JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 4, No. 5, 1969

SURVEY PAPER

Multiplier and Gradient Methods[1]

MAGNUS R. HESTENES[2]

Abstract. The main purpose of this paper is to suggest a method for finding the minimum of a function $f(x)$ subject to the constraint $g(x) = 0$. The method consists of replacing f by $F = f + \lambda g + \frac{1}{2}cg^2$, where c is a suitably large constant, and computing the appropriate value of the Lagrange multiplier. Only the simplest algorithm is presented. The remaining part of the paper is devoted to a survey of known methods for finding unconstrained minima, with special emphasis on the various gradient techniques that are available. This includes Newton's method and the method of conjugate gradients.

1. Introduction

About twenty years ago, the author became interested in computational methods for optimal control problems (Ref. 1). This interest was stimulated by an attempt to compute the time-optimal path of an airplane from take-off to level flight at a prescribed position and velocity. At that time, large-scale digital-computing machines were not available. Computing had to be carried out by analog computers or by mechanical desk computers. In Ref. 2, an attempt was made to compute the time-optimal path for an airplane by integrating the corresponding Euler–Lagrange equations on an analog computer (REAC). However, the differential equations were unstable and the results were unsatisfactory. However, a good estimate could be found by

[1] Paper received February 2, 1969. The preparation of this paper was sponsored by the U.S. Army Research Office, Grant No. DA-31-124-ARO(D)-355. This paper was presented at the Second International Conference on Computing Methods in Optimization Problems, San Remo, Italy, 1968.

[2] Professor, Department of Mathematics, University of California at Los Angeles, Los Angeles, California.

809/4/5-1

hand computation using special properties of the problem. This experience convinced the author that general procedures should be devised for obtaining solutions or for improving estimates of solutions. Accordingly, the author experimented with three methods, namely, Newton's method, the gradient method, and the method of penalty functions. Since he was restricted to the use of hand computation, the author considered only simple variational problems which possessed nonminimizing as well as minimizing extremals. It was found that Newton's method and the gradient method were very effective (Refs. 3–5).

The author, however, had difficulties with the method of penalty functions because of round-off errors. To obtain any accuracy to the solution of the problem considered required carrying more significant figures than were convenient in hand computation. Although the method of penalty functions has been used with reasonable success in recent years, the author has always felt that an improvement of the method could be made. The purpose of this paper is to suggest a modification of the method of penalty functions which we shall call the method of multipliers. In addition, we shall make some remarks concerning Newton's method, the method of gradients, and conjugate gradients that may be useful.

2. Constrained and Unconstrained Minima

Before describing the method of multipliers, it is instructive to recall a connection between constrained and unconstrained minima upon which the method is based. We shall consider only the simplest case, in which a point x_0 affords a minimum to a real-valued function $f(x) = f(x^1,..., x^m)$ subject to a single constraint

$$g(x) = 0 \tag{1}$$

The extension to the case in which g is vector-valued is immediate. We assume that f and g are of class C'' and that the gradient

$$g'(x) = (\partial g(x)/\partial x^i)$$

of g is not zero at x_0. Then, there exists a multiplier λ such that, if we set $G = f + \lambda g$, we have

$$G'(x_0) = 0, \qquad g(x_0) = 0 \tag{2}$$

$$G''(x_0, h) = \sum_{i,j=1}^{m} (\partial^2 G/\partial x^i\, \partial x^j)\, h^i h^j \geqslant 0 \tag{3}$$

for all $h \neq 0$ such that

$$g'(x_0, h) = \sum_{i=1}^{m} (\partial g(x_0)/\partial x^i)\, h^i = g'(x_0) \cdot h = 0 \tag{4}$$

Here, $g'(x_0, h)$ is the first differential of g at x_0 and is the Cartesian inner product of $g'(x_0)$ and h. Similarly, $G''(x_0, h)$ is the second differential of G at x_0. The point x_0 is said to be *nonsingular* in case

$$\begin{vmatrix} \partial^2 G/\partial x^i\, \partial x^j & \partial g/\partial x^i \\ \partial g/\partial x^j & 0 \end{vmatrix} \neq 0$$

at $x = x_0$. If x_0 is a nonsingular minimum point for f subject to $g = 0$, the equality in (3) holds only if $h = 0$. This implies the existence of a positive number c such that

$$G''(x_0, h) + c[g'(x_0, h)]^2 > 0$$

for all $h \neq 0$. Setting

$$F = f + \lambda g + \tfrac{1}{2} c g^2 = G + \tfrac{1}{2} c g^2$$

it is seen that, at $x = x_0$, we have

$$F'(x_0) = G'(x_0) = 0$$

$$F''(x_0, h) = G''(x_0, h) + c[g'(x_0, h)]^2 > 0, \qquad h \neq 0$$

Here, we have used the fact that $g(x_0) = 0$. In view of these relations, we see that x_0 affords an unconstrained local minimum to F. This yields the following theorem:

Theorem 2.1. If x_0 is a nonsingular minimum point of f subject to $g = 0$, there exists a multiplier λ and a constant c such that x_0 affords an unconstrained local minimum to the function

$$F = f + \lambda g + \tfrac{1}{2} c g^2$$

Conversely, if $g(x_0) = 0$ and x_0 affords a minimum to a function F of this type, then x_0 affords a minimum to f subject to $g = 0$.

3. Method of Penalty Functions and Method of Multipliers

One of the popular methods of finding a constrained minimum point is the method of penalty functions. For the problem considered in the last section, this method seeks a minimum point x_n of the function

$$f_n(x) = f(x) + \tfrac{1}{2} n g^2(x)$$

A limit point of the sequence $\{x_n\}$, if it exists, is then the solution x_0 to our problem. Moreover, inasmuch as

$$0 = f_n'(x_n) = f'(x_n) + ng(x_n)\, g'(x_n)$$

it is seen that, if $g'(x_0) \neq 0$, then $\lambda_n = ng(x_n)$ converges to the corresponding multiplier λ for x_0. It is a simple matter to formulate conditions which ensure the existence of a minimum point x_n for $f_n(x)$ on an open set S and the convergence of the sequence $\{x_n\}$ to a point x_0 in S which minimizes f on S subject to $g = 0$. Observe that

$$f_n(x_n) = f(x_n) + \tfrac{1}{2}ng^2(x_n) \leqslant f_n(x_0) = f(x_0) \tag{5}$$

If x_0 is a nonsingular solution, there are, by Theorem 2.1, a constant c, a multiplier λ, and a neighborhood N of x_0 such that

$$f(x_0) \leqslant f(x_n) + \lambda g(x_n) + \tfrac{1}{2}cg^2(x_n) \tag{6}$$

whenever x_n is in N. Combining (5) and (6), we see that

$$(n - c)\, g^2(x_n) \leqslant 2\lambda g(x_n)$$

whenever x_n is in N. In the event that $f'(x_0) = 0$, we have $\lambda = 0$, since $g'(x_0) \neq 0$, by virtue of the nonsingularity of x_0. In this event, $x_n = x_0$ if $n > c$ and x_n is in N. Thus, the method of penalty functions is very effective whenever $f'(x_0) = 0$ and should be reasonably effective when $f'(x_0)$ is near zero. However, in general, this is not the case, and the method becomes sensitive to round-off errors in the term $ng^2(x)$. For large values of n, it is difficult to obtain an appropriate numerical approximation of x_n and, hence, of x_0.

In order to circumvent the numerical difficulty that may arise in the method of penalty functions, the author suggests a simple modification. This modified method is based on the theorem stated in the last section and will be called the method of multipliers. In this method, we select a positive constant c and consider the function

$$F(x, \lambda) = f(x) + \lambda g(x) + \tfrac{1}{2}cg^2(x)$$

The constant c, if chosen suitably large, is held fast. Let λ_1 be an initial estimate of λ, and select a minimum point x_1 of $F(x, \lambda_1)$. In general, having obtained an estimate λ_n of λ, we select x_n to minimize $F(x, \lambda_n)$. Observe that

$$F'(x_n, \lambda_n) = f'(x_n) + (\lambda_n + cg(x_n))\, g'(x_n) = 0$$

This suggests that we select

$$\lambda_{n+1} = \lambda_n + c_n g(x_n)$$

where $0 < c_n \leqslant c$. Various rules can be given for selecting c_n. For example, we can choose $c_n = \gamma c$, where γ is a positive constant, normally $\leqslant 1$. Or we can choose c_n such that $g(x_n)\, g(x_{n+1}) > 0$. It is not difficult to give criteria which ensure the convergence of the method for problems of this type.

In order to illustrate this method, consider the special case in which

$$f(x) = \tfrac{1}{2} x^* A x - \alpha b^* x, \qquad g(x) = b^* x$$

where A is a nonsingular symmetric matrix, b and x are column vectors with $b \neq 0$, b^* is the transpose of b, and α is a positive number. We assume that $x^* A x > 0$ for all $x \neq 0$ such that $g(x) = b^* x = 0$. The point $x = 0$ minimizes f subject to $g = 0$, and $\lambda = \alpha$ is the corresponding Lagrange multiplier. Select c such that

$$x^* A x + c(b^* x)^2 > 0$$

for all $x \neq 0$ and set

$$F(x, \lambda) = f + \lambda g + \tfrac{1}{2} c g^2$$

The minimum point $\bar{x}$ of $F(x, \lambda)$ solves the equation

$$Ax - \alpha b + (\lambda + c b^* x)\, b = 0$$

The solution takes the form

$$\bar{x} = \gamma A^{-1} b, \qquad \gamma = (\alpha - \lambda)/(1 + c\beta), \qquad \beta = b^* A^{-1} b$$

If one uses the method of multipliers with $\lambda_1 = 0$ and $\lambda_{n+1} = \lambda_n + c g(x_n)$, it is found that

$$x_n = [\alpha/(1 + c\beta)^n]\, A^{-1} b, \qquad \lambda_{n+1} = \alpha - [\alpha/(1 + c\beta)^n]$$

Convergence is obtained if $| 1 + c\beta | > 1$. If the method of penalty functions is used, we have

$$x_n = [\alpha/(1 + n\beta)]\, A^{-1} b, \qquad n g(x_n) = \alpha n \beta/(1 + n\beta)$$

Of course, one would not use the method of multipliers or penalty functions for this problem, since the solution is easily obtained by the Lagrange mul-

tiplier rule. However, this example gives some indication of the nature of the two methods in the general case, provided that we are close to a solution.

The reader will find it instructive to consider the two-dimensional cases in which

$$f(x, y) = x^2 - y^2 - y, \qquad g(x, y) = y \qquad \text{or} \qquad g(x, y) = y + y^3$$

The method of multipliers has the advantage that the coefficient of g^2 need not be very large. It has the disadvantage that one needs to have an initial estimate of the multiplier λ. Perhaps, a combination of the method of multipliers and the penalty-function method would be most effective. Begin with the method of penalty functions so as to obtain an initial estimate of x_0 and, from this, deduce an estimate of λ. Then, switch to the method of multipliers. The author has not had time to experiment with the method of multipliers for general functions, but plans to do so in the near future.

4. Extensions to Variational and Optimal Control Problems

The method of penalty functions has been applied to a large class of variational problems. It has been used, for example, to eliminate terminal constraints or isoperimetric constraints. More recently, it has been used by Balakrishnan (Ref. 6) to eliminate dynamic constraints. In each case, one can modify the method of penalty functions to obtain a corresponding method of multipliers. We shall give a heuristic description of how to formulate the method of multipliers for the simple nonlinear optimal control problem with fixed terminal and initial states.

We consider the following problem: Let ξ denote a pair

$$\xi: \qquad x^i(t), \quad u^k(t), \qquad 0 \leqslant t \leqslant T$$

of state functions $x^i(t)$ and control functions $u^k(t)$ on a fixed interval $0 \leqslant t \leqslant T$. We consider ξ to be an arc in txu-space. We make the usual continuity and differentiability assumptions on $x(t)$ and $u(t)$. The class of arcs ξ whose elements $(t, x(t), u(t))$ lie in a prescribed region in txu-space is denoted by $\mathscr{U}$. We denote by $\mathscr{B}$ the subclass of $\mathscr{U}$ having prescribed initial and terminal states $x(0)$ and $x(T)$. Finally, we denote by $\mathscr{C}$ the class of all arcs ξ in $\mathscr{B}$ satisfying the differential constraint

$$\dot{x}^i = f^i(t, x, u)$$

We suppose that there is a unique arc

$$\xi_0 : \quad x_0(t), \quad u_0(t), \quad 0 \leqslant t \leqslant T$$

in $\mathscr{C}$ that minimizes a given integral

$$I(\xi) = \int_0^T L(t, x(t), u(t))\, dt$$

This problem becomes a classical problem of Lagrange with one variable endpoint if we introduce the auxiliary state variable

$$y(t) = \int_0^t u(s)\, ds, \quad 0 \leqslant t \leqslant T$$

so that $\dot{y}(t) = u(t)$.

It has been shown by the author (Ref. 7) that, if the arc ξ_0 satisfies certain standard sufficiency conditions for a strong relative minimum on $\mathscr{C}$, there exist multipliers $p_i(t)$ and a function $c(t, x, \dot{x}, u)$ such that ξ_0 affords a local minimum to the function

$$J = \int_a^b \{L + p_i(\dot{x}^i - f^i) + \tfrac{1}{2}c \mid \dot{x} - f \mid^2\}\, dt$$

where i is summed over its range. The functions $p_i(t)$ are the usual costate functions associated with ξ_0. In many cases, c can be chosen to be a constant. In particular, this is the case when one is concerned only with weak relative minima.

This result suggests the following method of multipliers. Having chosen c sufficiently large, hold c fast and proceed as follows: Let $p_q(t)$ be an estimate of $p(t)$ and let $J_q(\xi)$ be the integral obtained from J by setting $p(t) = p_q(t)$. We then seek a minimum ξ_q of $J_q(\xi)$ on $\mathscr{B}$. This problem is a classical minimum problem of the type considered by Tonelli and his school. Having obtained ξ_q, determine a new set of multipliers p_{q+1} by some reasonable rule. For example, one can select

$$p_{q+1} = p_q + c(\dot{x} - f)$$

evaluated along ξ_q. Or one could let P_{q+1} be a solution along ξ_q of

$$\dot{p}_i = L_{x^i} - p_j f^j_{x^i}$$

with suitable initial or terminal conditions.

The author has not determined conditions under which this method would be effective. It is to be expected that ξ_0 exists and, if suitably strong hypotheses are made, then ξ_q exists and converges to ξ_0. In the general case, it is expected that one would have to enlarge the problem so as to include generalized curves and relaxed controls.

5. Newton's Method and Gradient Methods

In the preceding pages, it was shown that problems with constraints often can be solved with the help of solutions of problems without constraints. This section is devoted to methods for obtaining unconstrained minima. To this end, let $F(x)$ be a real-valued function on a normed linear space $\mathscr{E}$. We normally consider $\mathscr{E}$ to be Euclidian. We assume that F possesses first and second Frechet differentials $F'(x, h)$ and $F''(x, h)$. We then have the Taylor formula

$$F(x + h) = P(x, h) + R(x, h) \tag{7}$$

where

$$P(x, h) = F(x) + F'(x, h) + \tfrac{1}{2}F''(x, h) \tag{8}$$

Newton's method for obtaining the minimum of F on an open set S of $\mathscr{E}$ can be described as follows: Having obtained an estimate x_n of the minimum point $\bar{x}$ of F on S, select h_n so as to minimize the function $P(x_n, h)$ and use the formula

$$x_{n+1} = x_n + h_n$$

to obtain a new estimate of $\bar{x}$. This method converges quadratically to $\bar{x}$ under the usual hypotheses on $F''(x, h)$ if a suitable initial point x_0 is chosen. The point h_n satisfies the relation

$$F'(x_n, h) + F''(x_n, h_n, h) = 0, \qquad h \text{ arbitrary} \tag{9}$$

where $F''(x, k, h)$ is the bilinear form associated with $F''(x, h)$, namely, the differential of $F'(x, k)$ for fixed k. In the Euclidean case,

$$h_n = -K_n F'(x_n) \tag{10}$$

where K_n is the inverse of the matrix $F''(x_n)$ of second derivatives of F at x_n. We have, accordingly, the iteration

$$x_{n+1} = x_n - K_n F'(x_n) \tag{11}$$

Quadratic convergence is assured if the matrix $F''(x_n)$, and, hence, also K_n, is positive definite and the initial point x_0 is suitably chosen.

Newton's method has the following geometric interpretation: Given the point x_n, approximate the level surface

$$F(x) = F(x_n)$$

by the ellipsoid

$$P(x_n, x - x_n) = F(x_n)$$

This ellipsoid is tangent to the level surface of F at x_n. Take the center x_{n+1} of this ellipsoid as the new estimate of the minimum point $\bar{x}$.

The difficulty encountered in Newton's method lies in the determination of the minimum point h_n of $P(x_n, h)$. In the finite-dimensional case, the matrix $F''(x_n)$ must be inverted. In infinite-dimensional cases, it involves the solution of a linear boundary-value problem. For this reason, it is often desirable to replace $P(x_n, h)$ by a simpler function

$$P_n(h) = F(x_n) + F'(x_n, h) + \tfrac{1}{2}Q_n(h)$$

where $Q_n(h)$ is a positive-definite quadratic form. We then choose h_n to minimize $Q_n(h)$. The ellipsoid

$$P_n(x - x_n) = F(x_n) \tag{12}$$

is tangent to the level surface of F at x_n and its center yields the desired new estimate $x_{n+1} = x_n + h_n$ of $\bar{x}$. Again, in the finite-dimensional case, the iteration takes the form (11), where K_n is the inverse of the matrix associated with Q_n.

If one selects Q_n to be of the form

$$Q_n(x) = c_n \| x \|^2$$

the surface (12) is a sphere. In this event, $K_n = c_n^{-1}I$ and the iteration (11) is the usual gradient method. If $Q_n(h) = F''(x_n, h)$, we have Newton's method.

If K is an arbitrary positive-definite matrix, the iteration

$$x_{n+1} = x_n - a_n KF'(x_n), \qquad a_n > 0 \tag{13}$$

can be considered to be a gradient method. This iteration is obtained if we select

$$\langle x, y \rangle = (K^{-1}x, y)$$

as our inner product in place of the Cartesian inner product (x, y). In this event, $Q_n(x) = a_n^{-1}\langle x, x\rangle$. The gradient g of F relative to the new inner product $\langle x, y\rangle$ is given by the identity

$$F'(x, h) = \langle g, h\rangle \tag{14}$$

Hence, $g = KF'(x)$. The iteration (13) is equivalent to the usual gradient method in a suitably chosen coordinate system.

In the infinite-dimensional case, one should choose the inner product $\langle g, h\rangle$ such that $\langle h, h\rangle$ has the essential properties of $F''(x, h)$. For example, for the variational integral

$$F(x) = \int_a^b L(t, x(t), \dot{x}(t))\, dt$$

one should select

$$\langle g, h\rangle = g(a)\, h(a) + \int_a^b \dot{g}(t)\, \dot{h}(t)\, dt$$

or its equivalent as the inner product instead of the more familiar

$$\langle g, h\rangle = \int_a^b g(t)\, h(t)\, dt$$

The reason for this choice becomes self evident when one attempts to construct a gradient method for minimizing $F(x)$.

6. Rayleigh–Ritz and Conjugate-Gradient Methods

The conjugate-gradient method can be introduced in many ways. In this section, we show that conjugate-gradient and conjugate-direction methods are variants of the Rayleigh–Ritz method. To this end, let $\mathscr{E}$ be a finite- or infinite-dimensional, real Hilbert space with (x, y) as its inner product and $\| x \| = \sqrt{(x, x)}$ as its norm. Let A be a positive-definite, self-adjoint, bounded operator on $\mathscr{E}$. Then,

$$(Ax, y) = (x, Ay), \qquad m \| x \|^2 \leqslant (Ax, x) \leqslant M \| x \|^2$$

where m, M are suitably chosen positive constants. We seek a solution of the linear system

$$Ax = b \tag{15}$$

where b is a given element in $\mathscr{E}$. The solution $\bar{x} = A^{-1}b$ of this equation affords a minimum on $\mathscr{E}$ to the function

$$F(x) = \tfrac{1}{2}(Ax, x) - (b, x)$$

Observe that

$$F'(x, h) = (Ax - b, h) \tag{16}$$

The gradient of F at x is therefore

$$F'(x) = Ax - b \tag{17}$$

The quantity $r = -F'(x) = b - Ax$ is called the residual at x and is also called the negative gradient of F at x.

If K is a second positive-definite, bounded, self-adjoint operator on $\mathscr{E}$, then

$$\langle x, y\rangle = (K^{-1}x, y) \tag{18}$$

is a second inner product topologically equivalent to the first. The negative gradient g of F at x relative to this new inner product is defined by the relation

$$F'(x, h) = -\langle g, h\rangle$$

for all h in $\mathscr{E}$. Hence,

$$g(x) = Kr(x) = -KF'(x)$$

The generalized gradient method for solving (15) accordingly takes the form

$$x_{n+1} = x_n + a_n K r_n\,, \qquad a_n > 0 \tag{19}$$

where $r_n = -F'(x_n)$ and a_n is a suitably chosen scale factor. The choice $K = A^{-1}$ would be the ideal choice for K. However, since A^{-1} is assumed to be unknown, this choice is impossible. The conjugate-gradient methods yield iterative methods for computing A^{-1}.

Before considering the Rayleigh–Ritz method, it is convenient to recall a theorem on the minimization of F on a set $z + \mathscr{B}$, where z is a fixed point of $\mathscr{E}$, $\mathscr{B}$ is a linear subspace of $\mathscr{E}$, and $z + \mathscr{B}$ is the set of all points $x = z + y$, where y is in $\mathscr{B}$.

Theorem 6.1. A point $\bar{x} = z + \bar{y}$ in $z + \mathscr{B}$ minimizes F on $z + \mathscr{B}$ if, and only if, the negative gradient $g(\bar{x}) = -KF'(\bar{x})$ is orthogonal to $\mathscr{B}$ relative to (18) or equivalently if, and only if, the residual $r(\bar{x}) = -F'(\bar{x})$ is orthogonal to $\mathscr{B}$ in the usual sense.

In the general Rayleigh–Ritz method, we select a basis $p_0, p_1, p_2, \ldots$, for $\mathscr{E}$. The linear subspace generated by $p_0, p_1, \ldots, p_{k-1}$ is denoted by $\mathscr{B}_k$. If $\mathscr{E}$ is n-dimensional, then $\mathscr{B}_n = \mathscr{E}$. Let x_0 be a point in $\mathscr{E}$. For example, we may select $x_0 = 0$. Denote by x_k the minimum point of F on $x_0 + \mathscr{B}_k$. Then, the sequence $\{x_k\}$ converges to the desired solution $\bar{x} = A^{-1}b$. Of course, if $\mathscr{E}$ is n-dimensional, then $x_n = A^{-1}b$. At each step, x_k is an estimate of $\bar{x}$ and $F(x_k)$ is an estimate of the minimum value $F(\bar{x})$. In applications, a wise choice of basis $\{p_j\}$ often yields good estimates $F(x_k)$ of $F(\bar{x})$ for small integers k.

The *conjugate-direction method* is the special case of the Rayleigh–Ritz method in which the basis $\{p_j\}$ is chosen to be a conjugate basis, in the sense that

$$(Ap_j, p_k) = 0, \qquad j \neq k$$

The advantage of this choice is that the point x_{k+1} is related to x_k by the simple formula

$$x_{k+1} = x_k + a_k p_k \tag{20}$$

where

$$a_k = c_k/d_k, \qquad d_k = (Ap_k, p_k), \qquad c_k = (r_k, p_k) = \langle g_k, p_k \rangle \tag{21}$$

$$r_k = -F'(x_k) = b - Ax_k, \qquad g_k = Kr_k \tag{22}$$

Moreover, r_{k+1} can be computed by the formula

$$r_{k+1} = r_k - a_k A p_k \tag{23}$$

These formulas greatly simplify computations. It is easy to see that

$$(r_{k+1}, p_j) = \langle g_{k+1}, p_j \rangle = 0, \qquad j \leqslant k \tag{24}$$

$$c_k = (r_j, p_k) = \langle g_j, p_k \rangle, \qquad j \leqslant k \tag{25}$$

In fact, the formula for a_k can be obtained from (22) and the relation

$$(r_{k+1}, p_k) = 0 \tag{26}$$

The conjugate-direction method yields an explicit formula for the inverse A^{-1} of A. To see this, we associate with each vector p the operator $pp^*/(Ap, p)$, which maps a vector x into the vector $p(p, x)/(Ap, p)$. We set

$$B_k = \sum_{j=0}^{k-1} (p_j p_j^*/d_j), \qquad d_j = (Ap_j, p_j)$$

where $\{p_j\}$ is our conjugate basis. The operator

$$P_k = B_k A$$

has the property that

$$P_k p_j = p_j, \qquad j < k$$
$$P_k p_j = 0, \qquad j \geqslant k$$

It follows that, if $\mathscr{E}$ is n-dimensional, then $B_n = A^{-1}$. If $\mathscr{E}$ is infinite-dimensional, then $\{B_k\}$ converges to A^{-1} in the sense that $\{B_k y\}$ converges to $A^{-1}y$ for each y in $\mathscr{E}$.

It is interesting to note that, if we set

$$\Delta x_k = x_{k+1} - x_k = a_k p_k$$
$$\Delta F_k' = F'(x_{k+1}) - F'(x_k) = a_k A p_k$$

then B_k can be put in the form

$$B_k = \sum_{j=0}^{k-1} [\Delta x_j \, \Delta x_j^* / (\Delta x_j, \Delta F_j')]$$

This formula is independent of the choice of the positive numbers $a_0, a_1, \ldots, a_{k-1}$. It follows that, if one is only concerned with the computation of A^{-1}, the point x_k need not minimize F on $x_0 + \mathscr{B}_k$. It must, however, be on the line $x = x_{k-1} + \alpha p_{k-1}$.

By $A\mathscr{B}$ we mean the class of all vectors $x = Ay$, where y is in $\mathscr{B}$. The orthogonal complement $\mathscr{C}$ of $A\mathscr{B}$ is called the A-orthogonal complement of $\mathscr{B}$. We have $(Ay, z) = 0$ whenever y is in $\mathscr{B}$ and z is in $\mathscr{C}$. Given a point x, by a (negative) conjugate gradient of F at x on $\mathscr{C}$ relative to $\langle y, z \rangle$ we mean a vector p in $\mathscr{C}$ such that

$$F'(x, h) = -(1/\beta)\langle p, h \rangle$$

holds for all h in $KA\mathscr{C}$, where β is a positive number. We have introduced the constant β for convenience in computations. It emphasizes that we are interested in the direction of p and not its magnitude. It follows from this definition that, if $r = -F'(x)$, then

$$(p - \beta K r, z) = 0, \qquad \text{all } z \text{ in } A\mathscr{C}$$

The vector $y = p - \beta Kr$ is therefore in $\mathscr{B}$. We have, accordingly, the simple formula

$$p = \beta Kr + y$$

for the conjugate gradient of p of F at K.

The Rayleigh–Ritz method becomes a conjugate-gradient method relative to K if, at each step, p_k is the conjugate gradient of F at x_k relative to the A-orthogonal complement $\mathscr{C}_k$ of $\mathscr{B}_k$. In this event, we have the convenient formula

$$p_{k+1} = \beta_k Kr_{k+1} + b_k p_k , \qquad p_0 = Kr_0$$

Here,

$$b_k = -\beta_k(Ap_k , Kr_{k+1})/d_k = \beta_k(r_{k+1} , Kr_{k+1})/c_k$$

and β_k is an arbitrary positive number. In practice, the choices $\beta_k = 1$ and $\beta_k = 1 - b_k$ are perhaps the preferred choices for β_k.

In the finite-dimensional case, any conjugate-direction method is a conjugate-gradient method with a suitable choice of K. In fact, K can be chosen so that $\beta_k = 1$. A similar result undoubtedly holds in the infinite-dimensional case.

7. Conjugate-Gradient Algorithms

In the present section, it is convenient to use the symbol x^*y for the inner product (x, y). In a conjugate-gradient algorithm, the conjugate basis $\{p_k\}$ is completely determined by the initial vector p_0, the positive-definite operator K, and the positive scale factor β_k for p_k. Starting with a vector $r_0 \neq 0$ and $p_0 = Kr_0$, $\beta_0 = 1$, we generate a sequence $\{p_k\}$ of conjugate vectors by the algorithm

$$s_k = Ap_k , \qquad g_k = Kr_k \tag{27}$$

$$r_{k+1} = r_k - a_k s_k , \qquad p_{k+1} = \beta_k g_k + b_k p_k \tag{28}$$

where

$$a_k = c_k/d_k , \qquad d_k = p_k{}^* s_k , \qquad c_k = p_k{}^* r_k = \beta_{k-1} g_k{}^* r_k \tag{29}$$

$$b_k = -\beta_k(s_k{}^* g_{k+1}/d_k) = \beta_k(r_{k+1}^* g_{k+1}/c_k) = c_{k+1}/c_k \tag{30}$$

The numbers a_k, b_k, c_k, d_k, β_k are positive. If $r_0 = -F'(x_0)$, the iteration

$$x_{k+1} = x_k + a_k p_k$$

yields a minimizing sequence for F. This sequence terminates in $m \leqslant n$ steps if $\mathscr{E}$ is of dimension n. The constant β_k in (27)–(30) is a scale factor for p_k. For example, we can choose β_k such that one has $\beta_k = 1$, $b_k = 1$, or $\beta_k = 1 - b_k$. This last relation is obtained by setting

$$\gamma_k = g_k{}^* r_k, \qquad \beta_k = c_k/(c_k + \gamma_{k+1}), \qquad b_k = \gamma_{k+1}/(c_k + \gamma_{k+1}) \tag{31}$$

We have the relations

$$c_k = p_k{}^* r_j, \qquad j \leqslant k \tag{32}$$

$$p_k{}^* r_j = 0, \qquad k < j \tag{33}$$

$$p_j{}^* A p_k = 0, \qquad j \neq k \tag{34}$$

$$r_j{}^* K r_k = 0, \qquad j \neq k \tag{35}$$

$$r_j{}^* K s_k = 0, \qquad j \neq k, k+1 \tag{36}$$

$$s_j{}^* K s_k = 0, \qquad j + 1 < k \tag{37}$$

In addition,

$$p_{k+1} = \lambda_k p_k - \mu_k K A p_k - \nu_k p_{k-1}, \qquad \nu_0 = 0 \tag{38}$$

where $\mu_k = \beta_k a_k$ is a scale factor and

$$\lambda_k = (\beta_k/\beta_{k-1}) + b_k = \mu_k(s_k{}^* K s_k/d_k) \tag{39}$$

$$\nu_k = \beta_k b_{k-1}/\beta_{k-1} = -\mu_k(s_k{}^* K s_{k-1}/d_{k-1}) \tag{40}$$

Equations (38)–(40) can be used in place of (27)–(30) to generate the conjugate basis $\{p_k\}$. The equations given here can be obtained from the case $K = I$ by a transformation of variables.

Given p_0, K, and $q_0 = Ks_0$, the conjugate sequence $\{p_k\}$ can also be generated by the algorithm

$$s_k = A p_k, \qquad d_k = p_k{}^* s_k, \qquad e_k = q_k{}^* s_k \tag{41}$$

$$p_{k+1} = p_k - \beta_k q_k, \qquad \beta_k = d_k/e_k \tag{42}$$

$$q_{k+1} = K s_{k+1} + \alpha_k q_k, \qquad \alpha_k = d_{k+1}/d_k = -(s_k{}^* K s_{k+1}/e_k) \tag{43}$$

Here, α_k and β_k are determined by the relations

$$p_{k+1}^* s_k = 0, \qquad q_{k+1}^* s_k = 0$$

If we wish, we can introduce a scale factor ρ_k for q_k by replacing the last equation by

$$q_{k+1} = \rho_{k+1} K s_{k+1} + \alpha_k q_k, \qquad \alpha_k = \rho_{k+1} d_{k+1}/\rho_k d_k$$

or by setting $s_k = \rho_k A p_k$, $d_k = p_k^* s_k/\rho_k$. This does not alter p_k. If $p_0 = Kr_0$, the iteration (35)–(37) is equivalent to (27)–(30) with $\beta_k + b_k = 1$. As described in the last section, the sequence of matrices $\{B_k\}$ generated by the algorithm

$$B_0 = 0, \qquad B_{k+1} = B_k + (p_k p_k^*/d_k) \tag{44}$$

determines A^{-1} unless the algorithm terminates prematurely.

The iteration (41)–(43) can be put in another form. Starting with $p_0 \neq 0$, $M_0 = K$, $B_0 = 0$, we generate p_k, q_k, M_k, B_k as follows:

$$s_k = Ap_k, \qquad q_k = M_k s_k, \qquad d_k = p_k^* s_k, \qquad e_k = q_k^* s \tag{45}$$

$$p_{k+1} = p_k - \beta_k q_k, \qquad \beta_k = d_k/e_k \tag{46}$$

$$M_{k+1} = M_k - (q_k q_k^*/e_k), \qquad B_{k+1} = B_k + (p_k p_k^*/d_k) \tag{47}$$

In this event, we have

$$M_k s_j = 0, \qquad j < k$$

$$M_k s_j = K s_j, \qquad j > k$$

If the iteration does not terminate prematurely, we have $M_n = 0$ if $\mathscr{E}$ is of dimension n and $\lim M_k x = 0$ for each x otherwise. Setting

$$H_k = M_k + B_k$$

and using the relation $B_k s_k = 0$, we see that the algorithm (45)–(47) yields the following algorithm. Starting with $p_0 \neq 0$ and $H_0 = K$, set

$$s_k = Ap_k, \qquad q_k = H_k s_k, \qquad d_k = p_k^* s_k, \qquad e_k = q_k^* s_k \tag{48}$$

$$p_{k+1} = p_k - \beta_k q_k, \qquad \beta_k = d_k/e_k \tag{49}$$

$$H_{k+1} = H_k - (q_k q_k^*/e_k) + (p_k p_k^*/d_k) \tag{50}$$

In the algorithms (45)–(47) and (48)–(50), one can replace s_k and d_k by $s_k = \rho_k A p_k$ and $d_k = p_k^* s/\rho_k$, $\rho_k > 0$, without altering p_{k+1}, M_{k+1}, B_{k+1}, or H_{k+1}.

In general, the sequence $\{x_k\}$ defined by the recursion formula

$$x_k = x_0 - B_k F'(x_0) \qquad \text{or} \qquad x_k = x_0 - H_k F'(x_0)$$

is a minimizing sequence for F. This is always true if $p_0 = -KF'(x_0)$. If no round-off errors occur, the iteration terminates in $m \leqslant n$ steps if $\mathscr{E}$ is of dimension n.

We shall give one final algorithm for generating a conjugate basis. This algorithm is equivalent to the preceding ones. We shall express the result in terms of the original function F to be minimized. Having chosen an initial point x_0 and the operators $M_0 = K$, $B_0 = 0$, we iterate as follows

$$p_k = -M_k F'(x_k), \qquad \rho_k > 0, \quad \rho_k \text{ arbitrary} \tag{51}$$

$$x_{k+1} = x_k + \rho_k p_k, \qquad s_k = F'(x_{k+1}) - F'(x_k) \tag{52}$$

$$q_k = M_k s_k, \qquad e_k = q_k{}^* s_k, \qquad d_k = p_k{}^* s_k / \rho_k \tag{53}$$

$$M_{k+1} = M_k - (q_k q_k{}^*/e_k), \qquad B_{k+1} = B_k + (p_k p_k{}^*/d_k) \tag{54}$$

The relations (52) can be replaced by

$$x_{k+1} = x_0 + \rho_k p_k, \qquad s_k = F'(x_{k+1}) - F'(x_0) \tag{55}$$

if one so desires, but we shall not do so here. The optimal choice for ρ_k is the minimum point $\rho = a_k$ of $F(x_k + \rho p_k)$. This algorithm with $\rho_k = a_k$ has been given by Kelley and Myers. If we select $\rho_k = a_k$ and replace M_k by $H_k = M_k + B_k$, we obtain Davidon's method as given by Fletcher and Powell. In these two cases, the sequence $\{x_k\}$ is a minimizing sequence for F. In all cases, the sequences $\{y_k\}$ or $\{z_k\}$ defined by

$$y_k = x_0 - B_k F'(x_0), \qquad z_k = x_0 - H_k F'(x_0)$$

are minimizing sequences for F. In the n-dimensional case, there is an integer $m \leqslant n$ such that $y_m = z_m$ is the solution to our problem.

In the n-dimensional case, the algorithm (51)–(54) with n steps applied to an arbitrary function F can be looked upon as one Newton iteration. In particular, if (55) is used in the iteration, the matrix B_n is an estimate $B(x_0)$ of the inverse of $F''(x_0)$. A repetition of the algorithm gives the sequence

$$\bar{x}_{k+1} = \bar{x}_k - B(\bar{x}_k) F'(\bar{x}_k), \qquad \bar{x}_0 = x_0$$

as a gradient method. Of course, at any stage, one can use $B(\bar{x}_k) = B(\bar{x}_{k-1})$ instead of computing $B(\bar{x}_k)$ by the algorithm (51)–(54).

The possibility of using an arbitrary scale factor ρ_k in (51)–(54) was suggested by a statement in the invited address by B. Pschenichniy, USSR, to the effect that he had devised a similar algorithm having this property. For a discussion of conjugate-gradient and Davidon's methods, the readers should consult Refs. 8–13. Further references can be found in these papers.

References

1. HESTENES, M. R., *A General Problem in the Calculus of Variations with Applications to Paths of Least Time*, The RAND Corporation, Research Memorandum No. RM-100, 1950.
2. MENGEL, A. S., *Optimum Trajectories*, The RAND Corporation, Report No. P-199, 1951.
3. HESTENES, M. R., *Numerical Methods for Obtaining Solutions of Fixed-Endpoint Problems in the Calculus of Variations*, The RAND Corporation, Research Memorandum No. RM-102, 1949.
4. STEIN, M. L., *On Methods for Obtaining Solutions of Fixed-Endpoint Problems in the Calculus of Variations*, Journal of Research of the National Bureau of Standards, Vol. 50, No. 5, 1953.
5. HESTENES, M. R., *Iterative Computational Methods*, Communications on Pure and Applied Mathematics, Vol. 8, No. 1, 1955.
6. BALAKRISHNAN, A. V., *On a New Computing Technique in Optimal Control and Its Application to Minimal-Time Flight Profile Optimization*, Journal of Optimization Theory and Applications, Vol. 4, No. 1, 1969.
7. HESTENES, M. R., *An Indirect Sufficiency Proof for the Problem of Bolza in Nonparametric Form*, Transactions of the American Mathematical Society, Vol. 62, No. 3, 1947.
8. HESTENES, M. R., and STIEFEL, E., *Methods of Conjugate Gradients for Solving Linear Systems*, Journal of Research of the National Bureau of Standards, Vol. 49, No. 6, 1952.
9. HESTENES, M. R., *The Conjugate Gradient Method for Solving Linear Systems*, Proceedings of the Sixth Symposium in Applied Mathematics, Edited by J. H. Curtiss, American Mathematical Society, Providence, Rhode Island, 1956.
10. HAYES, R. M., *Iterative Methods for Solving Linear Problems in Hilbert Space*, Contributions to the Solutions of Systems of Linear Equations and the Determinations of Eigenvalues, Edited by O. Tausky, National Bureau of Standards, Applied Mathematics Series, US Government Printing Office, Washington, D.C., 1954.
11. FLETCHER, R., and POWELL, M. J. D., *A Rapidly Convergent Descent Method for Minimization*, Computer Journal, Vol. 6, No. 2, 1964.
12. MYERS, G. E., *Properties of the Conjugate-Gradient and Davidon Methods*, Journal of Optimization Theory and Applications, Vol. 2, No. 4, 1968.
13. HORWITZ, L. B., and SARACHICK, P. E., *Davidon's Method in Hilbert Space*, SIAM Journal on Applied Mathematics, Vol. 16, No. 4, 1968.