INTERLACING INEQUALITIES FOR EIGENVALUES OF SUMS OF MATRICES

By

ERNEST AFARI

B. Sc., University of Cape Coast, 1982

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in
THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF MATHEMATICS

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
March 1989
© ERNEST AFARI, 1989
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of **MATHEMATICS**

The University of British Columbia
Vancouver, Canada

Date **MARCH 29, 1989**
abstract

We begin with some historical remarks in section 1, where we present the basic interlacing inequalities.

In the rest of the thesis we discuss R.C. Thompson's contributions to the subject. His main concern is to extend the inequalities to include more than just one of the \((n - 1) \times (n - 1)\) principal submatrices. It is clear that for each such submatrix, the interlacing inequalities are valid, but the resulting set of inequalities is not sufficient to guarantee the existence of a matrix whose principal submatrices have eigenvalues satisfying these inequalities.

We have included two examples to show that Thompson's inequalities are necessary.
Acknowledgement

I wish to express my thanks to Dr. Roy Westwick for his guidance, sound criticism and most of all, for his endless patience with me throughout the preparation of this work. I also thank Dr. G. Maxwell for reading the manuscript and providing valuable criticisms and comments. Thanks to my other teachers at U.B.C. from whom I benefitted much; especially Dr. E. Luft and Dr. K. Lam for their encouragement on many timely occasions and in words and deeds reminding me of the necessity of hard work. I express my gratitude to the Mathematics Department of U.B.C. for its generous financial support. I also thank my friend Mr. E. Akita for lending me his personal computer and giving me computer terminal tutorials for the typing of this work. I am also grateful to Ms. Sherry Armstrong for typing several drafts of this work. My appreciation to my beloved wife Comfort for accepting my emotional ups-and-downs, during my study, merrily providing a lovely home and constantly upholding me with prayers. This work is respectfully dedicated to Comfort Afari.
INTRODUCTION

The origins of the interlacing inequalities for the eigenvalues of a $n \times n$ matrix and its principal $(n - 1) \times (n - 1)$ submatrix is generally attributed to investigations by the renowned mathematician Augustin-Louis Cauchy (1789 - 1857) [3]. His results appeared as early as 1829. Cauchy found that if $H$ is an $n$ - square real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ ordered so that

$$
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \quad (1.1)
$$

and if $H(i\backslash i)$ denotes the principal $(n - 1)$ - square submatrix of $H$ obtained by deleting row $i$ and column $i$ with eigenvalues $\mu_1, \ldots, \mu_{n-1}$ ordered so that

$$
\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \quad (1.2)
$$

then

$$
\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n \quad (1.3)
$$

These are known as the interlacing inequalities.

In 1957, Ky Fan and Gordon Pall [5] published the converse of the above results and also showed that it holds for Hermitian matrices. The results by Ky Fan and Pall are as follows: Given real numbers $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_{n-1}$ satisfying (1.1), (1.2) and (1.3), there exists a $n$ -square Hermitian matrix $H$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and
Section 1. INTRODUCTION

(n - 1)-square principal submatrix with eigenvalues $\mu_1, \ldots, \mu_{n-1}$

Subsequently, there have been several papers extending these ideas in various ways. The principal contributor is R.C. Thompson who has a series of papers which we will be describing.
Section 2

TRIVIAL AND NONTRIVIAL EIGENVALUES

Given $A$, a $n \times n$ diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and principal submatrices $A(1\setminus 1), \ldots, A(k\setminus k)$ ($k$ fixed, $1 \leq k \leq n$) we know that interlacing inequalities hold separately. These are

$$
\lambda_1 \leq \mu_{i1} \leq \lambda_2 \leq \mu_{i2} \leq \cdots \leq \lambda_{n-1} \leq \mu_{i,n-1} \leq \lambda_n
$$

(2.4)

Assume that real numbers $\lambda_i$ and $\mu_{ij}$ are given satisfying (2.4). It is not in general true that an $A$ can be found such that its first $k$ principal submatrices have these $\mu_{ij}$ as their eigenvalues.

In fact suppose for $\lambda_1, \ldots, \lambda_n$, we assume that every set of $k$ interlacing systems of inequalities can be realised by a matrix $A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $k$ principal submatrices $A(i\setminus i)$. This implies the strong condition that each of the eigenvalues has multiplicity at least $k$. It is stated precisely in the following theorem.

Theorem 2.1

Let $k$ be fixed, $1 \leq k \leq n$. As $U$ varies over all unitary matrices the eigenvalues

$$
\mu_{i1} \leq \cdots \leq \mu_{i,n-1}
$$

of the principal submatrices $H(i\setminus i)$ of $H = U \ D \ U^{-1}$ where $U$ is unitary and $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ $1 \leq i \leq k,$ independently assume all values permitted by the inequalities (2.4) if and only if each distinct eigenvalue of $H$ has multiplicity at least $k$. 
In order to address the problem of what extra conditions are needed on the eigenvalues of $A$ so that the inequalities are sufficient to guarantee the existence of $A$, Thompson [10] defined the trivial and nontrivial eigenvalues for matrices.

Let

$$
\lambda_1 < \lambda_2 < \cdots < \lambda_s
$$

be the distinct eigenvalues of $A$, where $\lambda_\alpha$ has multiplicity $e_\alpha$, $1 \leq \alpha \leq s$; $e_1 + \cdots + e_s = n$.

$A(i\backslash i)$ has $\lambda_\alpha$ as eigenvalues with multiplicity $e_\alpha - 1$ or larger. Thus $\lambda_\alpha$ with multiplicity $e_\alpha - 1$ are always eigenvalues of $A(i\backslash i)$, $1 \leq \alpha \leq s$. These are called the trivial eigenvalues of $A(i\backslash i)$. There exist $s - 1$ additional eigenvalues of $A(i\backslash i)$ denoted by

$$
\mu_{i1} \leq \mu_{i2} \leq \cdots \leq \mu_{i,s-1}
$$

and these are called the nontrivial eigenvalues of $A(i\backslash i)$.

They satisfy

$$
\lambda_1 \leq \mu_{i1} \leq \lambda_2 \leq \mu_{i2} \leq \cdots \leq \lambda_{s-1} \leq \mu_{i,s-1} \leq \lambda_s \quad (2.5)
$$

It may happen that the nontrivial eigenvalues of $A(i\backslash i)$ are not all distinct and that some of the nontrivial eigenvalues of $A(i\backslash i)$ equal some of the trivial eigenvalues.
Section 3

QUADRATIC AND LINEAR INEQUALITIES

Thompson [11] established a new set of inequalities involving the eigenvalues of a matrix A and its principal \((n - 1)\) - square submatrices. These inequalities are quadratic in form except in certain special cases where they degenerate into inequalities involving arithmetic means of the relevant eigenvalues.

Let \(H, H(i|j), \lambda_i, \mu_{ij}\) be defined as in section 2. We have of course the Cauchy inequalities (2.4).

Thompson inequalities are:

\[
\sum_{i=1}^{n} \frac{\lambda_j - \mu_{i,j-1}}{\lambda_j - \lambda_{j-1}} \frac{\mu_{ij} - \lambda_j}{\lambda_{j+1} - \lambda_j} \geq 1, \quad 1 \leq j \leq n. \tag{3.6}
\]

\[
\sum_{i=1}^{n} \frac{\lambda_j - \mu_{i,j-1}}{\lambda_j - \lambda_1} \frac{\mu_{ij} - \lambda_j}{\lambda_n - \lambda_j} \leq 1, \quad 1 \leq j \leq n. \tag{3.7}
\]

In (3.6) the factor

\[
\frac{\lambda_j - \mu_{i,j-1}}{\lambda_j - \lambda_{j-1}}
\]

and in (3.7) the factor

\[
\frac{\lambda_j - \mu_{i,j-1}}{\lambda_j - \lambda_1}
\]

are absent if \(j=1\).
Again, in (3.6) the factor
\[
\frac{\mu_{ij} - \lambda_j}{\lambda_{j+1} - \lambda_j}
\]
and in (3.7) the factor
\[
\frac{\mu_{ij} - \lambda_j}{\lambda_n - \lambda_j}
\]
are absent if \( j = n \).

Thompson referred to the inequality (3.6) as the upper quadratic inequality based on \( \lambda_j \) and he referred to the inequality (3.7) as the lower quadratic inequality based on \( \lambda_j \).

He said "quadratic" because of the fact that the left sides of (3.6) and (3.7) are polynomials of degree two in \( \mu_{ij} \).

**Notation:**

Let

\[
j A_{j+1} = n^{-1} \sum_{i=1}^{n} \mu_{ij}, \quad 1 \leq j < s.
\]

(3.8)

i.e. \( j A_{j+1} \) is the arithmetic mean \([3]\) of the nontrivial eigenvalues of the \( H(i \setminus i) \) belonging to the interval \( [\lambda_j, \lambda_{j+1}] \).

\[
(n - 1)n^{-1}\lambda_j + n^{-1}\lambda_{j+1} \leq j A_{j+1} \leq n^{-1}\lambda_j + (n - 1)n^{-1}\lambda_{j+1}, \quad 1 \leq j \leq n
\]

(3.9)

are referred to as Linear inequalities. \([5]\)

In order to get examples where the quadratic inequalities are seen to be necessary, we display the case \( n = 3 \) explicitly.

Let \( H \) be a 3-square Hermitian matrix with distinct eigenvalues

\[
\lambda_1 < \lambda_2 < \lambda_3
\]

(3.10)
Let $H(3 \setminus 3)$ denote the principal 2-square submatrix of $H$ obtained by deleting row 3 and column 3.

Let

$$
\begin{align*}
\mu_{11} &\leq \mu_{12} \\
\mu_{21} &\leq \mu_{22} \\
\mu_{31} &\leq \mu_{32}
\end{align*}
$$

be the eigenvalues of $H(3 \setminus 3)$.

The Cauchy interlacing inequalities assert that

$$
\begin{align*}
\lambda_1 &\leq \mu_{11} \leq \lambda_2 \leq \mu_{12} \leq \lambda_3 \\
\lambda_1 &\leq \mu_{21} \leq \lambda_2 \leq \mu_{22} \leq \lambda_3 \\
\lambda_1 &\leq \mu_{31} \leq \lambda_2 \leq \mu_{32} \leq \lambda_3.
\end{align*}
$$

For the upper quadratic inequalities we get

$$
\frac{\mu_{11} - \lambda_1}{\lambda_2 - \lambda_1} + \frac{\mu_{21} - \lambda_1}{\lambda_2 - \lambda_1} + \frac{\mu_{31} - \lambda_1}{\lambda_2 - \lambda_1} \geq 1
$$

$$
\Rightarrow \mu_{11} + \mu_{21} + \mu_{31} \geq 2\lambda_1 + \lambda_2
$$

$$
\frac{\lambda_2 - \mu_{11} \lambda_3 - \lambda_2}{\lambda_2 - \lambda_1} + \frac{\lambda_2 - \mu_{21} \lambda_3 - \lambda_2}{\lambda_2 - \lambda_1} + \frac{\lambda_2 - \mu_{31} \lambda_3 - \lambda_2}{\lambda_2 - \lambda_1} \geq 1
$$

$$
\frac{\lambda_3 - \mu_{12}}{\lambda_3 - \lambda_2} + \frac{\lambda_3 - \mu_{22}}{\lambda_3 - \lambda_2} + \frac{\lambda_3 - \mu_{32}}{\lambda_3 - \lambda_2} \geq 1
$$
For the lower quadratic inequality, we get

\[
\frac{\mu_{11} - \lambda_1}{\lambda_3 - \lambda_1} + \frac{\mu_{21} - \lambda_1}{\lambda_3 - \lambda_1} + \frac{\mu_{31} - \lambda_1}{\lambda_3 - \lambda_1} \leq 1
\]  

(3.16)

\Rightarrow \mu_{11} + \mu_{21} + \mu_{31} \leq 2\lambda_1 + \lambda_3

\[
\frac{\lambda_2 - \mu_{11} \mu_{12} - \lambda_2}{\lambda_2 - \lambda_1} + \frac{\lambda_2 - \mu_{21} \mu_{22} - \lambda_2}{\lambda_2 - \lambda_1} + \frac{\lambda_2 - \mu_{31} \mu_{32} - \lambda_2}{\lambda_2 - \lambda_1} \leq 1
\]  

(3.17)

\[
\frac{\lambda_3 - \mu_{12}}{\lambda_3 - \lambda_1} + \frac{\lambda_3 - \mu_{22}}{\lambda_3 - \lambda_1} + \frac{\lambda_3 - \mu_{32}}{\lambda_3 - \lambda_1} \leq 1
\]  

(3.18)

\Rightarrow \mu_{12} + \mu_{22} + \mu_{33} \geq \lambda_1 + 2\lambda_3

We note that (3.14) and (3.17) have the same left hand side so they imply equality (both equal to 1). While (3.13), (3.15), (3.16), (3.18) give rise to the linear inequalities.

We get

\[
2\lambda_1 + \lambda_2 \leq \mu_{11} + \mu_{21} + \mu_{31} \leq 2\lambda_1 + \lambda_3
\]  

(3.19)

\[
\lambda_1 + 2\lambda_3 \leq \mu_{12} + \mu_{22} + \mu_{32} \leq \lambda_2 + 2\lambda_3
\]  

(3.20)

Now, if we consider the example for the case \(\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3\) the linear inequalities become

\[
4 \leq \mu_{11} + \mu_{21} + \mu_{31} \leq 5
\]  

(3.21)
7 ≤ μ_{12} + μ_{22} + μ_{32} ≤ 8 \quad (3.22)

and (3.14) and (3.17) become

\[(2 - μ_{11})(μ_{12} - 2) + (2 - μ_{21})(μ_{22} - 2) + (2 - μ_{31})(μ_{32} - 2) = 1 \quad (3.23)\]

If

μ_{11} = μ_{21} = μ_{31} = λ_1

then clearly (3.12) is valid but (3.13) fails.

If

μ_{11} = μ_{12} = μ_{21} = μ_{22} = μ_{31} = μ_{32} = λ_2

then the linear inequalities (3.13), (3.15), (3.16), (3.18) are all valid but (3.14) is not.

Now for fixed k, with k < n, let Q_{nk} denote the set of all sequences

ω = \{i_1, i_2, \ldots, i_k\}

of integers such that

1 ≤ i_1 < i_2 < \cdots < i_k ≤ n.

The number of sequences in Q_{nk} is

\[\binom{n}{k}\]
Section 3. QUADRATIC AND LINEAR INEQUALITIES

Let $A[\omega \setminus \omega]$ denote the principal submatrix of $A$ lying at the intersection of rows and columns $i_1, i_2, \ldots, i_k$. The submatrix $A[\omega \setminus \omega]$ is Hermitian and let

$$
\mu_{\omega 1} \leq \mu_{\omega 2} \leq \cdots \leq \mu_{\omega k}
$$

be the eigenvalues of $A[\omega \setminus \omega]$.

The Cauchy interlacing inequalities asserts that

$$\lambda_j \leq \mu_{\omega j} \leq \lambda_{j+n-k}, \quad 1 \leq j \leq k \quad (3.24)$$

Thompson [15] established the following bounds on the arithmetic mean of the $\mu_{\omega j}$ as $\omega$ varies over $Q_{n,k}$ and $j$ is fixed:

$$\sum_{r=0}^{n-k} \Psi_r \lambda_{n-k+j-r} \leq \binom{n}{k}^{-1} \sum_{\omega \in Q_{n,k}} \mu_{\omega j} \leq \sum_{r=0}^{n-k} \Psi_r \lambda_{j+r} \quad (3.25)$$

where

$$\Psi_r = E_r(n-1, n-2, \ldots, k)\{\prod_{i=k+1}^{n} i\}^{-1}, \quad 0 \leq r \leq n-k \quad (3.26)$$

$$\sum_{r=0}^{n-k} \Psi_r = 1 \quad (3.27)$$

$E_r(n-1, n-2, \ldots, k)$ denotes the elementary symmetric function of degree $r$ on the integers $n-1, n-2, \ldots, k$. The significance of the inequalities (3.25) is that they provide convex combinations of

$$\lambda_j, \lambda_{j+1}, \ldots, \lambda_{j+n-k}$$

which are bounds for the arithmetic mean, over all $k \times k$ principal submatrices of $A$, of the $j^{th}$ eigenvalue of each of these submatrices. That (3.27) holds follows immediately by setting $\lambda = 1$ in the polynomial identity
Section 3. **QUADRATIC AND LINEAR INEQUALITIES**

\[
\Pi_{j=k}^{n-1}(\lambda + j) = \sum_{r=0}^{n-k} E_r(n-1, n-2, \ldots, k) \lambda^{n-k+r}.
\]  

(3.28)

In order to see what these inequalities look like, we produce here explicitly, the case

\[n = 4, \quad k = 1, 2, 3.\]

Now for \(n = 4, k = 1,\) we get

\[\omega = \{i_1\}, \quad 1 \leq i_1 \leq 4\]

\[i.e. \quad \omega = \{1\}, \quad \{2\}, \quad \{3\}, \quad \{4\}.\]

Let \(A\{1\}\backslash\{1\}, \quad A\{2\}\backslash\{2\}, \quad A\{3\}\backslash\{3\}, \quad A\{4\}\backslash\{4\}\) denote the principal submatrices of \(A.\)

Let

\[
\mu_{\{1\}1}, \quad \mu_{\{2\}1}, \quad \mu_{\{3\}1}, \quad \mu_{\{4\}1}
\]

be the eigenvalues of \(A[\omega \backslash \omega].\)

\[
\left(\begin{array}{c}
4
\end{array}\right)^{-1} \sum_{\omega \in Q_{nk}} \mu_{\omega 1} = \frac{1}{4} \{\mu_{\{1\}1} + \mu_{\{2\}1} + \mu_{\{3\}1} + \mu_{\{4\}1}\}
\]

(3.29)

\[
\sum_{r=0}^{3} \Psi_r \lambda_{1+r} = \Psi_0 \lambda_1 + \Psi_1 \lambda_2 + \Psi_2 \lambda_3 + \Psi_3 \lambda_4
\]

(3.30)

\[
\sum_{r=0}^{3} \Psi_r \lambda_{4-r} = \Psi_0 \lambda_1 + \Psi_1 \lambda_2 + \Psi_2 \lambda_3 + \Psi_3 \lambda_4
\]

(3.31)
\[ \Psi_r = \frac{1}{24} E_r(3, 2, 1) \]
\[ \Psi_0 = \frac{1}{24} E_0(3, 2, 1) \]
\[ \Psi_1 = \frac{1}{24} E_1(3, 2, 1), \quad 0 \leq r \leq 3 \] (3.32)
\[ \Psi_2 = \frac{1}{24} E_2(3, 2, 1) \]
\[ \Psi_3 = \frac{1}{24} E_3(3, 2, 1) \]

Now

\[ \prod_{j=1}^{3} (\lambda + j) = \sum_{r=0}^{3} E_r(3, 2, 1) \lambda^{3-r} \] (3.33)

\[ \Rightarrow \lambda^3 + 6\lambda^2 + 11\lambda + 6 = E_0(3, 2, 1)\lambda^3 + E_1(3, 2, 1)\lambda^2 + E_2(3, 2, 1)\lambda + E_3(3, 2, 1)\lambda^0 \]

\[ \Rightarrow \]
\[ E_0(3, 2, 1) = 1 \]
\[ E_1(3, 2, 1) = 6 \] (3.34)
\[ E_2(3, 2, 1) = 11 \]
\[ E_3(3, 2, 1) = 6 \]

From (3.32) and (3.34) we get
\[ \Psi_0 = \frac{1}{24}, \quad \Psi_1 = \frac{1}{4}, \quad \Psi_2 = \frac{11}{24}, \quad \Psi_3 = \frac{1}{4}. \]

From (3.30), (3.31), we get

\[ \sum_{r=0}^{3} \Psi_r \lambda_{1+r} = \frac{1}{24} \lambda_1 + \frac{1}{4} \lambda_2 + \frac{11}{24} \lambda_3 + \frac{1}{4} \lambda_4 \] (3.35)
Section 3. **QUADRATIC AND LINEAR INEQUALITIES**

\[ \sum_{r=0}^{3} \Psi_r \lambda_{4-r} = \frac{1}{4} \lambda_1 + \frac{11}{24} \lambda_2 + \frac{1}{4} \lambda_3 + \frac{1}{24} \lambda_4 \]  

(3.36)

Finally, we get

\[ \frac{1}{4} \lambda_1 + \frac{11}{24} \lambda_2 + \frac{1}{4} \lambda_3 + \frac{1}{24} \lambda_4 \leq \frac{1}{4} \{ \mu_{(1)1} + \mu_{(2)1} + \mu_{(3)1} + \mu_{(4)1} \} \leq \frac{1}{24} \lambda_1 + \frac{1}{4} \lambda_2 + \frac{11}{24} \lambda_3 + \frac{1}{4} \lambda_4 \]

(3.37)

For \( n = 4, k = 2 \)

\[ \omega = \{i_1, i_2\}, 1 \leq i_1 < i_2 \leq 4 \]

i.e. \( \omega = \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\} \).

Let \( A[\{1, 2\} \setminus \{1, 2\}], A[\{1, 3\} \setminus \{1, 3\}], A[\{1, 4\} \setminus \{1, 4\}], A[\{2, 3\} \setminus \{2, 3\}], A[\{2, 4\} \setminus \{2, 4\}], A[\{3, 4\} \setminus \{3, 4\}] \) denote the principal submatrices of \( A \).

Let

\[ \mu_{(1,2)1} \leq \mu_{(1,2)2}, \]
\[ \mu_{(1,3)1} \leq \mu_{(1,3)2}, \]
\[ \mu_{(1,4)1} \leq \mu_{(1,4)2}, \]
\[ \mu_{(2,3)1} \leq \mu_{(2,3)2}, \]
\[ \mu_{(2,4)1} \leq \mu_{(2,4)2}, \]
\[ \mu_{(3,4)1} \leq \mu_{(3,4)2}, \]

be its eigenvalues.
Section 3. QUADRATIC AND LINEAR INEQUALITIES

Now for $j = 1$

$$
\left( \begin{array}{c}
4 \\
2
\end{array} \right)^{-1} \sum_{\omega \in Q_{42}} \mu_{\omega} = \frac{1}{6} \{ \mu_{\{1,2\}} + \mu_{\{1,3\}} + \mu_{\{1,4\}} + \mu_{\{2,3\}} + \mu_{\{2,4\}} + \mu_{\{3,4\}} \} \tag{3.38}
$$

$$
\sum_{\tau=0}^{2} \Psi_\tau \lambda_{1+\tau} = \Psi_0 \lambda_1 + \Psi_1 \lambda_2 + \Psi_2 \lambda_3 \tag{3.39}
$$

$$
\sum_{\tau=0}^{2} \Psi_\tau \lambda_{3-\tau} = \Psi_2 \lambda_1 + \Psi_1 \lambda_2 + \Psi_0 \lambda_3 \tag{3.40}
$$

For $j = 2$, we get

$$
\left( \begin{array}{c}
4 \\
2
\end{array} \right)^{-1} \sum_{\omega \in Q_{42}} \mu_{\omega} = \frac{1}{6} \{ \mu_{\{1,2\}} + \mu_{\{1,3\}} + \mu_{\{1,4\}} + \mu_{\{2,3\}} + \mu_{\{2,4\}} + \mu_{\{3,4\}} \} \tag{3.41}
$$

$$
\sum_{\tau=0}^{2} \Psi_\tau \lambda_{2+\tau} = \Psi_0 \lambda_2 + \Psi_1 \lambda_3 + \Psi_2 \lambda_4 \tag{3.42}
$$

$$
\sum_{\tau=0}^{2} \Psi_\tau \lambda_{4-\tau} = \Psi_2 \lambda_2 + \Psi_1 \lambda_3 + \Psi_0 \lambda_4 \tag{3.43}
$$

From (3.25), (3.26), (3.27) we get

$$
\Psi_0 = \frac{1}{12},
$$

$$
\Psi_1 = \frac{5}{12},
$$

$$
\Psi_2 = \frac{1}{2}.
$$
Finally, we get for $j = 1$,

$$\frac{1}{2} \lambda_1 + \frac{5}{12} \lambda_2 + \frac{1}{12} \lambda_3 \leq \frac{1}{6} \{\mu_{(1,2)} + \mu_{(1,3)} + \mu_{(1,4)} + \mu_{(2,3)} + \mu_{(2,4)} + \mu_{(3,4)}\} \leq \frac{1}{12} \lambda_1 + \frac{5}{12} \lambda_2 + \frac{1}{2} \lambda_3$$

(3.44)

and for $j = 2$, we get

$$\frac{1}{2} \lambda_2 + \frac{5}{12} \lambda_3 + \frac{1}{12} \lambda_4 \leq \frac{1}{6} \{\mu_{(1,2)} + \mu_{(2,3)} + \mu_{(2,4)} + \mu_{(3,4)} + \mu_{(1,3)} + \mu_{(1,4)}\} \leq \frac{1}{12} \lambda_2 + \frac{5}{12} \lambda_3 + \frac{1}{2} \lambda_4.$$  

(3.45)

We now consider $n = 4, k = 3$.

$$\omega = \{i_1, i_2, i_3\}, \quad 1 \leq i_1 < i_2 < i_3 \leq 4$$

i.e. \quad \omega = \{1, 2, 3\}, \quad \{1, 2, 4\}, \quad \{1, 3, 4\}, \quad \{2, 3, 4\}.

For $j = 1$,

$$\left(\begin{array}{c} 4 \\ 3 \end{array}\right)^{-1} \sum_{\omega \in \mathcal{Q}_4} \mu_{\omega 1} = \frac{1}{4} \{\mu_{(1,2,3)} + \mu_{(1,2,4)} + \mu_{(1,3,4)} + \mu_{(2,3,4)}\}$$

(3.46)

$$\sum_{r=0}^{1} \Psi_r \lambda_{1+r} = \Psi_0 \lambda_1 + \Psi_1 \lambda_2$$

(3.47)

$$\sum_{r=0}^{1} \Psi_r \lambda_{2-r} = \Psi_1 \lambda_1 + \Psi_0 \lambda_2$$

(3.48)

For $j = 2$, we get

$$\left(\begin{array}{c} 4 \\ 3 \end{array}\right)^{-1} \sum_{\omega \in \mathcal{Q}_4} \mu_{\omega 2} = \frac{1}{4} \{\mu_{(1,2,3)} + \mu_{(1,2,4)} + \mu_{(1,3,4)} + \mu_{(2,3,4)}\}$$

(3.49)
\[
\sum_{\tau=0}^{1} \Psi_{\tau} \lambda_{2+\tau} = \Psi_{0} \lambda_{2} + \Psi_{1} \lambda_{3}
\] (3.50)

\[
\sum_{\tau=0}^{1} \Psi_{\tau} \lambda_{3-\tau} = \Psi_{1} \lambda_{2} + \Psi_{0} \lambda_{3}
\] (3.51)

and

\[
\Psi_{0} = \frac{1}{4},
\]

\[
\Psi_{1} = \frac{3}{4}.
\]

Finally, we get

\[
\frac{3}{4} \lambda_{1} + \frac{1}{4} \lambda_{2} \leq \frac{1}{4} \left\{ \mu_{\{1,2,3\}} + \mu_{\{1,2,4\}} + \mu_{\{1,3,4\}} + \mu_{\{2,3,4\}} \right\} \leq \frac{1}{4} \lambda_{1} + \frac{3}{4} \lambda_{2}
\] (3.52)

and

\[
\frac{3}{4} \lambda_{2} + \frac{1}{4} \lambda_{3} \leq \frac{1}{4} \left\{ \mu_{\{1,2,3\}} + \mu_{\{1,2,4\}} + \mu_{\{1,3,4\}} + \mu_{\{2,3,4\}} \right\} \leq \frac{1}{4} \lambda_{2} + \frac{3}{4} \lambda_{3}.
\] (3.53)
In the previous sections we mentioned the links between the roots of

\[ \text{det}(\lambda I_n - A) = 0 \]  \hspace{1cm} (4.54)

(the eigenvalues of \( A \)) and the roots of

\[ \text{det}(\lambda I_{n-1} - A(i\setminus i)) = 0, \quad i = 1, 2, \ldots, n, \]  \hspace{1cm} (4.55)

(the eigenvalues of \( A(i\setminus i) \)). Where \( A \) is an \( n \)-square Hermitian matrix and \( A(i\setminus i) \) denote the principal submatrix of \( A \) obtained by deleting from \( A \) both row \( i \) and column \( i \).

For fixed \( i \), the roots of (4.55) interlace the roots of (4.54) and this is the Cauchy interlacing inequalities.

Let \( C \) be an \( n \)-square positive definite Hermitian matrix. Thompson [17] obtained new links between the roots of the determinantal equations.

\[ \text{det}(\lambda C - A) = 0 \]  \hspace{1cm} (4.56)

and

\[ \text{det}(\lambda C(i\setminus i) - A(i\setminus i)) = 0, \quad i = 1, \ldots, n \]  \hspace{1cm} (4.57)

The roots of (4.56) are real (they are the eigenvalues of \( C^{-\frac{1}{2}}AC^{-\frac{1}{2}} \)).

It is a fact, not quite as well known as the Cauchy inequalities, that the roots of any one of the equations (4.57) interlace the roots of (4.56).
Section 4. GENERALIZATION TO FORMS

Some of the results obtained in the previous section concerning the Cauchy inequalities may be extended to the roots of (4.56) and (4.57).

Notation: In this section, we shall let

\[ \lambda_1 \leq \cdots \leq \lambda_n \]

be the roots of (4.56) and let

\[ \eta_{i1} \leq \cdots \leq \eta_{i,n-1} \]

be the roots of (4.57).

The numbers

\[ \lambda_1, \ldots, \lambda_n \]

need not be distinct. The ideas of trivial and nontrivial eigenvalues as discussed in section 2 are also employed here.

Let \( \mu_i \) with multiplicity \( e_i \), for \( i = 1, 2, \ldots, s \) be the distinct numbers among

\[ \lambda_1, \ldots, \lambda_n. \]

Let

\[ \gamma_1 \leq \cdots \leq \gamma_n \]

be the eigenvalues of \( C \).
Let

\[ f(\lambda) = \det(\lambda C - A) \]

and

\[ f_{(i)}(\lambda) = \det(\lambda C(i) - A(i)). \]

Hence the number \( \mu_j \) with multiplicity \( e_{j-1} \), \( 1 \leq j \leq s \) arranged such that

\[ \mu_1 < \mu_2 < \cdots < \mu_s \]

are always eigenvalues of \( f_{(i)}(\lambda) \), called the **trivial eigenvalues** of \( f_{(i)}(\lambda) \).

The remaining eigenvalues denoted by

\[ \xi_{i1} \leq \cdots \leq \xi_{i,s-1} \]

are called the **nontrivial eigenvalues**.

Thompson [17] established the first interlacing principle as stated in the following theorem.

**Theorem 4.1**

Let \( C \) be positive definite, and let

\[ \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \]

be the roots of (4.56). For fixed \( i \), let

\[ \eta_{i1} \leq \cdots \leq \eta_{i,n-1} \]

be the roots of (4.57).
Then

\[ \eta_1, \ldots, \eta_{n-1} \]

interlace

\[ \lambda_1, \ldots, \lambda_n; \]

that is,

\[ \lambda_3 \leq \eta_1 \leq \lambda_2 \leq \eta_2 \leq \cdots \leq \lambda_{n-1} \leq \eta_{n-1} \leq \lambda_n. \] (4.58)

Thompson developed new generalizations of the interlacing principle that was derived in the previous sections which he called the second interlacing principles.

Let \( j \) be fixed with \( 1 \leq j < s \). Define polynomials \( G_j(\lambda) \) and \( H_j(\lambda) \) by

\[
G_j(\lambda) = \sum_{t=1}^{j} (\gamma_{n-1}^{-1} + \cdots + \gamma_{n+1-\varepsilon_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t} + \sum_{t=j+1}^{s} (\gamma_{t-1}^{-1} + \cdots + \gamma_{s-1}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t}, \tag{4.59}
\]

\[
H_j(\lambda) = \sum_{t=1}^{j} (\gamma_{t-1}^{-1} + \cdots + \gamma_{s-1}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t} + \sum_{t=j+1}^{s} (\gamma_{n-1}^{-1} + \cdots + \gamma_{n+1-\varepsilon_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t}. \tag{4.60}
\]

Where

\[
\hat{f}(\lambda) = \det C\Pi_{j=1}^{s}(\lambda - \mu_j). \tag{4.61}
\]

\( G_j(\lambda) \) (respectively \( H_j(\lambda) \)) is alternately positive and negative when evaluated at \( \mu_2, \ldots, \mu_1 \). Hence \( G_j(\lambda) \) (resp. and \( H_j(\lambda) \) has exactly one root in each interval

\[(\mu_1, \mu_2), (\mu_2, \mu_3), \ldots, (\mu_{s-1}, \mu_s)\]

Let \( \alpha \) be the unique root of \( G_j(\lambda) \) in \([\mu_j, \mu_{j+1}]\) and let \( \beta \) be the unique root of \( H_j(\lambda) \) in \([\mu_j, \mu_{j+1}]\).
Lemma 4.2

\[ \alpha \leq \beta. \]

The second interlacing principle is stated in the following theorem.

**Theorem 4.3 (Second Interlacing Principle)**

Let \( j \) be fixed, \( 1 \leq j < s \). With \( \alpha \) and \( \beta \) as above, we have:

(i) either

\[ \xi_{1j} = \xi_{2j} = \cdots = \xi_{nj} = \alpha \]

or for at least one \( i \), we have

\[ \xi_{ij} > \alpha; \]

(ii) either

\[ \xi_{1j} = \xi_{2j} = \cdots = \xi_{nj} = \beta, \]

or for at least one \( i \), we have

\[ \xi_{ij} < \beta. \]

Thompson showed how the quadratic inequalities and linear inequalities developed in section 3 may be generalized to the situation under discussion in this section.

**Theorem 4.4**

Let \( s > 1 \), for \( 1 \leq j \leq s \), we have

\[
\sum_{i=1}^{n} \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}} \frac{\xi_{ij} - \mu_j}{\mu_{j+1} - \mu_j} \geq \gamma_1(\sum_{r=1}^{s} \gamma_{n+1-r}^{-1}) \tag{4.62}
\]

and

\[
\sum_{i=1}^{n} \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_1} \frac{\xi_{ij} - \mu_j}{\mu_{s} - \mu_j} \leq \gamma_n(\sum_{r=1}^{s} \gamma_r^{-1}) \tag{4.63}
\]
Remark: Formula (4.62) is called the upper quadratic inequality and (4.63) is called the lower quadratic inequality. When \( j = 1 \), the factors involving 

\[ \mu_j - \xi_{i,j-1} \]

are understood to be absent in these formulas, and when \( j = s \), the factors involving 

\[ \xi_{i,j} - \mu_j \]

are similarly understood to be absent. If we delete 

\[ \frac{\xi_{i,j} - \mu_j}{\mu_{j+1} - \mu_j} \]

we obtain the inequality (4.64) below. If we delete instead 

\[ \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}} \]

we obtain (4.65) below.

These inequalities (4.64) and (4.65) bound the arithmetic mean of the \( \xi_{ij} \).

**Theorem 4.5**

Let \( s > 1 \). Then:

\[ \frac{1}{n} \sum_{i=1}^{n} \xi_{i,j-1} \leq \phi_j \mu_{j-1} + (1 - \phi_j)\mu_j, \quad j = 2, 3, \ldots, s, \quad (4.64) \]

\[ \frac{1}{n} \sum_{i=1}^{n} \xi_{ij} \geq \phi_j \mu_{j+1} + (1 - \phi_j)\mu_j, \quad j = 1, 2, \ldots, s - 1. \quad (4.65) \]

Here 

\[ \phi_j = \frac{1}{n} \gamma_1 \sum_{r=1}^{s_j} \gamma_{n+1-\tau}^{-1}, \]

and
Remark:

In (4.64) we have a convex combination of $\mu_{j-1}$ and $\mu_j$ which serves as an upper bound for the arithmetic mean of the $\xi_{i,j-1}$ (by interlacing, we only know that this arithmetic mean lies in $[\mu_{j-1}, \mu_j]$). In (4.65) we have similar convex combination of $\mu_j$ and $\mu_{j+1}$ which bound from below the arithmetic mean of $\xi_{ij}$.

As our final result, we show how the bounds on the arithmetic mean of the $\eta_{ij}$ developed in section 3 may be generalized to the situation under discussion in this section.

Let $Q_{nk}$ denote the set of all sequences $\omega = \{i_1, \ldots, i_k\}$ of strictly increasing positive integers not exceeding $n$. The number of sequences in $Q_{nk}$ is

\[
\binom{n}{k}
\]

Let $A[\omega \setminus \omega]$ denote the principal submatrix of $A$ lying at the intersection of rows and columns $i_1, \ldots, i_k$.

Let

\[
\xi_{\omega 1} \leq \xi_{\omega 2} \leq \cdots \leq \xi_{\omega k}
\]

denote the roots of

\[
det(\lambda C[\omega \setminus \omega] - A[\omega \setminus \omega]) = 0 \tag{4.66}
\]

It follows from the theorem (4.5) that if $\tau \in Q_{n,k+1}$, and $\tau \supset \omega$, then the roots of

\[
det(\lambda C[\tau \setminus \tau] - A[\tau \setminus \tau]) = 0 \tag{4.67}
\]
are interlaced by the roots of (4.66). It therefore follows that

\[ \lambda_j \leq \xi_{\omega j} \leq \lambda_{j+n-k}, \quad j = 1, 2, \ldots, k \]

for \( \omega \in \mathcal{Q}_{nk} \).

Thompson [17] generalized the bounds on the arithmetic mean discussed in section 3 by finding convex combinations of

\[ \lambda_j, \lambda_{j+1}, \ldots, \lambda_{j+n-k}, \]

which serve as upper and lower bounds for the arithmetic mean

\[ \left( \binom{n}{k} \right)^{-1} \sum_{\omega \in \mathcal{Q}_{nk}} \xi_{\omega j} \]

of the \( j^{th} \) root (fixed \( j \)) of all the different equations (4.66).

Let

\[ E_r(a_1, \ldots, a_k) \]

denote the elementary symmetric function of degree \( r \) on \( k \) variables.

**Theorem 4.6**

For fixed \( j \) and \( k, 1 \leq k \leq n-1, 1 \leq j \leq k \), we have

\[ \sum_{r=0}^{n-k} \Psi_r \lambda_{n-k+j-r} \leq \left( \binom{n}{k} \right)^{-1} \sum_{\omega \in \mathcal{Q}_{nk}} \xi_{\omega j} \leq \sum_{r=0}^{n-k} \Psi_r \lambda_{j+r} \]

(4.68)

where

\[ \Psi_r = \frac{(\frac{n}{\gamma_n})^{n-k-r} E_r(n - \frac{\gamma_1}{\gamma_n}, n - 1 - \frac{\gamma_1}{\gamma_n}, \ldots, k + 1 - \frac{\gamma_1}{\gamma_n})}{\prod_{i=k+1}^{n} i}, \quad 0 \leq r \leq n - k, \]

(4.69)

\[ \sum_{r=0}^{n-k} \Psi_r = 1. \]

(4.70)
Section 5

GENERALIZATION TO SINGULAR VALUES

We examined, simultaneously all the \(k\)-square principal submatrices of an \(n\)-square matrix \(A\) in the previous sections.

Usually \(A\) has been symmetric or Hermitian, and much of our effort has centered around the well-known fact asserting that the eigenvalues of an \((n - 1)\)-square principal submatrix of Hermitian \(A\) always interlace the eigenvalues of \(A\).

In this section, we mention and discuss the singular values of the submatrices (not necessarily principal submatrices) of an arbitrary matrix \(A\) due to Thompson [18].

Now a brief summary of certain particular cases of the results developed by Thompson that merit special attention is given. Let \(A\) be an \(nxn\) real or complex matrix and let

\[
\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n
\]

be the singular value of \(A\), defined to be the eigenvalues of positive semi definite matrix

\[
(\overline{A}A^*)^{\frac{1}{2}}.
\]

Let \(B = A_{ij}\) be the \((n - 1)\)-square submatrix of \(A\) obtained by deleting row \(i\) and column \(j\), and let

\[
\beta_1 \geq \beta_2 \geq \cdots \geq \beta_{n-1}
\]

be singular values of \(B\).
The first theorem developed by Thompson yields, as a special case, these interlacing inequalities:

\[
\begin{align*}
\alpha_1 & \geq \beta_1 \geq \alpha_3, \\
\alpha_2 & \geq \beta_2 \geq \alpha_4, \\
\vdots \\
\alpha_t & \geq \beta_t \geq \alpha_{t+2}, & 1 \leq t \leq n - 2 \\
\vdots \\
\alpha_{n-2} & \geq \beta_{n-2} \geq \alpha_{n}, \\
\alpha_{n-1} & \geq \beta_{n-1}
\end{align*}
\] (5.71)

That inequalities (5.71) are the best that can be asserted is shown by (a special case of) Theorem 5.3. It follows from theorem 5.3 that, if arbitrary nonnegative numbers

\[
\beta_1 \geq \cdots \geq \beta_{n-1}
\]

are given satisfying (5.71), there will always exist unitary matrices U and V such that the singular values of

\[
(UAV)_{ij}
\]

are

\[
\beta_1, \ldots, \beta_{n-1}
\]

Of course, A and U A V always have the same singular values

\[
\alpha_1, \ldots, \alpha_n.
\]
Section 5. GENERALIZATION TO SINGULAR VALUES

We now give the definition of the singular values of a rectangular matrix.

**Definition 5.1** Let $A$ be an $m \times n$ matrix. The singular values

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{\min(m,n)}$$

of $A$ are the common eigenvalues of the positive semi definite matrices

$$(AA^*)^{\frac{1}{2}}$$

and

$$(A^*A)^{\frac{1}{2}}.$$ (Here $A^*$ is the Hermitian adjoint of $A$).

Since $AA^*$ is $m$-square and $A^*A$ is $n$-square, the eigenvalues of

$$(AA^*)^{\frac{1}{2}}$$

and

$$(A^*A)^{\frac{1}{2}}$$

do not coincide in full.

However, it is well known that the non zero eigenvalues (including multiplicities ) of these two matrices always coincide.

It is frequently convenient to define $\alpha_t$ to be zero for

$$\min(m,n) < t \leq \max(m,n).$$

Then

$$\alpha_1^2 \geq \cdots \geq \alpha_{\max(m,n)}^2$$

and the roots of $AA^*$ (respectively $A^*A$ ) are the first $m$ (respectively $n$) of these numbers.
Section 5. GENERALIZATION TO SINGULAR VALUES

We now state the first theorem due to Thompson.[18]

Theorem 5.2

Let A be an $m \times n$ matrix with singular values $(5.72)$. Let $B$ be a $p \times q$ submatrix of $A$, with singular values

$$\beta_1 \geq \beta_2 \geq \cdots \geq \beta_{\min(p,q)}.$$  \hspace{1cm} (5.73)

Then

$$\alpha_i \geq \beta_i, \quad i = 1, 2, \ldots, \min(p, q),$$ \hspace{1cm} (5.74)

$$\beta_i \geq \alpha_{i+(m-p)+(n-q)}, \quad i \leq \min(p + q - m, p + q - n).$$ \hspace{1cm} (5.75)

We state the second theorem, also due Thompson.

Theorem 5.3

Let $A$ be an $m \times n$ matrix with singular values $(5.72)$. Let arbitrary non-negative number $(5.73)$ be given, satisfying both $(5.74)$ and $(5.75)$. Then $m$-square unitary matrix $U$ and $n$-square unitary matrix $V$ exist such that the singular values of the $p \times q$ submatrix

$$(UAV)_{i_1, \ldots, i_p; j_1, \ldots, j_q}$$

of $UA \ V$ are the numbers $(5.73)$.

We remark that the nonincreasing condition $(5.73)$ is actually superfluous.

More precisely, we have Theorem 5.3(a).

Theorem 5.3(a)

Let arbitrary numbers

$$\beta_1, \ldots, \beta_{\min(p,q)}$$

be given, such that $(5.74)$ and $(5.75)$ hold. Then the conclusions of Theorem 5.3 are valid.
Bibliography


