CS 61

ON THE ASYMPTOTIC DIRECTIONS OF THE S-DIMENSIONAL OPTIMUM GRADIENT METHOD

BY

GEORGE E. FORSYTHE

TECHNICAL REPORT NO. CS 61

APRIL 13, 1967

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

On the Asymptotic Directions of the s-dimensional Optimum Gradient Method*

By

George E. Forsythe

Abstract

The optimum s-gradient method for minimizing a positive definite quadratic function $f(x)$ on E_n has long been known to converge for $s > 1$. For these s the author studies the directions from which the iterates x_k approach their limit, and extends to $s > 1$ a theory proved by Akaike for $s = 1$. It is shown that $f(x_k)$ can never converge to its minimum value faster than linearly, except in degenerate cases where it attains the minimum in one step.

Computer Science Department

Stanford University

*Research sponsored by the U. S. Office of Naval Research under contract Nonr 225 (37) = NR 044 211.

1. Introduction and Summary.

To minimize a smooth real-valued function f(x) of n real variables, the optimum s-gradient method has been described by Birman [3], Faddeev and Faddeeva [5], Khabaza [8], and others. We here consider the model function $f(x) = \frac{1}{2}x^T Ax$, where A is a positive definite matrix. Then each iterate x_k is equal to its error. The convergence of the method was proved long ago--see (2.14)--and the question now under study is to find the asymptotic manner in which the iterates $x_k \to \theta$, the null vector.

For $s = 1$ it was conjectured by Forsythe and Motzkin [7] and proved by Akaike [1]--see (4.12)--that the iterates x_k converge to θ by asymptotically alternating between two directions--the "cage" of Stiefel [10]. Thus the convergence of $f(x_k)$ to 0 for $s = 1$ is known to be linear, and no faster than linear, for any start x_0 that is not an eigenvector. Moreover, if coordinates are chosen so that A is a diagonal matrix, then the two asymptotic directions have only two nonzero components. Finally, any direction with only two nonzero components is invariant under two steps of the optimum 1-gradient method.

In the present paper the author has extended most of the known results to arbitrary $s > 1$. The main result (3.8) shows that the directions of the even iterates x_{2k} have as a limit set a continuum R (which might be a single direction). Moreover, each direction of R is invariant under two steps of the optimum s-gradient method. Let A be a diagonal matrix. It is then shown in (3.10) that in the optimum s-gradient process $f(x_k)$ converges to 0 no faster than linearly for any initial vector x_0 with at least $s + 1$ nonzero components.

Theorem (4.7) shows that all vectors of R have between $s + 1$ and $2s$ nonzero coordinates, inclusive, Theorem (4.8) says that any direction with $s + 1$ nonzero components is invariant under two steps of the method, for any s. Examples are shown in Sec. 4 of directions with this invariance and with as many as $2s$ nonzero components,

Experimental evidence from computer runs for $s = 2$ suggests strongly that R is always a single point, just it has been proved to be for $s = 1$. The author conjectures without proof that R is a single point for all s, so that $x_k \to \theta$ in an alternating manner completely analogous to the case with $s = 1$.

The author is aware that for minimizing quadratic functions f(x) in practice, the conjugate-gradient method of Hestenes and Stiefel (see [5]) may usually be expected to be superior to the optimum s-gradient methods, although Khabaza [8] claims superiority for the 3-gradient method in some cases. For nonquadratic functions f(x) the relative merits of the methods are less clear, The purpose of the present investigation was the intellectual one of trying to understand the asymptotic behavior of the various gradient methods for quadratic functions, The author expects that this information may have some useful application to the minimization of general smooth functions f(x).

2. The Optimum s-gradient Method for Quadratic Functions.

Let $f(x)$ be real for all x in real euclidean n-space E_n. Let $f(x)$ assume a minimum value for a unique x, which can be taken as θ, the origin of E_n, without loss of generality in the analysis. The advantage of using θ is that the iterate x_k is then also its own error $x_k - \theta$ as a minimizing vector. We wish to analyze certain asymptotic properties of a class of optimum gradient methods for finding the minimum of $f(x)$.

The simplest f to analyze is the quadratic function

$$f(x) = \tfrac{1}{2}x^T Ax, \tag{2.1}$$

where A is a symmetric, positive definite, nonderogatory matrix of order n. Moreover, (2.1) represents the local behavior at θ of $f(x) - f(0)$ for most sufficiently smooth functions f. The author conjectures that the theorems proved below for a quadratic function apply essentially also to any sufficiently smooth function f which is locally like (2.1). In this paper only quadratic functions will be studied. See Daniel [4] for an investigation comparing gradient methods for quadratic and nonquadratic-functions in Hilbert space.,

In the various gradient methods one starts with an arbitrary vector x_0, and computes a sequence $\{x_k\}$ converging to θ. We assume all arithmetic to be exact, and round-off error is not considered in this paper.

Let $z_k = \text{grad } f(x_k) = Ax_k$ denote the gradient of f at x_k. In the optimum 1-gradient method [5], x_{k+1} is taken to be the unique point on the line $L_1 = \{x_k + \alpha Ax_k : -\infty < \alpha < \infty\}$ for which $F(\alpha) = f(x_k + \alpha Ax_k)$ is a minimum. (The existence and uniqueness of x_{k+1} result from the easily proved fact that $F(\alpha)$ is a quadratic function of α with $F''(\alpha) > 0$.) The line L_1 through x_k is called the line of steepest descent of $f(x)$ at x_k .

For $x \in L_1$, grad $f(x) = A(x_k + a Ax_k) = Ax_k + a A^2 x_k$. We therefore consider the 2-dimensional plane through x_k,

$$L_2 = \{x_k + \alpha_1 Ax_k + \alpha_2 A^2 x_k : -\infty < \alpha_1 < \infty, -\infty < \alpha_2 < \infty\},$$

and call it the 2-plane of steepest descent of $f(x)$ at x_k .

By extension, for any integer s $(1 \leq s \leq n)$ let

$$L_s = \{x_k + \sum_{i=1}^{s} \alpha_i A^i x_k : -\infty < \alpha_i < \infty \quad \text{(all } i\text{)}\}$$

be the s-dimensional plane of steepest descent of $f(x)$ at x_k. Since A is not derogatory, $Ax_k, \ldots, A^n x_k$ are linearly independent, provided x_0 is a vector whose minimum polynomial is of degree n . In that case L_n is the whole space E_n .

In the optimum s-gradient method [5] for minimizing the quadratic function f of (2.1), the point x_{k+1} is defined to be the unique point y in L_s for which $f(y)$ is a minimum $(k = 0, 1, \ldots)$. (Again existence and uniqueness follow from the positive definiteness of A .) It is the optimum s-gradient methods that we shall analyze in this paper.

We now give two representations of the minimizing $\{\alpha_i\}$ which are useful in the analysis. Actual computing algorithms for the optimum s-gradient method often proceed differently, and find x_{k+1} by taking s steps of the conjugate gradient method, starting from x_k. See [5]. We concentrate on the gradients $z_k = Ax_k$.

First representation

Let

$$x_{k+1} = x_k + \gamma_1 Ax_k + \ldots + \gamma_s A^s x_k .$$

Then the gradient of $f(x)$ at x_{k+1} is

$$z_{k+1} = z_k + \gamma_1 Az_k + \ldots + \gamma_s A^s z_k . \tag{2.2}$$

Since x_{k+1} minimizes $f(y)$ for $y \in L_s$, the vector z_{k+1} must be orthogonal to L_s. For this it is necessary and sufficient that z_{k+1} be orthogonal to $z_k, Az_k, \ldots, A^{s-1}z_k$. Then $\gamma_1, \ldots, \gamma_s$ are determined by the s conditions

$$\left\{\begin{array}{l} (z_k, z_{k+1}) = (z_k, z_k) + \gamma_1(z_k, Az_k) + \ldots + \gamma_s(z_k, A^s z_k) = 0 \\ \quad \cdots\cdots\cdots \\ (A^{s-1}z_k, z_{k+1}) = (A^{s-1}z_k, z_k) + \gamma_1(A^{s-1}z_k, Az_k) + \ldots + \gamma_s(A^{s-1}z_k, A^s z_k) = 0 . \end{array}\right.$$

Here $(u, v) = u^T v + v^T u$ denotes the inner product of two column vectors. Since $(A^p z, A^q z) = (z, A^{p+q}z) = z^T A^{p+q} z$, we may write the above equations as

$$
(2.3) \quad \begin{cases}
z_k^T z_k + \gamma_1 z_k^T A z_k + \ldots + \gamma_s z_k^T A^s z_k = 0 \\
z_k^T A z_k + \gamma_1 z_k^T A^2 z_k + \ldots + \gamma_s z_k^T A^{s+1} z_k = 0 \\
\ldots\ldots \\
z_k^T A^{s-1} z_k + \gamma_1 z_k^T A^s z_k + \ldots \gamma_s z_k^T A^{2s-1} z_k = 0 \quad .
\end{cases}
$$

As long as $z_k, Az_k, \ldots . A^{s-1} z_k$ are linearly independent, the equations (2.3) determine the minimizing $\gamma_1, \ldots . \gamma_s$ uniquely.

Second representation

Let $q_s(t) = t^s + \beta_{s-1} t^{s-1} + \ldots + \beta_0$ denote any monic polynomial of degree s, with $\beta_0 \neq 0$. Then

$$q_s(A) z_k = A^s z_k + \beta_{s-1} A^{s-1} z_k + \ldots + \beta_0 z_k,$$

and

$$(2.4) \quad \frac{q_s(A) z_k}{q_s(0)} = \frac{1}{\beta_0} A^s z_k + \frac{\beta_{s-1}}{\beta_0} A^{s-1} z_k + \ldots + z_k \; .$$

Comparing (2.4) with (2.2), we see that we can write

$$(2.5) \quad z_{k+1} = \frac{p_s(A)}{p_s(0)} z_k,$$

where $p,(t)$ is the particular polynomial

$$(2.6) \quad p,(t) = t^s + \frac{\gamma_{s-1}}{\gamma_s} t^{s-1} \ldots + \frac{\gamma_1}{\gamma_s} t + \frac{1}{\gamma_s} \; .$$

Now $p,(t)$ is a certain orthogonal polynomial. Without loss of generality assume A to be the diagonal matrix

(2.7) $$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{bmatrix} \lambda_1 & & & \bigcirc \\ & \lambda_2 & & \\ & & \ddots & \\ \bigcirc & & & \lambda_n \end{bmatrix},$$

where $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n$ are its eigenvalues (distinct because A is not derogatory).

(2.8) Definition. In the coordinate system corresponding to (2.7), let the nonzero vector z be $(\zeta_1, \ldots, \zeta_n)^T$. Let orthogonality of two polynomials p(t), q(t) (relative to z) be defined by

$$\langle p(t), q(t) \rangle = \sum_{i=1}^{n} p(\lambda_i) q(\lambda_i) \zeta_i^2 = 0 .$$

(2.9) Definition. Let $P_s(t; z) = t^s + \ldots$ be the unique monic polynomial of degree s that, relative to z, is orthogonal in the sense of (2.8) to all polynomials of degree $\leq s-1$.

Note that $P_s(t; z)$ depends only on the direction of z, and not its magnitude. I.e., $P_s(t; z) = P_s(t; \alpha z)$, for all real $\alpha \neq 0$.

(2.10) Theorem. The polynomial $p_s(t)$ of (2.5), (2.6) is the orthogonal polynomial $P_s(t; z_k)$ defined in (2.9).

We shall not prove (2.10). For a related proof see, for example, p. 349 of [5]. The basic reason for (2.10) is the isomorphism, well expounded by Stiefel [11], between orthogonality of the polynomials

$p(t)$, $q(t)$ in the sense of (2.8) and geometric orthogonality of the vectors $p(A)z$, $q(A)z$ in E_n. That is,

$$\langle p(t),\ q(t)\rangle = (p(A)z,\ q(A)z)\ .$$

Hence the conditions (2.3) asserting the orthogonality of the vector $z_{k+1} = P_s(A;\ z_k)\ z_k\ /\ P_s(0;\ z_k)$ to z_k, Az_k, A^2z_k, ..., $A^{s-1}z_k$ are equivalently asserting the orthogonality of the polynomial $P_s(t;\ z_k)$ to the polynomials $1, t, t^2, \ldots, t^{s-1}$.

In summary z_{k+1} is uniquely determined from z_k by the formula

$$z_{k+1} = \frac{P_s(A;\ z_k)}{P_s(0;\ z_k)}\ z_k\ . \qquad (2.11)$$

Moreover,

$$x_{k+1} = \frac{P_s(A;\ z_k)}{P_s(0;\ z_k)}\ x_k\ . \qquad (2.12)$$

Relation (2.12) is the basis for a proof by Birman[3] that in the optimum s-gradient method $f(x_k)$ converges to 0 linearly, or faster. To be precise, let $\sigma = (\lambda_n + \lambda_1)(\lambda_n - \lambda_1)^{-1}$. Let $T_s(t)$ denote the Chebyshev polynomial on [-1, 1], normalized so that $\max_{-1 < t < 1}|T_s(t)| = 1$. Let

$$Q_s(u) = T_s\left(\frac{\lambda_n + \lambda_1 - 2u}{\lambda_n - \lambda_1}\right)\ .$$

Then $Q_s(0) = T_s(\sigma) > 1$, and $|Q_s(t)| \leq 1$, for $\lambda_1 \leq t \leq \lambda_n$. It is known that

$$T_s(\sigma) = \frac{(\sigma + \sqrt{\sigma^2-1})^s + (\sigma - \sqrt{\sigma^2-1})^s}{2} > 1 .$$

Birman's proof goes as follows:

$$f(x_{k+1}) = f\left(\frac{P_s(A;\, z_k)}{P_s(0;\, z_k)}\, x_k\right)$$

$$\leq f\left(\frac{Q_s(A)}{Q_s(0)}\, x_k\right), \text{ because } P_s(t;\, z_k) \text{ is the polynomial that minimizes } f(x_{k+1})$$

$$= \frac{1}{[Q_s(0)]^2}\, x_k^T Q_s(A)\; A\, Q_s(A)\, x_k$$

(2.13)
$$= \frac{1}{[Q_s(0)]^2} \sum_{i=1}^{n} \lambda_i [Q_s(\lambda_i)]^2 [\xi_i^{(k)}]^2$$

$$\leq \frac{1}{[Q_s(0)]^2} \sum_{i=1}^{n} \lambda_i [\xi_i^{(k)}]^2$$

$$= \frac{1}{[T_s(\sigma)]^2}\, f(x_k) .$$

Hence

(2.14)
$$\sqrt{f(x_k)} \leq \frac{1}{[T_s(\sigma)]^k} \sqrt{f(x_0)} ,$$

proving the convergence of $f(x_k)$ to 0 to be linear or faster.

(2.15) Definition. For $\alpha = 0, \pm 1, \pm 2, \ldots,$ let the <u>moments</u> μ_α of $z = (\zeta_1, \ldots, \zeta_n)^T$ be defined by

$$\mu_\alpha = \sum_{i=1}^{n} \lambda_i^\alpha \zeta_i^2 .$$

(2.16) Theorem. Fix $s > 1$. Except for a constant factor, the orthogonal polynomial $P_s(t; z)$ of (2.9) can be expressed by the determinant

$$(2.17) \qquad P_s(t; z) = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{s-1} & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_s & t \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mu_s & \mu_{s+1} & & \mu_{2s-1} & t^s \end{vmatrix} .$$

The proof is left to the reader.

In the next theorem we give an explicit representation for the ratio $f(x_{k+1})/(f(x_k)$ in terms of the moments of z_k.

(2.18) Theorem. Fix $s > 1$. Let x_k be any vector in the optimum s-gradient method, and let μ_α be the moments defined by (2.15) for the gradient vector $z_k = Ax_k$. Then

$$\frac{f(x_{k+1})}{f(x_k)} = \frac{\begin{vmatrix} \mu_{-1} & \mu_0 & \mu_1 & \cdots & \mu_{s-1} \\ \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_s \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mu_{s-1} & \mu_s & \mu_{s+1} & & \mu_{2s-1} \end{vmatrix}}{\mu_{-1} M_{-1}} ,$$

where M_{-1} is the minor determinant of μ_{-1} in the above determinant:

$$M_{-1} = \begin{vmatrix} \mu_1 & \mu_2 & \dots & \mu_s \\ \mu_2 & \mu_3 & \dots & \mu_{s+1} \\ \cdot & \cdot & \cdot & \cdot \\ \mu_s & \mu_{s+1} & \dots & \mu_{2s-1} \end{vmatrix} .$$

Proof. We have $2f(x_k) = x_k^T A x_k = z_k^T A^{-1} z_k = \mu_{-1}$. To simplify the notation, let $z_k = (\zeta_1, \dots, \zeta_n)^T$ and $z_{k+1} = (\zeta_1', \dots, \zeta_n')^T$. Then

$$\zeta_i' = \frac{P_s(\lambda_i;\, z_k)}{P_s(0;\, z_k)}\, \zeta_i, \qquad \text{by (2.11)}$$

$$= \frac{P_s(\lambda_i;\, z_k)}{(-1)^s M_{-1}}\, \zeta_i,$$

where we use the representation (2.17) for $P_s(t;\, z_k)$. Then

$$\text{(2.19)} \quad 2f(x_{k+1}) = z_{k+1}^T A^{-1} z_{k+1} = \frac{1}{M_{-1}^2} \sum_{i=1}^{n} [P_s(\lambda_i;\, z_k)]^2 \frac{1}{\lambda_i} \zeta_i^2$$

$$\frac{1}{M_{-1}^2} \sum_{i=1}^{n} P_s(\lambda_i;\, z_k) \frac{P_s(\lambda_i;\, z_k)}{\lambda_i} \zeta_i^2 .$$

Now $P_s(t;\, z_k)$ is orthogonal in the sense of (2.8) to all polynomials of degree $\leq s - 1$. Hence the only term of $P_s(\lambda_i;\, z_k)/\lambda_i$ that contributes anything nonzero to the sum (2.19) is the term $(-1)^s M_{-1}/\lambda_i$. Hence

$$2f(x_{k+1}) = \frac{(-1)^s}{M_{-1}} \sum_{i=1}^{n} P_s(\lambda_i;\, z_k)\, \zeta_i^2/\lambda_i$$

$$= \frac{(-1)^s}{M_{-1}} \sum_{i=1}^{n} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{s-1} & \zeta_i^2/\lambda_i \\ \mu_1 & \mu_2 & \cdots & \mu_s & \zeta_i^2 \\ & & \cdots\cdots & & \\ \mu_s & \mu_{s+1} & \cdots & \mu_{2s-1} & \zeta_i^2 \lambda_i^{s-1} \end{vmatrix}$$

$$= \frac{(-1)^s}{M_{-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_{s-1} & \mu_{-1} \\ \mu_1 & \mu_2 & \cdots & \mu_s & \mu_0 \\ & & & & \\ \mu_s & \mu_{s+1} & \cdots & \mu_{2s-1} & \mu_{s-1} \end{vmatrix} .$$

Dividing $2f(x_{k+1})$ by $2f(x_k) = \mu_{-1}$ and rearranging the columns of the last determinant proves theorem (2.18).

(2.20) Corollary. _In the notation of theorem_ (2.18), _for_ $s = 1$,

$$\text{(2.21)} \qquad \frac{f(x_{k+1})}{f(x_k)} = \frac{\mu_1\mu_{-1} - \mu_0^2}{\mu_1 - 1} .$$

If $n = 2$ _and_ $s = 1$, _then_

$$\text{(2.22)} \qquad \frac{f(x_{k+1})}{f(x_k)} = \frac{\zeta_1^2\zeta_2^2(\lambda_2 - \lambda_1)^2}{(\lambda_2\zeta_1^2+\lambda_1\zeta_2^2)(\lambda_1\zeta_1^2+\lambda_2\zeta_2^2)} = c^2 = c^2(x_k) .$$

Proof. The second expression comes from the first by using (2.15) and (2.21), where $z_k = (\zeta_1, \zeta_2)^T$, with some algebraic manipulation.

(2.23) Corollary. _The expression_ (2.22) _for_ $f(x_{k+1})/f(x_k)$ _is unchanged, if_ $(\zeta_1, \zeta_2)^T$ _is changed to_ $(\zeta_2, -\zeta_1)^T$.

The inequality (2.13) yields an upper bound for the expression in (2.18). We may state this result in the form of the following inequality, valid for $s = 1, 2, \ldots$.

$$(2.24)\qquad \begin{vmatrix} \mu_{-1} & \mu_0 & \mu_1 & \cdots & \mu_{s-1} \\ \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_s \\ & & \cdots & & \\ \mu_{s-1} & \mu_s & \mu_{s+1} & \cdots & \mu_{2s-1} \end{vmatrix} \Bigg/ \mu_{-1} \times \begin{vmatrix} \mu_1 & \cdots & \mu_s \\ & \cdots & \\ \mu_s & \cdots & \mu_{2s-1} \end{vmatrix} < \frac{1}{\left[T_s\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right]^2}$$

This is essentially the inequality of Meinardus [8a], who derived it by the same argument for a slightly different iteration in which $\|x\|^2$ is minimized instead of $f(x)$.

The special case for $s = 1$,

$$(2.25)\qquad \begin{vmatrix} \mu_{-1} & \mu_0 \\ \mu_0 & \mu_1 \end{vmatrix} \Bigg/ \mu_{-1}\,\mu_1 \le \frac{1}{\left[T_1\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right]^2} = \left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2$$

is a well-known inequality of Kantorovich; see (8) on p. 410 of [5].

It was stated by Birman [3] that the bound (2.14) is sharp, in the sense that for each s and each given λ_1, λ_n $(s < n)$, one can find A and x_0 so that (2.14) is an equality for all k . This is done by finding a set of λ_i and $\xi_i^{(0)}$ so that the shifted Chebyshev polynomial $Q_s(t)$ is (up to a scalar factor) identical with $P_s(t;z_0)$ and so that $|Q_s(\lambda_i)| = 1$ for each eigenvalue λ_i . This is known to be possible because the Chebyshev polynomials, like cosines, are orthogonal with respect to summation over certain points.

However, Birman did not investigate the actual manner or rate of convergence of $f(x_k)$ to 0 in the optimum s-gradient method for a general given A and x_0 . He left open the question of whether the convergence of $f(x_k)$ to 0 might actually be faster than linear in certain nontrivial cases.

For s = 1 Forsythe and Motzkin [7] conjectured that if $\xi_1^{(0)} \xi_n^{(0)} \neq 0$, then $\xi_i^{(k)} = \underline{o}(\|x_k\|)$, as $k \to \infty$, for all i with $1 < i < n$. In words, $x_k \to \theta$ asymptotically in the 2-space $\pi_{1,n}$ spanned by the eigenvectors belonging to λ_1 and λ_n . The conjecture was proved by Forsythe and Motzkin (unpublished) only for n = 3 . Akaike [1] proved the conjecture for arbitrary n . In an unpublished manuscript Arms [2] had found a similar proof. We give a proof in (4.12) as a consequence of our result (3.8) for the s-gradient method.

Suppose the optimum 1-gradient process is performed entirely in the two-dimensional space $\pi_{1,n}$. Then, if $x_0 \in \pi_{1,n}$ and x_0 is not an eigenvector, it is easy to prove that:

(i) $x_0, x_2, x_4, \cdots$ are all collinear vectors, and that $x_1, x_3, x_5, \cdots$ are also collinear in another direction. Furthermore, $x_{2k+2} = c^2 x_{2k}$ and $x_{2k+1} = c^2 x_{2k-1}$, for all k . Here c^2 is given by (2.22). The basic reason why these vectors are collinear is that the gradients z_{k+1} and z_k must always be perpendicular in any optimum gradient method.

(ii) Moreover, for each k = 0, 1, ... $f(x_{k+1}) = c^2 f(x_k)$. This is an immediate consequence of Corollary (2.23). Hence $f(x_k) \searrow 0$ in a strictly linear fashion, like the k-th term of a convergent geometric series, even though the vectors x_k alternate between two fixed directions.

It is a consequence of the Forsythe-Motzkin-Arms-Akaike result on the manner of convergence of x_k to θ in E_n for s = 1 that the iteration behaves asymptotically, as $k \to \infty$, as though it were entirely in the two-space $\pi_{1,n}$. The vectors x_k behave ultimately as though they had resulted from an iteration started with some x_0^* in $\pi_{1,n}$. In particular, we find that $f(x_k) \searrow 0$ linearly, in the sense that

$$\lim_{k \to \infty} \frac{f(x_{k+1})}{f(x_k)} = c^2(x_0^*) .$$

However, the vector x_0^* is an extremely complex function of x_0.

Till now, the asymptotic nature of the optimum s-gradient method has not been described for $s > 1$. This problem, posed on p. 314 of Forsythe [6], is studied in the next section

3. Asymptotic Behavior of the s-gradient Method.

We are still assuming A to have distinct positive eigenvalues $\lambda_1 < \lambda_2 < \ldots < \lambda_n$. Fix any s with $1 < s$. Motivated by (2.11) and by Akaike's approach [1] for $s = 1$, we shall consider the transformation

$$w' = P_s(A;\, w)\; w \; . \tag{3.1}$$

Here $w \neq \theta$ and $P_s(t;\, w) = t^s + \ldots$ is the orthogonal polynomial defined in (2.9). Let

$$\varphi(w) = \frac{\|w'\|^2}{\|w\|^2} = \frac{\|P_s(A;\, w)\|^2}{\|w\|^2} , \tag{3.2}$$

where $\|u\|$ denotes the euclidean length of u .

Similarly, if $w' \neq \theta$, let $w'' = P_s(A;\, w')\, w'$, so that

$$\varphi(w') = \frac{\|w''\|^2}{\|w'\|^2} .$$

The following theorem is of basic importance to our analysis of the asymptotic behavior of the s-gradient method.

(3.3) Theorem. Let ψ be the angle between w and w'' . For any w such that $w'' \neq \theta$, we have

$$\varphi(w) = \frac{\|w'\|^2}{\|w\|^2} = \cos^2\psi \, \frac{\|w''\|^2}{\|w'\|^2} \leq \frac{\|w''\|^2}{\|w'\|^2} = \varphi(w'),$$

and there is equality if and only if $w'' = cw$ _for some scalar_ $c > 0$.

Proof. By the Cauchy-Schwarz inequality and the definition of ψ,

$$(3.4) \qquad (w^T w'')^2 = \cos^2\psi \|w\|^2 \, \|w''\|^2 \leq \|w\|^2 \, \|w''\|^2,$$

with equality if and only if $w = cw''$, for $c \neq 0$.

Now

$$\begin{aligned} \|w'\|^2 - w^T w'' &= \|P_s(A; w) \ w\|^2 - w^T P_s(A; w') \ P_s(A; w) \ w \\ &= w^T [P_s(A; w)]^2 \ w - w^T P_s(A; w') \ P_s(A; w) \ w \\ &= w^T P_s(A; w) \ \{P_s(A; w) - P_s(A; w')\} \ w \\ &= w^T P_s(A; w) \ D(A) \ w \\ &= 0, \end{aligned}$$

by (2.3), because $D(t)$ is a polynomial of degree at most $s - 1$, since the leading terms t^s cancel. Hence $\|w'\|^2 = w^T w''$, whence

$$(3.5) \qquad \|w'\|^4 = (w^T w'')^2 .$$

Combining (3.4) with (3.5), we have

$$\|w'\|^4 = \cos^2\psi \|w\|^2 \, \|w''\|^2 \leq \|w\|^2 \, \|w''\|^2,$$

with equality if and only if $w'' = cw$. That $c > 0$ follows from the fact that $w^T w'' = \|w'\|^2 > 0$. This proves theorem (3.3).

(3.6) Definition. Fix s with $1 \leq s \leq n - 1$. Fix a euclidean coordinate system in E_n so A takes the form (2.7). Let Σ be the unit sphere in E_n. Define $\Sigma^* \subset \Sigma$ to consist of all unit vectors y with at least $s + 1$ nonzero components. We define a transformation $T: \Sigma^* \to \Sigma^*$, as follows: For each y in Σ^*, let $y' = Ty = w/\|w\|$, where $w = P_s(A; y)\, y$. (That $w \neq \theta$ and $y' \in \Sigma^*$ are proved in theorem (5.1).)

(3.7) Definition. By a continuum we mean a closed connected set in E_n, with the understanding that a single point is a continuum.

(3.8) Theorem. Fix s with $1 \leq s \leq n - 1$. Let $y_0 = (\eta_1^{(0)}, \ldots, \eta_n^{(0)})^T$ be any vector in Σ^* with $\eta_i^{(0)} \neq 0$ $(i = 1, \ldots, n)$. For $k = 0, 1, \ldots,$ define $y_{k+1} = Ty_k$, where T was defined in (3.6). Then the set of limit points of the sequence $\{y_{2k} : k = 0, 1, 2, \ldots\}$ of normalized gradients is a continuum $R \subset \Sigma^*$. Moreover, for any point r in R, we have $r = T^2 r = T(Tr)$.

Proof. Let $w_0 = y_0$. For $k = 0, 1, \ldots,$ let $w_{k+1} = P_s(A; y_k)\, w_k$, where $P_s(t; y)$ was defined in (2.9). It is easily shown that $y_k = w_k/\|w_k\|$, for all k. Since $n \geq s + 1$ components of w_0 are nonzero, it follows from theorem (5.1) that at least $s + 1$ components of w_k are nonzero for $k = 1, 2, \ldots$. Hence no $w_k = \theta$.

Let $w_k = (\omega_1^{(k)}, \ldots, \omega_n^{(k)})^T$. By theorem (3.3),

$$\varphi(w_0) \leq \varphi(w_1) \leq \ldots \leq \varphi(w_k) \leq \ldots .$$

But for each k the s zeros of $P_s(t; w_k)$ lie in the interval (λ_1, λ_n) . Hence $|P_s(t; w_k)| \leq (\lambda_n - \lambda_1)^s$, for $\lambda_1 \leq t \leq \lambda_n$, and so

$$\varphi(w_k) = \frac{\|w_{k+1}\|^2}{\|w_k\|^2} = \frac{\|P_s(A; w_k)\|^2}{\|w_k\|^2}$$

$$= \frac{\sum_{i=1}^{n} [P_s(\lambda_i; w_k)]^2 [\omega_i^{(k)}]^2}{\sum_{i=1}^{n} [\omega_i^{(k)}]^2}$$

$$\leq (\lambda_n - \lambda_1)^s, \quad \text{for all } k .$$

As a monotone bounded sequence, $\{\varphi(w_k)\}$ has a limit L . Hence

$$\varphi(w_{k+1}) - \varphi(w_k) \to 0 \qquad (\text{as } k \to \infty). \tag{3.9}$$

But, by theorem (3.3),

$$\varphi(w_{k+1}) - \varphi(w_k) = \frac{\|w_{k+2}\|^2}{\|w_{k+1}\|^2} - \frac{\|w_{k+1}\|^2}{\|w_k\|^2}$$

$$= \frac{\|w_{k+2}\|^2}{\|w_{k+1}\|^2} [1 - \cos^2 \psi_k], \tag{3.10}$$

where ψ_k is the angle between w_k and w_{k+2} . Then, by (3.9) , $\cos^2\psi_k \to 1$, and $\psi_k \to 0$, as $k \to \infty$. (Since $c > 0$ in (3.3), $\psi_k \not\to \pi$.)

Now consider the set Y of unit vectors $\{y_{2k} : k = 0, 1, 2, \ldots\}$. As an infinite subset of the compact unit sphere Σ, $\{y_{2k}\}$ has limit points; let R be the set of all limit points of Y . Since $\psi_k \to 0$, as $k \to \infty$, we have $\|y_{2k+2} - y_{2k}\| \to 0$, as $k \to \infty$. Then, as Ostrowski shows on p. 203 of [9], the set R must be a continuum in the sense of (3.7).

Let r be any point of R. Then there is a subsequence $\{y_{2k_i}\}$ converging to r. Since $\|y_{2k_i+2} - y_{2k_i}\| \to 0$, we have also that $y_{2k_i+2} = T^2 y_{2k_i} \to r$. But T is a continuous transformation. Hence $T^2 y_{2k_i} \to T^2 r$, and $T^2 r = r$. Since $T^2 r = r$, we see from theorem (5.1) that $r \in \Sigma^*$. Hence $R \subset \Sigma^*$. This completes the proof of theorem (3.8).

The author has programmed a number of test cases with s = 2, to investigate the nature of the set R. In every case, R appeared to be a single point. <u>The author conjectures</u> that R <u>is</u> always <u>a</u> single <u>point in theorem</u> (3.8). So far, this has been proved only for s = 1, and we give the proof in (4.12).

The following theorem shows one way in which one might be able to prove that R consists always of a single point.

(3.11) Theorem, *Suppose in the proof of theorem (3.8) that $\varphi(w_k)$ were to converge to L so rapidly that, for some $\alpha < 1$.*

(3.12) $$0 \le \varphi(w_{k+1}) - \varphi(w_k) \le \alpha[\varphi(w_k) - \varphi(w_{k-1})], \qquad \text{for all } k .$$

Then R would consist of a single point.

Proof. If (3.12) held, then the following infinite series would be convergent:

(3.13) $$\sum_{1}^{\infty} [\varphi(w_{k+1}) - \varphi(w_k)]^{\frac{1}{2}} < \infty$$

as is seen from (3.12), by the ratio test. It is shown in (3.10) that

(3.14) $$[\varphi(w_{k+1}) - \varphi(w_k)]^{\frac{1}{2}} \sim \sin |\psi_k|, \qquad \text{as} \quad k \to \infty,$$

where ψ_k is the angle between the vectors w_k and w_{k+2}. Then, from (3.13) and (3.14), we would have

(3.15) $$\sum_{1}^{\infty} |\psi_k| < \infty .$$

Now, let $y_k = w_k/\|w_k\|$ be the unit vector in the direction of w_k. It would follow from (3.15) that

$$\sum_{k=0}^{\infty} \|y_{2k+2} - y_{2k}\| < \infty,$$

whence

(3.16) $$\sum_{k=0}^{\infty} (y_{2k+2} - y_{2k})$$

would be an absolutely convergent series of vectors. Since

$$y_{2k} = \sum_{h=0}^{k-1} (y_{2h+2} - y_{2h}) + y_0,$$

we see that the sequence $\{y_{2k}\}$ would then have one limit point. This proves the theorem (3.11).

However, the author sees no way to prove (3.12) nor the conjecture.

The following theorem proves that, whether R has one point or an infinite number, $f(x_k) \to 0$ no faster than linearly.

(3.17) Theorem. Fix s with $1 \le s \le n - 1$. Given any A in the form (2.7). Let $x_0 = (\xi_1^{(0)}, \ldots, \xi_n^{(0)})^T$ be any vector in E_n with m nonzero components. Then in the $f(x_k)$mum s-gradient method converges to 0 in the following ways:

(i) If $m \le s$, then $x_1 = \theta$, $f(x_1) = 0$, and the iteration terminates in one step.

(ii) If $s + 1 \le m$, then the convergence of $f(x_k)$ to 0 is asymptotically linear, in the sense that there exist constants c_1, c_2 depending on x_0, with

$$(3.18) \qquad 0 < c_1 \le \frac{f(x_{2k+2})}{f(x_{2k})} \le c_2 < 1 \qquad \text{for all } k .$$

Proof. We may ignore any zero components of x_0, as they remain zero throughout the iteration. We are thus minimizing $f(x)$ in E_m.

Proof of (i): If $m \le s$, then the subspace L_s defined in Sec. 2 is E_m. Hence $x_1 = \theta$ and $f(x_1) = 0$, the minimum of $f(x)$ in E_m.

Proof of (ii): That

$$\frac{f(x_{2k+2})}{f(x_{2k})} \le c_2 < 1$$

follows from the chain of inequalities (2.13). We have to prove the inequalities involving c_1.

Given x_0 with at least s + 1 nonzero components. By theorem (5.1) all other vectors x_k have at least s + 1 nonzero components, so that no $x_k = \theta$. By theorem (3.8), the normalized gradient vectors y_{2k} have as a limit set a continuum R. For each point r in R, we have $T^2 r = r$. Suppose a position vector x were such that $r = Ax/\|Ax\| \in R$. That is, x would be in the direction of $A^{-1}r$. Let x" be the result of two steps of the optimum s-gradient method applied to x. Since $T^2 r = r$, we see that x" would be in the same direction as x. Hence

$$x'' = \gamma x \quad \text{and so} \quad f(x'') = \gamma^2 f(x), \tag{3.19}$$

for some γ with $0 < \gamma = \gamma(r) < 1$.

I.e., for each point r of R there is a positive real number $\gamma(r)$ such that whenever the gradient of a vector x lies in the direction of r, then (3.19) holds.

Let C be the minimum of $\gamma(r)$ for $r \in R$. Since R is compact, the minimum is assumed and $C > 0$. Hence

$$0 < C^2 \leq \frac{f(x'')}{f(x)}, \tag{3.20}$$

for all x such that $Ax/\|Ax\| \in R$.

Now the ratio $f(x'')/f(x)$ is a continuous function of x. Let $N(R) \subset \Sigma$ be such a neighborhood of R that

$$\tfrac{1}{2}C^2 \leq \frac{f(x''}{f\ x\ f} \tag{3.21}$$

for all x with $Ax/\|Ax\|$ in N(R) . Consider the sequence $\{x_{2k}\}$. Let $z_{2k} = Ax_{2k}$, and let $y_{2k} = z_{2k}/\|z_{2k}\|$. By theorem (3.8), the $\{y_{2k}\}$ have R as a limit set. Hence there is a K such that for $k \geq K$, all y_{2k} lie in N(R) . By (3.21) then for $k \geq K$ we have

$$\tfrac{1}{2}C^2 \leq \frac{f(x_{2k+2})}{f(x_{2k})} .$$

Letting $c_1 = \frac{1}{2}C^2$ completes the proof of the theorem.

Actually we could have taken $c_1 = C^2 - \epsilon$, for any $\epsilon > 0$.

(3.22) Corollary. With the hypotheses of theorem (3.17), there exist constants d_1, d_2 with

$$0 < d_1 \leq \frac{f(x_{k+1})}{f(x_k)} \leq d_2 < 1, \qquad \text{for all } k .$$

Proof. The corollary follows from theorem (3.17), the inequalities (2.13), and the fact that $f(x_k) \searrow 0$, as $k \nearrow \infty$.

(3.23) Theorem . Fix $s \geq 1$. Let x_0 be any vector such that x_2 is parallel to x_0 in the optimum s-gradient method. In other words, $z_0/\|z_0\|$ is in the set F(A) of (4.5), where $z_0 = Ax_0$. Then

$$\frac{f(x_{k+1})}{f(x_k)} = c^2 \qquad (k = 0, 1, 2, \ldots), \tag{3.24}$$

where c^2 depends on A and on x_0 .

Remark. The import of this theorem is that, although the x_k

alternate between two fixed directions, as $k \to \infty$, the ratio (3.24) is constant for all k, and does not alternate.

Proof of (3.23). We first note from Corollary (2.23) that the theorem is true for s = 1, and that (2.22) gives a formula for c^2 in terms of the two nonzero components ζ_1, ζ_2 of z_0 .

For any fixed $s > 1$, let π be the 2-space spanned by x_0 and x_1 . Let $f_\pi(x)$ be the restriction of $f(x) = \frac{1}{2}x^T Ax$ to the subspace π . Then the vectors $x_0, x_1, x_2, \ldots$ can be shown by a geometrical argument to be the successive iterates of the optimum 1-gradient method for finding the minimum of $f,(x)$ in π, starting with x_0 . Then (3.24) for s = 1 states that

$$\frac{f_\pi(x_{k+1})}{f_\pi(x_k)} = c^2,$$

for some constant c^2 depending on the eigenvalues of f_π , Since $f_\pi(x) = f(x)$ in π, this proves the theorem for s .

Presumably theorem (3.23) could somehow be proved from theorem (2.18), just as the case s = 1 follows from (2.22).

Corollary (3.22) could also be proved from theorem (3.23).

4. Nature of the Asymptotic Directions.

We should like to characterize as well as we can the possible limiting vectors $r \in R$ of the (normalized) gradient vectors y_{2k} of theorem (3.8). Since $T^2 r = r$, for r in R, we have

$$(4.1) \qquad cr = P_s(A; Tr)\, P_s(A; r)$$
$$= Q_{2s}(A)\, r,$$

where $c > 0$ is a constant and $Q_{2s}(t)$ is the product of the two polynomials $P_s(t; Tr)$ and $P_s(t; r)$. Letting $r = (\rho_1, \ldots, \rho_n)^T$, we have

$$(4.2) \qquad c\rho_i = Q_{2s}(\lambda_i)\, \rho_i \qquad (i = 1, \ldots, n).$$

Recall from p. 44 of [12] that $P_s(t; Tr) = t^s + \ldots$ and $P_s(t; r) = t^s + \ldots$ are polynomials of degree s, each with s distinct real zeros in the open interval (λ_1, λ_n). Hence $Q_{2s}(t) = t^{2s} + \ldots$ is a polynomial of degree 2s with 2s real zeros in the interval (λ_1, λ_n), counting double zeros twice, if any. Now $c > 0$ in (4.2), which implies that for each i

$$(4.3) \qquad Q_{2s}(\lambda_i) = c > 0 \quad \underline{\text{or}} \quad \rho_i = 0 \qquad (i = 1, \ldots, n).$$

Since $Q_{2s}(t)$ vanishes for some t in (λ_1, λ_n), the equation $Q_{2s}(t) = c > 0$ can have 2, 3, 4, ... or 2s distinct real roots, which we call μ_j $(j = 1, \ldots, m)$, and number so that

$$\mu_1 < \mu_2 < \ldots < \mu_m .$$

(Here we count a multiple root of Q,,(t) = c only once.) Thus

$$Q_{2s}(\mu_j) = c \qquad (j = 1, \ldots . m) .$$

By (4.3) each λ_i for which $\rho_i \neq 0$ is one of the μ_j.

(4.4) Definition. Given x_0 . Let R be the set of limiting points of the normalized gradients $\{y_{2k} : k = 0' 1' \ldots\}$ of the optimum s-gradient method starting from x_0 . For any vector $r = (\rho_1, \ldots . \rho_n)^T$ in R, let S be the set of λ_i for which $\rho_i \neq 0$. Any such set is called an asymptotic spectrum of the optimum s-gradient method for the given x_0 . Any r in R is called an asymptotic gradient vector of the same iteration.

Note that R depends on A and x_0, and we occasionally write $R(x_0, A)$ to make the dependence explicit. Note that S is a property of r only, and only indirectly of x_0 .

-(4.5) Definition. For a given A, we define the invariant set F(A) of the optimum s-gradient method to be the set of unit vectors r such that $T^2 r = r$.

We have shown in theorem (3.8) that, for any x_0, $R(x_0, A) \subseteq F(A)$. It is never true that $R(x_0, A) = F(A)$. However, it is true that

$$F(A) = \bigcup_{x_0 \in E_n} R(x_0, A) .$$

For, if $r \in F(A)$, then $T^2 r = r$, so that $r = R(r, A)$.

(4.6) Theorem. Given $x_0 = (\xi_1^{(0)}, \ldots, \xi_n^{(0)})^T$ with $\xi_i^{(0)} \neq 0$ $(i = 1, \ldots, n)$. Assume $s < n$. Then both eigenvalues λ_1 and λ_n belong to all asymptotic spectra S of the optimum s-gradient method starting with x_0 .

Proof. Assume that λ_q $(q < n)$ is the largest eigenvalue in the asymptotic spectrum S corresponding to an asymptotic vector r of $R(x_0, A)$. The zeros of each $P_s(t; z_k)$ $(k = 0, 1, \ldots)$ lie in the open interval (λ_1, λ_n) . Hence $P_s(\lambda_n; z_k) \neq 0$ for all k . Hence $\eta_n^{(2k)} \neq 0$ for all k, where the $\eta_i^{(2k)}$ are the components of $y_{2k} = z_{2k}/\|z_{2k}\|$.

Let τ be the largest zero of $P_s(t; Tr)\, P_s(t; r)$. Since the zeros of both $P_s(t; Tr)$ and $P_s(t; r)$ lie in the open interval (λ_1, λ_q), we see that $P_s(t; Tr)\, P_s(t; r) \nearrow$, as $t \nearrow$, for $t > \tau$. Hence

$$c^2 = P_s(\lambda_q; Tr)\, P_s(\lambda_q; r) < P_s(\lambda_n; Tr)\, P_s(\lambda_n; r) .$$

-But then, by continuity,

$$P_s(\lambda_q; z_{2k+1})\, P_s(\lambda_q; z_{2k}) \leq \sigma P_s(\lambda_n; z_{2k+1})\, P_s(\lambda_n; z_{2k})$$

for some $\sigma < 1$ and all $k > K$. Since all $\eta_n^{(2k)} \neq 0$, and since $\eta_q^{(2k_j)} \to \rho_q \neq 0$, for a certain subsequence k_j, this means that $|\eta_n^{(2k_j)}| \to \infty$, as $j \to \infty$. This is impossible, since all $y^{(2k)}$ lie on the unit sphere. Hence $q = n$, and λ_n is in the asymptotic spectrum S .

The proof that λ_1 is in S is analogous.

(4.7) Theorem. Given x_0 with $\xi_i^{(0)} \neq 0$ $(i = 1, \ldots, n)$; assume that $s > n$. Let m be the number of eigenvalues in any asymptotic spectrum S of the optimum s-gradient method. Then

$$s + 1 \leq m \leq 2s .$$

Proof. Let $r \in R$ be an asymptotic gradient vector corresponding to a given S . As shown after (4.3), the asymptotic spectrum S is a subset of the set of t for which $P_s(t; Tr)\ P_s(t; r) = c$, and the number of such t is between 2 and $2s$.

However, if $m \leq s$, one step of the optimum gradient method would carry r into θ, and so r could not belong to R . Hence $s + 1 \leq m \leq 2s$.

(4.8) Theorem. Suppose $s < n$. Let $x_0 = (\xi_1^{(0)}, \ldots, \xi_n^{(0)})^T$ be any vector in E_n with exactly $s + 1$ nonzero components $\xi_i^{(0)}$. Then x_0, x_2, x_4, are all collinear vectors. That is, the normalized gradient vector $y_0 = Ax_0/\|Ax_0\|$ is in the invariant set $F(A)$ of (4.5).

Proof. Let $z_0 = Ax_0$. It will suffice to prove that $z_2 = c_0 z_0$, for some positive constant c_0 . Without loss of generality we may assume that $n = s + 1$, since the components for which $\xi_i^{(0)} = 0$ remain zero.

By (2.2)

$$z_1 = z_0 + \gamma_1 A z_0 + \ . \ . \ . + \gamma_s A^s z_0, \tag{4.9}$$

and $\gamma_1, \ldots, \gamma_s$ are so chosen that z_1 is orthogonal to z_0, Az_0, ..., $A^{s-1}z_0$. Because s + 1 components of z_0 are nonzero, the s vectors z_0, Az_0, ... $A^{s-1}z_0$ are linearly independent. Hence the set $\{z_0, Az_0, \ldots, A^{s-1}z_0\}$ forms a basis for the subspace of E_{s+1} orthogonal to z_1.

Next, z_2 is formed as a linear combination of z_1, Az_1, ..., $A^s z_1$ which is orthogonal to z_1, Az_1, ..., $A^{s-1}z_1$. Since z_2 is orthogonal to z_1, it is expressible in terms of the basis z_0, ..., $A^{s-1}z_0$:

$$z_2 = c_0 z_0 + c_1 A z_0 + \ . \ . \ . + c_{s-1} A^{s-1} z_0 \ . \tag{4.10}$$

We shall prove that $c_1 = c_2 = \ . \ . \ . = c_{s-1} = 0$.

Take the inner product of (4.10) with Az_1 :

$$z_1^T A z_2 = c_0 z_1^T A z_0 + c_1 z_1^T A^2 z_0 + \ . \ . \ . + c_{s-2} z_1^T A^{s-1} z_0 + c_{s-1} z_1^T A^s z_0 \ . \tag{4.11}$$

But $z_1^T A z_2 = z_2^T A z_1 = 0$ because z_2 is orthogonal to Az_1. And $z_1^T A z_0 = z_1^T A^2 z_0 = \ . \ . \ . = z_1^T A^{s-1} z_0 = 0$ because z_1 is orthogonal to Az_0, $A^2 z_0$, ... $A^{s-1}z_0$. And $z_1^T A^s z_0 \neq 0$, since otherwise by (4.11) z_1 would be θ . It then follows from (4.11) that $c_{s-1} = 0$.

Next, taking the inner product of (4.10) with $A^2 z_1$ and using the same argument and the fact that $c_{s\,1} = 0$, we show that $c_{s\,2} = 0$. After taking the inner product of (4.10) with $Az_1, A^2z_1, \ldots, A^{s-1}z_1$, we will have proved that $c_{s-1} = \ldots = c_2 = c_1 = 0$. Then, from (4.10), $z_2 = c_0 z_0$. That $c_0 > 0$ follows from the proof of (3.3). This completes the proof of theorem (4.8).

Theorem (4.8) implies that any $s + 1$ eigenvalues of A can be in the asymptotic spectrum for some start x_0. Moreover, any vector r with exactly $s + 1$ nonzero components can be an asymptotic gradient vector of an iteration. This extends to $s \geq 2$ the known fact for the ordinary optimum 1-gradient method in 2 dimensions that any initial gradient direction is repeated at every other step of the iteration. See the end of Sec. 2 above, or p. 214 of Ostrowski [9].

That for all s the period of the iteration in theorems (3.8) and (4.8) is 2, and not higher than 2, was a surprising fact to the author. However, the experiments of Khabaza [8] for $s = 3$ suggest the period 2.

For $s = 1$ we have $s + 1 = 2s = 2$, and then by theorem (4.7) all the vectors invariant under two steps of the optimum 1-gradient method are of the type covered in theorem (4.8). From this we can now show for $s = 1$ that the limiting set R of theorem (3.8) is actually a single point. The following is a modification of Akaike's proof in [1] of the Forsythe-Motzkin conjecture [7].

(4.12) Theorem (Akaike). Let $s = 1$. Let $y_0 = (\eta_1^{(0)}, \ldots, \eta_n^{(0)})^T$ be any vector in Σ^* with $\eta_i^{(0)} \neq 0$ $(i = 1, \ldots, n)$. Then the sequence $\{y_{2k} : k = 0, 1, \ldots\}$ of normalized gradients converges to a single point r whose spectrum is $\{\lambda_1, \lambda_n\}$. Moreover, $T^2 r = r$.

Proof. By theorem (3.8) the set of unit vectors $\{y_{2k} : k = 0, 1, \ldots\}$ has a continuum R as a limit set. By theorem (4.7), for any $r \in R$ the corresponding spectrum S of r has only 2 eigenvalues in it (for $s + 1 = 2s = 2$). Now by theorem (4.6) the two eigenvalues in S must be λ_1 and λ_n. Let r be any point of R; let $r = (\rho_1, 0, \ldots\ldots 0, \rho_n)^T$, with $\rho_1^2 + \rho_n^2 = 1$. Then $P_1(t; r) = t - \mu$, where $\mu = \lambda_1\rho_1^2 + \lambda_n\rho_n^2$. Hence $P_1(A; r)\, r = ((\lambda_1 - \mu)\rho_1, 0, \ldots, 0, (\lambda_n - \mu)\rho_n)^T$.

By the proof of theorem (3.8),

$$L = \lim_{k \to \infty} \varphi(w_k) = \varphi(r) = \|P_1(A; r)\, r\|^2, \qquad \text{since } \|r\|^2 = 1$$
$$= (\lambda_1 - \mu)^2\rho_1^2 + (\lambda_n - \mu)^2\rho_n^2,$$

or

$$L = (\lambda_n - \lambda_1)^2 \rho_1^2 \rho_n^2 . \tag{4.13}$$

Now L is a number determined by the iteration, λ_1 and λ_n are given eigenvalues, and $\rho_1^2 + \rho_n^2 = 1$. Hence the pair ρ_1^2, ρ_n^2 are determined by (4.13), up to an interchange at most. Hence the set R can have at most eight vectors in it, if all permutations of signs are considered. But then, since R is a continuum, it must consist of a single point, which we call r. Then $y_{2k} \to r$, as $k \to \infty$. This proves theorem (4.12).

Actually, if $r = (\rho_1, \ldots, \rho_n)^T$, then $Tr = (\rho_n, \ldots\ldots -\rho_1)^T$, where all components $\rho_i = 0$ for $1 < i < n$. Then $r = \lim y_{2k}$ and $Tr = \lim y_{2k+1}$, as $k \to \infty$. So, the directions of the gradient vectors z_k alternately approach the directions of r and Tr, as $k \to \infty$.

The reason we cannot extend our proof of theorem (4.12) to $s > 1$ is that the equation analogous to (4.13) involves between $s + 1$ and $2s$ unknown components of r, and we do not see how to limit r to a finite number of vectors. Even for $s = 2$, theorem (4.8) shows that all vectors r with 3 nonzero components are invariant under T^2. Prescribing the vector r to have unit length and prescribing the value of L reduce the number of free parameters in r to 1. But, so far as the author can see, there remain ∞^1 possible limiting vectors r in R.

Moreover, for an even number $s > 1$, there are asymptotic spectra containing more than $s + 1$ eigenvalues, as will now be demonstrated. We shall consider only spectra with symmetry about a midpoint. We do not know whether there are asymptotic spectra with more than $s + 1$ eigenvalues without such a symmetry.

We shall first examine possible asymptotic spectra with an even number $2q$ of eigenvalues. Let the eigenvalues in S be $a - \mu_q$, $a - \mu_{q-1}, \ldots, a - \mu_1, a + \mu_1, \ldots, a + \mu_{q-1}, a + \mu_q$, where $0 < a - \mu_q$ and $0 < \mu_1 < \ldots < \mu_q$. Let us consider unit vectors r with symmetric components $\rho_q, \ldots, \rho_1, \rho_1, \ldots, \rho_q$, corresponding to the respective -points of the spectrum.

Because of the symmetry about the point $t = a$, the orthogonal polynomials $P_{2k}(t; r)$, $P_{2k+1}(t; r)$ associated with S and the $\{\rho_i^2\}$ satisfy the conditions

$$P_{2k}(t; r) = g_k((t - a)^2), \tag{4.14}$$

where g_k is a monic polynomial of degree k;

$$(4.15) \qquad P_{2k+1}(t; r) = (t - a)\, h_k((t - a)^2),$$

where h_k is a monic polynomial of degree k .

By symmetry, the even and odd polynomials $P_k(t; r)$ are automatically orthogonal. By (4.14) orthogonality of the $P_{2k}(t; r)$ among themselves can be expressed in the form

$$(4.16) \qquad \sum_{i=1}^{q} g_j(\mu_i^2)\, g_k(\mu_i^2)\, \rho_i^2 = 0 \qquad (j, k = 0, 1, \ldots, \text{[illegible]})$$

Thus the $g_k(t)$ are themselves orthogonal polynomials over the set $\mu_1^2, \ldots, \mu_q^2$ with the weight factors $\rho_1^2, \ldots, \rho_q^2$. Moreover, $\hat{g}_k(t) = (-1)^k g_k(a^2 - t)$ are monic orthogonal polynomials over the transformed set $\hat{S} = \{a^2 - \mu_q^2, \ldots, a^2 - \mu_1^2\}$ with the same weights $\rho_1^2, \ldots, \rho_n^2$. Note that $|\hat{g}_k(0)| = |P_{2k}(0; r)|$ and that $|\hat{g}_k(a^2 - \mu_i^2)| = |P_{2k}(a \pm \mu_i; r)|$ for i = 1, ..., q . Hence $|\hat{g}_k(t)/\hat{g}_k(0)|$ has the same constant value over the set $\hat{S}$ that $|P_{2k}(t; r)/P_{2k}(0; r)|$ has over the set S .

By (4.15) the orthogonality of the P_{2k+1} among themselves can be expressed as

$$(4.17) \qquad \sum_{i=1}^{q} h_j(\mu_i^2)\, h_k(\mu_i^2)\, \mu_i^2\, \rho_i^2 = 0 \qquad (j, k = 0, 1, \ldots ; j \neq k) .$$

Thus the $\hat{h}_k(t) = (-1)^k h_k(a^2 - t)$ are monic orthogonal polynomials over the set $\hat{S} = \{a^2 - \mu_q^2, \ldots, a^2 - \mu_1^2\}$ with the different weights $\mu_1^2 \rho_1^2, \ldots, \mu_q^2 \rho_q^2$. Note that $|\hat{h}_k(0)| = |h_k(a^2)| = |P_{2k+1}(0; r)|/a$, and that

$$|\hat{h}_k(a^2 - \mu_i^2)| = |h_k(\mu_i^2)| = |P_{2k+1}(a \pm \mu_i; r)| / \mu_i .$$

Thus constancy of $|P_{2k+1}(t; r)|$ over S does not imply constancy of $|\hat{h}_k(t)|$ over 5 . The even and odd polynomials transform differently.

By means of these orthogonal polynomials $\hat{g}_k$ we can reduce the problem of the invariance of the r under two steps of the optimum 2s-gradient method over S to the problem of the invariance of an optimum s-gradient method over $\hat{S}$ in a space of half the dimension.

To be precise, the above relations imply the following result, which we do not prove.

(4.18) Theorem. If s is even and $s + 1 < 2q \leq 2s$, then the vector $r = (\rho_q, \ldots, \rho_1, \rho_1, \ldots, \rho_q)^T$ (with no $\rho_i = 0$) is in the invariant set (4.5) for the optimum s-gradient method for the diagonal matrix of 2q nonzero elements

$$\text{diag}(a - \mu_q, \ldots, a - \mu_1, a + \mu_1, \ldots, a + \mu_q)$$

if and only if the vector $\hat{r} = (\rho_1, \ldots, \rho_q)^T$ (with no $\rho_i = 0$) is in the invariant set for the optimum (s/2)-gradient method for the diagonal matrix of q nonzero elements

$$\text{diag}(a^2 + \mu_1^2, \ldots, a^2 + \mu_q^2) .$$

Moreover, when iterations exist with these invariance properties, if $z_0 = r/\|r\|$ and $\hat{z}_o = \hat{r}/\|\hat{r}\|$, then $\|z_k\| = \|\hat{z}_k\|$ for $k = 0, 1, 2, \ldots$. where z_k and $\hat{z}_k$ are the gradient vectors of the respective iterations.

We do not know a comparable theorem for odd integers s .

As an application of theorem (4.18), we can show that for any s of the form $s = 2^p$ $(p = 0, 1, 2, \ldots)$, there exist vectors with $2s$ nonzero components that are in the invariant set of some optimum s-gradient method. For $p = 0$ this is theorem (4.12), and is true for any diagonal matrix of two positive elements $\mathrm{diag}(a^2 + \mu_1^2, a^2 + \mu_2^2)$ and any vector $r = (\rho_1, \rho_2)^T$. Application of the first sentence of (4.18) leads to $s = 2$ with any matrix of form $\mathrm{diag}(b^2 + \nu_1^2, b^2 + \nu_2^2, b^2 + \nu_3^2, b_2 + \nu_4^2)$ where $b^2 + \nu_1^2 = a - \mu_2$, $b^2 + \nu_2^2 = a - \mu_1$, $b^2 + \nu_3^2 = a + \mu_1$, $b^2 + \nu_4^2 = a + \mu_2$, and corresponding vector $c(\rho_2, \rho_1, \rho_1, \rho_2)^T$. Another application of (4.18) leads to $s = 4$ with the matrix

$$\mathrm{diag}(b - \nu_4, \ldots . \; b - \nu_1, b + \nu_1, \ldots . \; b + \nu_4)$$

and corresponding vector $c' (\rho_2, \rho_1, \rho_1, \rho_2, \rho_2, \rho_1, \rho_1, \rho_2)^T$. It is clear that the process may be continued to $s = 2^p$ for any p .

Note from theorem (4.7) that $2s$ is the maximal number of nonzero components in any vector in the invariant set for an optimum s-gradient method. Our above example illustrates the maximal case.

We next consider 'symmetric asymptotic spectra with an odd number $2q + 1$ of eigenvalues $a - \mu_q, \ldots, a - \mu_1, a, a + \mu_1, \ldots, a + \mu_q$ and a corresponding symmetric vector

$$(\rho_q, \ldots, \rho_1, \rho_o, \rho_1, \ldots, \rho_q)^T$$

invariant under T^2 . Then again the orthogonal polynomials take the forms (4.14) and (4.15). The odd polynomials are still defined by the condition (4.17), but the condition (4.16) must be replaced by

(4.19) $$2 \sum_{i=1}^{q} g_j(\mu_i^2)\, g_k(\mu_i^2)\, \rho_i^2 + g_j(0)\, g_k(0)\, \rho_o^2 = 0$$

$$(j,\ k = 0,\ 1,\ \ldots\ j \neq k)\ .$$

The analog of theorem (4.18) is now stated, but not proved:

(4.20) Theorem. If s is even and $s + 1 \leq 2q + 1 < 2s$, then the vector $r = (\rho_q, \ldots, \rho_1, \rho_o, \rho_1, \ldots \rho_q)^T$ (with no $\rho_i = 0$) is in the invariant set (4.5) for the optimum s-gradient method for the diagonal matrix of $2q + 1$ nonzero elements

$$\text{diag}(a - \mu_q,\ \ldots,\ a - \mu_1,\ a,\ a + \mu_1,\ \ldots,\ a + \mu_q)$$

if and only if the vector $\hat{r} = (\rho_0/\sqrt{2}, \rho_1, \ldots \rho_q)^T$ (with no $\rho_i = 0$) is in the invariant set for the optimum (s/2)-gradient method for the diagonal matrix of $q + 1$ elements

$$\text{diag}(a^2,\ a^2 + \mu_1^2,\ \ldots\ a^2 + \mu_q^2)\ .$$

Moreover, when iterations exist with these invariance properties, if $z_0 = r/\|r_0\|$ and $\hat{z}_0 = \hat{r}/\|\hat{r}\|$, then $\|z_k\| = \|\hat{z}_k\|$ for $k = 0, 1, 2, \ldots$. where z_k and $\hat{z}_k$ are the gradient vectors of the respective iterations.

If s is odd, then the set of $2q + 1$ eigenvalues $\{a - \mu_q, \ldots, a - \mu_1, a, a + \mu_1, \ldots a + \mu_q\}$ can never be the asymptotic spectrum of an optimum s-gradient iteration.

The first two sentences are strict analogs of theorem (4.18). The third is true because $P_{2k+1}(a; r) = 0$ for all k .

The signs of the ρ_i are of no importance in theorems (4.18) and (4.20), and any ρ_i could be left alone or replaced by $-\rho_i$ independently at any place it is mentioned.

5. Singular and Derogatory Quadratic Forms; Zero Components.

Two restrictions placed on A above are really irrelevant--that A be regular and nonderogatory. If A is singular, then for some $p \geq 1$, we have $\lambda_1 = \ldots = \lambda_p = 0 < \lambda_{p+1} < \ldots < \lambda_n$. Then it follows from (2.12) that

$$\xi_i^{(k+1)} = \xi_i^{(k)}, \qquad \text{for } 1 \leq i \leq p;\ k = 0, 1, 2, \ldots,$$

while all components $\xi_i^{(k)} \to 0$, as $k \to \infty$, for $p + 1 \leq i \leq n$. On the other hand $f(x) = \frac{1}{2}x^T Ax = \sum_{i=1}^{n} \lambda_i \xi_i^2 = \sum_{i=p+1}^{n} \lambda_i \xi_i^2$. Thus $f(x)$ is minimized for all vectors in the subspace N where $\xi_1 = \ldots = \xi_p = 0$, and the gradient methods proceed from x_0 to the closest point x_∞ of N, with all $x_k - x_\infty$ and all gradients z_k located in the orthogonal complement of N .

If A is derogatory, it has multiple eigenvalues but a complete set of eigenvectors (because A is symmetric). Suppose, for example, that $0 < \lambda_1 = \lambda_2 = \ldots = \lambda_r < \lambda_{r+1} < \ldots < \lambda_n$, and suppose that

$$x_0 = (\xi_1^{(0)}, \ldots, \xi_r^{(0)}, \xi_{r+1}^{(0)}, \ldots, \xi_n^{(0)})^T .$$

Now the orthogonal basis of eigenvectors belonging to $\lambda_1, \ldots \lambda_r$ is not uniquely defined. Our preceding analysis required at various places (e.g., in the proof of (4.8) that the λ_i be distinct for each nonzero component $\xi_1^{(0)}$, but zero components $\xi_i^{(0)}$ were ignored. If any of $\xi_2^{(0)}, \ldots, \xi_r^{(0)}$ are nonzero, make an orthogonal transformation of the eigenvector basis so that x_0 takes the form

$$x_0 = ((\xi_1^{(0)2} + \ldots + \xi_r^{(0)2})^{\frac{1}{2}}, 0, \ldots, 0, \xi_{r+1}^{(0)}, \ldots, \xi_n^{(0)})^T .$$

Then drop the new zero components $\xi_2, \ldots \xi_r$ entirely, and effectively reduce A to a nonderogatory matrix A of order $n - r + 1$.

Thus, in effect, only the set and number of distinct nonzero eigenvalues of A have a real relevance to the gradient methods for quadratic functions $\frac{1}{2}x^TAx$.

Moreover, zero components of any x_k should be ignored, and the order of A reduced by unity for each zero component $\xi_i^{(k)}$ that occurs at any stage of the iteration.

If fewer than $s + 1$ components of any x_k are nonzero, then $x_{k+1} = \theta$ and the iteration terminates at once. Hence we have always insisted that at least $s + 1$ components of x_0 be nonzero. Even so, one may ask, might not enough $P_s(\lambda_i; z_h)$ be "accidentally" zero, so that for some later x_k fewer than $s + 1$ components are nonzero? The answer is negative, as the following theorem shows:

(5.1) Theorem. Assume $s + 1 < n$. Assume $\xi_i^{(k)} \neq 0$ for $i = 1, \ldots, n$. Then at least $s + 1$ components $\xi_i^{(k+1)} \neq 0$.

Proof. By (2.12), $\xi_i^{(k+1)} = P_s(\lambda_i; z_k)\, \xi_i^{(k)}$, up to a multiplicative constant that does not matter, where $P_s(t; z_k)$ is the orthogonal polynomial of degree s over the set $\{\lambda_1, \ldots, \lambda_n\}$ with weights $\left[\xi_i^{(k)}\right]^2$. We shall prove that there exist $s + 1$ eigenvalues out of the λ_i :

$$\lambda_1' < \lambda_2' < \ldots < \lambda_{s+1}' , \tag{5.2}$$

such that $P_s(\lambda'_i; z_k)\, P_s(\lambda'_{i+1}; z_k) < 0$ for $i = 1, \ldots, s$. A fortiori, $P_s(\lambda_i; z_k) \neq 0$ for $i = 1, 2, \ldots, s+1$, and the theorem will have been proved.

If the above sign-alternation property is false, then let $q < s$ be the largest integer such that we can find $\{\lambda'_i\}$ with

$$(5.3) \qquad P_s(\lambda'_i; z_k)\, P_s(\lambda'_{i+1}; z_k) < 0 \qquad \text{for } i = 1, \ldots, q-1.$$

(Clearly some $q \geq 2$ exists, or else $P_s(\lambda_i; z_k)$ would always be of one sign and hence P_s could not be orthogonal to $P_0 = 1$. Then pick $\mu_1, \ldots, \mu_{q-1}$ with

$$\lambda'_1 < \mu_1 < \lambda'_2 < \mu_2 < \ldots < \lambda'_{q-1} < \mu_{q-1} < \lambda'_q,$$

so that, if $Q(t) = (t - \mu_1) \ldots (t - \mu_{q-1})$, then $P_s(\lambda_i; z_k) Q(\lambda_i) > 0$ for all $i = 1, \ldots, n$. (We omit details of the construction.) Then

$$\langle P_s(t; z_k), Q(t)\rangle = \sum_{i=1}^{n} P_s(\lambda_i; z_k)\, Q(\lambda_i) \left[\zeta_i^{(k)}\right]^2 > 0,$$

so that P_s and Q are not orthogonal. But, since Q is of degree $q - 1 \leq s - 1$, P_s must be orthogonal to Q. This contradiction completes the proof of theorem (5.1).

References

[1] Akaike, Hirotugu: On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. Ann. Inst. Statist. Math. Tokyo 11, l-16 (1959).

[2] Arms, R. J.: On the asymptotic behavior of the method of steepest descent. Unpublished manuscript, 5 March 1957. Presented to Association for Computing Machinery, Houston, Texas, June 1957.

[3] Birman, M. Š.: Nekotorye ocenki dlja metoda naiskoreišego spuska, Uspehi Matem. Nauk (N. S.) 5, 152-155 (1950).

[4] Daniel, James W.: The conjugate gradient method for linear and nonlinear operator equations. Dissertation, Mathematics Department, Stanford University, Stanford, Calif., August 1965, 87 p.

[5] Faddeev, D. K. and Faddeeva, V. I.: Computational methods of linear algebra. Translated by Robert C. Williams. San Francisco: W. H. Freeman and Co. 1963.

[6] Forsythe, George E.: Solving linear algebraic equations can be interesting. Bull. Amer. Math. Soc. 59, 299-329 (1953).

[7] Forsythe, G. E. and Motzkin, T. S.: Asymptotic properties of the optimum gradient method (abstract). Bull. Amer. Math. Soc. 57, 183 (1951).

[8] Khabaza, I. M.: An iterative least-square method suitable for solving large sparse matrices. Comp. J. 6, 202-206 (1963).

[8a] Meinardus, Günther: Über eine Verallgemeinerung einer Ungleichung von L. V. Kantorowitsch. Num. Math. 5, 14-23 (1963).

[9] Ostrowski, A. M.: Solution of equations and systems of equations. Second edition. New York and London: Academic Press 1966.

[10] Stiefel, Eduard: Über einige Methoden der Relaxationsrechnung. Z. Angew. Math. Physik 3, l-33 (1952).

[11] Stiefel, Eduard L.: Kernel polynomials in linear algebra and their numerical applications. Pp. 1-22 of Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues, National Bureau of Standards Applied Mathematics Series 49, Washington, D. C.: U. S. Govt. Printing Office 1958.

[12] Szegö, Gábor: Orthogonal Polynomials. Revised edition. New York: American Mathematical Society 1959.