数理解析研究所講究録
第 483 巻 1983 年 27-39

A MINIMIZATION METHOD FOR SIMULTANEOUS COMPUTATION OF SEVERAL EIGENVECTORS

Shigeru Ando

§1. Introduction

In this paper we propose an iterative method for copmputing simultaneously several eigenvectors with the least eigenvalues based on the notion of function minimization. The problem we treat is $A x = \lambda B x$ with the matrix A symmetric and the matrix B symmetric positive definite, where the large dimensionality and sparceness of A and B inhibits those methods which require modifications of the whole entries of A and B. This kind of problems arise naturally from finite-element discretization of linear oscillation of continuums.

As we are not solving linear systems of equation in the iteration processes, band structure of A and B is out of our concern. Operations needed in the iterations are, matrix-by-vector multiplications, inner-product operations and small size eigen-computations whose dimensionality does not exceed twice the number of eigenpairs to be found out.

We proceed the iteration seeking to minimize a criterion function J, a generalization of the Rayleigh quotient, which applies to linear subspaces rather than to single vectors. So, our method is another kind of "subspace iteration method", althogh that word has so far been confined to that method based on inverse iteration due to Bathe and Wilson [1].

This criterion function is one that has been used extensively in the field of statistical discriminant analysis as the scatter

criterion of extracted features, and part of the results in §2 is, though in somewhat different form, found in Fukunaga [2]. An interesting relationship between the optimal Bayes features and the scatter criterions applied to general nonlinear features has been reported in Fukunaga and Ando [3].

§2. The criterion J .

We denote by V the n-dimensinal real linear space of all column n-vectors, and by $\langle ., . \rangle$ the inner product in V . We assume that A and B are symmetric n-by-n matrices and B is positive definite. In the subsequent discussions we shall simply refer as "eigenvalue" and "eigenvctor" to the solution λ, x of the generalized eigenproblem $A x = \lambda B x$.

When $x_1, x_2, .., x_k$ are members of V, we denote by $A(x_1,..,x_k)$ and $B(x_1,..,x_k)$ the k-by-k matrix whose i-j component is $\langle Ax_i, x_j \rangle$ and $\langle Bx_i, x_j \rangle$ respectively. Obviously, $A(x_1,..,x_k)$ and $B(x_1,..,x_k)$ are symmetric and, if $x_1,..,x_k$ are linearly independent, $B(x_1,..,x_k)$ is positive definite. We also write A(X), B(X) for $A(x_1,..,x_k)$, $B(x_1,..,x_k)$ with the n-by-k matrix $X = [x_1,..,x_k]$. Then, in matrix notations, $A(X) = {}^tX A X$, $B(X) = {}^tX B X$.

For U, a k-dimensional linear subspace of V, we define

$$J(U) = \mathrm{trace}(B(x_1,..,x_k)^{-1} A(x_1,..,x_k)) ,$$

where $x_1,..,x_k$ is a basis of U. Note that J is just the Rayleigh quotient of x_1 when $k = 1$.

That this J does not depend on the choice of basis could be stated as the following PROPOSITION 1. The proof is straightforward and so is omitted.

PROPOSITION 1.

Let $Y = X H$ where X and Y are n-by-k rank k matrix and H is a k-by-k non-singular matrix. Then

$$\mathrm{trace}(B(X)^{-1} A(X)) = \mathrm{trace}(B(Y)^{-1} A(Y)) \quad .$$

The following two PROPOSITIONs are essential not only in the proof of the THEOREM 1 but also in establishing minimization algorithms.

PROPOSITION 2.

When X is a n-by-k rank k matrix,

$$d\,\mathrm{trace}(B(X)^{-1} A(X)) = 2\,\mathrm{trace}({}^t dX\,(A X - B X B(X)^{-1} A(X))\, B(X)^{-1}).$$

Proof

Let $P = A(X)$, $Q = B(X)$.

As $dP = d(Q\,Q^{-1} P) = dQ\,Q^{-1} P + Q\,d(Q^{-1} P)$,

$d(Q^{-1} P) = Q^{-1}\,(dP - dQ\,Q^{-1} P)$.

Since 'trace' is a linear operation,

$$d\,\mathrm{trace}(Q^{-1} P) = \mathrm{trace}(d(Q^{-1} P))$$
$$= \mathrm{trace}((dP - dQ\,Q^{-1} P)\,Q^{-1}).$$

While, we have

$$d\,A(X) = {}^t dX\,A\,X + {}^t X\,A\,dX = 2\,{}^t dX\,A\,X$$

and similarly

$$d\,B(X) = 2\,{}^t dX\,B\,X$$

which are to be substituted into the last equation.

PROPOSITION 3.

Any k-dimensional subsapace U has such a basis $x_1,..,x_k$ that $B(x_1,..x_k)$ is the identity and $A(x_1,..,x_k)$ is diagonal.

Proof

Let $\tilde{A} = A(y_1,..,y_k)$, $\tilde{B} = B(y_1,..,y_k)$ with some basis $y_1,...y_k$ of U. As $\tilde{A}$ and $\tilde{B}$ are symmetric and $\tilde{B}$ is positive definite, we can choose such k linearly independent column k-vecteos $s_1,..,s_k$ that

$$\langle\tilde{A}s_i,s_j\rangle = \mu_i \delta_{ij} \quad \text{and} \quad \langle\tilde{B}s_i,s_j\rangle = \mu_i \delta_{ij}$$

as the solution of the "smaller" eigenproblem $\tilde{A}\, s_i = \mu_i \tilde{A}\, s_i$.

Let $x_i = [y_1,..,y_k]s_i$ for i = 1,..,k,

then

$$\begin{aligned}\langle Ax_i,x_j\rangle &= \langle A[y_1,..,y_k]s_i,[y_1,..,y_k]s_j\rangle \\ &= \langle {}^t[y_1,..,y_k]A[y_1,..,y_k]s_i,s_j\rangle \\ &= \langle\tilde{A}s_i,s_j\rangle = \mu_i \delta_{ij} .\end{aligned}$$

Similarly

$$\langle Bx_i,x_j\rangle = \langle\tilde{B}s_i,s_j\rangle = \delta_{ij} .$$

The purport of Propositin 3 is what is known as "Rayleigh-Ritz analysis". We shall subsequently call "A-B-eigenbasis" or simply "eigenbasis" of U, such a basis of U that possesses the property treated in PROPOSITION 3. You should note here that the diagonal elements of the diagonal matrix $A(x_1,..,x_k)$ when $x_1,..,x_k$ is an eigenbasis of U are the Rayleigh quotients of $x_1,..,x_k$.

THEOREM 1.

J(U) is stationary if and only if U is spanned by some k linearly independent eigenvectors, and then its value is just the sum of the corresponding eigenvalues.

Proof

As trace(^{t}P Q) is equal to the sum of the products of corresponding entries of matrix P and Q, trace(^{t}dP Q) vanishes if

and only if Q is the zero matrix. PROPOSITION 2, hence, tells us that $d\,\mathrm{trace}(B(X)^{-1}A(X)) = 0$ if and only if

$$A\,X - B\,X\,B(X)^{-1}A(X) = 0 .$$

If we choose an A-B-eigenbasis $x_1,..,x_k$ of U then

$$B(x_1,..,x_k)^{-1}A(x_1,..,x_k) = \mathrm{diag}(\lambda_1,..,\lambda_k) ,$$

where $\lambda_1,..,\lambda_k$ are the Rayleigh quotients of $x_1,..,x_k$. So the above reads

$$A\,x_j - \lambda_j B\,x_j = 0 \qquad \text{for} \quad j=1,..,k .$$

Thus, $d\,J(U) = 0$ means that $x_1,..,x_k$ are eigenvectors with eigenvalues $\lambda_1,..,\lambda_k$,

and then

$$J(U) = \mathrm{trace}(\mathrm{diag}(\lambda_1,..,\lambda_k)) = \lambda_1+\lambda_2+...+\lambda_k .$$

COROLLARY

J(U) takes on its minimum value when U is spanned by the k eigenvectors with the k least eigenvalues and the minimum value is just the sum of the k least eigenvalues.

Proof Immediate from THEOREM 1.

Now we have reached to the conclusion that, to find the subspace spanned by the k eigenvectors with the least k eigenvalues, it suffices to minimize J . Our task now is to find out an algorithm for minimization of J .

§3. Minimization algorithm

Since $\mathrm{trace}({}^t dX\,R)$ is the sum of the products of the corresponding entries of dX and R, PROPOSITION 2 tells us that, in terms of some basis $x_1,..,x_k$ of U and the matrix $X=[x_1,..,x_k]$, the gradient (precisely half the gradient but as we

concern only to the direction it makes no difference) of J(U) is represented by

$$R = (A\,X - B\,X\,B(X)^{-1}\,A(X))\,B(X)^{-1} \quad .$$

By that R represents the gradient of J , we mean that slight modification of U in the steepest descend direction of J(U) is realised by slight displacement of each basis of U proportional to each column vector of R .

In particular, if we choose an eigenbasis $x_1,..,x_k$ of U, the gradient reduces to

$$R = A\,X - B\,X\,\mathrm{diag}(\lambda_1,..,\lambda_k) \ ,$$

where $\lambda_1,..,\lambda_k$ are the Rayleigh quotients of $x_1,..,x_k$. One can write down this with each column vector r_j of R as the following.

$$r_j = Ax_j - \langle Ax_j,x_j\rangle\, Bx_j \qquad \text{for } j=1,..,k \quad .$$

This may be expressed that the gradient of J(U) is represented by the matrix whose each column vectors are the residuals (in the eigenproblem sense) of the each elements of an eigenbasis of U. With this terminology, THEOREM 1 may be re-expressed that J(U) is stationary if and only if the residual of each element of the eigenbasis of U vanishes.

Our minimization algorithm is to be worked out based on this information. For example, if the steepest-descend method were to be adopted, the above information is interpreted as the following steps 0 - 4.

0) Choose some lineraly independent vectors $y_1,..,y_k$ as the basis of the starting subspace U.

1) Find an eigenbasis $x_1,..,x_k$ of U by solving the k-dimensional eigenproblem introduced in the proof of PROPOSITION 3 .

2) Calculate the residuals $r_1,..,r_k$ of $x_1,..,x_k$ and if all

of them are sufficiently small then end.

3) Let $[x_1^*,..,x_k^*] = [x_1,..,x_k] + t\,[r_1,..,r_k]$, where t is to be chosen so as to minimize

$J = \mathrm{trace}(B(x_1^*,..,x_k^*)^{-1}\, A(x_1^*,..,x_k^*))$.

4) Let the next U be that subspace spanned by $x_1^*,..,x_k^*$ and go to 1.

Although we shall not adopt the steepest-descend method in itself, we shall use the above steps 0 to 4 as the point of departure, revising them step by step. First, step 3 is to be revised as follows.

3′) Let $[x_1^*,..,x_k^*] = [x_1,..,x_k,r_1,..,r_k]\,C$, where the 2k-by-k matrix

$$C = \begin{bmatrix} c_{1,1},\cdots, c_{1,k} \\ .. \quad .. \quad .. \\ c_{2k,1},\cdots, c_{2k,k} \end{bmatrix}$$ is to be chosen so that J be minimized.

It may be obvious that, with this revised step, one goes closer to the goal in each iteration than with the original step 3. The main purpose of this revision, however, is to make simpler the minimization task in each iteration, which may sound paradoxical since the number of parameters to be determined increases so much. In fact the revised minimization step 3′ is itself an eigenproblem, and so is handled in a uniform manner.

Since the purport of step 3′ is to seek for a k-dimensional subspace which minimizes J within the 2k-dimensional subspace spanned by x_j′s and r_j′s, we can make use of THEOREM 1 and its COROLLARY in the inverse direction. Namely, we can solve this problem by solving 2k-dimensional eigenproblem when 2k is small

enough to apply some known full-matrix algorithm.

Let $\tilde{A} = A(x_1,..,x_k,r_1,..,r_k)$ and $\tilde{B} = B(x_1,..,x_k,r_1,..,r_k)$. Then $A(x_1*,..,x_k*) = \tilde{A}(C)$ and $B(x_1*,..,x_k*) = \tilde{B}(C)$ and $J(x_1*,..,x_k*) = \mathrm{trace}(\tilde{B}(C)^{-1}\ \tilde{A}(C))$.

As is known from the THEOREM 1 and its COROLLARY, J is minimized when and only when $C = [c_1,..,c_k]$ where $c_1,..,c_k$ are the eigenvectors with the least k eigenvalues $\lambda_1<..<\lambda_k$ of the eigenproblem

$$\tilde{A}\ c_j = \lambda_j \tilde{B}\ c_j .$$

This 2k-dimensional eigenproblem can be handled with, say, the generalized Jacobi method [1].

In addition, with $x_1*,..,x_k*$ thus constructed, the next step 1 is no longer necessary as $x_1*,..,x_k*$ itself make an eigenbasis of the new U, since

$$\langle Ax_i*,x_j*\rangle = \langle \tilde{A}c_i,c_j\rangle = \mu_i \delta_{ij} \quad \text{and}$$
$$\langle Bx_i*,x_j*\rangle = \langle \tilde{B}c_i,c_j\rangle = \delta_{ij} \quad .$$

Another revision is one that is motivated from the congruent-gradient method. As is well known, steepest-descend method is in general rather poorly convergent, which is still the case in our revised step 3′. So we shall adopt as the search direction, not $R = [r_1,..,r_j]$, but $P = [p_1,..,p_k]$ as is defined below, and let $\tilde{A}$ and $\tilde{B}$ be $A(x_1,..,x_k,p_1,..,p_k)$ and $B(x_1,..,x_k,p_1,..,p_k)$ respectively.

We define $p_j = r_j + \beta_j \Delta x_j$ for j=1,..,k where Δx_j is the last correction of x_j , i.e. the difference between the current and the previous x_j, and β_j is to be determined so that p_j be orthogonal to Δr_j, the difference between the current and the previous r_j. Namely, $\beta_j = -\langle r_j,\Delta r_j\rangle/\langle \Delta x_j,\Delta r_j\rangle$. (You know this can be applied only from the second iteration, so the first p_j′s

are to be just the r_j's.)

According to the mean-value theorem, Δr_j is an approximation of the current Hessian by the last correction, and so the $P = [p_1,..,p_k]$ thus determined is an approximation of the congruent-gradient direction.

The next revision is, as it should be, the deflation. Convergence speeds of $x_1,..,x_k$ are not equal. In our experiences, smaller eigenvalues and their corresponding eigenvectors converge faster in most cases. At the point when some x_j has almost converged, p_j is so small that positive-definiteness of $\tilde{B} = B(x_1,..,x_k,p_1,..,p_k)$ becomes numerically uncertain and so the generalized-Jacobi process might fail. So, as soon as some x_j has reached to some prescribed level of convergence, p_j should be eliminated from the construction of $\tilde{A}$ and $\tilde{B}$. Then the dimensionality of $\tilde{A}$ and $\tilde{B}$ reduces by one, serving also to save the subsequent processing time.

The last account is about the choice strategy of the starting subspace. As we are treating 2k-dimensional eigenproblems in each iteration, for no reasons should we spare the same effort at the beginning. 2k-dimensional Rayleigh-Ritz analysis leads us to a fairly good initial approximation of the lowest-spectral k-dimensional subspace we are seeking for if the 2k-dimmensional vectors are chosen carefully.

In the experiments, we adopted trigonometrical functions of lower frequencies. Considerations of the shape of the domain and the boundary conditions gave us useful guidelines to determine parameters of these 2k trigonometrical functions. If the global shapes of the eigenfunctions have been already approximated to some extent, local modifications along with the iterations are

very fast to converge, as our matrices (linear operators) A, B are "local".

Note that, after this Rayleigh-Ritz analysis, an eigenbasis of the starting subspace is already found.

The following is the resultant algorithm of what we have set forth.

0) Choose some linearly indepent vectors $y_1,..,y_{2k}$. Carry out Rayleigh-Ritz analysis for these y_j's , i.e. with $\tilde{A} = A(y_1,..,y_{2k})$ and $\tilde{B} = B(y_1,..,y_{2k})$, solve the eigenproblem $\tilde{A}\, c_j = \lambda_j \tilde{B}\, c_j$ where $\lambda_1 < .. < \lambda_{2k}$ and c_j's are to be normalized so that $\langle \tilde{B} c_j, c_j \rangle = 1$, and let $x_j = [y_1,..,y_{2k}]c_j$ for $j = 1,..,k$.

Initialize the "deflation counter" m to be 1.

1) Calculate the residuals $r_j = Ax_j - \lambda_j Bx_j$ for $j = m,..,k$, where $\lambda_j = \langle Ax_j, x_j \rangle / \langle Bx_j, x_j \rangle$ are already at hand with the preceding eigenvalue computation. Let new m be the least j such that magnitude of r_j exceeds some prescribed small number ε. If no such j is found then stop.

2) If for the first time here then let $p_j = r_j$ else let $p_j = r_j + \beta_j \Delta x_j$ for $j = m,..,k$, where $\beta_j = -\langle r_j, \Delta r_j \rangle / \langle \Delta x_j, \Delta r_j \rangle$ and Δx_j = current x_j - previous x_j and Δr_j = current r_j - previous r_j .

3) Let $\tilde{A} = A(x_1,..,x_k,p_m,..,p_k)$ and $\tilde{B} = B(x_1,..,x_k,p_m,..,p_k)$ and solve the eigenproblem $\tilde{A}\, c_j = \lambda_j \tilde{B}\, c_j$ $\lambda_1 < .. < \lambda_k < .. < \lambda_{2k-m+1}$, normalising c_j's so that $\langle \tilde{B} c_j, c_j \rangle = 1$.

4) Let $x_j^* = [x_1,..,x_k,p_m,..,p_k]c_j$ for $j = 1,..,k$, and go to 1 with the x_j^*'s as the new x_j's.

Note that, if we retain Ax_j's, Bx_j's, Ap_j's and Bp_j's each time in somewhere, there is no need of matrix multiplication in computing Ax_j^*'s and Bx_j^*'s, since
$Ax_j^* = [Ax_1,..,Ax_k,Ap_m,..,Ap_k]c_j$ and
$Bx_j^* = [Bx_1,..,Bx_k,Bp_m,..,Bp_k]c_j$ respectively. Only A by p_j's and B by p_j's are necessary, the number of which is reducing as deflation proceeds.

§4. An experiment

We applied the above algorithm to a simple test problem. The test problem is the 2-dimensional Helmholtz equation

$$u_{xx} + u_{yy} + \lambda u = 0 \ .$$

The domain and the boundary condition are as depicted in Fig 1. We adopted, for simplicity, discretization via triangular linear elements, as are shown also in Fig 1, giving rize to the stiffness matrix A and the consistent mass matrix B. The degree of freedom, or the number of nodal points which are not on the Dirichlet boundary is 214. The number of the off-diagonal entries is 1158 in both A and B. We executed a FORTRAN program which implimented the algorithm described in the last section in three cases, i.e. with k = 8, 12, 20, where k is the number of eigenpairs to be found out.

We chose 2k trigonometrical functions at the nodal points and carried out a Rayleigh-Ritz analysis to aquire an initial approximation of the lowest-spectral k dimensinal subspace. The

parameters were chosen with a consideration of the feature of the boundary condition. The following is the 40 functions we used when k = 20.

$$u_{7i+j+1}(x,y) = \sin((2i+1)\pi(1-x)/4)\cos(j\pi(y-1)/2)$$

for i = 0,..,5 and j = 0,..,6, (u_{41} and u_{42} are not used). Only the numbering is different when k = 8 and k = 12.

We judged the convergence of an eigenpair when the square-sum of the components of each residual vector has become less than 10^{-5}. Though we cannot offer theoretical accounts for this convergence criterion, seven decimal digits in the eigenvalues and four decimal digits in the components of eigenvectors has become stationary at the point when this criterion has been satisfied.

The following is the number of iterations and the elapsed CPU time before the convergence of all the k eigenpaires. VAX11/780 with floating-point hardware option, somewhere around 1.0 mips performance, has been used in this experiment. The whole computations has been carried out in double precision floating-point arithmetic, with 56-bit fractional part.

	iteration	CPU time
k = 8	20	87 sec.
k = 12	17	165 sec.
k = 20	16	521 sec.

Resultant eigenvalues are listed below.

0.42791	2.77536	4.31649	9.54730	11.58713
19.01808	19.82062	22.95112	26.82124	28.02605
31.38743	35.79016	43.22550	43.61609	47.19846
52.99571	64.74309	65.43948	69.25672	73.13649

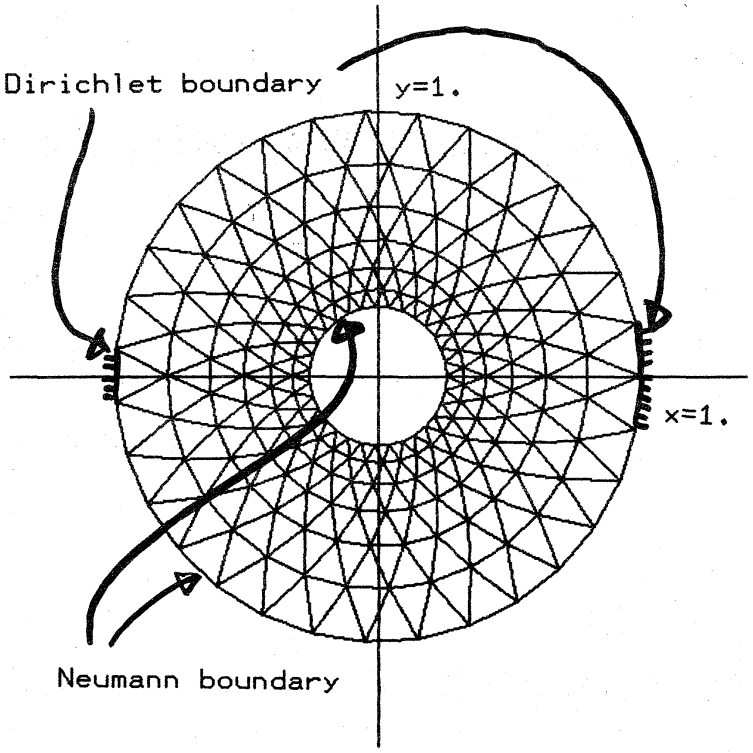

Fig 1.

REFERENCES

[1] K.J.Bathe and E.L.Wilson, Numerical Methods in Finite Element Analysis. Prentce-Hall,inc. Englewood Cliffs, New Jersey, 1976, Ch.10,11,12.

[2] K.Fukunaga, Introduction to Statistical Pattern Recognition. Academic Press, New York, 1972, Ch. 9.

[3] K.Fukunaga and S.Ando, "The Optimum Nonlinear Features for a Scatter Criterion in Discriminant Analysis," IEEE trans. Information Theory, vol.IT-23,NO.4,(1977),pp.453-459.