UNITARY JACOBI-METHOD FOR THE EIGENVALUE PROBLEM OF AN ARBITRARY NORMAL MATRIX

TIBOR FIALA

Dept. of Numerical Analysis and Computer Science, Eötvös University
1088 Budapest, Múzeum krt. 6–8.

(Received 24 February, 1982)

In his classical paper [4] Jacobi has given a constructive proof for the following well-known theorem.

Given an nxn real symmetric matrix, A, there exists a real orthogonal matrix U, such that UAU^{x} is diagonal.

Jacobi constructs the orthogonal matrix U, as an infinite product of orthogonal matrices of the form

$$\left|\begin{array}{ccccccccc}
1 & & & & & & & & \\
 & 1 & & & & & & & \\
 & & 1 & & & & & & \\
 & & & \cos\alpha & & & \sin\alpha & & \\
 & & & & 1 & & & & \\
 & & & & & 1 & & & \\
 & & & -\sin\alpha & & & \cos\alpha & & \\
 & & & & & & & 1 & \\
 & & & & & & & & 1
\end{array}\right| .$$

His method has the following advantageous properties:

- It is always convergent;
- It is a simple iteration procedure;
- It gives all the eigenvalues simultaneously;
- It is quadratically convergent, whenever all the eigenvalues of A are distinct;
- Every step of the algorithm is a similarity transformation with optimal error bound.

The last property means the following:

For an invertible matrix V, the similarity transformation

$$A \to VAV^{-1}$$

is a linear transformation on the space of nxn matrices. The norm of this transformation is $\|V\|\cdot\|V^{-1}\|$, the condition number of V, and this bounds the error-growth in the following manner

$$\|VAV^{-1}-V\tilde{A}V^{-1}\|\leq\|V\|\cdot\|V^{-1}\|\cdot\|A-\tilde{A}\|.$$

It is easy to prove, that for every matrix $\|V\|\cdot\|V^{-1}\|\geq 1$, and equality holds iff $V = cU$, where U is an unitary matrix. Thus it is worthwhile to use unitary transformations from the view of point of error-growth, too.

Since the late forties, when Jacobi method was rediscovered, several generalizations of this method have been proposed for the eigenvalue problem of not necessarily symmetric matrices (See [1], [2] and [3]).

K. Veselic solves the problem for real matrices, where the real parts of the eigenvalues haven't multiplicity more then double [2]. P. J. Anderson and G. Loizou solve the case of diagonaliable complex symmetric matrices [1]. However both of them use similarity transformations, which are not unitary, so the error-growth after k steps can be exponential in k. The class of matrices for which convergence proof is available doesn't contain the class of normal matrices.

The aim of this paper is to extend Jacobi's method for arbitrary (complex) normal matrices, preserving the advantages mentioned above. In addition we prove, that a special group generated by two-dimensional rotations is dense in the group of unitary matrices.

Definitions, notations

Let us denote by C^{nxn} the set of nxn matrices with complex elements. For $A\in C^{nxn}$ a_{ji} is the i-th element in the j-th row of A. A^x is the Hermitian transpose of A, with elements $a^x_{ij} = \bar{a}_{ji}$, where $\bar{a}_{ji}$ is the complex conjugate of a_{ji}.

The matrix $A\in C^{nxn}$ is called

Hermitian	if	$A^x = A$
anti-Hermitian	if	$A^x = -A$
normal	if	$A^xA = A A^x$
unitary	if	$AA^x = I$, where I is the unit matrix.

The set of Hermitian, anti-Hermitian, normal and unitary matrices from C^{nxn} will be denoted by H^{nxn}, AH^{nxn}, N^{nxn}, U^{nxn} respectively. For $A\in C^{nxn}$ we denote by $S^2(A)$ the sum of squares of absolute values of all subdiagonal elements of A, that is

$$S^2(A) = \sum_{j<i} |a_{ij}|^2 .$$

The symbol i will be used for the imaginary unit vector and for row or column-indexes for matrices; we hope, this will not cause any confusion. For $j>i$, and for arbitrary real α and β we introduce the following unitary matrices

$$(1)\qquad U^{i,j}_{\alpha,\beta} = \begin{Bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \cos\alpha & & & e^{i\beta}\cdot\sin\alpha & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & -\sin\alpha & & & e^{i\beta}\cdot\cos\alpha & & & \\ & & & & & & 1 & & \\ & & & & & & & 1 \end{Bmatrix} \begin{matrix} \\ \\ -i \\ \\ \\ -j \\ \\ \\ \end{matrix} .$$

Jacobi method for Hermitian and anti-Hermitian matrices

The following algorithm is an easy extension of Jacobi's classical method. We formalize it for the sake of explicit formulas and for the purpose of the further sections.

Algorithm I. An $n*n$ matrix A is given, which is either Hermitian or anti- Hermitian.

$k:=0$

$A_0:=\mathrm{A}$

$$c:=\begin{cases} 1 & \text{if } A \text{ is Hermitian} \\ i & \text{if } A \text{ is anti-Hermitian} \end{cases}$$

1. Let j_k and i_k be the row and column index respectively of the off-diagonal element of A_k with maximal absolute value.
If $a_{j_k i_k}=0$, the algorithm is finished.
If not, we choose β_k such that

$$e^{i\beta_k} a_{j_k i_k} = c\cdot |a_{j_k i_k}|$$

and choose α_k in the following way

$$\alpha_k := \frac{\pi}{4} \quad \text{if} \quad a_{i_k i_k} = a_{j_k j_k}$$

else

$$\alpha_k := \frac{1}{2}\,\mathrm{Arctg}\,\frac{2c\cdot|a_{j_k i_k}|}{a_{i_k i_k}-a_{j_k j_k}}$$

$A_{k+1}:=U^{i_k,j_k}_{\alpha_k,\beta_k} A_k U^{i_k,j_k *}_{\alpha_k,\beta_k}$

$k:=k+1$

Go to 1.

Theorem I. Let $A \in H^{nxn} \cup AH^{nxn}$. Algorithm I generates a sequence of matrices $A_0, A_1, \ldots$, which are similar to A. If this sequence is finite, then the final A_k is diagonal. If the sequence is infinite, then it converges to a fixed diagonal matrix with the eigenvalues of A in the main diagonal.

A detailed proof of this theorem can be found in [3] for the case of real symmetric matrices, Hovewer all the arguments remain valid in the case of Hermitian and anti-Hemitian matrices, so we don't repeat them, just notice, that the proof is based on the relation

$$S^2(A_{k+1}) = S^2(A_k) - |a_{j_k i_k}|^2 \leq \left(1 - \frac{2}{n(n-1)}\right) \cdot S^2(A_k),$$

and on Gerschgorin-theorem.

The Jacobi-group

For fixed n let us consider for all $j > i$, and for all real α and β the matrices $U^{i,j}_{\alpha,\beta}$. It is easy to compute, that

$$\left(U^{i,j}_{\alpha,\beta}\right)^{-1} = U^{i,j}_{0,-\beta} \cdot U^{i,j}_{-\alpha,0},$$

so the finite products of these matrices form a group, which we call Jabobi-group. The aim of this section is to prove that the Jacobi-group is dense in the group of unitary matrices.

Theorem II. For $U \in U^{nxn}$ there exist sequences of indexes j_s, i_s, and real numbers α_s, β_s such that $j_s > i_s$ and

$$\lim_{k \to \infty} \left\| \prod_{s=1}^{k} U^{i_s, j_s}_{\alpha_s, \beta_s} - U \right\| = 0.$$

Proof. First of all we remark, that whenever a matrix has all distinct eigenvalues, its eigen-subspaces are one-dimensional, and these one-dimensional subspaces depend continuously on the elements of the matrix.

Now let us denote by $u_1, u_2, \ldots, u_n$ the columns of U. We introduce the following Hermitian matrix

$$A = \sum_{k=1}^{n} k \cdot u_k \cdot u_k^*.$$

Matrix A is Hermitian, and has eigenvalues $1, 2, \ldots, n$ with eigenvectors $u_1, u_2, \ldots, u_n$. Applying the diagonalization process of Algorithm I to the matrix A we obtain a sequence of products

$$P_k = \prod_{s=1}^{k} U^{i_s, j_s}_{\alpha_s, \beta_k}$$

such that

$$P_k A P_k^* \to \begin{Bmatrix} 1 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & n \end{Bmatrix} = \Lambda .$$

Matrix $P_k A P_k^*$ has eigenvalues $1, 2, \ldots, n$ with eigenvectors $P_k u_1, P_k u_2, \ldots, P_k u_n$. The limit-matrix Λ has the same eigenvalues with eigenvectors $e_1, e_2, \ldots, e_n$, where e_i is the ith column of the unit-matrix.

Λ has distinct eigenvalues, so the set of its unit-eigenvectors depends continuously on its elements up to a constant factor. So there are real numbers φ_i^k $(i = 1, \ldots, n)$ and a permutation matrix P, such that

$$\lim_{k \to \infty} \left\| P_k u_i - e^{i \cdot \varphi_i^k} P e_i \right\| = 0 \quad i = 1, \ldots, n .$$

This means that $P_k^* P \Lambda_k \to U$, where Λ_k is a diagonal-matrix with elements $e^{i \cdot \varphi_i^k}$ in the main diagonal.

It is easy to show that Λ_k and P are finite products of matrices of the form (1), so the theorem is proved.

Diagonalization of normal matrices

In this section we give an algorithm for the diagonalization of an arbitrary normal matrix.

Theorem III. Let $C \in N^{n x n}$. An unitary matrix U, for which UAU^x is diagonal can be given as an infinite product of matrices of the form (1).

We give a constructive proof. The unitary matrix U will be constructed in the form

(2) $$\begin{pmatrix} U_1 & & & \\ & U_2 & & \\ & & \ddots & \\ & & & U_r \end{pmatrix} \begin{pmatrix} P \end{pmatrix} \begin{pmatrix} Q \end{pmatrix},$$

where Q is a unitary matrix, P is a permutation matrix, and the third multiplier is a block-diagonal matrix with unitary blocks $U_1, \ldots, U_r$. (For the sake of theoretical completeness we remark, that an arbitrary permutation matrix is a finite product of matrices of the form (1) with $\alpha = -\frac{\pi}{2}$ and $\beta = \pi$),

Our algorithm for constructing U consists of three phases. Phase I gives us the matrix Q, Phase II gives the permutation P, and Phase III gives the blocks $U_1, U_2, \ldots, U_r$.

Phase I. $A := C + C^x$

Compute the unitary matrix Q by Algorithm I, for which QAQ^x is diagonal

$D := QAQ^x$

Before describing Phase II we state the following

Proposition

$$d_{ij} + \overline{d_{ji}} = 0 \quad \text{for} \quad i \neq j \tag{3}$$

and

$$d_{ij} = 0 \quad \text{whenever} \quad Red_{ii} \neq Red_{jj}. \tag{4}$$

Proof. (3) is an immediate consequence of the fact, that $Q(C+C^x)\,Q^x$ is diagonal. Using the normality of D

$$\sum_{s=1}^{n} d_{si}\overline{d_{sj}} = \sum_{s=1}^{n} \overline{d_{is}}\, d_{js}.$$

However by (3)

$$d_{si}\overline{d_{sj}} = (-\overline{d_{is}})(-d_{js}) = \overline{d_{is}}\, d_{js} \quad \text{for} \quad s \neq i, j.$$

Thus

$$d_{ii}\,\overline{d_{ij}} + d_{ji}\,\overline{d_{jj}} = \overline{d_{ii}}\, d_{ji} + \overline{d_{ij}}\, d_{jj},$$

or equivalently

$$d_{ji}(\overline{d_{jj}} + d_{jj}) = (d_{ii} + \overline{d_{ii}})\, d_{ji}.$$

The final relation implies (4).
This Proposition means that D can be transformed to a block diagonal form by permuting coordinates.

Phase II. Compute the permutation matrix P, such that for $B := PDP^x$ the values Reb_{ii} $i = 1, 2, \ldots, n$ form a monotone increasing sequence.

Because of (4) B is a block-diagonal matrix with blocks $B_1, \ldots, B_r$ with dimensions $n_1, n_2, \ldots, n_r$, where

$$\sum_{s=1}^{r} n_s = n$$

$$b_{ij} = -\overline{b_{ij}} \quad \text{for} \quad i \neq j$$

and Reb_{ii} is constant in each block.

Phase III. For $s = 1, 2, \ldots, r$ $A_s := B_s - k_s \cdot I_{n_s x n_x}$ where B_s is the s-th block of B and k_s is the constant value of Reb_{ii} in the s-th block. For $s = 1, 2, \ldots, r$ compute the matrices $U_s \in U^{n_s x n_s}$ for which $U_s A_s U_s^x$ is diagonal. These matrices can be calculated by Algorithm I (being A_s anti-Hermitian).

We have finished the description of our algorithm. Clearly the similarity transformation, which diagonalizes A_s, will diagonalize B_s too. So Theorem III and the correctness of the algorithm are immediate consequences of the correctness of algorithm I, of the Proposition and remarks between Phase II and Phase III.

REFERENCES

[1] *P. J. Anderson* and *G. Loizou;* A Jacobi type method for complex symmetric matrices. Numer. Math. 25, 347 – 363 (1976).

[2] *K. Veselic:* A convergent Jacobi method for solving the eigenvalueproblem or arbitrary real matrices. Numer Math. 25, 179 – 184 (1976).

[3] *Wilkinson, J. H:* Algebraic eigenvalue problem. Clarendon, Oxford, 1965.

[4] *Jacobi C. G. J.:* 1846, Über ein leichtes Verfahren die in der theorie der Sacilarstörungen vorkommenden Gleichungen numerisch aufzulösen. Crelle's J. 30, 51 – 94.