A Comparison of Two Proofs of Yamamoto's Theorem Relating Eigenvalue moduli and Singular Values of a Matrix

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

Melinda Pell

Certificate of Approval:

Randall R. Holmes
Associate Professor
Mathematics

Tin-Yau Tam, Chair
Professor
Mathematics

Krystyna Kuperberg
Professor
Mathematics

Greg Harris
Associate Professor
Mathematics

Stephen L. McFarland
Dean
Graduate School

A Comparison of Two Proofs of Yamamoto's Theorem Relating Eigenvalue moduli and Singular Values of a Matrix

Melinda Pell

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Science

Auburn, Alabama
15 December, 2006

A Comparison of Two Proofs of Yamamoto's Theorem Relating Eigenvalue moduli and Singular Values of a Matrix

Melinda Pell

Signature of Author

Date of Graduation

Vita

Melinda Pell, daughter of Lonnie and Lucinda Tilley, was born in Oceana, West Virginia on December 20, 1971. Upon graduation from Oceana High School, Melinda began college, attending Concord College in Athens, WV. After only a year of schooling she and her husband of two years relocated to Georgia. It was some time and a state later when Melinda returned to her academic adventure in Florida by attending Florida Community College at Jacksonville. This was also a brief visit in the world of academia due to relocation of the family in 1994 to Phenix City, Alabama. In 1998, she decided to return to her scholarly goals. She enrolled at Columbus State University in Columbus, Georgia where she completed her B.S. (Math) in May, 2002. She began her graduate study in the Department of Mathematics of Auburn University in August, 2002. During the 2004-2005 year Melinda was a temporary instructor for the Department of Mathematics of Columbus State University.

Thesis Abstract

A Comparison of Two Proofs of Yamamoto's Theorem Relating Eigenvalue moduli and Singular Values of a Matrix

Melinda Pell

Master of Science, 15 December, 2006
(B.S., Columbus State University, 2002)

39 Typed Pages

Directed by Tin-Yau Tam

We examine two proofs of Yamamoto's theorem regarding the asymptotic relationship between singular values and eigenvalue moduli of a matrix. The first proof is by T. Yamamoto in 1967 and makes use of compound matrices. The second is by R. Mathias in 1990 through utilization of an interlacing theorem for singular values. We compare the two proofs.

Acknowledgments

The author would like to thank Dr. Tin-Yau Tam for his never-ending patience and excellent guidance throughout the course of this research, Carlos Almada for his generous contribution of time and assistance. Without the support of her family this task would have been unattainable, and it is with great sincerity that she sends gratitude out to them. Her three children have been most patient and kind throughout this experience. And most of all she would like to express her appreciation to her husband Tommy Pell for his loving patience and support through the many difficult days and long endless nights during the trials of this research.

Style manual or journal used Transactions of the American Mathematical Society (together with the style known as "aums"). Bibliograpy follows van Leunen's *A Handbook for Scholars.*

Computer software used The document preparation package TeX (specifically LaTeX) together with the departmental style-file `aums.sty`.

Table of Contents

Chapter 1

Introduction

In this thesis, we investigate two proofs of Yamamoto's theorem which provides the relationship between eigenvalue moduli and singular values of a square matrix in an asymptotic way.

Tools such as singular value decomposition, compound matrices, interlacing inequalities for eigenvalues and singular values, and Jordan form are briefly discussed. Some treatment is provided for developing notions of norms and other features that are necessary for completing the proofs. We denote by $\mathbb{C}_{n\times n}$ the vector space of all $n \times n$ complex matrices and $\mathbb{C}^n$ the vector space of complex n-tuples. The theorem of Yamamoto [9] is stated below:

Theorem 1.1 (Yamamoto)
Let $A \in \mathbb{C}_{n\times n}$. Then

$$\lim_{p\to\infty}[\sigma_i(A^p)\ ]^{\frac{1}{p}} = |\lambda_i(A)|, \qquad i = 1, 2, \ldots, n, \tag{1.1}$$

where $\sigma_1(A) \geq \cdots \geq \sigma_n(A) \geq 0$ are the singular values of A and $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A which are arranged in the non-increasing order $|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|$ with respect to their moduli.

The case $i = 1$ of Theorem 1.1 is a special case of Gelfand's Spectral Radius Theorem (1941) which may take the following form.

Theorem 1.2 (Gelfand)

Let $\|A\| := \max_{\|x\|=1} \|Ax\|$ be a matrix norm induced by a vector norm $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ where $A \in \mathbb{C}_{n\times n}$. Then

$$\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A), \tag{1.2}$$

where $\rho(A) := |\lambda_1(A)|$ is the spectral radius of A.

We remark that Gelfand's Theorem is also valid for Hilbert space bounded operators but the proof requires more advanced tools [2].

In Chapter 2, we will present and briefly discuss some concepts, in order to aid the reader and clarify any confusion. Then, in Chapter 3 we will focus on the original proof by Yamamoto [9], and in Chapter 4 we will examine and discuss the more recent proof by Mathias [3]. Once these proofs are each fully analyzed, a comparison of the two proofs will be presented in Chapter 5.

We finally remark that very recently Yamamoto's theorem is extended in the context of semi-simple Lie group by Tam and Huang [7].

Chapter 2

Notations and Theorems

The eigenvalues of $A \in \mathbb{C}_{n\times n}$ are the numbers λ such that $Ax = \lambda x$, for some nonzero vector $x \in \mathbb{C}^n$. The vector x is known as an eigenvector corresponding to the eigenvalue λ for the matrix A. According to the Fundamental Theorem of Algebra, each $A \in \mathbb{C}_{n\times n}$ has n eigenvalues $\lambda_1(A), \ldots, \lambda_n(A) \in \mathbb{C}$, counting multiplicities. We order them in such a way to have

$$|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|.$$

The spectral radius of A is the largest eigenvalue modulus and is denoted by

$$\rho(A) := |\lambda_1(A)|.$$

A matrix $A \in \mathbb{C}_{n\times n}$ is said to be Hermitian if $A^* = A$ where A^* is the complex conjugate transpose of A. It is said to be positive semi-definite (p.s.d.) if it is Hermitian and has nonnegative eigenvalues, and it is said to be positive definite (p.d.) if it is Hermitian and has positive eigenvalues. A matrix $U \in \mathbb{C}_{n\times n}$ is said to be unitary if it satisfies the condition $U^* = U^{-1}$. A matrix $A \in \mathbb{C}_{n\times n}$ is said to be nilpotent if $A^p = 0$ for some $p \in \mathbb{N}$.

The singular values $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_n(A) \geq 0$ of the matrix $A \in \mathbb{C}_{n\times n}$ are the square roots of the corresponding eigenvalues of the p.s.d. matrix A^*A, i.e.,

$$\sigma_i(A) := \sqrt{\lambda_i(A^*A)}, \qquad i = 1, \ldots, n.$$

One can use AA^* to define singular values: $\sigma_i(A) := \sqrt{\lambda_i(AA^*)}$ because AB and BA have the same spectrum.

A norm on a vector space X is a map $\|\cdot\| : X \to \mathbb{R}$ satisfying the following properties

1. $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$,
2. (Triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in X$,
3. $\|\alpha x\| = |\alpha|\|x\|$ for all scalars α and $x \in X$.

Example 2.1 Let $X = \mathbb{C}^n$ and $x \in \mathbb{C}^n$.

The 2-norm is defined as

$$\|x\|_2 = \sqrt{x^*x} = (\sum_{i=1}^{n} |x_i|^2)^{1/2}$$

and $\|\cdot\|_2$ is a special case of the p-norms

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p < \infty.$$

For example $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ and $\|x\|_\infty = \max_{1\leq i\leq n}|x_i|$.

Let $\|\cdot\|$ be a vector norm on $\mathbb{C}^n$. The map $\|\cdot\| : \mathbb{C}_{n\times n} \to \mathbb{R}$:

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

is called the induced matrix norm on $\mathbb{C}_{n\times n}$ (induced by the vector norm $\|\cdot\|$). It is easy to verify that an induced matrix norm is a norm and satisfies

1. $\|Ax\| \le \|A\|\|x\|$ for all $A \in \mathbb{C}_{n\times n}$, $x \in \mathbb{C}^n$.

2. (submultiplicative) $\|AB\| \le \|A\|\|B\|$, for all $A, B \in \mathbb{C}_{n\times n}$.

3. $\|I\| = 1$.

Example 2.2 Let $A \in \mathbb{C}_{n\times n}$. Then

1. $\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p}$, $1 \le p < \infty$.

2. $\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$ (column sum norm).

3. $\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$ (row sum norm) and thus $\|A\|_1 = \|A^*\|_\infty$.

4. $\|A\|_2$ is the square root of the largest eigenvalue of A^*A (or AA^*). It is also called the largest singular value.

The induced matrix norm $\|\cdot\|_2$ is a very important norm on $\mathbb{C}_{n\times n}$ and is called the spectral norm. It is well-known [4] that for all $A \in \mathbb{C}_{n\times n}$,

1. $\|A^*\|_2 = \|A\|_2$, and

2. $\|A^*A\|_2 = \|A\|_2^2$.

We will specify which norm we use if there is a need.

Determinants are used often throughout this paper, and it is worthwhile to make note of the properties they possess. Some useful properties of determinants [1] are:

1. A matrix $A \in \mathbb{C}_{n\times n}$ is singular if and only if $\det A = 0$. If A is nonsinuglar, then $\det(A^{-1}) = (\det A)^{-1}$.

2. For an upper triangular matrix A the determinant is the product of the diagonal entries of A, $\det A = \prod_{i=1}^{n} a_{ii}$. This property holds true for lower triangular matrices as well. A particular case is the identity matrix I, for which $\det I = 1$.

3. The determinant of the product of two matrices A and B is equal to the product of the determinant of A and the determinant of B, $\det(AB) = \det A \det B$.

4. If $U \in \mathbb{C}_{n\times n}$ is unitary, then $\det U = 1$.

A Jordan block [4] $J_k(\lambda)$ is a $k \times k$ upper triangular matrix of the form

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{C}_{k\times k}$$

which can be expressed as $J_k = \lambda I_k + N_k$ where I_k is the $k \times k$ identity matrix and

$$N_k = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{C}_{k\times k}.$$

A Jordan matrix $J \in \mathbb{C}_{n\times n}$ is a direct sum of Jordan blocks and has the form:

$$\mathbf{J} = \begin{pmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_k}(\lambda_k) \end{pmatrix} \in \mathbb{C}_{n\times n},$$

where $n_1 + n_2 + \cdots + n_k = n$. Neither the values λ_i and the orders n_i need be distinct.

Theorem 2.3 (Jordan Canonical Form Theorem)

Let $A \in \mathbb{C}_{n\times n}$. Then there exists a nonsingular matrix $M \in \mathbb{C}_{n\times n}$ such that

$$M^{-1}AM = J,$$

where J is a Jordan matrix.

The matrix J in the above theorem is called a Jordan form of A and is unique up to permutation of the Jordan blocks.

We now prove the following classical result which gives a necessary and sufficient condition for the powers of a given matrix to tend to zero. It will be used in the proof of Gelfand's result.

Theorem 2.4 Let $B \in \mathbb{C}_{n\times n}$. Then $\lim_{p\to\infty} B^p = 0$ if and only if $\rho(B) < 1$.

Proof: ($\Rightarrow$) Suppose $\lim_{p\to\infty} B^p = 0$. Let λ be any eigenvalue of B, that is, there exists a nonzero $x \in \mathbb{C}^n$ such that $Bx = \lambda x$. Then for any $p \in \mathbb{N}$,

$$B^p x = \lambda^p x.$$

Now $B^p \to 0$ implies that $\lambda^p x \to 0$ and thus $\lambda^p \to 0$. So $|\lambda| < 1$ for all eigenvalues λ of B and hence $\rho(B) < 1$.

($\Leftarrow$) Let $B \in \mathbb{C}_{n\times n}$ such that $\rho(B) < 1$. Let

$$J := \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

be a Jordan form of B. Then by Theorem 2.3 there is a nonsingular matrix $M \in \mathbb{C}_{n\times n}$ such that $B = MJM^{-1}$. Each block $J_i \in \mathbb{C}_{n_i \times n_i}$ $(i = 1, \ldots, k)$ can be written as the sum of λI for some eigenvalue λ of B and a nilpotent matrix N, namely,

$$N = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{C}_{n_i \times n_i}.$$

Taking the p^{th} power of B yields $B^p = MJ^pM^{-1}$ where

$$J^p = \begin{pmatrix} J_1^p & & & \\ & J_2^p & & \\ & & \ddots & \\ & & & J_k^p \end{pmatrix}.$$

To show that $B^p \to 0$ it suffices to show $J_i^p \to 0$ as $p \to \infty$. Since each J_i is of the form $\lambda I + N$, we need to show that $\lim_{p\to\infty}(\lambda I + N)^p = 0$ under the assumption that $|\lambda| < 1$. Now the binomial expansion gives

$$(\lambda I + N)^p = \lambda^p + \binom{p}{1}\lambda^{p-1}N + \binom{p}{2}\lambda^{p-2}N^2 + \cdots + \binom{p}{p-1}\lambda^1 N^{p-1} + N^p.$$

Since N is nilpotent, $N^m = 0$ for some $m \in \mathbb{N}$. Then for $p \geq m$ we have

$$(\lambda I + N)^p = \lambda^p + \binom{p}{1}\lambda^{p-1}N + \binom{p}{2}\lambda^{p-2}N^2 + \cdots + \binom{p}{m-1}\lambda^{p-m+1}N^{m-1}.$$

It is sufficient to show that for each $j = 1, \ldots, m-1$,

$$\lim_{p\to\infty}\binom{p}{j}\lambda^{p-j} = 0$$

where $|\lambda| < 1$. Now

$$\lim_{p\to\infty} \frac{|\binom{p}{j}\lambda^{p-j}|}{|\binom{p-1}{j}\lambda^{p-1-j}|} = \lim_{p\to\infty} \left|\frac{p}{p-j}\lambda\right| = |\lambda| \lim_{p\to\infty} \left|\frac{p}{p-j}\right| = |\lambda| < 1$$

for any $j = 0, \ldots, m-1$. The ratio test implies $\lim_{p\to\infty} \binom{p}{j}\lambda^{p-j} = 0$ and we have the desired result. □

There are other methods of matrix decomposition, two of which are defined in the following theorems:

Theorem 2.5 [10] (Schur's Triangularization Theorem)

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $A \in \mathbb{C}_{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}_{n\times n}$ such that U^*AU is an upper triangular matrix, that is,

$$U^*AU = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

where the order of $\lambda_1, \ldots, \lambda_n$ can be arbitrarily fixed.

Theorem 2.6 [10] (Singular Value Decomposition)

Let $A \in \mathbb{C}_{m\times n}$ and let $\sigma_1, \sigma_2, \ldots, \sigma_r$ be the nonzero singular values of A. Then there exist unitary matrices $U \in \mathbb{C}_{m\times m}$ and $V \in \mathbb{C}_{n\times n}$ such that

$$A = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V,$$

where $D = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$. Thus $\operatorname{rank} A = r$ which is the number of nonzero singular values of A.

Due to Theorem 2.6 the singular values remain the same under unitary equivalence, i.e., A and UAV have the same singular values if U and V are unitary matrices.

The rank of a matrix also has some important properties that will be needed. They include [1]:

1. $\text{rank}\,(AB) \leq \text{rank}\,(A)$.

2. $\text{rank}\,(AB) \leq \text{rank}\,(B)$.

3. $\text{rank}\,(A^*A) = \text{rank}\,(A)$.

Suppose $A \in \mathbb{C}_{n\times n}$ and $1 \leq k \leq n$. Then the k^{th} compound of A is defined as the $\binom{n}{k} \times \binom{n}{k}$ complex matrix $C_k(A)$ whose elements are defined by

$$C_k(A)_{\alpha,\beta} = \det A[\alpha|\beta]. \tag{2.1}$$

Here $A[\alpha|\beta]$ is the $k \times k$ submatrix of A obtained by choosing the rows indexed by α and the columns indexed by β, where $\alpha, \beta \in Q_{k,n}$ and

$$Q_{k,n} := \{\omega = (\omega(1), \ldots, \omega(k)) : 1 \leq \omega(1) < \cdots < \omega(k) \leq n\}$$

is the set of increasing sequences of length k chosen from $1, \ldots, n$. For example, if $n = 3$ and $k = 2$, then

$$C_2(A) = \begin{pmatrix} \det A[1,2|1,2] & \det A[1,2|1,3] & \det A[1,2|2,3] \\ \det A[1,3|1,2] & \det A[1,3|1,3] & \det A[1,3|2,3] \\ \det A[2,3|1,2] & \det A[2,3|1,3] & \det A[2,3|2,3] \end{pmatrix}.$$

In general $C_1(A) = A$ and $C_n(A) = \det A$.

Some properties of the compound matrix [8] are listed in the following result.

Theorem 2.7 Let $A, B \in \mathbb{C}_{n\times n}$. Then

1. $C_k(AB) = C_k(A)C_k(B)$.

2. $[C_k(A)]^* = C_k(A^*)$.

3. $C_k(A^{-1}) = [C_k(A)]^{-1}$ if A is nonsingular.

4. If A is normal, Hermitian, positive definite (or nonnegative) or unitary, then so is $C_k(A)$.

5. The eigenvalues of $C_k(A)$ are the $\binom{n}{k}$ numbers $\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}$ for $(i_1,\ldots,i_k) \in Q_{k,n}$, where $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of A. In particular the eigenvalue of maximal modulus of $C_k(A)$ is $|\lambda_1(C_k(A))| = |\lambda_1(A)|\cdots|\lambda_k(A)|$.

6. The singular values of $C_k(A)$ are the $\binom{n}{k}$ numbers $\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k}$ for $(i_1,\ldots,i_k) \in Q_{k,n}$, where $\sigma_1,\ldots,\sigma_n$ are the singular values of A. In particular, the largest singular value of $C_k(A)$ is $\sigma_1(C_k(A)) = \sigma_1(A)\cdots\sigma_k(A)$.

Principal submatrices are used in Chapter 4, and the relationship between the eigenvalues of a matrix A and the principal submatrices of A is described in the following theorem.

Theorem 2.8 [10] (Interlacing Inequalities for Eigenvalues)

Let H be an $n \times n$ Hermitian matrix partitioned as

$$H = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}$$

where A is an $m \times m$ principal submatrix of H, $1 \le m \le n$. Then

$$\lambda_k(H) \ge \lambda_k(A) \ge \lambda_{k+n-m}(H), \qquad k = 1, 2, \ldots, m.$$

In particular, when $m = n - 1$,

$$\lambda_1(H) \geq \lambda_1(A) \geq \lambda_2(H) \geq \cdots \geq \lambda_{n-1}(H) \geq \lambda_{n-1}(A) \geq \lambda_n(H).$$

Proof: [10] It is sufficient to prove the $m = n - 1$ case. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of H and let $\mu_1 \geq \cdots \geq \mu_{n-1}$ be the eigenvalues of A. By the Spectral Theorem of Hermitian matrices, there is a unitary $U \in \mathbb{C}_{n\times n}$ such that

$$H = U^* \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} U.$$

Then

$$tI - H = U^* \begin{pmatrix} t-\lambda_1 & & & \\ & t-\lambda_2 & & \\ & & \ddots & \\ & & & t-\lambda_n \end{pmatrix} U,$$

and thus for $t \neq \lambda_i$, $i = 1, 2, \ldots, n$,

$$(tI - H)^{-1} = U^* \begin{pmatrix} \frac{1}{t-\lambda_1} & & & \\ & \frac{1}{t-\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{t-\lambda_n} \end{pmatrix} U. \qquad (2.2)$$

When $t \neq \lambda_i$, $i = 1, 2, \ldots, n$,

$$(tI - H)^{-1} = \frac{\operatorname{adj}(tI - H)}{\det(tI - H)}, \tag{2.3}$$

where $\operatorname{adj} A$ denotes the adjugate of $A \in \mathbb{C}_{n \times n}$. Upon computation, the (n,n)-entry of $(tI - H)^{-1}$ by using (2.2) is

$$\frac{|u_{1n}|^2}{t - \lambda_1} + \frac{|u_{2n}|^2}{t - \lambda_2} + \cdots + \frac{|u_{nn}|^2}{t - \lambda_n}$$

and the (n,n)-entry of $\operatorname{adj}(tI - H)$ is $\det(tI - A)$. Thus by (2.3)

$$\varphi(t) := \frac{\det(tI - A)}{\det(tI - H)} = \frac{|u_{1n}|^2}{t - \lambda_1} + \frac{|u_{2n}|^2}{t - \lambda_2} + \cdots + \frac{|u_{nn}|^2}{t - \lambda_n}. \tag{2.4}$$

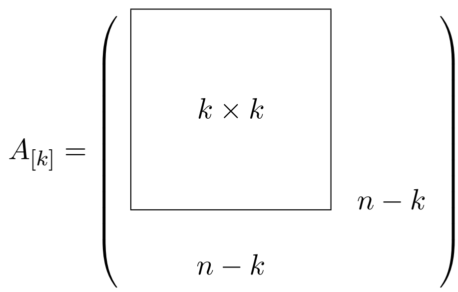
Assume that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. Note that $\varphi(t)$ is continuous whose roots are $\mu_1 \geq \ldots \geq \mu_{n-1}$, which must interlace $\lambda_1, \ldots, \lambda_n$, i.e.,

$$\mu_i \in [\lambda_{i+1}, \lambda_i], \qquad i = 1, 2, \ldots, n-1.$$

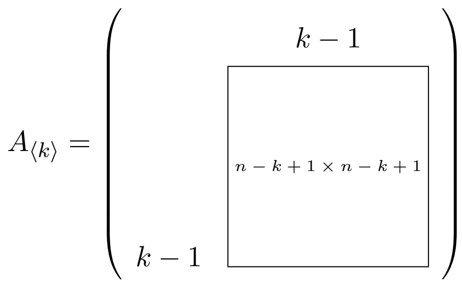
By continuity argument, we have the same conclusion for $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

□

Notations: Let $1 \leq k \leq n-1$. We denote by $A_{[k]}$ the submatrix formed by selecting the first k rows and columns of A. In other words, $A_{[k]}$ is upper-left corner principal $k \times k$ submatrix of A:



Denote by $A_{\langle k \rangle}$ the submatrix generated by deleting the first $k-1$ rows and columns of A. In other words, $A_{\langle k \rangle}$ is lower-right corner principal $(n-k+1)\times(n-k+1)$ submatrix of A:



Notice that the k^{th} entry of the diagonal is a member in each submatrix.

Theorem 2.9 [3] (Interlacing Inequalities for Singular Values)

Let $A \in \mathbb{C}_{n\times n}$ be given and let $A_p \in \mathbb{C}_{n\times(n-p)}$ (respectively $A_p \in \mathbb{C}_{(n-p)\times n}$) denote a submatrix of A obtained by deleting any p columns (or respectively any p rows) from A. Then

$$\sigma_i(A) \geq \sigma_i(A_p) \geq \sigma_{i+p}(A), \qquad i = 1, 2, \ldots, n-p.$$

Proof: [10] It is sufficient to establish the $p = 1$ case and for definiteness suppose that A_1 is obtained by deleting the last column from A. Then A_1 is a submatrix of A and

$A_1^*A_1$ is a principal submatrix of A^*A. By Theorem 2.8 we have

$$\lambda_i(A^*A) \geq \lambda_i(A_1^*A_1) \geq \lambda_{i+1}(A^*A).$$

By taking square roots, we obtain $\sigma_i(A) \geq \sigma_i(A_1) \geq \sigma_{i+1}(A)$. □

Lemma 2.10 [3] Let $A \in \mathbb{C}_{n\times n}$ be given, and consider $B = (0 \mid A) \in \mathbb{C}_{n\times(n+p)}$ and $C = \begin{pmatrix} A \\ 0 \end{pmatrix} \in \mathbb{C}_{(n+p)\times n}$ obtained by adjoining p zero columns (respectively, rows) to A. Then

$$\sigma_i(A) = \sigma_i(B) = \sigma_i(C), \qquad i = 1, 2, \ldots, n.$$

Proof: The nonzero singular values of A are the square roots of the positive eigenvalues of A^*A or AA^*. Now $BB^* = AA^*$ and $C^*C = A^*A$. So $\sigma_i(A) = \sigma_i(B) = \sigma_i(C)$, $i = 1, 2, \ldots, n$, when zero singular values are also counted. □

Chapter 3

Yamamoto's original proof

Recall Yamamoto's theorem as it was stated earlier in Theorem 1.1:

Let $A \in \mathbb{C}_{n\times n}$. Then

$$\lim_{p\to\infty} [\sigma_i(A^p)\,]^{\frac{1}{p}} = |\lambda_i(A)|, \qquad i = 1, 2, \ldots, n.$$

The largest singular value and largest eigenvalue modulus fall under $i = 1$ case and is a special case of Gelfand's Spectral Radius Theorem (Theorem 1.2) where the induced matrix norm is the spectral norm. We may use the spectral norm and its properties to achieve our goal in this case. In his proof, Yamamoto [9] first provides the following lemma in order to establish Gelfand's result which gives the case $k = 1$.

Lemma 3.1 Let $\|\cdot\| : \mathbb{C}_{n\times n} \to \mathbb{R}$ be a matrix norm induced by a vector norm $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$. Given $A \in \mathbb{C}_{n\times n}$, for every $p \in \mathbb{N}$, we have

$$\rho(A) \le \|\, A^p\|^{\frac{1}{p}} \le \|\, A\|,$$

where $\rho(A)$ is the spectral radius of A.

Proof: Let $A \in \mathbb{C}_{n\times n}$. For any eigenvalue λ of A, let $x \in \mathbb{C}^n$ be a unit eigenvector corresponding to λ. That is, $\|x\| = 1$ and

$$Ax = \lambda x.$$

So we have

$$A^p x = \lambda^p x.$$

By taking the norm of this equality we have $\|A^p x\| = \|\lambda^p x\|$. The homogeneous and sub-multiplicative property yield

$$\|A\|^p = \|A\|^p \, \|x\| \geq \|A^p\| \, \|x\| \geq \|A^p x\| = \|\lambda^p x\| = |\lambda^p| \, \|x\| \geq |\lambda|^p \, \|x\| = |\lambda|^p.$$

By taking the p^{th}-root on both sides we obtain

$$\|A\| \geq \|A^p\|^{\frac{1}{p}} \geq |\lambda| = \rho(A)$$

since $\|x\| = 1$. In particular it is true for $\lambda = \lambda_1$. □

We now have the tools to prove Gelfand's Spectral Radius Theorem which states that for any induced matrix norm $\|\cdot\|$ on $\mathbb{C}_{n\times n}$ and any $A \in \mathbb{C}_{n\times n}$,

$$\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A). \tag{3.1}$$

Proof: By Lemma 3.1 the sequence $\{\|A^p\|^{\frac{1}{p}}\}_{p\in\mathbb{N}}$ is contained in the closed and bounded interval $[\rho(A), \|A\|]$. Let α be any limit point of the sequence $\{\|A^p\|^{\frac{1}{p}}\}_{p\in\mathbb{N}}$. So there is a convergent subsequence in $[\rho(A), \|A\|]$, denoted by $\{\|A^{p_i}\|^{\frac{1}{p_i}}\}_{i\in\mathbb{N}}$ where $1 \leq p_1 \leq p_2 \leq \cdots$, such that $\|A^{p_i}\|^{\frac{1}{p_i}}$ converges to the limit point $\alpha \in [\rho(A), \|A\|]$. We claim that $\alpha = \rho(A)$.

Suppose on the contrary $\rho(A) < \alpha$. There would exist some positive number α' such that $0 \leq \rho(A) < \alpha' < \alpha$. Then

$$\frac{\rho(A)}{\alpha'} = \rho(\frac{A}{\alpha'}) < 1$$

which implies $\left(\frac{A}{\alpha'}\right)^{p_i} \to 0$ as $p_i \to \infty$ by Theorem 2.4. So

$$\| \left(\frac{A}{\alpha'}\right)^{p_i} \| \to 0 \quad \text{as} \quad p_i \to \infty.$$

Thus, for a fixed constant $\epsilon > 0$, there exists a positive integer $N(\epsilon)$ such that

$$\| \left(\frac{A}{\alpha'}\right)^{p_i} \| < \epsilon$$

for every $i > N(\epsilon)$. Then

$$1 < \frac{\alpha}{\alpha'} = \frac{1}{\alpha'} \lim_{i\to\infty} \|A^{p_i}\|^{\frac{1}{p_i}} = \lim_{i\to\infty} \| \left(\frac{A}{\alpha'}\right)^{p_i} \|^{\frac{1}{p_i}} < \lim_{i\to\infty} \epsilon^{\frac{1}{p_i}} = 1,$$

a contradiction! So $\rho(A) = \alpha$. Since α is arbitrary we have established that there is only one limit point, $\rho(A)$.

Suppose that $\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}}$ were not equal to $\rho(A)$. Then there would exist an $\epsilon > 0$ such that for every $j \in \mathbb{N}$, there would exist $p_j > m$ with $|\|A^{p_j}\|^{\frac{1}{p_j}} - \rho(A)| \geq \epsilon$. So there would exist a subsequence $\{\|A^{p_j}\|^{\frac{1}{p_j}}\}_{j\in\mathbb{N}}$ of $\{\|A^p\|^{\frac{1}{p}}\}_{p\in\mathbb{N}}$ such that for all $j \in \mathbb{N}$

$$|\|A^{p_j}\|^{\frac{1}{p_j}} - \rho(A)| \geq \epsilon. \tag{3.2}$$

But $\{\|A^{p_j}\|^{\frac{1}{p_j}}\}_{j\in\mathbb{N}}$ is a subsequence of $\{\|A^p\|^{\frac{1}{p}}\}_{p\in\mathbb{N}}$ which is contained in the closed and bounded interval $[\rho(A), \|A\|]$. So there is a convergent subsequence of $\{\|A^{p_j}\|^{\frac{1}{p_j}}\}_{j\in\mathbb{N}}$, namely $\{\|A^{p_{j_k}}\|^{\frac{1}{p_{j_k}}}\}_{k\in\mathbb{N}}$, and this convergent subsequence must converge to the limit point $\rho(A)$ since $\rho(A)$ is the only limit point of $\|A^p\|^{\frac{1}{p}}$, i.e. $\lim_{k\to\infty} \|A^{p_{j_k}}\|^{\frac{1}{p_{j_k}}} = \rho(A)$. So for the same $\epsilon > 0$, there exists $N(\epsilon) \in \mathbb{N}$ such that

$$|\|A^{p_{j_k}}\|^{\frac{1}{p_{j_k}}} - \rho(A)| < \epsilon$$

whenever $k > N(\epsilon)$, contradicting (3.2). □

Hence we have $\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A)$. When $\|\cdot\| = \|\cdot\|_2$ we have the $k = 1$ case of Yamamoto's Theorem as a corollary since $\|A\|_2 = \sigma_1(A)$.

Corollary 3.2 Let $A \in \mathbb{C}_{n\times n}$. Then $\lim_{p\to\infty}[\sigma_1(A^p)]^{\frac{1}{p}} = |\lambda_1(A)|$.

We now prove the remaining cases $k = 2, \ldots, n$ of (1.1).

Proof: The properties of compound matrices allow the use of the $k = 1$ case (Corollary 3.2) to show the result true for the finishing case. By Theorem 2.7 we have

$$|\lambda_1(C_k(A))| = \prod_{i=1}^{k} |\lambda_i(A)|, \quad \sigma_1(C_k(A)) = \prod_{i=1}^{k} \sigma_i(A). \tag{3.3}$$

Apply Corollary 3.2 on the kth compound $C_k(A)$ of A:

$$\lim_{p\to\infty} [\prod_{i=1}^{k} \sigma_i(A^p)]^{\frac{1}{p}} = \prod_{i=1}^{k} |\lambda_i(A)|. \tag{3.4}$$

Case 1: A is nilpotent, i.e., $|\lambda_1(A)| = 0$. Then for all $j = 2, \ldots, n$,

$$0 = |\lambda_1(A)| = \lim_{p\to\infty}[\sigma_1(A^p)]^{1/p} \geq \lim_{p\to\infty}[\sigma_j(A^p)]^{1/p} \geq 0.$$

So

$$\lim_{p\to\infty}[\sigma_j(A^p)]^{1/p} = |\lambda_j(A)|.$$

Case 2: A is not nilpotent, i.e., $|\lambda_1(A)| \neq 0$. Let A have k nonzero eigenvalues for some $1 \leq k \leq n$, i.e., $|\lambda_1| \geq \cdots \geq |\lambda_k| > |\lambda_{k+1}| = \cdots = |\lambda_n| = 0$. Utilizing Theorem 2.5 we have $U^*AU = T$ where T is upper triangular with $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the diagonal.

Raising both sides to the power p, we have

$$(U^*AU)^p = U^*A^pU = T^p$$

and due to the upper triangular form of T we have $\lambda_1^p, \lambda_2^p, \ldots, \lambda_n^p$ on the diagonal of T^p. Since k of the eigenvalues are nonzero, $\operatorname{rank}(T^p) \geq k$. Notice

$$\operatorname{rank}(T^p) = \operatorname{rank}(U^*A^pU) = \operatorname{rank}(A^p).$$

So there are at least k nonzero singular values of A^p since the rank of a matrix is the number of nonzero singular values. Then we have for any p,

$$\prod_{i=1}^{t} \sigma_i(A^p) > 0, \qquad 1 \leq t \leq k. \tag{3.5}$$

Then for $1 \leq j \leq k+1$ we have

$$\begin{aligned}
\lim_{p\to\infty} [\sigma_j(A^p)]^{\frac{1}{p}} &= \lim_{p\to\infty} \left[\frac{\prod_{i=1}^{j} \sigma_i(A^p)}{\prod_{i=1}^{j-1} \sigma_i(A^p)} \right]^{\frac{1}{p}} \\
&= \frac{\lim_{p\to\infty} \left[\prod_{i=1}^{j} \sigma_i(A^p)\right]^{\frac{1}{p}}}{\lim_{p\to\infty} \left[\prod_{i=1}^{j-1} \sigma_i(A^p)\right]^{\frac{1}{p}}} && \text{(nonzero denominator by (3.5))} \\
&= \frac{\prod_{i=1}^{j} |\lambda_i(A)|}{\prod_{i=1}^{j-1} |\lambda_i(A)|} && \text{(by Corollary 3.2)} \\
&= |\lambda_j(A)|.
\end{aligned}$$

In particular when $j = k+1$

$$\lim_{p\to\infty} [\sigma_{k+1}(A^p)]^{\frac{1}{p}} = 0. \tag{3.6}$$

If $j > k+1$ we have by (3.6)

$$0 \leq [\sigma_j(A^p)]^{\frac{1}{p}} \leq [\sigma_{k+1}(A^p)]^{\frac{1}{p}} \to 0.$$

So, $\lim_{p\to\infty}[\sigma_j(A^p)]^{\frac{1}{p}} = |\lambda_j(A)| = 0$. Then for any $i = 1, \ldots, n$,

$$\lim_{p\to\infty}[\sigma_i(A^p)]^{\frac{1}{p}} = |\lambda_i(A)|.$$

□

Chapter 4

Mathias' proof

Mathias [3] provides a different method of proving Yamamoto's theorem. Like Yamamoto, he also makes use of Gelfand's Spectral Radius Theorem to show the $k = 1$ case. Recall Gelfand's result (Theorem 1.2):

Let $\|A\| := \max_{\|x\|=1} \|Ax\|$ be a matrix norm induced by a vector norm $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ where $A \in \mathbb{C}_{n\times n}$. Then

$$\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A),$$

where $\rho(A) := |\lambda_1(A)|$ is the spectral radius of A.

The proof of Gelfand's result provided by Mathias proceeds as follows.

Proof: For any eigenvalue λ associated with $A \in \mathbb{C}_{n\times n}$, let x be a unit eigenvector corresponding to λ. Now

$$Ax = \lambda x \Rightarrow A^p x = \lambda^p x.$$

By taking the norm of both sides we obtain:

$$\|A^p\|\|x\| \geq \|A^p x\| = \|\lambda^p x\| = |\lambda|^p \|x\|.$$

Now since $\|x\|$ is a unit vector, we have $\|A^p\|^{\frac{1}{p}} \geq |\lambda|$. In particular this is true for the largest eigenvalue λ_1. So $\rho(A) \leq \|A^p\|^{\frac{1}{p}}$. Now define

$$\tilde{A} = \frac{A}{\rho(A) + \epsilon}$$

for an arbitrary $\epsilon > 0$. Clearly $\rho(A) + \epsilon > 0$. Then we have

$$\rho(\tilde{A}) = \rho\left(\frac{A}{\rho(A)+\epsilon}\right) = \frac{\rho(A)}{\rho(A)+\epsilon} < 1$$

and $\|\tilde{A}^p\| \to 0$ as $p \to \infty$ by Theorem 2.4. So in particular there exists an integer $N(\epsilon, A)$ such that $\|\tilde{A}^p\| \leq 1$ whenever $p > N(\epsilon, A)$. This actually provides the upper bound we need, since

$$1 \geq \|\tilde{A}^p\| = \|\left(\frac{A}{\rho(A)+\epsilon}\right)^p\| = \frac{\|A^p\|}{(\rho(A)+\epsilon)^p}.$$

Thus $\|A^p\| \leq (\rho(A)+\epsilon)^p$ and by taking the p^{th} root we obtain

$$\|A^p\|^{\frac{1}{p}} \leq \rho(A) + \epsilon.$$

Now putting all of this together provides a nice small interval around $\|A^p\|^{\frac{1}{p}}$, that is,

$$\rho(A) \leq \|A^p\|^{\frac{1}{p}} \leq \rho(A) + \epsilon.$$

As $\epsilon \to 0$, then $p \to \infty$ and thus

$$\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A) = |\lambda_1(A)|.$$

□

In particular since $\|A\|_2 = \sigma_1(A)$ we have

$$\lim_{p\to\infty} [\sigma_1(A^p)]^{\frac{1}{p}} = |\lambda_1(A)|. \tag{4.1}$$

Our next step is to examine the case $k = n$ for

$$\lim_{p\to\infty}[\sigma_k(A^p)]^{\frac{1}{p}} = |\lambda_k(A)|. \tag{4.2}$$

When $k = n$ we consider two possibilities for the matrix A, the singular and nonsingular cases. Recall that when $k = n$, we are dealing with the smallest of the singular values and eigenvalue moduli of A.

Proof: of (4.2).

Case1: A is singular. So $\det A = \prod_{i=1}^{n}\lambda_i(A) = 0$. For each $p \in \mathbb{N}$ by Theorem 2.6 there are unitary matrices U and V such that

$$A^p = UDV = U\text{diag}\,(\sigma_1, \sigma_2, \ldots, \sigma_n)V.$$

So we have

$$0 = |\det A|^p = |\det A^p| = |\det(UDV)| = |\det U||\det D||\det V| = |\det D| = \prod_{i=1}^{n}\sigma_i(A^p).$$

So at least one of each of the eigenvalues and singular values is equal to zero, namely the smallest ones, $|\lambda_n(A)|$ and $\sigma_n(A^p)$, respectively. Then we have

$$\lim_{p\to\infty}[\sigma_n(A^p)]^{\frac{1}{p}} = |\lambda_n(A)| = 0.$$

Case 2: Now consider the case when A is nonsingular. Of course nonsingularity indicates that A has an inverse and we can use this to our advantage as follows:

$$Ax = \lambda x \Rightarrow \frac{1}{\lambda}x = A^{-1}x.$$

Then the smallest eigenvalue of A, with respect to modulus, $|\lambda_n|$, is the reciprocal of the largest eigenvalue of A^{-1} with respect to modulus, that is, $\frac{1}{|\lambda_n(A)|} = \rho(A^{-1})$. We also have $\sigma_1(A) = [\lambda_1(A^*A)]^{\frac{1}{2}}$ so that

$$\sigma_1(A^{-1}) = \left(\frac{1}{[\lambda_n(A^*A)]}\right)^{\frac{1}{2}} = \frac{1}{[\lambda_n(A^*A)]^{\frac{1}{2}}} = \frac{1}{\sigma_n(A)}.$$

By (4.1),

$$\frac{1}{\lim_{p\to\infty}[\sigma_n(A^p)]^{\frac{1}{p}}} = \lim_{p\to\infty}[\frac{1}{\sigma_n(A^p)}]^{\frac{1}{p}} = \lim_{p\to\infty}[\sigma_1(A^{-p})]^{\frac{1}{p}} = |\lambda_1(A^{-1})| = \frac{1}{|\lambda_n(A)|}.$$

By reciprocating each side of the equality we obtain

$$\lim_{p\to\infty}[\sigma_n(A^p)]^{\frac{1}{p}} = |\lambda_n(A)|. \tag{4.3}$$

□

So far we have established Yamamoto's result (1.1) when $k = 1$ and when $k = n$. To show the case for $1 < k < n$, Mathias employs Theorem 2.9 and Lemma 2.10 and uses the properties of principle submatrices. In particular there are two principle submatrices that we shall consider: $A_{[k]}$, the upper-left corner principal $k \times k$ submatrix of A and $A_{\langle k\rangle}$ lower-right corner principal $(n-k+1) \times (n-k+1)$ submatrix of A . The last case of his proof proceeds as follows:

Proof: By Theorem 2.5 there exist a unitary matrix U such that $U^*AU = T$ where T is an upper triangular matrix T having $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ on the diagonal. Then

$$U^*A^pU = T^p. \tag{4.4}$$

Since T is an upper triangular matrix,

$$(T_{[k]})^p = (T^p)_{[k]}, \qquad (T_{\langle k \rangle})^p = (T^p)_{\langle k \rangle}. \tag{4.5}$$

So by Theorem 2.9 and (4.5) we have

$$\sigma_k(A^p) = \sigma_k(T^p) \geq \sigma_k((T^p)_{[k]}) = \sigma_k((T_{[k]})^p). \tag{4.6}$$

Hence $\sigma_k(A^p)$ is bounded below by $\sigma_k((T_{[k]})^p)$.

In order to construct an upper bound for $\sigma_k(A^p)$ employ Theorem 2.9, Lemma 2.10, and (4.5) to obtain

$$\begin{aligned}
\sigma_k(A^p) &= \sigma_k(T^p) \qquad \text{(by unitary invariance (4.4))} \\
&= \sigma_{1+(k-1)}(T^p) \\
&\leq \sigma_1(L) \qquad \text{(by Theorem 2.9)} \\
&= \sigma_1((T^p)_{\langle k \rangle}) \qquad \text{(by Lemma 2.10)} \\
&= \sigma_1((T_{\langle k \rangle})^p) \qquad \text{(by (4.5))},
\end{aligned} \tag{4.7}$$

where L is the submatrix of T^p by deleting the first $k-1$ rows of T^p. This establishes the upper bound we were looking for. Putting (4.6) and (4.7) together we have

$$\sigma_k((T_{[k]})^p) \leq \sigma_k(A^p) \leq \sigma_1((T_{\langle k \rangle})^p).$$

By taking the p^{th} root we obtain

$$[\sigma_k((T_{[k]})^p)]^{\frac{1}{p}} \leq [\sigma_k(A^p)]^{\frac{1}{p}} \leq [\sigma_1((T_{\langle k \rangle})^p)]^{\frac{1}{p}}.$$

Taking the limit yields

$$\lim_{p\to\infty}[\sigma_k((T_{[k]})^p)]^{\frac{1}{p}} \leq \lim_{p\to\infty}[\sigma_k(A^p)]^{\frac{1}{p}} \leq \lim_{p\to\infty}[\sigma_1((T_{\langle k\rangle})^p)]^{\frac{1}{p}}. \tag{4.8}$$

By applying (4.3) on the principal submatrix $T_{[k]} \in \mathbb{C}_{k\times k}$ of T,

$$\lim_{p\to\infty}[\sigma_k((T_{[k]})^p)]^{\frac{1}{p}} = |\lambda_k(T_{[k]})|. \tag{4.9}$$

Likewise, applying (4.1) on the principal submatrix $T_{[k]} \in \mathbb{C}_{(n-k+1)\times(n-k+1)}$ of T

$$\lim_{p\to\infty}[\sigma_1((T_{\langle k\rangle})^p)]^{\frac{1}{p}} = |\lambda_1(T_{\langle k\rangle})|. \tag{4.10}$$

Then by putting together (4.8), (4.9), (4.10) we have bounds on the limit of $\sigma_k(A^p)$:

$$|\lambda_k(T_{[k]})| \leq \lim_{p\to\infty}[\sigma_k(A^p)]^{\frac{1}{p}} \leq |\lambda_1(T_{\langle k\rangle})|.$$

But $|\lambda_k(T_{[k]})| = |\lambda_k(T)| = |\lambda_k(A)|$ and $|\lambda_1(T_{\langle k\rangle})| = |\lambda_k(T)| = |\lambda_k(A)|$. Thus

$$|\lambda_k(A)| \leq \lim_{p\to\infty}[\sigma_k(A^p)^{\frac{1}{p}}] \leq |\lambda_k(A)|$$

Hence, $\lim_{p\to\infty}[\sigma_k(A^p)^{\frac{1}{p}}] = |\lambda_k(A)|$ for $1 \leq k \leq n$. □

Chapter 5

Comparing Approaches

About 23 years after T. Yamamoto introduced his proof that

$$\lim_{p\to\infty}[\sigma_i(A^p)]^{\frac{1}{p}} = |\lambda_i(A)|,$$

R. Mathias introduced a different proof. The proofs share some common characteristics but the main ingredient in each is different and that leads to several distinguishing characteristics. We will consider the development of the proofs by comparing the cases $k = 1$ and $1 \leq k \leq n$ for each author.

Initially both Yamamoto and Mathias use Gelfand's Spectral Radius Theorem. They offer proofs for Gelfand's result in which both make use of Theorem 2.4. Yamamoto leads into the theorem by introducing a lemma (3.1) which bounds $\|A^p\|^{\frac{1}{p}}$ by $\rho(A)$ and $\|A\|$. This is followed by the use of the closed and bounded interval to manipulate subsequences that show that

$$\lim_{p\to\infty} \|A^p\|]^{\frac{1}{p}} = \rho(A).$$

Mathias defines a matrix $\tilde{A}$ in order to bound $\|A^p\|^{\frac{1}{p}}$ below and above by $\rho(A)$ and $\rho(A)+\epsilon$ respectively. Then let $\epsilon \to 0$ so that $\lim_{p\to\infty} \|A^p\|^{\frac{1}{p}} = \rho(A)$. Both authors use properties of norms, and the fact that $\|A\| = \sigma_1(A)$. Since Mathias approach is very different from Yamamoto we will discuss each separately from this point forward.

For Yamamoto the procession from the $k = 1$ case to the $1 < k \leq n$ case is a natural step eased by the use of the compound matrix $C_k(A)$. Once the $k = 1$ case is applied to $C_k(A)$, only a few details need to be checked for the completion of the proof. If A is nilpotent ($|\lambda_i(A)| = 0$, $i = 1, \ldots, n$) and the $k = 1$ case $\lim_{p\to\infty}[\sigma_1(A^p)]^{\frac{1}{p}} = |\lambda_1(A)| = 0$ to obtain

$$\lim_{p\to\infty}[\sigma_j(A^p)]^{\frac{1}{p}} = |\lambda_j(A)| \ = 0$$

since the eigenvalues and singular values are ordered in non-increasing order. Now when A is not nilpotent with k nonzero eigenvalues, Yamamoto applies Schur's Triangularization Theorem (Theorem 2.5) and uses the rank argument to show that the number of nonzero singular values is at least k. Then by manipulation of the product of the first k singular values he is able to establish that

$$\lim_{p\to\infty}[\sigma_i(A^p)]^{\frac{1}{p}} = |\lambda_i(A)|, \quad i = 1, \ldots, n$$

Of course through this final step by necessity he divides the product up in order to isolate the j^{th} term. The fact that k of the singular values are nonzero guarantees the denominator is nonzero and then application of Corollary 3.2 transforms the notation to eigenvalues and the desired result falls out.

In comparison Mathias' method requires him to split up the case where $1 < k \leq n$. He establishes the $k = n$ case easily by considering the singular and nonsingular cases. Using the fact that for a singular matrix $\lambda_i(A) = 0$ and employing Singular Value Decomposition (Theorem 2.6) he gets that $\sigma_i(A^p) = 0$. The invertibility of a nonsingular matrix allows the use of the $k = 1$ result to be applied to A^{-1} to obtain the equality for the smallest eigenvalue and singular value. He continues with the $1 < k < n$ case only after finishing the $k = 1$ and $k = n$ cases. Now armed with the two limits obtained in the cases for $k = 1$ and $k = n$, Mathias incorporates the use of principal submatrices to show the desired result is true. Since Schur's Triangularization Theorem (Theorem 2.5) allows conversion of A^p into an upper triangular matrix through unitary similarity, he can simplify the process by engaging the Singular Value Interlacing Theorem (Theorem 2.9) as well as Lemma 2.10 to obtain upper and lower bounds on $[\sigma_i(A^p)]^{\frac{1}{p}}$. Having these bounds basically finishes the proof since the triangular form of T together with the application of the results for $k = 1$ and $k = n$ gives the desired result.

Clearly the tools needed to finish the proofs for the case(s) where $1 < k \leq n$ are varied for each approach. Yamamoto's use of compound matrices reduces the additional tools he is required to use to properties of nilpotent matrices and Schur's Triangularization Theorem. The approach Mathias chose to take requires more steps since he has separate cases for $k = n$ and $1 < k < n$. Also he needs to have at hand the benefits of Schur's Triangularization Theorem (2.5), Singular Value Decomposition (Theorem 2.5), properties of principal submatrices as well as the Singular Value Interlacing Theorem (Theorem 2.9). While the approach of R. Mathias is nice, the original method of Yamamoto is a more elegant approach, since it relies on only a few tools. Mathias' reasoning is easy to follow and understand, but requires a wider base of knowledge to establish the result.

Bibliography

[1] S. Friedberg, A. Insel, and L. Spence, Linear Algebra, 3^{rd} Ed. , Prentice Hall, New Jersey, 1997.

[2] P.R. Halmos, A Hilbert Space Problem Book, Springer-Verlag, New York, 1974.

[3] R. Mathias, Two theorems on singular values and eigenvalues, Amer. Math. Monthly, 97:47–50, 1990.

[4] R. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[5] R. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.

[6] P. Nylen and T.Y. Tam, On the spread of a Hermitian matrix and a conjecture of Thompson, Linear and Multilinear Algebra 37:3–11, 1994.

[7] T.Y. Tam and H. Huang, An extension of Yamamoto's theorem on the eigenvalues and singular values of a matrix, J. of Math. Soc. Japan, to appear.

[8] J. Todd, Survey of Numerical Analysis, McGraw-Hill, New York, 1962.

[9] T. Yamamoto, On the extreme values of the roots of matrices, J. Math. Soc. Japan, 19:173–178, 1967.

[10] F. Zhang, Matrix Theory: basic results and techniques, Springer, New York, 1999.