A note on convergence and \textit{a posteriori} error estimates of the classical Jacobi method

\textit{Takuya Tsuchiya}\textsuperscript{1a)} and Kensuke Aishima\textsuperscript{2}

\textsuperscript{1} Graduate School of Science and Engineering, Ehime University
2-5 Bunkyo-cho, Matsuyama 790-8577, Japan

\textsuperscript{2} Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

\textsuperscript{a)} tsuchiya@math.sci.ehime-u.ac.jp

Received October 18, 2014; Revised January 15, 2015; Published July 1, 2015

\textbf{Abstract:} We consider convergence and \textit{a posteriori} error estimates of the classical Jacobi method for solving symmetric eigenvalue problems. The famous convergence proof of the classical Jacobi method consists of two phases. First, it is shown that all the off-diagonal elements converge to zero. Then, from a perturbation theorem, Parlett or Wilkinson shows convergence of the diagonal elements in the textbooks. Ciarlet also gives another convergence proof based on a discussion about a bounded sequence corresponding to a diagonal element. In this paper, we simplify the Ciarlet’s convergence proof. Our proof does not use any perturbation theory. Moreover, employing this approach, we obtain \textit{a posteriori} error estimates for eigenvectors.

\textbf{Key Words:} eigenvalues, classical Jacobi method, error estimations

\section{Introduction}
Let $A = (a_{ij})_{i,j=1}^n$ be a real symmetric matrix of order $n \geq 2$ with the eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \subset \mathbb{R}$. The Jacobi method is the most basic method for computing the all eigenvalues and the corresponding eigenvectors of $A$, and is based on consecutive rotations for annihilating off-diagonal elements. Suppose that a certain off-diagonal element $a_{pq}$, $(1 \leq p < q \leq n)$ is chosen as the pivot. Then, we define the orthogonal matrix $Q_1 = (s_{ij}^{(1)})_{i,j=1}^n$ by

For the method iterates the procedure until all off-diagonal elements are sufficiently annihilated. We denote block versions of the Jacobi methods. Convergence analysis for the block versions are currently.

This means that by this procedure. The Jacobi method iterates the procedure until all off-diagonal elements are sufficiently annihilated. We denote the kth matrix by \( A_k = (a^{(k)}_{ij})_{i,j=1}^n \). There are several ways of choosing the pivot \( a_{pq}^{(k)} \). The most basic manner is taking the pivot such that \( |a_{pq}^{(k)}| = \max_{i,j=1,\ldots,n} |a_{ij}^{(k)}| \), and in this case the algorithm is called the classical Jacobi method.

**Algorithm 1 (The classical Jacobi method)**

**Input:** a real symmetric matrix \( A \).

1. \( A_0 := A \), \( S_0^2 := \sum_{i\neq j} a_{ij}^2 \).
2. For \( A_k \), \( k \geq 0 \), do the following:
   1. Find \( a_{pq}^{(k)} \), \( p < q \), such that \( |a_{pq}^{(k)}| = \max_{i,j=1,\ldots,n} |a_{ij}^{(k)}| \).
   2. For such \( p, q \), find \( \theta \) so that
      \[
      \tan 2\theta = \frac{2a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}}, \quad \theta \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right],
      \]
      
      and define \( Q_{k+1} \) by (1).
   3. \( A_{k+1} := Q_{k+1}^T A_k Q_{k+1} = Q_{k+1}^T A Q_1 \cdots Q_{k+1} \).
   4. \( S_{k+1}^2 := \sum_{i\neq j} (a_{ij}^{(k+1)})^2 \).
3. If \( S_{k+1}^2 \) is sufficiently small, STOP.

**Output:** \( a_{11}^{(k+1)}, \ldots, a_{nn}^{(k+1)} \), \( R_{k+1} := Q_1 \cdots Q_{k+1} \).

The output of the classical Jacobi method \( a_{11}^{(k+1)}, \ldots, a_{nn}^{(k+1)} \) are approximations of eigenvalues of \( A \) and the column vectors of \( R_{k+1} \) are eigenvectors of corresponding eigenvalues.

Recently, the Jacobi methods attract attention because they are inherently parallel. There are block versions of the Jacobi methods. Convergence analysis for the block versions are currently
well investigated. For example, see [9] for the recent research about global convergence and quadratic convergence of the classical block Jacobi method. Regarding the cyclic block Jacobi method, Drmač [3] clarified a sufficient condition for the block rotations to guarantee global convergence for a special type of pivoting strategies.

In this paper, we go back to the classical Jacobi method. Our analysis especially focuses on the convergence proof for the diagonal elements. We believe such analysis is also useful for the recent block versions. The convergence of the classical Jacobi method has been understood very well (see [4–8] and the references therein) and its proof can be found in many textbooks. However, regarding the convergence of the diagonal elements, it seems that we are still able to consider the proof from a different point of view. In this paper, we simplify the proof for the diagonal elements given in Ciarlet [2] (see the proofs of Theorem 6.1-2 and 6.1-3). Also note that Parlett [7, Section 9.3] uses the Wielandt-Hoffman theorem for the diagonal elements. Wilkinson [8, Section 5] also needs a similar perturbation theorem for the convergence of the diagonal elements. However, we do not use any perturbation theory in the following proof.

2. Convergence and a posteriori error estimates

In this section, we show the main results. Before that, we briefly summarize the convergence proof for diagonal elements in Ciarlet [2]. First, Ciarlet [2] shows that the sequence of each diagonal element \( \{a_{ii}^{(k)}\}_{k=0}^{\infty} \) has a finite number of cluster points \( \lambda_1, \ldots, \lambda_n \). Then, it is proved that \( a_{ii}^{(k)} \) are convergent to one of its cluster points in view of \( \lim_{k \to \infty} (a_{ii}^{(k+1)} - a_{ii}^{(k)}) = 0 \).

In our proof, the convergence of \( a_{ii}^{(k)} \) for \( i = 1, \ldots, n \) are derived rather directly as follows. Ciarlet’s mathematical discussion about the multiple cluster points is not needed.

**Theorem 2** As \( k \to \infty \), \( A_k \) converges to a diagonal matrix \( D \) and the diagonal elements \( \{\lambda_1, \cdots, \lambda_n\} \) are eigenvalues of \( A \).

**Proof.** Recall the following notations: \( A_k = (a_{ij}^{(k)}) \), \( \max_{i \neq j} |a_{ij}^{(k)}| = |a_{pq}^{(k)}| \) and

\[
\alpha := \left(1 - \frac{2}{n^2 - n}\right)^{1/2}, \quad S = S_0 := \left( \sum_{i \neq j} (a_{ij}^{(k)})^2 \right)^{1/2}, \quad S_k := \left( \sum_{i \neq j} (a_{ij}^{(k)})^2 \right)^{1/2}.
\]

From (2) and \( \sum_{i \neq j} (a_{ij}^{(k)})^2 \leq (n^2 - n) (a_{pq}^{(k)})^2 \), we have

\[
S_{k+1}^2 = \sum_{i \neq j} (a_{ij}^{(k+1)})^2 \leq \sum_{i \neq j} (a_{ij}^{(k)})^2 - 2 \left( a_{pq}^{(k)} \right)^2 \leq \alpha^2 S_k^2 \leq \alpha^{2k+1} S^2,
\]

\[S_{k+1} \leq \alpha^{k+1} S, \quad \lim_{k \to \infty} S_k = 0.\]

Therefore, we see that \( A_k \) converges to a diagonal matrix as \( k \to \infty \). However, we still have to show that each diagonal element is convergent.

As is shown in Ciarlet [2, p194], we have

\[
a_{ii}^{(k+1)} - a_{ii}^{(k)} = \begin{cases} 0 & i \neq p, q, \\ a_{pq}^{(k)} \tan \theta & i = p, \\ -a_{pq}^{(k)} \tan \theta & i = q. \end{cases}
\]

Since \( -\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4} \), we have

\[|a_{ii}^{(k+1)} - a_{ii}^{(k)}| \leq |a_{pq}^{(k)}| \leq S_k \leq \alpha^k S.\]

Hence, for any positive integer \( m \), we obtain the estimate

\[|a_{ii}^{(k+m)} - a_{ii}^{(k)}| \leq S_{k+m-1} + \cdots + S_k \leq (\alpha^{m-1} + \cdots + 1)S_k < \frac{S_k}{1-\alpha} \leq \frac{S_k \alpha^k}{1-\alpha},\]
Proof. Suppose that $a_{pp}$, $a_{qq}$ converge to the eigenvalues $\lambda_p$, $\lambda_q$ as $k \to \infty$, we have, for sufficiently large $N_1 > 0$,  

$$|a_{pp} - \lambda_p| < \frac{\mu_p}{3}, \quad |a_{qq} - \lambda_q| < \frac{\mu_p}{3}, \quad |a_{pq} - a_{qp}| < \frac{\mu_p}{3} \quad \forall k \geq N_1.$$ 

Using these estimates and $|a_{pq}| \leq S_k \leq a^k \cdot S$, we obtain, for any $k \geq N_1$,  

$$2|\theta| \leq |\tan 2\theta| \leq \frac{6}{\mu_p} S_k \leq \frac{6S}{\mu_p} \alpha^k \quad \text{or} \quad |\theta| \leq \frac{3}{\mu_p} S_k \leq C_1 \alpha^k, \quad C_1 := \frac{3S}{\mu_p},$$

$$1 - \cos \theta \leq \frac{\theta^2}{2} \leq \frac{9}{2\mu_p^2} S_k^2 \leq \frac{C_1^2}{2} \alpha^{2k}. \quad (3)$$

We now denote $R_k = [r_1^{(k)}, r_2^{(k)}, \ldots, r_n^{(k)}]$. By the definitions, we see  

$$r_p^{(k+1)} = r_p^{(k)} \cos \theta + r_q^{(k)} \sin \theta, \quad r_q^{(k+1)} = -r_p^{(k)} \sin \theta + r_q^{(k)} \cos \theta.$$ 

By the Cauchy-Schwarz inequality and (3), we have, for any $k \geq N_1$,  

$$\left| r_p^{(k+1)} - r_p^{(k)} \right| \leq \left( |r_p^{(k)}|^2 + |r_q^{(k)}|^2 \right)^{1/2} \left( 2(1 - \cos \theta) \right)^{1/2} \leq \frac{3\sqrt{2}}{\mu_p} S_k \leq \sqrt{2} C_1 \alpha^k.$$ 

Here, we use the fact $|r_p^{(k)}| = |r_q^{(k)}| = 1$, because $R_k$ is orthogonal. If the pivot is $a_{ij}^{(k)}$, $i, j \neq p$, $r_p^{(k+1)} = r_p^{(k)}$. Combining these estimates, we obtain, for any positive integer $m$ and any $k \geq N_1$,  

$$\left| r_p^{(k+m)} - r_p^{(k)} \right| \leq \frac{\sqrt{2} C_1 \alpha^k}{1 - \alpha}.$$ 

This means that $\{r_p^{(k)}\}$ is a Cauchy sequence and $\{r_p^{(k)}\}$ converges to a unit vector $r_p$. Since the set of orthogonal matrices of order $n$ is compact, one can choose a convergent subsequence $\{R_{k_i}\} \subset \{R_k\}$. Let $\bar{R}$ be the limit of $\{R_{k_i}\}$. Of course, the $p$th column vector of $\bar{R}$ is $r_p$. Since we have  

$$R^T A R = \lim_{k_i \to \infty} R_{k_i}^T A R_{k_i} = D,$$ 

de each column vector $R$ is an eigenvector of $A$. Thus, we conclude that $r_p$ is the eigenvector of $\lambda_p$. □

Next, we focus on the eigenvectors. The following theorem is proved in the same manner as the convergence proof of the diagonal elements.

**Theorem 3** Let an index $p \in \{1, \ldots, n\}$ be taken and fixed. Suppose that $\lambda_p = \lim_{k \to \infty} a_{pp}^{(k)}$ and the eigenvalue $\lambda_p$ is simple. Then, as $k \to \infty$, the $p$-th column vector of $R_k := Q_1 Q_2 \cdots Q_k$ converges to the eigenvector of $\lambda_p$.

If all eigenvalues of $A$ are simple, we conclude from Theorem 2.3 that $R_k$ itself converges to $R$. Actually in fact, [1] proves $\lim_{k \to \infty} R_k = R$ in a similar manner to the above proof, if the eigenvalues of $A$ are all distinct.
From now on, we focus on the error estimate of the eigenpairs. First of all, it is easy to see that
\[
\|a_{pp}^{(k+m)} - a_{pp}^{(k)}\| \leq \frac{S_k}{1 - \alpha}, \quad |\lambda_p - a_{pp}^{(k)}| \leq \frac{S_k}{1 - \alpha}, \quad \forall p \in \{1, 2, \cdots, n\}
\] (4)
from the proof of Theorem 2.1. This is an \textit{a posteriori} error estimate, which is useful to obtain error bounds of computed eigenvalues. A similar \textit{a posteriori} error estimate for an eigenvector is obtained in the following corollary.

**Corollary 4** Suppose that the eigenvalue \(\lambda_p\) is simple and the followings hold for sufficiently large \(k\):
\[
\mu_p^{(k)} := \min_{i \neq p} |a_{ii}^{(k)} - a_{ii}^{(k)}| > 0, \quad \mu_p^{(k)} - \frac{2S_k}{1 - \alpha} > 0.
\] (5)
Then, we have the following \textit{a posteriori} error estimate for the eigenvector \(r_p^{(k)}\):
\[
|r_p^{(k)} - r_p^{(k)}| \leq \frac{\sqrt{2S_k}}{(1 - \alpha)\mu_p^{(k)} - 2S_k}.
\] (6)

**Proof.** From (4) and (5), we conclude that
\[
\mu_p = \min_{i \neq p} |\lambda_p - \lambda_i| \geq \min_{i \neq p} \left( |a_{ii}^{(k)} - a_{ii}^{(k)}| - |\lambda_p - a_{pp}^{(k)}| - |\lambda_i - a_{il}^{(k)}| \right) \geq \mu_p^{(k)} - \frac{2S_k}{1 - \alpha} > 0.
\]
We also notice that, for any positive integer \(m\),
\[
|a_{pp}^{(k+m)} - a_{pp}^{(k+m)}| \geq |a_{pp}^{(k)} - a_{pp}^{(k)}| - |a_{pp}^{(k+m)} - a_{pp}^{(k+m)}| - |a_{il}^{(k)} - a_{il}^{(k+m)}| \geq \mu_p^{(k)} - \frac{2S_k}{1 - \alpha} > 0, \quad p \neq l.
\]
Noting
\[
|2\theta| \leq |\tan 2\theta| \leq \frac{2S_{k+m-1}}{\mu_p^{(k+m-1)} - 2S_{k+m-1}/(1 - \alpha)},
\]
where \(\theta\) is an angle used in the \((k + m - 1)\)th step, we have
\[
|r_p^{(k+m)} - r_p^{(k+m-1)}| \leq \frac{\sqrt{2S_{k+m-1}}}{\mu_p^{(k)} - 2S_{k+m-1}/(1 - \alpha)} \leq \frac{\sqrt{2}\alpha^{m-1}S_k}{\mu_p^{(k)} - 2S_k/(1 - \alpha)},
\]
\[
|r_p^{(k+m)} - r_p^{(k)}| \leq \frac{\sqrt{2S_k}}{(1 - \alpha)\mu_p^{(k)} - 2S_k}.
\]
Thus, letting \(m \to \infty\) in the second inequality, we obtain (6). \( \square \)

Although we obtain the \textit{a posteriori} error estimate based on the convergence analysis of the Jacobi method, the above error bounds are not as sharp as the bounds obtained by famous perturbation theory. It is proved in [4] that if \(\lim_{k \to \infty} a_{pp}^{(k)} = \lambda_p\), we have
\[
|\lambda_p - a_{pp}^{(k)}| \leq S_k, \quad \forall p, \quad (7)
\]
which is a corollary of the Wielandt-Hoffman theorem [7, Section 9.3]. The error bound \(S_k\) is smaller than the right-hand side of (4). Regarding the eigenvector, let \(\phi^{(k)} := \angle(r_p, r_p^{(k)})\). Then we have
\[
|\sin(\phi^{(k)})| \leq \frac{|A r_p^{(k)} - a_{pp}^{(k)} r_p^{(k)}|}{\min_{i \neq p} (\lambda_i - a_{pp}^{(k)})} = \frac{|R_k^T A r_p^{(k)} - R_k^T a_{pp}^{(k)} r_p^{(k)}|}{\min_{i \neq p} (\lambda_i - a_{pp}^{(k)})} \leq \frac{\sum_{l \neq p} |a_{lp}^{(k)}|^2}{\min_{i \neq p} (\lambda_i - a_{pp}^{(k)})} \frac{\mu_p^{(k)} - S_k}{\mu_p^{(k)} - S_k} = \mu_p^{(k)} - S_k.
\]
from the so-called sin $\theta$ theorem [7, Theorem 11.7.1]. It is easy to see that $|r_p - r_p^{(k)}| \leq \sqrt{2} |\sin(\phi^{(k)})|$. Hence, we obtain

$$|r_p - r_p^{(k)}| \leq \sqrt{2} \sum_{l \neq p} a_{lp}^{(k)} \mu_p^{(k)} - S_k,$$

which is sharper than (6).

Although we can not obtain new error bounds from our approach based on the convergence analysis of the Jacobi method, we would like to emphasize here that our analysis shows several important properties of the Jacobi method without any perturbation theory. In particular, we simplify the Ciarlet’s convergence proof of the diagonal elements.

References