf{H}^+D + S(\mathbf{H}^+)^{-1}S - 2S \cdot D \tag{4}$$

is to be minimized. This quantity vanishes when $S = \mathbf{H}^+D$. The a which satisfies these requirements, together with the corresponding Δ, as functions of $N = \nabla^+ H \nabla^+$ and $M = \nabla^+ H \nabla$, are as follows:[2]

(5)

Range of M	a	Δ
$M < -N/(R-1)$	$1/(M-N)$	0
$-N/(R-1) < M < N/(R+1)$	$(1/RN) - (1/N)$	$(N - M + MR)^2/RN$
$N/(R+1) < M < NR/(R+1)$	$(N-2M)/N(M-N)$	$4M(N-M)/N$
$NR/(R+1) < M < NR/(R-1)$	$(R/N) - (1/N)$	$(M + NR - MR)^2/RN$
$NR/(R-1) < M$	$1/(M-N)$	0

The dependence of Δ on M is bell-shaped, symmetric about a maximum at $M = N/2$, for which $a = 0$ and $\Delta = N$.

[1] The following method is a description of a simplified method embodying some of the ideas of the procedure presented in this report.

[2] When the function is known to be quadratic, the first condition can be dispersed with, in which case $a = (M-N)^{-1}$, $\Delta = 0$.

After forming the new trial matrix $\mathbf{H}^+$, the next step is taken in accordance with (2) and the process repeated, provided that $N = \nabla^+ H \nabla^+$ is greater than some preassigned ε. When the $\mathbf{G}$ is constant, Δ can be written as

$$\nabla = \mathbf{G}(x - \xi). \tag{6}$$

If u is an eigenvector of $\mathbf{HG}$ with eigenvalue one, then it will be an eigenvector of $\mathbf{H}^+\mathbf{G}$ with eigenvalue one as well, since

$$\begin{aligned} \mathbf{H}^+\mathbf{G}u &= \mathbf{HG}u + a\mathbf{H}\nabla^+(\nabla^+\mathbf{HG}u) \\ &= u + a\mathbf{H}\nabla^+[\nabla\mathbf{HG}(1-\mathbf{HG})u] \\ &= u. \end{aligned} \tag{7}$$

Furthermore, when $\Delta = 0$,

$$\mathbf{H}^+\mathbf{G}S = \mathbf{H}^+ D = S, \tag{8}$$

so that S becomes another such eigenvector. After no more than N steps (for which $\Delta = 0$), $\mathbf{H}$ will equal $\mathbf{G}^{-1}$ and the following step will be to the exact minimum.

The entire procedure is covariant under an arbitrary linear coordinate transformation. Under these transformations of x, ∇ transforms as a covariant vector, $\mathbf{G}$ transforms as a covariant tensor of second rank, and $\mathbf{H}$ transforms as a contravariant tensor of second rank. The intrinsic characteristics of a particular hunting calculation are determined by the eigenvalues of the mixed tensor $\mathbf{HG}$, and the components of the initial value of $(x - \xi)$ along the direction of the corresponding eigenvectors. Since successive steps will bring $\mathbf{HG}$ closer to unity, convergence will be rapidly accelerating even when $\mathbf{G}$ itself is ill-conditioned. Constraints of the form $b \cdot x = c$ can be improved by using an initial $\mathbf{H}$ which annuls b, i.e.,

$$\mathbf{H} \cdot b = 0,$$

and choosing the initial vector $\mathbf{x}$ such that it satisfies $b \cdot x = c$. Then all steps taken will be perpendicular to b and this inner product will be conserved. For example, if it is desired to hold one component of x constant, all the elements of $\mathbf{H}$ corresponding to that component are initially set equal to zero.

REFERENCES

[1] A. CAUCHY, *Méthode générale pour la résolution des systèmes d'équations simultanées*, Compt. Rend., 25, 536 (1847).

[2] M. R. HESTENES AND C. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409-436.

[3] F. B. HILDEBRAND, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956.

[4] W. A. NIERENBERG, Report UCRL-3816, University of California Radiation Laboratory, Berkeley, CA, 1957.

[5] R. L. GARWIN AND H. A. REICH, *An efficient iterative least squares method* (to be published).

[6] H. CRAMER, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, NJ, 1946.

[7] G. J. PERLOW AND A. F. STEHNEY, *Halogen delayed-neutron activities*, Phys. Rev., 113 (1959), pp. 1269-1276.

[8] E. FERMI AND N. METROPOLIS, Los Alamos unclassified report LA-1492, Los Alamos National Laboratory, Los Alamos, NM, 1952.

[9] H. L. ANDERSON, W. C. DAVIDON, M. G. GLICKSMAN, AND U. E. KRUSE, *Scattering of positive pions by hydrogen at* 189 *MeV*, Phys. Rev., 100 (1955), pp. 279-287.

[10] H. L. ANDERSON AND W. C. DAVIDON, *Machine analysis of pion scattering by the maximum likelihood method*, Nuovo Cimento, 5 (1957), pp. 1238-1255.

[11] R. FLETCHER AND M. J. D. POWELL, *A rapidly convergent descent method for minimization*, Comput. J., 6 (1963), pp. 163-168.