Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

Mathematica Balkanica

New series Vol.8, 1994, Fasc.1

Algorithm for Solving Spectral Problem of Complex Matrix in Real Arithmetic

I.G. Ivanov

Presented by P. Kenderov

1. Introduction

Let $D = L+iM$ be $n\times n$ complex diagonalizable (matrix D has a complect system of eigenvectors), where $L, M \in \mathbb{R}^{n\times n}$ and let $\mathbb{R}^{p\times q}$ be the set of real $p\times q$ matrices. The spectral problem for matrix D

$$(L+iM)(x+iy) = (\lambda+i\mu)(x+iy), \tag{1}$$

where $x, y \in \mathbb{R}^{n\times 1}$ and $\lambda, \mu \in \mathbb{R}^{1\times 1}$, is equivalent to the real problem (with real solutions)

$$\begin{pmatrix} L & -M \\ M & L \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda I & -\mu I \\ \mu I & \lambda I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{2}$$

where $I \in \mathbb{R}^{n\times n}$ is a unit matrix.

In this paper we propose an effective algorithm for solving the eigenproblem of real matrix

$$H = H(A,B) = \begin{pmatrix} L & -M \\ M & L \end{pmatrix}.$$

2. A description of the algorithm and a convergence theorem.

For the matrix H we construct the sequence

$$H_{k+1} = U_k^{-1} H_k U_k = (h_{rs}^{(k+1)}), \ H_1 = H, \quad k = 1,2,3,\ldots, \tag{3}$$

where each $H_k = H(A_k, B_k)$ and $U_k = U(p_k, q_k, \varphi_k)$. For each k the matrix U_k depending from three parameters p_k, q_k, φ_k.

Let denote

$$H_k = H(A_k, B_k) = (h_{rs}^{(k)}), \quad A_k = (a_{\beta\gamma}^{(k)}) \quad \text{and} \quad B_k = (b_{\beta\gamma}^{(k)}),$$

$$C_k = C(H_k) = H_k H_k^T - H_k^T H_k = H(F_k, E_k) = (c_{rs}^{(k)}),$$

where

$$F_k = F_k^T = A_k A_k^T + B_k B_k^T - A_k^T A_k - B_k^T B_k = (f_{\beta\gamma}^{(k)})$$

and

$$E_k = -E_k^T = B_k A_k^T - A_k B_k^T + B_k^T A_k - A_k^T B_k = (e_{\beta\gamma}^{(k)}),$$

$$\|H_k\|^2 = \sum_{r,s=1}^{2n} \left|h_{rs}^{(k)}\right|^2 .$$

For symmetric matrix $H_k + H_k^T$, we have

$$H_k + H_k^T = H(A_k + A_k^T, B_k - B_k^T).$$

The strategy determining U_k from (3), p_k, q_k, φ_k is the following. For the matrix H_k, find

$$c^{(k)} = \max_{r\neq s} |c_{rs}^{(k)}|^{1/2} \quad \text{and} \quad h^{(k)} = \max_{r\neq s} |h_{rs}^{(k)} + h_{sr}^{(k)}|.$$

If $c(k) \geq h(k)$ then we choose the matrix U_k to decrease Euclidean norm of H_{k+1}. If $c(k) < h(k)$ then we choose the matrix U_k to innihilate the off-diagonal elements $h_{p_k q_k}^{(k)} + h_{q_k p_k}^{(k)}$ and $h_{q_k p_k}^{(k)} + h_{p_k q_k}^{(k)}$ of matrix $\frac{1}{2}(H_k + H_k^T)$.

Then there are four possible cases

<u>**A.1**</u> $|f_{pq}^{(k)}|^{1/2} = c^{(k)} \geq h^{(k)}$ and $1 \leq p = p_k < q = q_k \leq n$, $\varphi = \varphi_k$. In this case $U_k = U_{pq}(\varphi) = (u_{rs})$ is different from the $2n \times 2n$ unit matrix by elements

$$\text{(4)} \qquad \begin{cases} u_{pp} = u_{qq} = u_{p+np+n} = u_{q+nq+n} = \cosh\varphi \\ u_{pq} = u_{qp} = u_{p+nq+n} = u_{q+np+n} = \sinh\varphi \end{cases}$$

Define φ by

$$\text{(5)} \qquad \tanh\varphi = f_{pq}^{(k)} \ / \ (G_A + G_B + 2D_A^2 + 2D_B^2 + 2\xi_A^2 + 2\xi_B^2),$$

where

$$D_A = a_{pp}^{(k)} - a_{qq}^{(k)}, \quad D_B = b_{pp}^{(k)} - b_{qq}^{(k)}, \quad \xi_A = a_{pq}^{(k)} - a_{qp}^{(k)}, \quad \xi_B = b_{pq}^{(k)} - b_{qp}^{(k)}$$

$$G_A = \sum_{j\neq p,q} (a_{pj}^{(k)})^2 + (a_{jp}^{(k)})^2 + (a_{qj}^{(k)})^2 + (a_{jq}^{(k)})^2$$

$$G_B = \sum_{j\neq p,q} (b_{pj}^{(k)})^2 + (b_{jp}^{(k)})^2 + (b_{qj}^{(k)})^2 + (b_{jq}^{(k)})^2$$

A.2 $|e_{pq}^{(k)}|^{1/2} = c^{(k)} \geq h^{(k)}$ and $1 \leq p = p_k < q = q_k \leq n$, $\varphi = \varphi_k$. In this case $U_k = U_{pq}(\varphi) = (u_{rs})$ is of the type

$$(6)\qquad \begin{cases} u_{rs} = \delta_{rs}, (r,s) \notin \{(p,p),\ (q,q),\ (p+n,p+n),\ (q+n,q+n), \\ \qquad\qquad (p,q+n),\ (q+n,p),\ (q,p+n),\ (p+n,q)\} \\ u_{pp} = u_{qq} = u_{p+np+n} = u_{q+nq+n} = \cosh\varphi \\ u_{pq+n} = u_{q+np} = -\,u_{qp+n} = -u_{p+nq} = \sinh\varphi \end{cases}$$

In this case, we define φ by

$$(7)\qquad \tanh\varphi = e_{pq}^{(k)} \ / \ (G_A + G_B + 2D_A^2 + 2D_B^2 + 2\eta_A^2 + 2\eta_B^2),$$

where

$$D_B = b_{qq}^{(k)} - b_{pp}^{(k)},\quad \eta_A = a_{pq}^{(k)} + a_{qp}^{(k)},\quad \eta_B = b_{pq}^{(k)} + b_{qp}^{(k)}.$$

Remark 1. In cases **A.1** and **A.2** when U_k choose from (4) or (6) and φ choose from (5) or (7) respectively we have

$$(8)\qquad \|H_k\|^2 - \|H_{k+1}\|^2 \geq \frac{1}{3}|c^{(k)}|^4 \ / \ \|H_k\|^2.$$

A.3 $|a_{pq}^{(k)} + a_{qp}^{(k)}| = h^{(k)} > c^{(k)}$ and $1 \leq p = p_k < q = q_k \leq n$, $\varphi = \varphi_k$. In this case we choose the matrix $U_k = U_{pq}(\varphi) = (u_{rs})$ of the form

$$(9)\qquad \begin{cases} u_{rs} = \delta_{rs}, (r,s) \notin \{(p,p),\ (q,q),\ (p+n,p+n),\ (q+n,q+n), \\ \qquad\qquad (p,q),\ (q,p),\ (p+n,q+n),\ (q+n,p+n)\} \\ u_{pp} = u_{qq} = u_{p+np+n} = u_{q+nq+n} = \cos\varphi \\ u_{pq} = -u_{qp} = u_{p+nq+n} = -u_{q+np+n} = \sin\varphi \end{cases}$$

From the condition $|a_{pq}^{(k)} + a_{qp}^{(k)}| = 0$ for φ we have

$$(10)\qquad \tan 2\varphi = (a_{pq}^{(k)} + a_{qp}^{(k)}) \ / \ (a_{pp}^{(k)} - a_{qq}^{(k)}).$$

A.4 $|-b_{pq}^{(k)} + b_{qp}^{(k)}| = h^{(k)} > c^{(k)}$ and $1 \leq p = p_k < q = q_k \leq n$, $\varphi = \varphi_k$. In this case we choose the matrix $U_k = U_{pq}(\varphi) = (u_{rs})$ of the form

$$(11)\qquad \begin{cases} u_{rs} = \delta_{rs}, (r,s) \notin \{(p,p),\ (q,q),\ (p+n,p+n),\ (q+n,q+n), \\ \qquad\qquad (p,q+n),\ (q+n,p),\ (q,p+n),\ (p+n,q)\} \\ u_{pp} = u_{qq} = u_{p+np+n} = u_{q+nq+n} = \cos\varphi \\ u_{pq+n} = -u_{q+np} = u_{qp+n} = -u_{p+nq} = -\sin\varphi \end{cases}$$

From the condition $|b_{qp}^{(k)} - b_{pq}^{(k)}| = 0$ for φ we have

$$(12) \qquad \tan 2\varphi = (b_{qp}^{(k)} - b_{pq}^{(k)}) \ / \ (a_{pp}^{(k)} - a_{qq}^{(k)}).$$

Remark 2. In cases **A.3** and **A.4** the similar transformations $U_k^{-1} H_k U_k$ are steps from Jacobi's method for the symmetric matrix $H_k + H_k^T$ [1].

Lemma . *Let α and β be nature numbers, $1 \le \alpha, \ \beta \le 2n, \ \alpha \ne \beta$, $|\alpha - \beta| \ne n$. Let $\tilde{H} = U^{-1} H U$, where U determine following*

If $1 \le \alpha, \ \beta \le n$, then $p = \min(\alpha, \beta)$, $q = \max(\alpha, \beta)$, φ is computing by (5) and U is of the type (4).

If $n + 1 \le \alpha, \ \beta \le 2n$, then $p = \min(\alpha - n, \beta - n)$, $q = \max(\alpha - n, \beta - n)$, φ is computing by (5) and U is of the type (4).

If $1 \le \alpha \le n, \ n + 1 \le \beta \le 2n$, then $p = \min(\alpha, \beta - n)$, $q = \max(\alpha, \beta - n)$, φ is computing by (7) and U is of the type (6).

If $n + 1 \le \alpha \le 2n, \ 1 \le \beta \le n$, then $p = \min(\alpha - n, \beta)$, $q = \max(\alpha - n, \beta)$, φ is computing by (7) and U is of the type (6).

Then

$$|\tilde{h}_{rs} - h_{rs}| \le \left(\frac{3}{2}\right)^{1/2} |c_{\alpha\beta}|^{1/2} \quad \textit{for all} \quad r, \ s$$

where $H = H(A, B) = (h_{rs})$, $\tilde{H} = H(\tilde{A}, \tilde{B}) = (\tilde{h}_{rs})$, $C(H) = H(F, E) = (c_{rs})$.

This lemma is proved analogy with Lemma 1 from [4].

Theorem . *For sequence (3) we have*

I. $C(H_k) \to 0, \ k \to \infty$

II. $\frac{1}{2}(H_k + H_k^T)$ *tends to matrix* $diag[\lambda_1, \ldots, \lambda_n, \lambda_1, \ldots, \lambda_n]$

III. If $\beta \ne \gamma \in (1, \ldots, n)$ and $\lambda_\beta \ne \lambda_\gamma$ then

$$a_{\beta,\gamma}^{(k)} \to 0, \qquad k \to \infty,$$
$$b_{\beta,\gamma}^{(k)} \to 0, \qquad k \to \infty.$$

IV. If for a fixed nature number $m \in [1, n]$ and each nature number $t \in [1, n]$, for which $t \ne m$ we have $\lambda_t \ne \lambda_m$, then

$$b_{mm}^{(k)} \to \mu_m, \quad k \to \infty.$$

Proof.

I. We consider a sequence $\|H_k\|^2$ $(k = 1, 2, 3, \ldots)$. The similar transformations with matrices of form (9) and (11) preserve the Euclidean norm and similar transformations with matrices of form (4) and (6) decrease or preserve (from (8)) Euclidean norm implies that the sequence $\|H_1\|^2, \|H_1\|^2, \ldots$ is monotonically decreasing. Let for each matrix H_k we introduce a number α_k such that

$$\alpha_k = \begin{cases} 0 \text{ if } U_k \text{ from (9)} & \text{or (11)} \\ 1 \text{ if } U_k \text{ from (4)} & \text{or (6)} \end{cases}.$$

The sequence of matrices $\{H_k\}$ is bounded. From this sequence we choose a convergent subsequence $\{H_s\}$ where $s \in S \subset \mathbb{N}$ and $\mathbb{N}$ is the set of nature numbers. Suppose that α_s contains the infinitely many of ones and H_m is a subsequence of $\{H_s\}$ that for each $m \in M \subset S$ we have $\alpha_m = 1$. Then from (8) we obtain

$$\|H_m\|^2 - \|H_{m+1}\|^2 \geq \frac{1}{3}(c^{(m)})^4 \;/\; \|H_m\|^2 \geq \frac{1}{3}(c^{(m)})^4 \;/\; \|H\|^2$$

It follows that $c^{(m)} \to 0,\ m \to \infty$. And from inequality $h^{(m)} \leq c^{(m)}$ follows that $h^{(m)} \to 0,\ m \to \infty$.

Let $H = H(A, B) = (h_{rs})$ is a limit a sequence $\{H_m\}$ and for $H + H^T$, $C(H)$ we have

$$H + H^T = H(A + A^T, B - B^T) = (h_{rs} + h_{sr}),$$

$$C = C(H) = HH^T - H^T H = H(F, E) = (c_{rs}),$$

where

$$F = F^T = AA^T + BB^T - A^T A - B^T B = (f_{\beta\gamma})$$

and

$$E = -E^T = BA^T - AB^T + B^T A - A^T B = (e_{\beta\gamma}).$$

Hence H is a limit if $r \neq s$ we obtain $c_{rs} = 0$ and $h_{rs} + h_{sr} = 0$.

From $h_{rs} + h_{sr} = 0$ $(r \neq s)$ follows that

(13) $$a_{\beta\gamma} + a_{\gamma\beta} = 0 \text{ for } \beta \neq \gamma$$

(14) $$b_{\beta\gamma} - b_{\gamma\beta} = 0 \text{ for all } \beta, \gamma.$$

For diagonal elements of $C(H)$ we have

$$f_{\beta\beta} = \sum_{j<\beta} \left((a_{\beta j}^2 - a_{j\beta}^2) + (b_{\beta j}^2 - b_{j\beta}^2)\right) + \sum_{j>\beta} \left((a_{\beta j}^2 - a_{j\beta}^2) + (b_{\beta j}^2 - b_{j\beta}^2)\right)$$

Formulas (13) and (14) implies that $f_{\beta\beta} = 0$ for $\beta = 1, \ldots, n$.

Analogously for $f_{\beta\gamma}$, $\beta \neq \gamma$, we have

$$f_{\beta\gamma} = 2a_{\beta\gamma}(a_{\gamma\gamma} - a_{\beta\beta}) = 0. \tag{15}$$

The diagonal elements of E are equal to zero because $E = -E^T$. For off-diagonal elements $e_{\beta\gamma}$ of E we find

$$e_{\beta\gamma} = 2b_{\beta\gamma}(a_{\gamma\gamma} - a_{\beta\beta}) = 0 \text{ for } \beta \neq \gamma. \tag{16}$$

Then formulas (15) and (16) are true for all β, γ. This implies that $c_{rs} = 0$ for all r and s, thus H is a normal matrix and $H + H^T$ is diagonal. Then the diagonal elements $a_{11}, \dots, a_{nn}$ are the real parts of the eigenvalues of H and H^T.

Since H is normal, we have

$$C(H_m) \to 0, \quad m \to \infty$$

and

$$\|H_m\|^2 \to 2\sum_{j=1}^{n} |\nu_j|^2, \quad (\nu_j = \lambda_j + i\mu_j).$$

Since the whole sequence $\|H_k\|$ is non-increasing, it follows that

$$\|H_k\|^2 \to 2\sum_{j=1}^{n} |\nu_j|^2$$

or equivalently

$$C(H_k) \to 0, \quad k \to \infty.$$

Hence the whole sequence $\{H_k\}$ tends to normality.

Let the sequence $\{\alpha_s\}$ contains only the finitely many of ones. Then there is a natural number s_0 such that for $s \geq s_0 U_s$ are from (9) or (11) and $\alpha_s = 0$. Then the process becomes a Jacobi's methods for the symmetric matrix $H_{s_0} + H_{s_0}^T$. Then

$$h^{(s)} \to 0, \quad s \to \infty.$$

Since $c^{(s)} \leq h^{(s)}$ then

$$c^{(s)} \to 0, \quad s \to \infty.$$

Every convergent subsequence of $\{H_k\}$ has a limit $H = H(A, B)$ with properties (15) and (16). This proves I.

II. We shall prove that $h^{(k)} \to 0$, $k = 1, 2, 3, \dots$ for the sequence $\{H_k\}$. In proving I we have found a subsequence $\{H_s\}$ of $\{H_k\}$ for which II is true. We consider the case when the sequence $\{\alpha_k\}$ contains infinitely many of both zero and ones. We know that for the subsequence $\{\alpha_p\}$ of $\{\alpha_k\}$ and $\alpha_p = 1$ then $h^{(p)} \to 0$, $p \to \infty$. We consider the sequence from all indices $k_1, \ldots, k_s, \ldots$ such that $\alpha_{k_s} = 0$, $\alpha_{k_s - 1} = 1$ for $s = 1, 2, \ldots$. In this case for $m = k_s$ and from lemma for $h^{(m)}$ we have

$$|h^{(m)}| \leq |h^{(m-1)}| + |h^{(m)} - h^{(m-1)}| \leq |h^{(m-1)}| + 2\left(\frac{3}{2}\right)^{1/2} c^{(m-1)}.$$

Since $c^{(m-1)} \to 0$ and $h^{(m-1)} \to 0$ for $m = k_s$ and $s = 1, 2, \ldots$, it follows that $h^{(m)} \to 0$. Let σ_k denote a sum from squares of off-diagonal elements in blocks of symmetric matrix $\frac{1}{2}(H_k + H_k^T)$. Then

$$h^{(k)} \leq \sigma_k^{1/2} \leq 2n\,(2n-1)\,h^{(k)}.$$

For the subsequence $\{H_m\}$ from $h^{(m)} \to 0$ it follows that $\sigma_m \to 0$.

Consider those indices $m+t$ of $\{\alpha_k\}$, $(m = k_s)$ with $\alpha_{m-1} = 1$, $\alpha_{m-1+t} = 0$ for $t = 1, 2, \ldots, p$ and $\alpha_{m+p} = 1$ for $s = 1, 2, \ldots$. For those $m+t$ σ_{m+t} is a monotonically decreasing because for matrix $H_{m+t} + H_{m+t}^T$ is used a step from the classical Jacoby's methods. It follows that $\sigma_k \to 0$ for $k = 1, 2, \ldots$ and $h^{(k)} \to 0$ for same k. Then from I, we conclude that every convergent subsequence of $\{H_k\}$ has a limit with the properties (13), (14), (15), (16) and its diagonal elements are the real parts of the eigenvalues of H.

III. In II we have show that $h^{(k)} \to 0$, $c^{(k)} \to 0$ as $k \to \infty$. From (13), (14), (15) and (16)

$$a_{\beta\gamma}^{(k)}(a_{\beta\beta}^{(k)} - a_{\gamma\gamma}^{(k)}) \to a_{\beta\gamma}(\lambda_\beta - \lambda_\gamma) = 0 \text{ and } \beta \neq \gamma, k \to \infty$$
$$b_{\beta\gamma}^{(k)}(a_{\beta\beta}^{(k)} - a_{\gamma\gamma}^{(k)}) \to b_{\beta\gamma}(\lambda_\beta - \lambda_\gamma) = 0 \text{ and } \beta \neq \gamma, k \to \infty,$$

where

$$a_{\beta\gamma}^{(k)} \to a_{\beta\gamma}, \quad b_{\beta\gamma}^{(k)} \to b_{\beta\gamma}, \quad k \to \infty$$

and $a_{\beta\beta} = \lambda_\beta$.

If $\beta \neq \gamma$ and $\lambda_\beta \neq \lambda_\gamma$ then

$$a_{\beta\gamma}^{(k)} \to 0$$

$$b_{\beta\gamma}^{(k)} \to 0.$$

IV. Let $t \neq m$ be nature numbers from set $(1, 2, \ldots, n)$. Let $\{H_m\}$ be a convergent subsequence of $\{H_k\}$ with limit $H = H(A, B)$ and $H = (h_{rs})$. Since the limit H satisfies (13), (14), (15), (16). In rows and columns with numbers m and $m+n$ of matrix H will have only four nonzero elements $h_{mm} = h_{m+nm+n} = \lambda_m$ and $h_{mm+n} = -h_{m+nm} = -b_{mm}$. This implies that $\lambda_m \pm ib_{mm}$ are eigenvalues of D.

Remark 3. From the proof of Theorem 1, it follows that the Algorithm computes always the real parts of all eigenvalues of the complex matrix $A+iB$. Let $\nu = \lambda + i\mu$ be eigenvalue of $A+iB$. If no other eigenvalue (real or complex) with the real part equal to λ, then Algorithm computes the imaginary part μ of ν.

References

[1] P. J. Eberlein. A Jacobi method for the automatic ecomputation of eigenvalues and eigenvectors of an arbitrary matrix. *SIAM J.*, **10**, 1962, 74-88.

[2] P. J. Eberlein, J. Boothroyd. Solution to the eigenproblem by a norm reducing Jacobi type method. *Numer. Math.*, **11**, 1968, 1-12.

[3] M. Petkov. Numerical solution of linear algebric problems in complex case. *Annual of Sofia University "Kl. Okhridski"*, **71 (Part II)**, 1976/1977, 15-26 (in Bulgarian).

[4] K. Veselic. A convergent Jacobi method for solving the eigenproblem of arbitrary real matrices. *Numer. Math.*, **25**, 1976, 179-184.

Pedagogical Institute
Shumen
BULGARIA

Received 16.09.1992