



ELSEVIER

Journal of Computational and Applied Mathematics 54 (1994) 99–106

---

JOURNAL OF  
COMPUTATIONAL AND  
APPLIED MATHEMATICS

---

# A parallel processed scheme for the eigenproblem of positive definite matrices

M.A. Shalaby

*Department of Engineering Mathematics, Faculty of Engineering, Alexandria University, Alexandria, Egypt*

Received 19 June 1992; revised 18 November 1992

---

## Abstract

This paper deals with the eigenproblem of positive definite matrices. A numerical algorithm, to find the largest eigenvalues of a full positive definite matrix using Householder reflections, is described. The proposed algorithm can be used to find all the eigenvalues of a symmetric matrix or at least the first few largest ones. The scheme is proved to be convergent and the convergence rate is calculated. The full matrix is operated upon, at each iteration; hence one could use the APL programming language to write down a very brief code to implement the program.

*Keywords:* Eigenvalues; Parallel processing

---

## 1. Introduction

Let  $A$  be an  $n \times n$  full positive matrix (i.e., the leading principal minors are strictly positive) with eigenvalues  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{n-1} < \lambda_n$ . In this article, we present a numerical scheme to obtain the eigenvalues of  $A$ , and in particular the largest eigenvalue or the first few largest ones. This basic idea is to apply a sequence of Householder reflections [1,2] so that, in the limit, the first column of  $A$  converges to a multiple of  $e_1$  (the first column of  $I_n$ ). The proposed method is a special case of the power method where the starting vector is  $e_1$ . In contrast of the power method itself, our method transforms the whole matrix  $A$  at each step of the iteration process. As a consequence, the presented scheme converges linearly with convergence rate  $(\lambda_{n-1}/\lambda_n)^2$ .

It is well known that the Jacobi method for obtaining the eigenvalues of a full symmetric matrix is well-suited for sequential machines, so is our algorithm. However, our scheme is eminently suitable for a parallel processing machine and should be superior to Jacobi's method presented in [3] on such machines.

## 2. The proposed scheme

Let  $A$  be a full positive definite matrix. Express  $A$  as

$$A = A_0 = \begin{bmatrix} a_0^{11} & \boldsymbol{\nu}_0^T \\ \boldsymbol{\nu}_0 & \mathbf{B}_0 \end{bmatrix},$$

where  $a_0^{11}$  is the first element of  $A_0$ ,  $\boldsymbol{\nu}_0$  is an  $(n-1)$ -vector and  $\mathbf{B}_0$  is an  $(n-1) \times (n-1)$  positive definite matrix. We denote by  $\mathbf{H}_0$  an orthonormal matrix consisting of the eigenvectors of  $A$  and by  $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  the matrix of the eigenvalues of  $A$  arranged in ascending order.

Therefore, we can write the similarity transformation

$$A_0 = \mathbf{H}_0 \boldsymbol{\Lambda} \mathbf{H}_0^T.$$

The vector  $\mathbf{H}_0^T \mathbf{e}_1$  will be denoted by  $\mathbf{f}_0$ .

The proposed algorithm is defined as follows. Starting with  $A_0$  we apply an orthogonal transformation  $\mathbf{U}_0$  and define  $A_1$  by

$$A_1 = \mathbf{U}_0 A_0 \mathbf{U}_0 = \begin{bmatrix} a_0^{11} & \boldsymbol{\nu}_1^T \\ \boldsymbol{\nu}_1 & \mathbf{B}_1 \end{bmatrix},$$

where

$$\mathbf{U}_0 = \mathbf{I}_n - \beta_0 \mathbf{u}_0 \mathbf{u}_0^T, \quad \beta_0 = \frac{2}{\|\mathbf{u}_0\|^2},$$

and the vector  $\mathbf{u}_0$  is chosen so that  $\mathbf{U}_0$  transforms the vector  $[a_0^{11}, \boldsymbol{\nu}_0^T]^T$  into  $k\mathbf{e}_1$  with

$$k = -\sqrt{(a_0^{11})^2 + \|\boldsymbol{\nu}_0\|^2}.$$

As is well known [1], we can write  $\mathbf{u}_0$  as

$$\mathbf{u}_0 = \left[ a_0^{11} + \sqrt{(a_0^{11})^2 + \|\boldsymbol{\nu}_0\|^2}, \boldsymbol{\nu}_0^T \right]^T.$$

The process is repeated; at a typical stage we have

$$A_{r+1} = \mathbf{U}_r A_r \mathbf{U}_r = \begin{bmatrix} a_{r+1}^{11} & \boldsymbol{\nu}_{r+1}^T \\ \boldsymbol{\nu}_{r+1} & \mathbf{B}_{r+1} \end{bmatrix}, \quad (1)$$

where

$$\mathbf{u}_r = \mathbf{I}_n - \beta_r \mathbf{u}_r \mathbf{u}_r^T, \quad \beta_r = \frac{2}{\|\mathbf{u}_r\|^2}, \quad (2)$$

and

$$\mathbf{u}_r = \left[ a_r^{11} + \sqrt{(a_r^{11})^2 + \|\boldsymbol{\nu}_r\|^2}, \boldsymbol{\nu}_r^T \right]^T. \quad (3)$$

In the following section, it will be shown that in the limit (i.e., as  $r \rightarrow \infty$ ), we obtain a matrix of the form

$$A_\infty = \begin{bmatrix} \lambda & 0 \\ 0 & B_\infty \end{bmatrix},$$

which means

$$\lim_{r \rightarrow \infty} \|\mathbf{v}_r\|^2 = 0 \quad \text{and} \quad \lim_{r \rightarrow \infty} a_r^{11} = \lambda.$$

So, this method provides a very reliable and economical criterion to deflate a matrix.

### 3. Convergence study

In this section, we prove the convergence of the algorithm together with rate of convergence. Before stating the theorem, we verify the following lemmas.

**Lemma 1.**

$$\beta_r = \frac{1}{(a_r^{11} + d)d}, \quad d = \sqrt{(a_r^{11})^2 + \|\mathbf{v}_r\|^2}. \quad (4)$$

**Proof.** From (2) and (3),

$$\begin{aligned} \beta_r &= \frac{2}{\|\mathbf{u}_r\|^2} = \frac{2}{\left(a_r^{11} + \sqrt{(a_r^{11})^2 + \|\mathbf{v}_r\|^2}\right)^2 + \|\mathbf{v}_r\|^2} \\ &= \frac{2}{2(a_r^{11})^2 + 2\|\mathbf{v}_r\|^2 + 2a_r^{11}\sqrt{(a_r^{11})^2 + \|\mathbf{v}_r\|^2}} \\ &= \frac{1}{(d + a_r^{11})d}, \quad d = \sqrt{(a_r^{11})^2 + \|\mathbf{v}_r\|^2}. \quad \square \end{aligned}$$

It can easily be proved that

$$d = \|A_r \mathbf{e}_1\|. \quad (5)$$

**Lemma 2.**

$$U_r = \begin{bmatrix} -a_r^{11}/d & -\mathbf{v}_r^T/d \\ -\mathbf{v}_r/d & \mathbf{I}_{n-1} - \beta_r \mathbf{v}_r \mathbf{v}_r^T \end{bmatrix}. \quad (6)$$

**Proof.** From (2) and (3),

$$\begin{aligned} U_r &= I_n - \beta_r u_r u_r^T = I_n - \frac{1}{(a_r^{11} + d)d} \begin{bmatrix} (a_r^{11} + d)^2 & (a_r^{11} + d) \mathbf{v}_r^T \\ (a_r^{11} + d) \mathbf{v}_r & \mathbf{v}_r \mathbf{v}_r^T \end{bmatrix} \\ &= \begin{bmatrix} 1 - (a_r^{11} + d)/d & -\mathbf{v}_r^T/d \\ -\mathbf{v}_r/d & I_{n-1} - \beta_r \mathbf{v}_r \mathbf{v}_r^T \end{bmatrix} = (6). \quad \square \end{aligned}$$

It can easily be proved that

$$U_r e_1 = -\frac{A_r e_1}{\|A_r e_1\|}. \quad (7)$$

**Lemma 3.**

$$a_{r+1}^{11} = a_r^{11} + \frac{a_r^{11} \|\mathbf{v}_r\|^2 + \mathbf{v}_r^T B_r \mathbf{v}_r}{d^2}. \quad (8)$$

**Proof.** Since  $A_{r+1} = U_r A_r U_r$ , then

$$\begin{bmatrix} a_{r+1}^{11} & \mathbf{v}_{r+1}^T \\ \mathbf{v}_{r+1} & B_{r+1} \end{bmatrix} = \begin{bmatrix} -a_r^{11}/d & -\mathbf{v}_r^T/d \\ -\mathbf{v}_r/d & I_{n-1} - \beta_r \mathbf{v}_r \mathbf{v}_r^T \end{bmatrix} \begin{bmatrix} a_r^{11} & \mathbf{v}_r^T \\ \mathbf{v}_r & B_r \end{bmatrix} \begin{bmatrix} a_r^{11}/d & \mathbf{v}_r^T/d \\ -\mathbf{v}_r/d & I_{n-1} - \beta_r \mathbf{v}_r \mathbf{v}_r^T \end{bmatrix},$$

from which

$$\begin{aligned} a_{r+1}^{11} &= \frac{a_r^{11}}{d} \left[ \frac{(a_r^{11})^2}{d} + \frac{\mathbf{v}_r^T \mathbf{v}_r}{d} \right] + \frac{a_r^{11}}{d^2} \mathbf{v}_r^T \mathbf{v}_r + \frac{\mathbf{v}_r^T B_r \mathbf{v}_r}{d^2} = \frac{a_r^{11}}{d^2} \left[ (a_r^{11})^2 + 2 \mathbf{v}_r^T \mathbf{v}_r \right] + \frac{\mathbf{v}_r^T B_r \mathbf{v}_r}{d^2} \\ &= \frac{a_r^{11}}{d^2} \left[ (a_r^{11})^2 + 2 \|\mathbf{v}_r\|^2 \right] + \frac{\mathbf{v}_r^T B_r \mathbf{v}_r}{d^2} = a_r^{11} + \frac{1}{d^2} \left[ a_r^{11} \|\mathbf{v}_r\|^2 + \mathbf{v}_r^T B_r \mathbf{v}_r \right]. \end{aligned}$$

**Theorem 4.**  $\lim_{r \rightarrow \infty} a_r^{11} = \lambda$  exists and is an eigenvalue of  $A$ . Moreover, if  $e_n^T f_0 \neq 0$ , then  $\lambda = \lambda_n$ ; if in addition  $e_{n-1}^T f_0 \neq 0$ , then

$$\lim_{r \rightarrow \infty} \left| \frac{a_{r+1}^{11} - \lambda}{a_r^{11} - \lambda} \right| = \left( \frac{\lambda_{n-1}}{\lambda_n} \right)^2. \quad (9)$$

**Proof.** Since  $A_r$  is positive definite, then, using Lemma 3,  $a_{r+1}^{11} \geq a_r^{11}$ . Thus the sequence  $\{a_r^{11}\}$  is a monotonically increasing sequence bounded from above by  $\|A\|$ . It follows that  $\lim_{r \rightarrow \infty} a_r^{11} = \lambda$  exists. From Lemma 3 also, it follows that  $\lim_{r \rightarrow \infty} \|\mathbf{v}_r\|^2 = 0$ .

Let  $H_r = U_{r-1} H_{r-1}$  denote the matrix of eigenvectors of  $A_r$ , and let  $f_r = H_r^T e_1$ . Then,

$$f_1 = H_1^T e_1 = (U_0 H_0)^T e_1 = H_0^T U_0 e_1.$$

Since from (7),  $U_0 e_1 = -(Ae_1)/\|Ae_1\|$ , then

$$f_1 = -\frac{H_0^T A e_1}{\|A e_1\|} = -\frac{\Lambda H_0^T e_1}{\|H_0 \Lambda H_0^T e_1\|} = -\frac{\Lambda f_0}{\|\Lambda f_0\|}.$$

Similarly,

$$f_2 = H_2 e_1 = (U_1 H_1)^T e_1 = H_1^T U_1 e_1.$$

Since from (7),  $U_1 e_1 = -(A_1 e_1)/\|A_1 e_1\|$ , then

$$f_2 = -\frac{\Lambda f_1}{\|\Lambda f_1\|} = \frac{(-1)^2 \Lambda^2 f_0}{\|\Lambda f_1\| \|\Lambda f_0\|}.$$

Since

$$f_1 = -\frac{\Lambda f_0}{\|\Lambda f_0\|},$$

then

$$\Lambda f_1 = \frac{\Lambda^2 f_0}{-\|\Lambda f_0\|} \quad \text{and} \quad \|\Lambda f_1\| = \left\| \frac{\Lambda^2 f_0}{-\|\Lambda f_0\|} \right\| = \frac{\|\Lambda^2 f_0\|}{\|\Lambda f_0\|}.$$

Then,

$$f_2 = (-1)^2 \frac{\Lambda^2 f_0}{\|\Lambda^2 f_0\|}.$$

By induction, it follows that

$$f_r = (-1)^r \frac{\Lambda^r f_0}{\|\Lambda^r f_0\|}.$$

Since, by hypothesis,  $e_n^T f_0 \neq 0$ , then

$$(-1)^r f_r \rightarrow (0, 0, \dots, 1)^T.$$

Since  $a_r^{11} = f_r^T \Lambda f_r$ , it follows that  $\lim_{r \rightarrow \infty} a_r^{11} = \lambda_n$ , and if  $e_{n-1}^T f_0 \neq 0$ , then

$$\lim_{r \rightarrow \infty} \left| \frac{a_r^{11} - \lambda_n}{a_{r-1}^{11} - \lambda_n} \right| = \left( \frac{\lambda_{n-1}}{\lambda_n} \right)^2.$$

It means that our proposed method converges linearly with convergence rate  $(\lambda_{n-1}/\lambda_n)^2$ .  $\square$

#### 4. The connection between the power method and the proposed one

The connection between the proposed method and the power method could be clarified by the following theorem.

**Theorem 5.** Consider the power method defined by

$$\mathbf{x}_0 = \mathbf{e}_1, \quad \mathbf{x}_r = \frac{A\mathbf{x}_{r-1}}{\|A\mathbf{x}_{r-1}\|}, \quad \rho_r = \mathbf{x}_r^T A \mathbf{x}_r, \quad r \geq 1.$$

Then,

$$\mathbf{x}_r = (-1)^r U_0 U_1 \cdots U_{r-1} \mathbf{e}_1 \quad \text{and} \quad \rho_r = a_1^{11}, \quad \text{for every } r \geq 1.$$

**Proof.** (i) For  $r = 1$ :

$$\mathbf{x}_1 = \frac{A\mathbf{x}_0}{\|A\mathbf{x}_0\|} = \frac{A\mathbf{e}_1}{\|A\mathbf{e}_1\|},$$

using (7); then,

$$\mathbf{x}_1 = -U_0 \mathbf{e}_1;$$

consequently,

$$\rho_1 = \mathbf{x}_1^T A \mathbf{x}_1 = \mathbf{e}_1^T U_0 A U_0 \mathbf{e}_1 = \mathbf{e}_1^T A_1 \mathbf{e}_1 = a_1^{11}.$$

So the theorem is valid for  $r = 1$ .

(ii) For  $r = m + 1$ : assuming that the theorem is valid for  $r = m$ , we will prove it for  $r = m + 1$ . By the induction hypothesis,

$$\mathbf{x}_m = (-1)^m U_0 U_1 \cdots U_{m-1} \mathbf{e}_1;$$

then,

$$\begin{aligned} A\mathbf{x}_m &= (-1)^m A U_0 U_1 \cdots U_{m-1} \mathbf{e}_1 = (-1)^m U_0 A_1 U_1 \cdots U_{m-1} \mathbf{e}_1 = (-1)^m U_0 U_1 A_2 \cdots U_{m-1} \mathbf{e}_1 \\ &= \cdots = (-1)^m U_0 U_1 U_2 \cdots U_{m-1} A_m \mathbf{e}_1. \end{aligned}$$

Using (7),

$$A\mathbf{x}_m = (-1)^m U_0 U_1 U_2 \cdots U_{m-1} (-\|A_m \mathbf{e}_1\| U_m \mathbf{e}_1) = (-1)^{m+1} \|A_m \mathbf{e}_1\| U_0 U_1 \cdots U_m \mathbf{e}_1.$$

Since  $\|A\mathbf{x}_m\| = \|A_m \mathbf{e}_1\|$ , then

$$\mathbf{x}_{m+1} = \frac{A\mathbf{x}_m}{\|A\mathbf{x}_m\|} = (-1)^{m+1} U_0 U_1 \cdots U_m \mathbf{e}_1.$$

Finally, the Rayleigh quotient at  $r = m + 1$  is

$$\rho_{m+1} = \mathbf{x}_{m+1}^T A \mathbf{x}_{m+1} = \mathbf{e}_1^T U_m U_{m-1} \cdots U_0 A U_0 U_1 \cdots U_m \mathbf{e}_1 = \mathbf{e}_1^T A_{m+1} \mathbf{e}_1 = a_{m+1}^{11},$$

and the proof is complete.  $\square$

## 5. Numerical examples

**Example 6.** Here we consider a  $6 \times 6$  Hilbert matrix  $A = [a^{ij}]$  defined by

$$a^{ij} = \frac{1}{i+j-1}, \quad 1 \leq i, j \leq 6.$$

Then,

$$\begin{aligned}
 A_1 &= \begin{bmatrix} 599.308 & -0.076195 & -0.081235 & -0.078978 & -0.075334 & -0.060945 \\ -0.076195 & 0.046760 & 0.050223 & 0.044605 & 0.040991 & 0.038012 \\ -0.081235 & 0.050223 & 0.061876 & 0.059875 & 0.055632 & 0.049568 \\ 0.078978 & 0.044605 & 0.059875 & 0.059807 & 0.057837 & 0.053895 \\ -0.075334 & 0.040991 & 0.055632 & 0.057837 & 0.055818 & 0.054132 \\ -0.060945 & 0.038012 & 0.049568 & 0.053895 & 0.054132 & 0.051716 \end{bmatrix}, \\
 A_2 &= \begin{bmatrix} 1.618453 & 0.009665 & 0.013066 & 0.012165 & 0.011752 & 0.009898 \\ 0.009665 & 0.44398 & 0.047510 & 0.042899 & 0.037976 & 0.033785 \\ 0.013066 & 0.047510 & 0.057290 & 0.055761 & 0.051852 & 0.047680 \\ 0.012165 & 0.042899 & 0.055761 & 0.056502 & 0.054069 & 0.050811 \\ 0.011752 & 0.037976 & 0.051852 & 0.054069 & 0.052789 & 0.050405 \\ 0.009898 & 0.033785 & 0.047680 & 0.050811 & 0.050405 & 0.048709 \end{bmatrix}, \\
 A_3 &= \begin{bmatrix} 1.618889 & -0.001438 & -0.001797 & -0.001842 & -0.001740 & -0.001592 \\ -0.001438 & 0.044377 & 0.048121 & 0.043001 & 0.037897 & 0.033722 \\ -0.001797 & 0.048121 & 0.057202 & 0.055655 & 0.051760 & 0.047631 \\ -0.001842 & 0.043001 & 0.055655 & 0.056403 & 0.053969 & 0.050729 \\ -0.001740 & 0.037897 & 0.051760 & 0.053969 & 0.052707 & 0.050320 \\ -0.001592 & 0.033722 & 0.047631 & 0.050729 & 0.050320 & 0.048634 \end{bmatrix}, \\
 A_4 &= \begin{bmatrix} 1.68889 & 0.000212 & 0.000259 & 0.000261 & 0.000258 & 0.000240 \\ 0.000212 & 0.044375 & 0.048119 & 0.043000 & 0.037896 & 0.033720 \\ 0.000259 & 0.048119 & 0.057200 & 0.055653 & 0.051758 & 0.047629 \\ 0.000261 & 0.043000 & 0.055653 & 0.056401 & 0.053967 & 0.050726 \\ 0.000258 & 0.037896 & 0.051758 & 0.053967 & 0.052705 & 0.050318 \\ 0.000240 & 0.033720 & 0.047629 & 0.050726 & 0.050318 & 0.048632 \end{bmatrix}.
 \end{aligned}$$

The convergence to the largest eigenvalue is quite fast as can be seen from the fact that  $\|\nu\|^2$  decreases to zero rapidly.

It has to be noticed that  $a_0^{11} = 1$ ,  $\nu_0 = [\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}]^T$ ,  $B_0 = [b_0^{ij}]$  where  $b_0^{ij} = 1/(i+j+1)$ ,  $1 \leq i, j \leq 5$ .

**Example 7.** This example was designed to show the effect of the shifts. A reasonable choice of the shift is  $\sigma = \frac{1}{2}(\lambda_1 + \lambda_{n-1})$ .

In this example, the eigenvalues are close to successive integers with  $\lambda_n \cong 10$  and  $\lambda_{n-1} \cong 9$ .

The matrix  $A$  is arranged such that  $a^{11} \geq a^{22} \geq \dots \geq a^{66}$  and after the  $k$ th iteration we use the shift

$$\sigma_k = \frac{1}{2}(a^{22}(k) + a^{66}(k)).$$

Then,

$$A = \begin{bmatrix} 10 & 1 & 0.5 & 0.25 & 0.125 & 0.1 \\ 1 & 9 & 0.4 & 0.2 & 0.1 & 0.05 \\ 0.5 & 0.4 & 8 & 0.3 & 0.15 & 0.1 \\ 0.25 & 0.2 & 0.3 & 7 & 0.5 & 0.4 \\ 0.125 & 0.1 & 0.15 & 0.5 & 6 & 1 \\ 0.1 & 0.05 & 0.1 & 0.4 & 1 & 5 \end{bmatrix}.$$

Without shifts, at the end of twenty iterations, we get

$$A_{20} = \begin{bmatrix} 10.798224 & 0.009873 & 0.003921 & 0.001318 & 0.000675 & 0.000481 \\ 0.009873 & 8.365704 & 0.098197 & 0.056740 & 0.023043 & -0.003464 \\ 0.003291 & 0.098197 & 7.858157 & 0.229358 & 0.114023 & 0.073459 \\ 0.001318 & 0.056740 & 0.229358 & 6.966996 & 0.477646 & 0.388721 \\ 0.000675 & 0.023043 & 0.114023 & 0.477646 & 5.990780 & 0.992488 \\ 0.000481 & -0.003464 & 0.073459 & 0.388721 & 0.992488 & 4.995558 \end{bmatrix}.$$

With shifts, at the end of seven iterations, we get

$$A_7 = \begin{bmatrix} -10.798302 & -0.004275 & -0.001436 & -0.000785 & -0.000200 & -0.000509 \\ -0.004275 & 8.365618 & 0.098312 & 0.056086 & 0.022983 & -0.003138 \\ -0.001436 & 0.098312 & 7.858186 & 0.229513 & 0.113912 & 0.073660 \\ -0.000785 & 0.056086 & 0.229513 & 6.967011 & 0.477631 & 0.388801 \\ -0.000200 & 0.022983 & 0.113912 & 0.477631 & 5.990887 & 0.992498 \\ -0.000509 & -0.003138 & 0.073660 & 0.388801 & 0.992498 & 4.995503 \end{bmatrix}.$$

## 6. Conclusion

To find all the eigenvalues of a symmetric matrix — or at least the first few largest ones — and the largest eigenvalue of a full positive definite matrix, a convergent numerical algorithm is proposed. The presented procedure could be easily implemented using APL programming language with a very brief code.

## Acknowledgement

The author thanks the referee of this paper for his valuable comments and suggestions.

## References

- [1] P. Deift and C. Tomei, Ordinary differential equations and the symmetric eigenvalue problem, *SIAM J. Numer. Anal.* **20** (1983) 1–22.
- [2] Li Baoxin et al., Eigenvalues of tridiagonal matrices, *Comput. Math. Appl.* **19** (4) (1990) 89–94.
- [3] B.N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).