AN ESTIMATE OF THE DIFFERENCE BETWEEN A DIAGONAL ELEMENT AND THE CORRESPONDING EIGENVALUE OF A SYMMETRIC TRIDIAGONAL MATRIX*

JIANG ER-XIONG (蒋尔雄)
(*Fudan University, Shanghai, China*)

Abstract

A sharper upperbound of the difference between a diagonal element and the corresponding eigenvalue of a symmetric tridiagonal matrix is given. The bound can be used in the QL and QR algorithms and Rayleigh quotient approximation. The change of eigenvalues is estimated when the first off-diagonal element β_1 is replaced by zero and when two neighboring off-diagonal elements β_{i-1}, β_i are replaced by zeros.

§ 1. Introduction

Let

$$T=T_{1,n}=\begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & 0 \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix}$$

be an unreduced symmetric tridiagonal matrix. Let

$$\lambda_1<\lambda_2<\cdots<\lambda_n$$

denote its eigenvalues. Let $\widetilde{T}_{1,n}$ be a matrix obtained by replacing β_1 in $T_{1,n}$ with zero. So α_1 is an eigenvalue of $\widetilde{T}_{1,n}$. Let

$$\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_n$$

denote n eigenvalues of $\widetilde{T}_{1,n}$ and $\alpha_1=\mu_j$. Hence

$$\mu_1,\ \mu_2,\ \cdots,\ \mu_{j-1},\ \mu_{j+1},\ \cdots,\ \mu_n$$

are $n-1$ eigenvalues of

$$T_{2,n}=\begin{pmatrix} \alpha_2 & \beta_2 & & 0 \\ \beta_2 & \alpha_3 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ 0 & & \beta_{n-1} & \alpha_n \end{pmatrix}.$$

* Received March 4, 1986.

How close is the diagonal element α_1 to an eigenvalue of T ? The question is important for the shifted QL algorithm. There are several results on this topic, see [1—5]. In [5], there is an eigenvalue λ of T, and $a=\min\limits_{\lambda_i\neq\lambda}|\alpha_1-\lambda_i|$ is the gap, then

$$|\alpha_1-\lambda|\leqslant\beta_1^2/a(1-\beta_1^2/a^2).$$

In [4], a result is

$$|\alpha_1-\lambda|\leqslant\beta_1^2/a.$$

In [1], for β_1^2 is sufficiently small, the estimation is

$$|\alpha_1-\lambda_j|\leqslant\beta_1^2/b,$$

where $b=\min\limits_{i\neq j}|\alpha_1-\mu_i|$.

In this paper a two-sided estimate of $\lambda_j-\alpha_1$ is given as follows:

$$\beta_1^2\sum_{k=j+1}^{n}\frac{s_{1k}^2}{\xi_2-\mu_k}\leqslant\lambda_j-\alpha_1\leqslant\beta_1^2\sum_{k=1}^{j-1}\frac{s_{1k}^2}{\xi_1-\mu_k},\tag{1}$$

where

$$\xi_1=\frac{1}{2}\{(\alpha_1+\mu_{j-1})+\sqrt{(\alpha_1-\mu_{j-1})^2+4\beta_1^2s_{1j-1}^2}\},$$

$$\xi_2=\frac{1}{2}\{(\alpha_1+\mu_{j+1})-\sqrt{(\alpha_1-\mu_{j+1})^2+4\beta_1^2s_{1j+1}^2}\},$$

s_k is a unit eigenvector of $T_{2,n}$ corresponding to the eigenvalue μ_k and s_{ik} is the first component of s_k. Because $\xi_1>\alpha_1$, $\xi_2<\alpha_1$ and $\sum\limits_{\substack{k=1\\k\neq j}}^{n}s_{1k}^2=1$, the result (1) of this paper is always sharper than the result in [1]. Furthermore, $s_{1,j-1}^2$ and $s_{1,j+1}^2$ are often small when n is big. They can offset the influence of a small gap such as $\xi_1-\mu_{j-1}$ and $\xi_2-\mu_{j+1}$. Even when $\alpha_1=\mu_j=\mu_{j-1}$ or $\alpha_1=\mu_{j+1}$, the result (1) is still available. This is different from the result in [1].

In Section 2, we also discuss the difference between $\mu_i(i\neq j)$, and the eigenvalue of T.

In Section 3, we consider the matrix

$$\hat{T}=\begin{pmatrix}T_{1,i-1} & & 0\\ & \alpha_i & \\ 0 & & T_{i+1,n}\end{pmatrix}.$$

If α_i is the j-th eigenvalue of $\hat{T}$, then a similar estimate of $\lambda_j-\alpha_i$ is given.

Define the vector $x=(x_1, x_2, \cdots, x_n)^T$ and $y=(y_1, y_2, \cdots, y_n)^T$. Hereafter th norm will be denoted by $\|x\|=(x_1^2+x_2^2+\cdots+x_n^2)^{\frac{1}{2}}$ and the inner-product by $(x, y)=x_1y_1+\cdots+x_ny_n$.

§ 2. Estimate for α_1 or α_n

Let

$$T=T_{1,n}=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix},$$

$$T_{2,n}=\begin{pmatrix} \alpha_2 & \beta_2 & & 0 \\ \beta_2 & \alpha_3 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ 0 & & \beta_{n-1} & \alpha_n \end{pmatrix}$$

be unreduced matrices.

Let

$$\tilde{T}=\begin{pmatrix} \alpha_1 & \\ 0 & T_{2,n} \end{pmatrix}$$

and let

$$\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_n$$

denote the eigenvalues of $\tilde{T}$. Obviously α_1 is an eigenvalue of $\tilde{T}$. If $\alpha_1=\mu_j$, then μ_1, $\mu_2, \cdots, \mu_{j-1}, \mu_{j+1}, \cdots, \mu_n$ are eigenvalues of $T_{2,n}$. Let s_k be a unit eigenvector of $T_{2,n}$ corresponding to the eigenvalue μ_k. Denote by s_{ik} the ith component of s_k. Therefore

$$\begin{pmatrix} 0 \\ s_1 \end{pmatrix}, \begin{pmatrix} 0 \\ s_2 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ s_{j-1} \end{pmatrix}, e_1, \begin{pmatrix} 0 \\ s_{j+1} \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ s_n \end{pmatrix}$$

are the unit eigenvectors of $\tilde{T}$ corresponding to the eigenvalues $\mu_1, \mu_2, \cdots, \mu_n$ and they are a basis of R^n. Let

$$E=\begin{pmatrix} 0 & \beta_1 & 0 \\ \beta_1 & 0 & \\ 0 & & \ddots & \\ & & & 0 \end{pmatrix}.$$

We have $T=\tilde{T}+E$. It is clear that the eigenvalues of E are

$$-\beta_1, 0, 0, \cdots, 0, \beta_1.$$

Denote the corresponding unit eigenvectors by $g_1, g_2, \cdots, g_n$. We know

$$g_1=(e_1-e_2)/\sqrt{2}, \quad g_n=(e_1+e_2)/\sqrt{2}$$

and can choose

$$g_i=e_{i+1}, \quad i=2, 3, \cdots, n-1,$$

where e_i is the i-th column of an identity matrix.

Theorem 1. *If α_1 is the j-th eigenvalue of $\tilde{T}$, then*

$$\beta_1^2 \sum_{k=j+1}^{n} \frac{s_{1k}^2}{\xi_2-\mu_k} \leqslant \lambda_j-\alpha_1 \leqslant \beta_1^2 \sum_{k=1}^{j-1} \frac{s_{1k}^2}{\xi_1-\mu_k}, \tag{1}$$

where

$$\xi_1=((\alpha_1+\mu_{j-1})+\sqrt{(\alpha_1-\mu_{j-1})^2+4\beta_1^2 s_{1j-1}^2})/2,$$
$$\xi_2=((\alpha_1+\mu_{j+1})-\sqrt{(\alpha_1-\mu_{j+1})^2+4\beta_1^2 s_{1j+1}^2})/2.$$

Proof. From the Courant-Fischer theorem[4, p.190],

$$\lambda_j=\min_{V_j}\max_{x\in V_j}\rho(T,x),$$

where V_j is a j-dimensional subspace and $\rho(T,x)=(Tx,x)/(x,x)$ is the Rayleigh quotient of x, T. First we suppose $j>1$.

Denote a special j-dimensional subspace by

$$\tilde{V}_j=\left\{\begin{pmatrix}0\\ s_1\end{pmatrix},\ \begin{pmatrix}0\\ s_2\end{pmatrix},\ \cdots,\ \begin{pmatrix}0\\ s_{j-1}\end{pmatrix},\ e_1\right\}.$$

So

$$\lambda_j\leqslant\max_{x\in\tilde{V}_j}\rho(T,x)=\max_{x\in\tilde{V}_j}\{\rho(\tilde{T},x)+\rho(E,x)\}.$$

For any $x\in\tilde{V}_j$, $\|x\|=1$, $x=\sum\limits_{k=1}^{j-1}\delta_k\begin{pmatrix}0\\ s_k\end{pmatrix}+\delta_j e_1$ and $\sum\limits_{k=1}^{j}\delta_k^2=1$.

$$\rho(\tilde{T},x)=(\tilde{T}x,x)=(\delta_1,\delta_2,\cdots,\delta_j)\begin{pmatrix}\mu_1 & & 0\\ & \mu_2 & \\ 0 & & \ddots \\ & & & \mu_j\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\\ \vdots\\ \delta_j\end{pmatrix}.$$

Let $x=\sum\limits_{i=1}^{n}\eta_i g_i$, so

$$\rho(E,x)=-\beta_1\eta_1^2+\beta_1\eta_n^2.$$

From

$$\sum_{i=1}^{n}\eta_i g_i=\sum_{k=1}^{j-1}\delta_k\begin{pmatrix}0\\ s_k\end{pmatrix}+\delta_j e_1,$$

we have

$$(\eta_1+\eta_n)/\sqrt{2}=\delta_j,\quad(\eta_n-\eta_1)/\sqrt{2}=\sum_{k=1}^{j-1}\delta_k s_{1k}.$$

Therefore

$$\eta_1=\Big(\delta_j-\sum_{k=1}^{j-1}\delta_k s_{1k}\Big)\frac{\sqrt{2}}{2},$$
$$\eta_n=\Big(\delta_j+\sum_{k=1}^{j-1}\delta_k s_{1k}\Big)\frac{\sqrt{2}}{2}$$

and

$$\rho(E,x)=-\beta_1\eta_1^2+\beta_1\eta_n^2$$
$$=(\delta_1,\delta_2,\cdots,\delta_j)\begin{pmatrix} & & & \beta_1 s_{11}\\ & 0 & & \vdots\\ & & & \beta_1 s_{1j-1}\\ \beta_1 s_{11}, & \beta_1 s_{12}, & \cdots,\ \beta_1 s_{1j-1}, & 0\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\\ \vdots\\ \delta_j\end{pmatrix}.$$

Hence

$$\rho(T, x)=(\delta_1, \delta_2, \cdots, \delta_j)\begin{pmatrix} \mu_1 & & & & \beta_1 s_{11} \\ & \mu_2 & & 0 & \beta_1 s_{12} \\ & & \ddots & & \vdots \\ & 0 & & \mu_{j-1} & \beta_1 s_{1j-1} \\ \beta_1 s_{11} & \beta_1 s_{12} & \cdots & \beta_1 s_{1j-1} & \alpha_1 \end{pmatrix}\begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_j \end{pmatrix}.$$

Let $\max\limits_{x\in\tilde{V}} \rho(T, x)=\Omega_j$, where Ω_j is the biggest eigenvalue of matrix

$$G=\begin{pmatrix} \mu_1 & & & 0 & \beta_1 s_{11} \\ & \mu_2 & & & \beta_1 s_{12} \\ & & \ddots & & \vdots \\ & 0 & & \mu_{j-1} & \beta_1 s_{1j-1} \\ \beta_1 s_{11} & \beta_1 s_{12} & \cdots & \beta_1 s_{1j-1} & \alpha_1 \end{pmatrix}.$$

It is easy to show

$$\mathrm{Det}\,(\lambda I-G)=\prod_{k=1}^{j-1}(\lambda-\mu_k)\Big(\lambda-\alpha_1-\sum_{p=1}^{j-1}\frac{\beta_1^2 s_{1p}^2}{\lambda-\mu_p}\Big).$$

Because Det $(\mu_i I-G)\neq 0$, $i=1, 2, \cdots, j-1$, the eigenvalues of G satisfy an equation

$$\lambda-\alpha_1-\sum_{p=1}^{j-1}\frac{\beta_1^2 s_{1p}^2}{\lambda-\mu_p}=0. \tag{2}$$

Ω_j is the biggest root of equation (2) and the roots of (2) are the points of intersection of line

$$y=\lambda-\alpha_1 \tag{3}$$

and curve

$$y=\sum_{p=1}^{j-1}\frac{\beta_1^2 s_{1p}^2}{\lambda-\mu_p}. \tag{4}$$

The biggest branch of the curve (4) lies on $\lambda>\mu_{j-1}$. Ω_j is the abscissa of the point of intersection of the branch and line (3). The biggest branch of the curve

$$y=\frac{\beta_1^2 s_{1j-1}^2}{\lambda-\mu_{j-1}} \tag{5}$$

also lies on $\lambda>\mu_{j-1}$. The biggest root of

$$\lambda-\alpha_1=\frac{\beta_1^2 s_{1j-1}}{\lambda-\mu_{j-1}}$$

is

$$\xi_1=((\alpha_1+\mu_{j-1})+\sqrt{(\alpha_1-\mu_{j-1})^2+4\beta_1^2 s_{1j-1}^2})/2.$$

Because on $\lambda>\mu_{j-1}$, $\sum\limits_{p=1}^{j-1}\frac{\beta_1^2 s_{1p}^2}{\lambda-\mu_p}>\frac{\beta_1^2 s_{1,j-1}^2}{\lambda-\mu_{j-1}}$, $\xi_1\leqslant\Omega_j$. We have

$$\Omega_j=\alpha_1+\sum_{k=1}^{j-1}\frac{\beta_1^2 s_{1k}^2}{\Omega_j-\mu_k}\leqslant\alpha_1+\sum_{k=1}^{j-1}\frac{\beta_1^2 s_{1k}^2}{\xi_1-\mu_k}$$

and

$$\lambda_j-\alpha_1\leqslant\sum_{k=1}^{j-1}\frac{\beta_1^2 s_{1k}^2}{\xi_1-\mu_k}.$$

When $j=1$, obviously $\lambda_1-\alpha_1\leqslant 0$. Therefore the right side of inequality (1) holds.

From

$$\lambda_j=\max_{V_{n-j+1}}\min_{x\in V_{n-j+1}}\rho(T,x),$$

the left side of inequality (1) can be proved analogously.

Corollary. If $\beta_1\neq 0$, then

$$|\lambda_j-\alpha_1|<\beta_1^2/b, \tag{6}$$

where $b=\min_{i\neq j}|\mu_i-\alpha_1|$.

Proof. Since $s_{1k}\neq 0$, $\xi_1>\alpha_1$ and $\xi_2<\alpha_1$. We have

$$\sum_{k=1}^{j-1}\frac{s_{1k}^2}{\xi_1-\mu_k}<\sum_{k=1}^{j-1}\frac{s_{1k}^2}{\alpha_1-\mu_k}\leqslant\frac{1}{\alpha_1-\mu_{j-1}},$$

$$\sum_{k=j+1}^{n}\frac{s_{1k}^2}{\xi_2-\mu_k}>\sum_{k=j+1}^{n}\frac{s_{1k}^2}{\alpha_1-\mu_k}\geqslant\frac{1}{\alpha_1-\mu_{j+1}}.$$

Hence the corollary holds.

Example 1.

$$T_{1,6}=\begin{pmatrix} 8 & 0.01 & 0 & 0 & 0 & 0\\ 0.01 & 1 & 1 & 0 & 0 & 0\\ 0 & 1 & 3 & 1 & 0 & 0\\ 0 & 0 & 1 & 5 & 1 & 0\\ 0 & 0 & 0 & 1 & 7 & 1\\ 0 & 0 & 0 & 0 & 1 & 9\end{pmatrix}.$$

Estimate the difference between $\alpha_1=8$ and the corresponding eigenvalue of $T_{1,6}$.

The eigenvalues of $T_{2,6}$ are

$$\mu_1=0.549129335,\quad \mu_2=2.953112038,$$
$$\mu_3=5.000000000,\quad \mu_4=7.046887962,$$
$$\mu_6=9.450870665,$$

and the corresponding s_{1k} are

$$s_{1,1}=0.907370,\quad s_{1,2}=0.405306,$$
$$s_{1,3}=0.109764,\quad s_{1,4}=0.0188497,$$
$$s_{1,6}=0.00178075.$$

$\alpha_1=8=\mu_5$, $j=5$, $\xi_1>\alpha_1=8$, $\xi_2<\alpha_1=8$,

$$\beta_1^2\sum_{k=1}^{4}\frac{s_{1k}^2}{\alpha_1-\mu_k}=10^{-4}*0.14744,$$

$$\beta_1^2\sum_{k=6}^{6}\frac{s_{1k}^2}{\alpha_1-\mu_k}=-10^{-10}*2.1856.$$

Hence by Theorem 1 we have

$$-2.1856*10^{-10}<\lambda_5-\alpha_1<1.4744*10^{-5}.$$

According to the result of [1]

$$|\lambda_5-\alpha_1|\leqslant\beta_1^2/(\alpha_1-\mu_4)=1.0492*10^{-4}.$$

Example 2. Let

$$A=\begin{pmatrix}1 & \varepsilon & \varepsilon\\ \varepsilon & 2 & \varepsilon\\ \varepsilon & \varepsilon & 3\end{pmatrix},$$

$\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of A. When $\varepsilon=10^{-5}$, the estimates of λ_i-i given in [6, p. 47] are

$$-2\varepsilon^2/(1-\sqrt{2}\,\varepsilon)\leqslant\lambda_1-1\leqslant 0,$$
$$-2\varepsilon^2/(1-\sqrt{2}\,\varepsilon)\leqslant\lambda_2-2\leqslant 2\varepsilon^2/(1-\sqrt{2}\,\varepsilon),$$
$$0\leqslant\lambda_4-3\leqslant 2\varepsilon^2/(1-\sqrt{2}\,\varepsilon).$$

Using similarity transformation by rotations and permutations, we have

$$A\backsim\begin{pmatrix}1 & \sqrt{2}\,\varepsilon & 0\\ \sqrt{2}\,\varepsilon & 2.5+\varepsilon & 0.5\\ 0 & 0.5 & 2.5-\varepsilon\end{pmatrix}\backsim\begin{pmatrix}2 & \sqrt{2}\,\varepsilon & 0\\ \sqrt{2}\,\varepsilon & 2+\varepsilon & 1\\ 0 & 1 & 2-\varepsilon\end{pmatrix}$$
$$\backsim\begin{pmatrix}3 & \sqrt{2}\,\varepsilon & 0\\ \sqrt{2}\,\varepsilon & 1.5+\varepsilon & -0.5\\ 0 & -0.5 & 1.5-\varepsilon\end{pmatrix}$$

and from Theorem 1 we get

$$-\frac{1.5\varepsilon^2}{1-\varepsilon^2}<\lambda_1-1\leqslant 0,\quad 0<\varepsilon<1,$$

$$-(1+\varepsilon)\varepsilon^2<\lambda_2-2<\frac{\varepsilon^2}{1+\varepsilon},\quad 0<\varepsilon,$$

$$0\leqslant\lambda_3-3<\frac{1.5+\varepsilon}{1-1.5\varepsilon^2}\,\varepsilon^2,\quad 0<\varepsilon<\sqrt{\frac{1}{1.5}}.$$

These estimates of λ_i-i are better than those in [6].

Theorem 1 can be applied to the Rayleigh quotient approximation. For any $n\times n$ symmetric matrix A and a unit vector $q_1\in R^n$, the Lanczos process is as follows:

$$\beta_1q_2=Aq_1-\alpha_1q_1,\quad \alpha_1=(Aq_1, q_1),\quad \beta_1=\|Aq_1-\alpha_1q_1\|,$$
$$\beta_iq_{i+1}=Aq_i-\alpha_iq_i-\beta_{i-1}q_{i-1},\quad \alpha_i=(Aq_i,q_i),$$
$$\beta_i=\|Aq_i-\alpha_iq_i-\beta_{i-1}q_{i-1}\|,\quad i=2, 3, \cdots, m.$$

If $\beta_m=0$, $m\leqslant n$, then $Q=(q_1, q_2, \cdots, q_m)$ with unit orthogonal columns and the process produces a symmetric tridiagonal matrix

$$T_{1,m}=\begin{pmatrix}\alpha_1 & \beta_1 & & & \mathbf{0}\\ \beta_1 & \alpha_2 & \beta_3 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{m-1}\\ \mathbf{0} & & & \beta_{m-1} & \alpha_m\end{pmatrix}.$$

Denote such a matrix by

$$T_{1,m}=L(A, q_1).$$

Theorem 2. *For any unit vector y, if the Rayleigh quotient*

$$y^T Ay = (Ay, y) = \alpha_1$$

and $r = Ay - \alpha_1 y$, *let* $T_{1,m} = L(A, y)$. *Then there is an eigenvalue* λ_{i_j} *of* A, *such that*

$$\|r\|^2 \sum_{k=j+1}^{n} \frac{s_{1k}^2}{\xi_2 - \mu_k} \leqslant \lambda_{i_j} - \alpha_1 \leqslant \|r\|^2 \sum_{k=1}^{j-1} \frac{s_{1k}^2}{\xi_1 - \mu_k}, \tag{7}$$

where μ_k *is the* k*-th eigenvalue of* $T_{2,m}$, s_k *is the unit eigenvector corresponding to* μ_k, s_{1k} *is the first component of* s_k *and* ξ_1, ξ_2 *are defined as in Theorem* 1.

Proof. From $T_{1,m} = L(A, y)$, we know $\beta_1 = \|r\|$. The eigenvalues of $T_{1,m}$ are the eigenvalues of A. Let these eigenvalues be

$$\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_m}.$$

Applying Theorem 1 to $T_{1,m}$ we get the estimate (7) immediately.

Now let us discuss the approximations of another eigenvalue rather than α_1 of $\tilde{T}$.

Theorem 3. *For any* $k \neq j$, *if*

$$\hat{T}_{1,m} = L\left(T, \begin{pmatrix} 0 \\ s_k \end{pmatrix}\right),$$

$\hat{\mu}_i$ *is the eigenvalue of* $\hat{T}_{2,m}$, $\hat{s}_i$ *is the unit eigenvector corresponding to* $\hat{\mu}_1$, *and* $\hat{s}_{1i}$ *is the first component of* $\hat{s}_i$, *then there is an eigenvalue* λ *of* T *such that*

$$\beta_1^2 s_{1k}^2 \sum_{i=j+1}^{m} \frac{\hat{s}_{1i}^2}{\xi_2 - \hat{\mu}_i} \leqslant \lambda - \mu_k \leqslant \beta_1^2 s_{1k}^2 \sum_{i=1}^{j-1} \frac{\hat{s}_{1i}^2}{\xi_1 - \hat{\mu}_i}, \tag{8}$$

where ξ_1 *and* ξ_2 *are defined as in Theorem* 1.

Proof. Let us denote the orthogonal matrix by

$$Q = \left(\begin{pmatrix} 0 \\ s_1 \end{pmatrix}, \begin{pmatrix} 0 \\ s_2 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ s_{j-1} \end{pmatrix}, e_1, \begin{pmatrix} 0 \\ s_{j+1} \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ s_n \end{pmatrix}\right)$$

and let

$$\hat{T} = Q^T T Q = \begin{pmatrix} \mu_1 & & & & \beta_1 s_{11} & & & \\ & \mu_2 & & 0 & \beta_1 s_{12} & & 0 & \\ & 0 & \ddots & & \vdots & & & \\ & & & \mu_{j-1} & \beta_1 s_{1j-1} & & & \\ \beta_1 s_{11} & \beta_1 s_{12} & \cdots & \beta_1 s_{1j-1} & \alpha_1 & \beta_1 s_{j+1} & \cdots & \beta_1 s_{1n} \\ & & & & \beta_1 s_{1j+1} & \mu_{j+1} & & 0 \\ & & 0 & & \vdots & & \ddots & \\ & & & & \beta_1 s_{1n} & 0 & & \mu_n \end{pmatrix}.$$

$\hat{T}$ is similar to T, and they have the same eigenvalues.

Let $y = e_k$, its Rayleigh quotient is

$$(\hat{T}y, y) = \mu_k.$$

We have

$$\|\hat{T}y - \mu_k \hat{y}\| = |\beta_1 s_{1k}|.$$

Denote

$$\hat{T}_{1,m} = L(\hat{T}, e_k).$$

Applying Theorem 2 we get the estimate (8).

On the other hand

$$L(\hat{T}, e_k)=L\left(T, \begin{pmatrix} 0 \\ s_k \end{pmatrix}\right).$$

Hence we have proved the theorem.

In Theorem 3 there is no guarantee that the λ in inequality (8) just equals λ_k. But if

$$\min_{i\neq k}|\lambda_i-\lambda_k|>\omega_k+|\beta_1| \tag{9}$$

then $\lambda=\lambda_k$, where

$$\omega_k=\max\left(\beta_1^2 s_{1k}^2 \sum_{i=1}^{j-1}\frac{\hat{s}_{1i}^2}{\xi_1-\hat{\mu}_i},\ \beta_1^2 s_{1k}^2 \sum_{i=j+1}^{m}\frac{\hat{s}_{1i}^2}{\xi_2-\hat{\mu}_i}\right).$$

Indeed from Theorem 3

$$|\lambda-\mu_k|\leqslant\omega_k.$$

By Wielandt-Hoffman theorem[5, p.104], we have

$$|\lambda_k-\mu_k|\leqslant|\beta_1|.$$

If $\lambda\neq\lambda_k$, we have

$$|\lambda-\lambda_k|=|\lambda-\mu_k+\mu_k-\lambda_k|\leqslant|\beta_1|+\omega_k.$$

It is in contradiction with (9), so $\lambda=\lambda_k$.

Now we consider the case when β_{n-1} of T is replaced by zero. Let

$$\tilde{\tilde{T}}=\begin{pmatrix} T_{1,n-1} & 0 \\ 0 & \alpha_n \end{pmatrix}, \qquad T_{1,n-1}=\begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & 0 \\ & \beta_2 & \ddots & \ddots & \\ 0 & & \ddots & \ddots & \beta_{n-2} \\ & & & \beta_{n-2} & \alpha_{n-1} \end{pmatrix}$$

and let

$$\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_n$$

be the eigenvalues of $\tilde{\tilde{T}}$. If $\alpha_n=\mu_j$ and s_k is the unit eigenvector of $T_{1,n-1}$ corresponding to the eigenvalue μ_k, $k\neq j$, we have

Theorem 4. *If the j-th eigenvalue of T is λ_j, then*

$$\beta_{n-1}^2 \sum_{k=j+1}^{n}\frac{s_{n-1,k}^2}{\xi_2-\mu_k}\leqslant\lambda_j-\alpha_n\leqslant\beta_{n-1}^2 \sum_{k=1}^{j-1}\frac{s_{n-1,k}^2}{\xi_1-\mu_k},$$

where

$$\xi_1=\left((\alpha_n+\mu_{j-1})+\sqrt{(\alpha_n-\mu_{j-1})^2+4\beta_{n-1}^2 s_{n-1,j-1}^2}\right)/2,$$
$$\xi_2=\left((\alpha_n+\mu_{j+1})-\sqrt{(\alpha_n-\mu_{j+1})^2+4\beta_{n-1}^2 s_{n-1,j+1}^2}\right)/2.$$

Proof. It is similar to the proof of Theorem 1.

§ 3. The Estimate for Another α_i

Let

$$\hat{T}=\begin{pmatrix} T_{1,i-1} & & \\ & \alpha_i & \\ 0 & & T_{i+1} \end{pmatrix},$$

$$\hat{E}=\begin{pmatrix} 0 & & & & & \\ & 0 & & & & \\ & & \ddots & \beta_{i-1} & & \\ & & \beta_{i-1} & 0 & \beta_i & \\ & & & \beta_i & \ddots & \\ & & & & & 0 \end{pmatrix},$$

$T=\hat{T}+\hat{E}$. We know the eigenvalues of $\hat{E}$ are

$$-p,\ 0,\ 0,\ \cdots,\ 0,\ p,$$

where $p=\sqrt{\beta_{i-1}^2+\beta_i^2}$. Denote the unit eigenvectors corresponding to the eigenvalues by $g_1, g_2, \cdots, g_n$. It is easy to know

$$g_1=\left(-\frac{\beta_{i-1}}{p}e_{i-1}+e_i-\frac{\beta_i}{p}e_{i+1}\right)\Big/2,$$

$$g_2=\frac{\beta_i}{p}e_{i-1}-\frac{\beta_{i-1}}{p}e_{i+1},$$

$$g_n=\left(\frac{\beta_{i-1}}{p}e_{i-1}+e_i+\frac{\beta_i}{p}e_{i+1}\right)\Big/2,$$

$$g_k=e_{k-2},\qquad k=3,\ 4,\ \cdots,\ i,$$

$$g_k=e_{k+1},\qquad k=i+1,\ i+2,\ \cdots,\ n-1.$$

Let

$$\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_n$$

be the eigenvalues of $\hat{T}$ and $\alpha_i=\mu_j$ be the j-th eigenvalue. Let $y_1, y_2, \cdots, y_{i-1}$ be $i-1$ unit orthogonal eigenvectors of $T_{1,i-1}$ and $z_{i+1}, z_{i+2}, \cdots, z_n$ be $n-i$ unit orthogonal eigenvectors of $T_{i+1,n}$. So

$$\begin{pmatrix} y_1\\0\\0\end{pmatrix},\ \begin{pmatrix} y_2\\0\\0\end{pmatrix},\ \cdots,\ \begin{pmatrix} y_{i-1}\\0\\0\end{pmatrix},\ e_i,\ \begin{pmatrix} 0\\0\\z_{i+1}\end{pmatrix},\ \cdots,\ \begin{pmatrix} 0\\0\\z_n\end{pmatrix} \tag{10}$$

are n unit orthogonal eigenvectors of $\hat{T}$. Let us reorder these n eigenvectors as $s_1, s_2, \cdots, s_n$ such that

$$\hat{T}s_k=\mu_k s_k.$$

So $e_i=s_j$. Since $T_{1,i-1}$ and $T_{i+1,n}$ are both unreduced,

$$s_{i-1,k}^2+s_{i+1,k}^2\neq 0,\quad k\neq j.$$

By (10)

$$s_{i-1,k}s_{i+1,k}=0.$$

Let λ_j be the j-th eigenvalue of T. We have

$$\lambda_j=\min_{V_j}\max_{x\in V_j}(T,\ x).$$

Denote the j-dimensional subspace

$$\tilde{V}_j=\{s_1,\ s_2,\ \cdots,\ s_j\};$$

so

$$\lambda_j\leqslant\max_{x\in\tilde{V}_j}\rho(T,\ x)=\max_{x\in\tilde{V}_j}(\rho(\hat{T},\ x)+\rho(\hat{E},\ x)).$$

For any $x\in\tilde{V}_j$, $\|x\|=1$, we have

$$x=\delta_1 s_1+\delta_2 s_2+\cdots+\delta_j s_j,\quad \sum_{k=1}^{j}\delta_k^2=1$$

and

$$\rho(\hat{T},\ x)=\sum_{k=1}^{j}\mu_k\delta_k^2.$$

Because $g_1, g_2, \cdots, g_n$ form a basis of R^n,

$$x=\sum_{k=1}^{n}\eta_k g_k$$

and

$$\rho(\hat{E},\ x)=-p\eta_1^2+p\eta_n^2.$$

By comparing the $i-1$–th, i–th, $i+1$–th components of the equality

$$\sum_{k=1}^{n}\eta_k g_k=\sum_{k=1}^{j}\delta_k s_k,$$

we have three equations as follows:

$$(\eta_1+\eta_n)/\sqrt{2}=\delta_j, \tag{11}$$

$$-\frac{\beta_{i-1}}{\sqrt{2}\ p}\eta_1+\frac{\beta_i}{p}\eta_2+\frac{\beta_{i-1}}{\sqrt{2}\ p}\eta_n=\sum_{k=1}^{j-1}\delta_k s_{i-1,k}, \tag{12}$$

$$-\frac{\beta_i}{\sqrt{2}\ p}\eta_1-\frac{\beta_{i-1}}{p}\eta_2+\frac{\beta_i}{\sqrt{2}\ p}\eta_n=\sum_{k=1}^{j-1}\delta_k s_{i+1,k}. \tag{13}$$

Multiplying (12) by β_{i-1} and (13) by β_i and adding the results, we get

$$-\frac{1}{\sqrt{2}}p\eta_1+\frac{1}{\sqrt{2}}p\eta_n=\sum_{k=1}^{j-1}\delta_k(\beta_{i-1}s_{i-1,k}+\beta_i s_{i+1,k}). \tag{14}$$

From (11) and (14)

$$\eta_1=\frac{\sqrt{2}}{2}\Big(\delta_j-\sum_{k=1}^{j-1}\delta_k\omega_k\Big),$$

$$\eta_n=\frac{\sqrt{2}}{2}\Big(\delta_j+\sum_{k=1}^{j-1}\delta_k\omega_k\Big),$$

$$\omega_k=\frac{\beta_{i-1}}{p}s_{i-1,k}+\frac{\beta_i}{p}s_{i+1,k}.$$

Then as in Theorems 1 and 3, we have

Theorem 5. *If α_i is the j-th eigenvalue of $\hat{T}$, then*

$$p^2\sum_{k=j+1}^{n}\frac{\omega_k^2}{\hat{\xi}_2-\mu_k}\leqslant\lambda_j-\alpha_i\leqslant p^2\sum_{k=1}^{j-1}\frac{\omega_k^2}{\hat{\xi}_1-\mu_k},$$

where

$$\hat{\xi}_1=((\mu_{j-1}+\alpha_i)+\sqrt{(\alpha_i-\mu_{j-1})^2+4p^2\omega_{j-1}^2})/2,$$
$$\hat{\xi}_2=((\mu_{j+1}+\alpha_i)-\sqrt{(\alpha_i-\mu_{j+1})^2+4p^2\omega_{j+1}^2})/2.$$

Theorem 6. *For any eigenvalue μ_k of $\hat{T}$ $k\neq j$, there is an eigenvalue λ of T such that*

$$(p\ \omega_k)^2\sum_{q=j+1}^{m}\frac{\hat{s}_{1,q}^2}{\tilde{\xi}_2-\hat{\mu}_q}\leqslant\lambda-\mu_k\leqslant(p\omega_k)^2\sum_{q=1}^{j-1}\frac{\hat{s}_{1,q}^2}{\tilde{\xi}_1-\hat{\mu}_q},$$

where $\hat{s}_{1,q}$ *is the first component of* $\hat{s}_q$, $\hat{s}_q$ *is a unit eigenvector of* $\hat{\hat{T}}_{2,m}$ *corresponding to the eigenvalue* $\hat{\mu}_q$ *and*

$$\hat{\hat{T}}_{1,m} = L(T, s_k).$$

Example 3.

$$T = \begin{pmatrix} 1 & 1 & & & & & & & & \\ 1 & 3 & 1 & & & & & & & \\ & 1 & 5 & 1 & & & & & & \\ & & 1 & 7 & x & & & & & \\ & & & x & 9 & x & & & & \\ & & & & x & 11 & 1 & & & \\ & & & & & 1 & 13 & 1 & & \\ & & & & & & 1 & 15 & 1 & \\ & & & & & & & 1 & 17 & \\ & & & & & & & & & 1 \end{pmatrix}.$$

$$T_{1,4} = \begin{pmatrix} 1 & 1 & & \\ 1 & 3 & 1 & \\ & 1 & 5 & 1 \\ & & 1 & 7 \end{pmatrix}$$

has eigenvalues and $s_{4,k}$ as follows:

$$\mu_1 = 0.5491550231, \qquad s_{4,1} = 0.01476,$$
$$\mu_2 = 2.955168392, \qquad s_{4,2} = 0.10907,$$
$$\mu_3 = 5.044831607, \qquad s_{4,3} = 0.40562,$$
$$\mu_4 = 7.450844376, \qquad s_{4,4} = 0.90739.$$

$T_{6,9} = 10I + T_{1,4}$; so

$$\mu_k = 10 + \mu_{k-5}, \quad k = 6, 7, 8, 9.$$
$$s_{1,6} = 0.90739, \qquad s_{1,7} = 0.40562,$$
$$s_{1,8} = 0.10907, \qquad s_{1,9} = 0.01476.$$

$\alpha_5 = 9 = \mu_5$, therefore using Theorem 5 we obtain

$$-0.575x^2 \leqslant \lambda_5 - \alpha_5 \leqslant 0.575x^2.$$

References

[1] Willam W. Hager, Perturbations in Eigenvalues, *Linear Algebra Appl.*, **42** (1982), 39—55.
[2] W. Kahan, When to neglect off-diagonal elements of symmetric tridiagonal matrices, Computer Science Department, Stanford Univ. Report CS42, July 1966.
[3] C. C. Paige, Eigenvalues of perturbed Hermitian matrices, *Linear Algebra Appl.*, **8** (1974), 1—10.
[4] B. N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall Inc., Englewood Cliffs, 1980.
[5] J. H. Wilkinson, The Algebraic Eigenvalue Problem, Oxford U. P., London, 1965.
[6] F. Chatelin, Spectral Approximation of Linear Operators, Academic Press, 1983.