Solutions of $Ax=\lambda Bx$ [1]

M. R. Hestenes [2] and W. Karush [3]

The problem is to determine characteristic numbers and vectors for the problem $Ax=\lambda Bx$, where A, B are $n\times n$ Hermitian matrices. A generalized gradient η is defined. From a first approximation x_0, a second approximation $x_1=x_0+\alpha\eta$ is determined. Successive approximations, with appropriate alphas, converge to a solution.

1. Introduction

Let A, B, be Hermitian matrices of order n with B positive definite. Then the characteristic vectors of the equation

$$Ax=\lambda Bx \tag{1}$$

are the critical points of the "Rayleigh quotient"

$$\mu(x)=\frac{(x,Ax)}{(x,Bx)}, \qquad x\neq 0, \tag{2}$$

and the corresponding values of the quotient are the characteristic values λ. In particular the minimum (maximum) of μ is the least (greatest) characteristic value of (1). Our purpose is to discuss a method of finding the solutions of (1) that is based upon this observation and that avoids a transformation of the problem.

The method is an iterative one that may be described briefly as follows.[4] With each non-null vector x we associate a vector $\eta(x)$ that is in a certain sense the gradient of μ at x. We then pass from one approximation x to the next x' by means of the formula

$$x'=x-\alpha\eta, \qquad \alpha>0,$$

where the scalar α may depend upon x. The gradient used here is determined by the equation

$$G\eta=Ax-\mu(x)Bx, \tag{3}$$

where G is an arbitrary positive definite Hermitian matrix. In computational practice G would be selected so that its inverse G^{-1} is known (e. g., $G=I$, the identity matrix). In section 4 it will be shown that this method is convergent if the scalars $\alpha(x)$ are appropriately chosen, and in section 6 two feasible schemes for this choice will be described. In general, convergence is established only to some, possibly intermediate, characteristic value (and vector). Under special hypotheses this will be the least characteristic value (see section 5).

The method has several computational advantages. It avoids a transformation of problem (1).[5] It minimizes round-off errors by beginning each step with a new initial vector. The calculations at each stage of the iteration are simple and identical in form with those of the preceding one. The method is thereby particularly suited to high-speed automatic computing machines. However, it appears to converge too slowly to be of use for hand calculation.

When one or more characteristic vectors are known the method may be modified so as to yield a new characteristic vector (see section 7). This is achieved by appropriately altering eq. 3 for the gradient.

For arbitrary complex matrices A, B it is of interest to know when the problem

$$Cx=\lambda Dx \tag{4}$$

may be transformed to one of type (1). Several characterizations [6] are given in section 8, some of which are of computational value.

2. Preliminary Results

In this section we shall state some definitions and assemble some well-known facts on matrices. No proofs will be given.

By a vector we understand an n-tuple $x=(a_1, a_2, \ldots, a_n)$ of complex numbers. We deal with the space of such vectors over the scalar field of complex numbers. We let

$$(x,y)=a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n,\ y=(b_1, b_2, \ldots, b_n),$$

where $\bar{c}$ denotes the complex conjugate of the scalar c. Thus $(x,y)=\overline{(y,x)}$. The length of x is $[x]=(x,x)^{\frac{1}{2}}$. If C is an arbitrary matrix then

$$(x,Cy)=(C^*x,y)$$

where Cy has the usual meaning, and C^* is the conjugate transpose of C. A matrix H is Hermitian if and only if

$$H^*=H.$$

In this case (x,Hx) is a real number. We shall say

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research.

[2] Univ. of California at Los Angeles and NBS at Los Angeles.

[3] Univ. of Chicago and NBS at Los Angeles.

[4] It is an extension of one used by the authors in the case that A is real symmetric and B is the identity matrix. See for example, A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix, J. Research NBS **47**, 45 (1951) RP2227.

[5] Such a transformation may, for example, involve finding the inverse of A or B. A more feasible scheme computationally is to write $B=LL^*$ with L triangular (L^*=conjugate transpose of L). The latter method is discussed on p. 159 to 160 of Fox, Huskey, and Wilkinson, Notes on the solution of algebraic linear simultaneous equations, Quart. J. Mech. and Applied Math. p. 147 to 173 (1948).

[6] These results are closely related to some of H. Wielandt, Zur Abgrenzung der selbstadjungierten Eigenwertaufgaben. I. Raume endlicher Dimension, Math. Nachrichten **2**, No. 6, 328 (1949).

that two vectors x and y are H-orthogonal in case

$$(x,Hy)=(y,Hx)=0.$$

Two sets of vectors are H-orthogonal in case each vector of one set is H-orthogonal to each vector of the other. By orthogonality is meant I-orthoganality, with I the identity matrix.

A matrix G is positive definite in case it is Hermitian and

$$(x,Gx)>0 \qquad \text{whenever } x\neq 0.$$

Let G be positive definite. There exist positive numbers $m(G)$ and $M(G)$ such that

$$m(G)|x|^2\leq (x,Gx)\leq M(G)|x|^2. \tag{5}$$

Also, we have the inequality

$$|(x,Gy)|^2\leq (x,Gx)(y,Gy). \tag{6}$$

the equality holding if and only if x and y are linearly dependent. Further, the matrix G^{-1} is positive definite, and there exists a positive definite matrix G_1 such that $G=G_1^2$.

We turn now to problem (4), where C and D are arbitrary matrices. The number λ' is a characteristic number (root, value) of (5) in case there is a non-null vector y' such that

$$Cy'=\lambda' Dy'.$$

We allow the characteristic value $\lambda'=\infty$; in this case $Dy'=0$. We say that y' is a characteristic vector belonging to λ'. For a problem of type (1), where A is Hermitian and B is positive definite, every characteristic value is finite and real. Let

$$\lambda_1<\lambda_2< \ . \ . \ . <\lambda_k$$

be the k distinct real characteristic roots of (1), and let $\boldsymbol{L}_j=\boldsymbol{L}(\lambda_j)$ be the characteristic manifold belonging to λ_j, that is, linear subspace spanned by the characteristic vectors belonging to λ_j. Then any two subspaces belonging to distinct λ's are B- and A-orthogonal, and have only the null vectors in common. Further, every vector z has a unique decomposition of the form

$$z=z_1+z_2+ \ . \ . \ . \ z_k, \qquad z_j\epsilon \boldsymbol{L}_j.$$

For problem (1) the important extremum principle is

$$\lambda_j=\min_{x\neq 0}\ \mu(x), \qquad x \ B\text{-orthogonal to } \boldsymbol{L}_1, \ . \ . \ . \ , \boldsymbol{L}_{j-1},$$

$$\lambda_j=\max_{x\neq 0}\ \mu(x), \qquad x \ B\text{-orthogonal to } \boldsymbol{L}_k, \ . \ . \ . \ , \boldsymbol{L}_{j+1}.$$

In particular,

$$\lambda_1\leq \mu(x)\leq \lambda_k, \qquad x\neq 0.$$

3. The Gradient

The direction for which the directional derivative of the function, μ given by (2), is a maximum will now be calculated. This optimal direction will be determined relative to the inner product (x,Gy) corresponding to an arbitrary, fixed positive definite matrix G. The generality of an arbitrary inner product has computational significance as well as theoretical interest; in practice it is limited to matrices G whose inverses are known. The iteration method and convergence theorems that are to follow later depend only upon the final formula that will be obtained for the maximizing direction, not upon the derivation of the formula; the derivation is intended to suggest the motivation for the method.

For fixed vectors $x\neq 0$ and $\delta x\neq 0$, consider the function $\mu(x+\epsilon\delta x)$ for real ϵ. By a simple calculation we find that at $\epsilon=0$,

$$\frac{d\mu}{d\epsilon}=\frac{2\boldsymbol{R}\{\delta x,\xi)\}}{(x,Bx)},$$

where

$$\xi=\xi(x)=Ax-\mu(x)\ Bx, \qquad x\neq 0, \tag{7}$$

and $\boldsymbol{R}\{c\}$ denotes the real part of c. We therefore seek that vector δx for which

$$\boldsymbol{R}\{(\delta x,\xi)\}=\max, \qquad (\delta x,G\delta x)=1. \tag{8}$$

η is defined by the equation

$$G\eta=\xi=Ax-\mu\ Bx. \tag{9}$$

Then, using (6),

$$\begin{aligned}\boldsymbol{R}\{(\delta x,\xi)\}=\boldsymbol{R}\{(\delta x,G\eta)\}&\leq|(\delta x,G\eta)|\\ &\leq(\delta x,G\delta x)^{\frac{1}{2}}\ (\eta,G\eta)^{\frac{1}{2}}=(\eta,G\eta)^{\frac{1}{2}}.\end{aligned}$$

It is an easy matter to verify that $\delta x=\eta/(\eta,G\eta)^{\frac{1}{2}}$ is the unique normalized vector for which equality holds between the first and last terms above. Hence this vector is the desired solution of (8). Introducing a change in normalization for convenience, η is termed the gradient of μ (with respect to G).

We shall have occasion to use the gradient relative to a side condition

$$(\delta x,z)=0, \qquad z \text{ fixed}. \tag{10}$$

Here we wish to solve (8) relative to (10). ζ is defined by the equation

$$G\zeta=\xi+hz=G\eta+hz$$

where h is determined so that $(\zeta,z)=0$. Thus

$$h=-\frac{(\eta,z)}{(G^{-1}z,z)}. \tag{12}$$

Then, in light of (10), $(\delta x,\xi)=(\delta x,G\zeta)$. As before, it follows that the maximizing vector is proportional to ζ, and this vector is chosen to be the gradient.

More generally, with several independent side conditions

$$(\delta x, z_1)=0,\ (\delta x, z_2)=0,\ \ldots$$

is as the gradient the vector ζ is obtained, where

$$G\zeta=\xi+h_1z_1+h_2z_2+\ \ldots, \tag{13}$$

with the h's determined so that ζ is orthogonal to each of the z's.
Thus the h's are the solutions of

$$\begin{aligned} h_1(G^{-1}z_1,z_1)+h_2(G^{-1}z_2,z_1)+\ \ldots &= -(\eta,z_1) \\ h_1(G^{-1}z_1,z_2)+h_2(G^{-1}z_2,z_2)+\ \ldots &= -(\eta,z_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{14}$$

in which the determinant of the h's is nonzero, by the positive definiteness of G^{-1} and the independence of the z's.

The change in μ when we pass from a vector $x\neq 0$ to the vector $x-\alpha\eta$, will now be computed where α is some real number and $\eta(x)$ is given by (9). Assume that x is not characteristic, i. e. $\eta\neq 0$. Direct calculation leads to

$$\mu(x)-\mu(x-\alpha\eta)=\frac{(\eta,G\eta)}{(x,Bx)}f(x,\alpha), \qquad x\neq 0, \eta\neq 0, \tag{15}$$

where

$$f(x,\alpha)=\frac{\alpha(2-ps\alpha)}{1-2q\alpha+r\alpha^2} \tag{16}$$

with

$$p=\frac{(\eta,B\eta)}{(\eta,G\eta)},\ q=\frac{\boldsymbol{R}\{(x,B\eta)\}}{(x,Bx)},\ r=\frac{(\eta,B\eta)}{(x,Bx)},\ s=\mu(\eta)-\mu(x).$$

Our iteration procedure takes the following form. An initial vector x_0 is given. Then the sequence $\{x_1\}$ is determined by

$$x_{i+1}=x_i-\alpha_i\,\eta_i, \qquad \eta_i=\eta(x_i), \tag{17}$$

where the real number α_i is to be specified at each step. In order to be sure that (17) determines a well-defined sequence we must verify that $x_i\neq 0$ for every i. To this end suppose that for a given j, $x_j\neq 0$; notice that from (9)

$$(x_j,G\eta_j)=(Gx_j,\ \eta_j)=0.$$

Hence by (17)

$$(x_{j+1},Gx_{j+1})=(x_j,Gx_j)+\alpha_j^2(\eta_j,G\eta_j)>0, \tag{18}$$

since G is positive definite. Hence $x_{j+1}\neq 0$. Since $x_0\neq 0$, the sequence is well defined.

From (15) we have

$$\mu(x_i)-\mu(x_{i+1})=\frac{(\eta_i,G\eta_i)}{(x_i,Bx_i)}f_i(\alpha_i), \qquad f_i(\alpha)=f(x_i,\alpha), \tag{19}$$

this equation holding whenever $\eta_i\neq 0$.

4. General Convergence Theorems

In this section and the next will be established convergence theorems under a certain general assumption on the real sequence $\{\alpha_i\}$. In the section following these two we shall describe two effective ways of meeting these conditions. For the present, we assume that the sequence has the property that there exist real positive constants b_2 and c such that

$$0\neq\alpha_i\leq b_2,\ \text{and}\ 0<c\leq f_i(\alpha_i)\ \text{whenever}\ \eta_1\neq 0. \tag{20}$$

To simplify our discussion we wish to dispose of the trivial case in which $\eta_j=0$ for some first index j. In this instance x_j is a characteristic vector belonging to the characteristic value μ_j, and $x_i=x_j$, $i\neq j$. The results to be given in this, and the next, section are now immediately verifiable. Hence we shall proceed on the basis that

$$\eta_i\neq 0, \qquad i=0,\ 1,\ 2,\ \ldots.$$

In particular (19) holds for every i.

Theorem 1. *Suppose that the sequence* $[\alpha_i]$ *satisfies* (20). *Then* $\mu_i=\mu(x_1)$ *is a decreasing sequence that converges to a characteristic value* λ *of* (1). *Also* $\lim_{i\to\infty}(x_i,\ Gx_i)=d>0$; *in particular the lengths* $|x_i|$ *are bounded and bounded away from zero. Every accumulation point of* $[x_i]$ *is a characteristic vector in* $\boldsymbol{L}(\lambda)$.

To make the proof, we notice first that by (19) and (20) that the sequence $\{\mu_i\}$ is decreasing, since B and G are positive definite. Since this sequence is bounded from below by the minimum characteristic value λ_1, it follows that it has a limit, call it μ'. By (5), (19), (20) there is a positive constant e such that

$$t_i\leq e(\mu(x_i)-\mu(x_{i+1})), \qquad t_i=\frac{(\eta_i,G\eta_i)}{(x_i,Gx_i)}>0.$$

Hence $t_i\to 0$; in fact

$$\sum_{i=0}^{\infty}t_i<\infty. \tag{21}$$

From (18) we derive

$$(x_{i+1},Gx_{i+1})=(x_0,Gx_0)\prod_{j=0}^{i}(1+\alpha_j^2t_j).$$

It is well known that the product on the right converges if $\sum_0^\infty\alpha_j^2t_j$ does. The latter condition holds by (20) and (21). This establishes the existence of the limit d and the asserted property of $|x_i|$.

It now follows that $\eta_i\to 0$ and hence, by (9),

$$Ax_i-\mu_iBx_i\to 0. \tag{22}$$

Let y' be any limit point of $\{x_i\}$; there exists at least one. Then y' is a non-null vector which, by (22), satisfies

$$Ay'-\mu'By'=0.$$

Thus μ' is a characteristic value and y' belongs to $\boldsymbol{L}(\mu')$. This completes the proof.

Theorem 2. *Let* (20) *hold. If the sequence* $\{x_i\}$ *has an isolated accumulation point* y, *then* $\{x_i\}$ *converges to* y. *Consequently if the characteristic root* λ *of the preceding theorem is simple* (i. e., *dim* $\boldsymbol{L}(\lambda)=1$), *then* $\{x_i\}$ *converges to a characteristic vector.*

Let y be an isolated accumulation point and let $\boldsymbol{P}$ be the set of remaining accumulation points. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ be open sets with disjoint closures with y in $\boldsymbol{S}_1$ and $\boldsymbol{P}$ in $\boldsymbol{S}_2$. There is an i' such that for $i\geq i'$ x_i lies in the union of the two open sets. Let $d'>0$ be the greatest lower bound of $|u-v|$ for u in $\boldsymbol{S}_1$, v in $\boldsymbol{S}_2$. Since $\eta_i\rightarrow 0$ by the preceding proof, we may, by (20) and (17), choose an $i''\geq i'$ such that $|x_{i+1}-x_i|<\frac{1}{2}d'$ for $i\geq i''$. Hence if x_i is in $\boldsymbol{S}_1$, $i\geq i''$, then x_{i+1} is in $\boldsymbol{S}_1$. It follows that for some j, x_i is in $\boldsymbol{S}_1$ for all $i\geq j$. Thus $\boldsymbol{P}$ is null. This establishes the first conclusion of the theorem.

Let λ be a simple root. From Theorem 1 every accumulation point y must satisfy $(y, Gy)=d$. There are exactly two vectors in $\boldsymbol{L}(\lambda)$ which satisfy this condition. By the first part of the theorem $\{x_1\}$ must converge to one of them. This completes the proof.

Theorem 3. *Let* (20) *hold, and let* λ *be the characteristic value of Theorem* 1. *Then there is a sequence of vectors* $\{y_i\}$ *in* $\boldsymbol{L}(\lambda)$ *such that*

$$\lim_{i\rightarrow\infty}(x_i-y_i)=0.$$

For the proof we utilize the decomposition

$$x_i=y_i+z_i,\ y_i \text{ in } \boldsymbol{L}(\lambda) \text{ and } z_i\ B\text{-orthogonal to } \boldsymbol{L}(\lambda).$$

Now suppose z_i does not converge to zero. Then some subsequence $\{z_i'\}$ converges to $z\neq 0$. The corresponding subsequence $\{x_i'\}$ has a further subsequence $\{x_i''\}$ which converges to y in $\boldsymbol{L}(\lambda)$, by Theorem 1. By the above decomposition, the corresponding subsequence $\{y_i''\}$ converges, necessarily to a vector y'' in $\boldsymbol{L}(\lambda)$. Hence,

$$z=\lim_{i\rightarrow\infty}(x_i''-y_i'')=y-y''.$$

Thus z is both in $\boldsymbol{L}(\lambda)$ and B-orthogonal to this subspace. Hence $z=0$. This contradiction completes the proof.

It is worthy of notice that insofar as the iteration described in this paper is to be used as a practicable numerical method for finding some characteristic vector of (1), then the conclusion of the preceding theorem is as effective as the assertion that the sequence $\{x_i\}$ actually converges. For the theorem asserts that the sequence will come, and remain, within an arbitrarily small distance of some characteristic vector, this vector possibly varying with x_i.

5. Convergence to the Least Characteristic Vector

Because our iteration method is a gradient procedure which decreases $\mu(x)$, it is to be expected that under appropriate hypotheses on the problem (1) and the matrix G the sequence μ_i will converge to λ_1, the minimum characteristic value. We shall show under a rather strong assumption that such convergence will take place, and further, that the sequence $\{x_i\}$ will converge, whether or not λ_1 is simple.

In passing, we remark that although for definiteness the iteration so as to produce a decreasing sequence μ_i has been formulated a slight modification in (20) produces an increasing sequence; the change is

$$0\leq -\alpha_i<b_2 \text{ and } f_i(\alpha_i)\leq c<0.$$

The results of the previous section hold in this case, and under the forthcoming additional hypothesis of this section, convergence will take place to λ_k, the the greatest characteristic value, and to a corresponding characteristic vector.

Lemma 1. *Suppose that*

$$AG^{-1}B=BG^{-1}A. \tag{23}$$

Then problem (1) *and the problem*

$$Gx=\nu Bx \tag{24}$$

have a common complete set of characteristic vectors y_1, y_2, . . . , y_n *with* $(y_p, By_q)=\delta_{pq}=$*Kronecker delta.*

To prove this $B=H^2$ is written with H positive definite and Hermitian. Then (1) and (24), respectively, are equivalent to $H^{-1}AH^{-1}z=\lambda z$ and $H^{-1}GH^{-1}z=\nu z$ where $z=Hx$. It is easily verified that the condition (23) is equivalent to the commutativity of the Hermitian matrices $H^{-1}AH^{-1}$ and $H^{-1}GH^{-1}$. It follows by standard theory that these matrices are simultaneously reducible to diagonal form by a unitary transformation; hence they share a complete ortho-normal set of characteristic vectors z_1, z_2, . . . , z_n. The desired vectors y_p are now given by $z_p=Hy_p$.

Theorem 4. *Assume that* (20) *and* (23) *hold. For a given initial vector* x_0, *let* m *be the smallest integer* j ($j=1, 2, \ldots, k$) *for which* x_0 *is not B-orthogonal to the characteristic manifold* $\boldsymbol{L}_j$. *Then*

$$\lim_{i\rightarrow\infty}\mu_i=\lambda_m,\qquad \lim_{i\rightarrow\infty}x_i=y\neq 0 \text{ with } y \text{ in } \boldsymbol{L}_m.$$

We employ the basis of Lemma 1 to write

$$x_0=a_{01}\,y_1+a_{02}\,y_2+\ldots+a_{0n}\,y_n.$$

It is assumed that the basis has been ordered so that the first r_1 vectors span $\boldsymbol{L}_1$, the next r_2-r_1 vectors span $\boldsymbol{L}_2$, etc. (We take $r_0=0$.) By multiplying each vector of the basis by $\pm$ 1 we may assume that

$$a_{0p}\geq 0,\qquad p=1, 2, \ldots, n.$$

Let

$$\tau_1\leq\tau_2\leq\ldots\leq\tau_n$$

be the characteristic numbers of (1) corresponding to the successive vectors of the basis. Thus $\tau_1 = \ldots = \tau_{r_1} = \lambda_1$, $\tau_{r_1+1} = \ldots = \tau_{r_2} = \lambda_2$, etc. Furthermore let ν_p be the characteristic number of (24) corresponding to y_p. Thus

$$\nu_p = (y_p, \ Gy_p) > 0$$

and

$$G^{-1}By_p = \nu_p^{-1}y_p, \qquad p = 1, 2, \ldots, n.$$

Using (17), (9), (1) and the preceding equality it is found that

$$x_i = a_{i1}y_1 + a_{i2}y_2 + \ldots$$

with

$$a_{i+1,p} = a_{ip}\{1 + \alpha_i \nu_p^{-1}(\mu_i - \tau_p)\}. \tag{25}$$

Also

$$a_{ip} = (y_p, Bx_i). \tag{26}$$

Since x_0 is by hypothesis B-orthogonal to each of the subspaces $\boldsymbol{L}_1, \ldots, \boldsymbol{L}_{m-1}$, we have $a_{0q} = 0$ for $q = 1, 2, \ldots, r_{m-1}$. By (25), $a_{iq} = 0$ for every i. Hence every x_i is B-orthogonal to the same subspaces and by the extremum principle of section 2 we have

$$\mu_i \geq \lambda_m.$$

Now for each $q = r_{m-1}+1, \ldots, r_m$ consider the sequence $\{a_{iq}\}$. For $i = 0$, $a_{0q} \geq 0$, with the strict inequality holding for at least one value of q. From (25) it follows that the sequence is nonnegative and nondecreasing, since the term in braces is not less than 1 for the present range of q by (20), $\nu_p > 0$, and the last displayed inequality. The sequence is also bounded, since $|x_i|$ is, by Theorem 1. Thus

$$\lim_{i \to \infty} a_{iq} = e_q \geq 0, \qquad q = r_{m-1}+1, \ldots, r_m, \tag{27}$$

with at least one limit positive.

Let y' be an arbitrary accumulation point of $\{x_i\}$ (there is at least one). Let $\{x_i'\}$ be a subsequence converging to y'. By Theorem 1, y' is a characteristic vector of (1). In addition it must belong to $\boldsymbol{L}_m$, for otherwise

$$a'_{1q} = (y_q, Bx_i') \to (y_q, By') = 0,$$

contrary to (27). It now follows that $a'_{ip} \to 0$ for p outside the range of q, as in (27). Thus

$$y' = \lim_{i - \infty} x_i' = \sum_{r_{m-1}+1}^{r_m} e_q y_q \neq 0 \text{ in } \boldsymbol{L}_m.$$

Since y' was an arbitrary limit point we have the desired convergence of x_i to a vector y in $\boldsymbol{L}_m$. Finally,

$$\lim_{i \to \infty} \mu_i = \lim_{i \to \infty} \frac{(x_i, Ax_i)}{(x_i, Bx_i)} = \frac{(y, Ay)}{(y, By)} = \lambda_m.$$

Corollary. The condition (23) *is satisfied if* (1) $G = B$, *or* (2) $G = I$ *and* $AB = BA$.

This is easily verified. The case $G = B = I$ with A real symmetric was studied in greater detail in the paper by the present authors referred to earlier.

6. Construction of the Sequence $\{\alpha_i\}$

We shall describe two methods of constructing this sequence so that condition (20) is satisfied.

Lemma 2. Let the real number b_2 satisfy

$$0 < b_2 < \frac{2}{\lambda_k - \lambda_1} \cdot \frac{m(G)}{M(B)}, \tag{28}$$

where $\lambda_k - \lambda_1$ is the spread of the characteristic values of (1) *and the other quantities are defined by* (5). *Then there is a constant $c_1 > 0$ such that for every $x \neq 0$ with $\eta \neq 0$ we have*

$$f(x, \alpha) \geq c_1 \alpha \quad \textit{for} \quad -b_2 \leq \alpha \leq b_2.$$

Let x and α satisfy the required conditions. Assume further that $\alpha \neq 0$. Then we may write (16) in the form

$$\frac{f(x,\alpha)}{\alpha} = \frac{(2 - ps\alpha)}{(x - \alpha\eta, B(x - \alpha\eta))/(x, Bx)}. \tag{29}$$

Now

$$|ps\alpha| \leq \frac{M(B)}{m(G)}(\lambda_k - \lambda_1)b_2 \equiv b_3 < 2,$$

where we have used (5) and the extremum property of λ_1 and λ_k. Thus the numerator on the right side of (29) exceeds the positive number $(2 - b_3)$. Our proof will be complete if we can show that the corresponding (positive) denominator is bounded uniformly in x and α. By (5) it is sufficient to show $|\eta|/|x|$ is bounded. But this is an immediate consequence of (9).

As a consequence of Lemma 2 we have the following result.

Theorem 5. Let the sequence $\{\alpha_i\}$ of real numbers be such that

$$0 < b_1 \leq \alpha_i \leq b_2 \tag{30}$$

for constants b_1, b_2 with the latter as in (28). *Then this sequence satisfies condition* (20).

Our second method of prescribing $\alpha = \alpha(x)$ stems from the idea of maximizing $f(\alpha) = f(x, \alpha)$ as a function of α, hence choosing α as a zero of $f'(\alpha)$. A simple calculation leads to

$$f'(\alpha) = 2\frac{(pqs - r)\alpha^2 - ps\alpha + 1}{(1 - 2q\alpha + r\alpha^2)^2}, \qquad x \neq 0, \eta \neq 0. \tag{31}$$

A function $\alpha(x)$ is now defined as follows. Choose an arbitrary fixed positive constant b_4. Let

$$\alpha(x) = \begin{cases} \text{first zero of } f'(\alpha) \text{ on } 0 \leq \alpha \leq b_4, \\ b_4, \text{ if no such zero exists.} \end{cases} \quad (x \neq 0, \eta \neq 0).$$

This function is computationally simple; its construction involves only the solution of a quadratic equation and a comparison of numbers.

Theorem 6. *For a given constant* $b_4>0$ *and a given initial vector* $x_0\neq 0$ *determine the sequence* $\{\alpha_i\}$ *by means of the iteration formula* (17) *and the equation*

$$\alpha_i=\begin{cases}\alpha(x_i) & \text{if } \eta_i\neq 0,\\ b_4 & \text{if } \eta_i=0,\end{cases}$$

where $\alpha(x)$ *is given by* (32). *Then this sequence satisfies condition* (20).

For the proof several properties of the function (32) are established. The coefficients of the quadratic expression in the numerator in (31) are uniformly bounded in x. This is a consequence of (9) and (5). It follows that its zeros are uniformly bounded away from 0. Hence, there is a number $b_3>0$ such that

$$b_3\leq\alpha(x)\leq b_4.$$

Thus the first condition of (20) is fulfilled. Now $f(\alpha)$ is a non-decreasing function on $0\leq\alpha\leq\alpha(x)$; for, by $f'(0)=2$ and (32) its derivative is non-negative on this interval. Select $b_2<b_3$ so that (28) holds. Then

$$f(\alpha(x))\geq f(b_2)\geq c_1b_2$$

uniformly in x. From this we see that the second condition of (20) also holds. The proof is complete.

7. Obtaining Further Characteristic Vectors.

Suppose that a characteristic vector y' with characteristic value λ' is known. We propose to show how the preceding iteration scheme may be modified so as to secure a new, independent characteristic vector. The procedure will be to start with an initial vector x_o which is B-orthogonal to y' and maintain this orthogonality at each step of the iteration.

Thus, let

$$z=By' \tag{33}$$

and suppose we have a vector $x\neq 0$ such that $(x,z)=0$. We wish our next approximation $x-\alpha\zeta$ to be orthogonal to z, i. e., we require

$$(\zeta, z)=0. \tag{34}$$

In order to select the direction ζ in an optimal manner, according to section 3, it is chosen proportional to the solution of (8) with the side condition (10), z as in (33). We thereby determine ζ by (11) and (12). Notice that by (11),

$$(\zeta, G\zeta)=(\zeta,\xi),\qquad (x, G\zeta)=0 \tag{35}$$

using (34) and (7). Now suppose $\zeta=\zeta(x)\neq 0$. Then a straightforward calculation using the first equation of (35) shows that equations (15) and (16) are valid with η everywhere replaced by ζ.

The iteration formula (17) is now replaced by

$$x_{i+1}=x_i-\alpha_i\zeta_i \tag{36}$$

with the corresponding formula (18) valid by the second equation of (35). Our present sequence however has the additional property

$$(x_i, By')=0. \tag{37}$$

An examination of section 4 shows that with one necessary verification, to be remarked on soon, the three theorems of that section remain valid. We may now add, however, that the characteristic accumulation vectors y in Theorem 1 and 2 are B-orthogonal to y', and that the vectors y_i of Theorem 3 have the same property. The required verification is to establish the equivalence of $\zeta(x)=0$ with $\eta(x)=0$ and $\zeta_i\rightarrow 0$ with $\eta_i\rightarrow 0$. Here, of course, η is given by (11). We shall prove only the second equivalence; this will suggest the proof of the first. That $\zeta_i\rightarrow 0$ when $\eta_i\rightarrow 0$ is immediate from (9), (11) and (12). Suppose $\zeta_i\rightarrow 0$. We note first that $(\xi_i,y')=0$; this follows from (7), (37) and the fact that y' is characteristic. Hence by (11)

$$(G\zeta_i,y')=h_i(By',y'),$$

where h_i of (12) has the obvious meaning. Hence $h_i\rightarrow 0$. It follows by (11) that $\xi_i\rightarrow 0$ and hence $\eta_i\rightarrow 0$, as desired.

The constructions of section 6 remain valid under the present iteration (36). It is only necessary to verify the uniform boundedness of $|\zeta|/|x|$, $x\neq 0$, $\zeta\neq 0$. This is an immediate consequence of (11), (12) and (9).

To maintain the validity of Theorem 4 of section 5 in the present context it is not necessary to modify the iteration procedure from (17) to (36). The earlier method is adequate. For, by (25), we see that if the initial vector x_0 is B-orthogonal to y', so is each vector x_i, and hence the limit vector y.

If several characteristic vectors $y', y'', \ldots$ with characteristic values $\lambda', \lambda'', \ldots$ are known then the iteration (36) is to be used with ζ determined by (13) and (14), where $z_1=By'$, $z_2=By''$, ... The resulting sequence $\{x_i\}$ will be B-orthogonal to the known characteristic vectors and hence all limit vectors will have this property. It may be easily verified that the preceding remarks concerning the validity of the results of the previous sections remain in force.

The iteration $x-\alpha\zeta$ determined by (13) and (14) is theoretically equivalent to the following procedure. First form the vector $x'=x-\alpha\eta\equiv x-\alpha G\text{--}1\xi$, and then determine $k_1, k_2, \ldots$ so that

$$x''=x'+k_1G^{-1}z_1+k_2G^{-1}z_2+\ \ldots \tag{38}$$

is B-orthogonal to y', y'', ... It is easy to verify that $x''=x-\alpha\zeta$. However the procedure just described has the computational advantage that the vector x_i formed at each stage is accurately B-orthogonal to the known characteristic vectors. In the alternative procedure (36), B-orthogonality of x_i may be gradually lost through round-off errors (although this may be remedied by the additional work of occasional B-orthogonalization).

It would also be extremely convenient for the determination of the k's in (38) (or the h's in (13)) to have $(G^{-1}z_j,z_m)=0$ for $j\neq m$. This may be achieved by successively orthogonalizing as additional characteristic vectors are accumulated. Suppose y'' has been calculated with y' known. Then with $z_1=By'$, define z_2 by

$$z_2=By''+lz_1$$

with l chosen so that $(z_2,G^{-1}z_1)=0$. Then the G^{-1}-orthogonal set (z_1,z_2) may be used in (38) (or (13)). Suppose now that a third independent characteristic vector y''' is calculated. Put

$$z_3-By'''+l_1z_1+l_2z_2$$

with l_1, l_2, chosen so that $(z^3,G^{-1}z_1)=(z_3,G^{-1}z_2)=0$; this determination is simplified by the G^{-1}-orthogonality of (z_1,z_2). The new G^{-1}-orthogonal set $(z_1,z_2,\ z_3)$ may now be used in (38). The extension to more vectors is clear.

8. Problems Equivalent to (1)

We leave now the calculation of characteristic vectors and values and raise the question of when a general problem of the type

$$Cx=\lambda Dx,\text{ with }|C-\lambda D|\neq 0\text{ in }\lambda, \tag{39}$$

is equivalent to one of type (1), that is, one with A Hermitian and B positive definite. For the moment we impose no additional conditions on the complex matrices C and D. Clearly (39) has at most n characteristic roots (including the possible real value $\lambda=\infty$), where n is the order of the matrices.

Consider a second problem

$$Rx=\lambda Sx. \tag{40}$$

By asserting that problems (39) and (40) are equivalent we mean that there is a nonsingular matrix K and one-to-one correspondence between the distinct characteristic values of (39) and those of (40) with the following property: if λ' of (39) corresponds to λ'' of (40), then y' is a characteristic vector of (39) belonging to λ' if and only if $y''=Ky'$ is a characteristic vector of (40) belonging to λ''. If b is a number such that $|C-bD|\neq 0$, then (39) is equivalent to $(C-bD)x=(\lambda-b)Dx$, and hence to

$$Ex=\nu Fx,\text{ with }E=D,F=C-bD,\nu=\frac{1}{\lambda-b},\quad |F|\neq 0. \tag{41}$$

Notice that $\nu=\infty$ is not a characteristic value of (41).

Lemma 2. Let λ_j, $j=1,\ 2,\ .\ .\ .\ ,\ k$, *be the distinct characteristic values of* (39) *and let* $\boldsymbol{H}$ *be the space spanned by the spaces* $\boldsymbol{L}_j=\boldsymbol{L}(\lambda_j)$. *Then*

$$\sum_{j=1}^{k}\dim \boldsymbol{L}(\lambda_j)=\dim \boldsymbol{H}.$$

It is sufficient to make the proof for the equivalent problem (41), where $\nu_j=1/(\lambda_j-b)$ and $\boldsymbol{L}_j=\boldsymbol{L}(\nu_j)$. It may be assumed that the characteristic values so ordered that if $\nu=0$ is a characteristic value, then $\nu_1=0$. Let $\boldsymbol{H}_j$ be the space spanned by $\boldsymbol{L}_1,\boldsymbol{L}_2,\ldots,\boldsymbol{L}_j$. Any two spaces $\boldsymbol{L}_j$ have only the null vector in common. It follows that the statement

$$\sum_{l=1}^{j}\dim \boldsymbol{L}_l=\dim \boldsymbol{H}_j \tag{42}$$

is valid for $j=2$. Assume that (42) holds for $j=m<k$; we shall show that it is then valid for $j=m+1$. If (42) were false for $j=m+1$, then there would exist a vector $y_{m+1}\neq 0$ of $\boldsymbol{L}_{m+1}$ such that

$$y_{m+1}=y_1+y_2+\ .\ .\ .\ +y_m,\qquad y_l\epsilon \boldsymbol{L}_l,$$
$$l=1,2,\ .\ .\ .\ ,m.$$

with not all the terms on the right null. From $Ey_l=\nu_lFy_l$ is obtained

$$Ey_{m+1}=F(\nu_1y_1+\ .\ .\ .\ +\nu_my_m).$$

But $Ey_{m+1}=\nu_{m+1}Fy_{m+1}$. Since $|F|\neq 0$,

$$\nu_{m+1}y_{m+1}=\nu_1y_1+\ .\ .\ .\ +\nu_my_m,$$

so that

$$\left(\frac{\nu_1}{\nu_{m+1}}-1\right)y_1+\ .\ .\ .\ +\left(\frac{\nu_m}{\nu_{m+1}}-1\right)y_m=0,$$

since $\nu_{m+1}\neq 0$. But (42) holds for $j=m$. Hence each term on the left above must vanish. At least one y_l, say y'_l, is not null. Then $\nu_{m+1}=\nu'_l$, $l'\leq m$. This contradiction completes the proof.

From the lemma it is clear that the condition

$$\sum_{j=1}^{k}\dim \boldsymbol{L}_j=n \tag{43}$$

for (39) is equivalent to the assertion that every vector may be written uniquely, apart from order of terms, as a sum of characteristic vectors belonging to distinct characteristic values.

Our main theorem is the following.

Theorem 8. For a problem (39) *each of the following conditions implies the other.*

I. *Problem* (39) *is equivalent to one of type* (1).

II. *The characteristic values of* (39) *are real and* (43) *holds.*

III. *There exists a positive definite matrix P such that CPD* is Hermitian.*

That II follows from I is a consequence of the well-known fact that for problem (1) condition II is valid. To show that II implies III we construct the nonsingular matrix Y whose columns are the n linearly independent (by Lemma 2) characteristic vectors of (41). Thus $EY=FY\Lambda$, where Λ is a real diagonal matrix comprising the characteristic roots of (41). Hence $\Lambda=Y^{-1}F^{-1}EY$ is Hermitian, that is

$$Y^{-1}F^{-1}EY=Y^*E^*F^{*-1}Y^{*-1}.$$

Hence

$$EPF^*=FPE^*, \qquad P=YY^*. \tag{44}$$

From the definition of E and F in (41) is obtained

$$DPC^*=CPD^* \tag{45}$$

as desired.

Finally, to show that III implies I we observe first that (45) implies (44). Problem (41) is equivalent to

$$EPF^*z=\nu FPF^*z, \qquad x=PF^*z,$$

which is of type (1). This completes the proof.

Corollary. The following condition may be added to those of Theorem 8.

IV. *There exists a positive definite matrix P such that C^*PD is Hermitian.*

For the proof we need only observe that II holds for (39) if and only if it holds for

$$C^*x=\lambda D^*x,$$

and then apply III to this problem.

If our theory is limited to real vectors over the field of real numbers, then the following specializations occur. The matrices A and B are to be taken real symmetric and the matrices C and D are to be taken real. The matrix P of III and IV is real symmetric and the condition of reality in II is superfluous.

If the matrix P of III or IV is known then the transformation of (39) to (1) involves only matrix multiplications and hence is computationally feasible. For example, in case of III with $|D|\neq 0$ we write

$$CPD^*z=\lambda DPD^*z, \qquad x=PD^*z.$$

From the solutions z of this problem we obtain the solutions x of (39) by a direct matrix transformation. If $|C|\neq 0$, we write

$$DPC^*z=\nu CPC^*z, \qquad \lambda=\frac{1}{\nu}, \qquad x=PC^*z.$$

In the case of IV with, say $|D|\neq 0$, we write

$$D^*PCx=\lambda D^*PDx$$

which is of type (1) with solutions exactly those of (39). In the case of either III or IV with $|C|=|D|=0$, we first transform to (41) and then apply the preceding technique indicated for $|D|\neq 0$.

Los Angeles, September 11, 1950.