The QR Algorithm for Eigenvalue Estimation: Theory and Experiments

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.

Wei Feng

Certificate of Approval:

Tin-Yau Tam Professor Mathematics and Statistics	Thomas H. Pate, Chair Professor Mathematics and Statistics
John P. Holmes, III Professor Mathematics and Statistics	George T. Flowers Dean Graduate School

The QR Algorithm for Eigenvalue Estimation: Theory and Experiments

Wei Feng

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
December 19, 2008

The QR Algorithm for Eigenvalue Estimation: Theory and Experiments

Wei Feng

Permission is granted to Auburn University to make copies of this thesis at its discretion, upon the request of individuals or institutions and at their expense. The author reserves all publication rights.

____
Signature of Author

____
Date of Graduation

Vita

Wei Feng, only son and the third child of Zhiguo Feng and Shulan He, was born in Tangshan city, Hebei Province, China on Nov. 28, 1976. He graduated from the First High School of Luanxian in 1996 and was enrolled in Shanghai Jiao Tong University the same year, where he earned both his Bachelor of Engineering and Master of Science degrees in electrical engineering in 2000 and 2003, respectively. He then worked for Magima Inc. in 2003 for several months before he went to the United States with his beloved wife, Shuang Feng. He then entered the PhD program in electrical engineering of Auburn University in Aug. 2004, where he is expected to acquire the degree in Spring, 2009. He was accepted in the Master's program in mathematics of Auburn University in Feb. 2008. He and his wife shared the same high school, the same university for their undergraduate and graduate years in China, and the same university in the United States in Auburn. They are blessed with a lovely daughter, Grace T. Feng, who was born on May 19, 2005.

Thesis Abstract

The QR Algorithm for Eigenvalue Estimation: Theory and Experiments

Wei Feng

Master of Science, December 19, 2008
(M.E., Shanghai Jiao Tong University, 2003)
(B.S., Shanghai Jiao Tong University, 2000)

57 Typed Pages

Directed by Thomas H. Pate

In this thesis, we explore one of the most important numerical algorithms ever invented, the QR algorithm for computing matrix eigenvalues. First, we describe out notations and mathematical symbols used throughout the thesis in Chapter 1. Then we lay the ground work by stating and proving some basic lemmas in Chapter 2. Then in Chapter 3, we prove the convergence of the QR algorithm under the assumption of distinct magnitudes for all eigenvalues. This constraint is relaxed in Chapter 4, where we prove the convergence of the QR algorithm under the assumption of possibly equal magnitude eigenvalues. Finally, in Chapter 5, we present some numerical experiments to validate the conclusions drawn in this thesis.

Acknowledgments

First of all, I would like to thank Dr. Thomas H. Pate, my principle adviser, for his teaching and guidance throughout the whole process. What is not shown in this work is his meticulous yet picture clear methodology toward mathematics that I have witnessed and hopefully learned from. I am also thankful to Dr. Gary M. Sampson, Dr. Jo W. Heath, Dr. Narendra K. Govil, Dr. Amnon J. Meir and Dr. Stephen E. Stuckwisch, from whose teachings I have enjoyed mathematics. I also want to thank Dr. Tin-Yau Tam and Dr. John P. Holms, for sitting on my committee and proof-reading my thesis.

I also want to thank Dr. Thomas S. Denney Jr. and Dr. Stanley J. Reeves from the Electrical and Computer Engineering department not only for the financial support they granted me during the years I spent in Auburn, but also the teaching and guidance that have honed my problem solving skills. I also appreciate the interactions that I had with the students in the image processing group.

Finally I want to thank my family members for providing the love and support all the time. I want to thank my wife not only for her love and care, but also for her sharing the burden for the years. I want to thank my baby daughter for being the greatest and happiest gift to me. In the end, I want to thank my parents and my sisters, who have raised me and supported me unconditionally all my life.

Style manual or journal used Journal of Approximation Theory (together with the style known as "aums"). Bibliography 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

List of Figures

Chapter 1

Introduction

The eigenvalues of an $n \times n$ matrix A are the roots of its characteristic polynomial $\det(A - \lambda I)$. The famous Abel-Ruffini theorem states there is no algebraic formula for the roots of a general polynomial of degree five or higher. This means that the best method for computing the eigenvalues of a general $n \times n$ ($n \geq 5$) matrix is likely to be iterative in nature. The most famous eigenvalue algorithm is the QR algorithm discovered by Francis [3, 4] and Kublanovskaya [5] independently. A convergence analysis of the QR algorithm was given by Wilkinson [6]. A brief sketch of the early days history of the development of the QR algorithm was given by Parlett [13].

In this thesis, we present a complete detailed proof of the convergence of the QR algorithm under mild assumptions. First, the convergence is proved assuming that the magnitudes of all eigenvalues are distinct. Then this assumption is loosened such that the magnitudes of some eigenvalues may be equal, under which the convergence is re-defined and proved. Huang and Tam [14] proved the convergence of the QR algorithm for real matrices with non-real eigenvalues. Since the magnitudes of such matrices are distinct except for the conjugate pairs, they fit the loosened assumption.

We employ certain standard notations throughout this thesis. We let $\mathbb{C}^{m\times n}$ denote the set of $m \times n$ complex matrices. Similarly, $\mathbb{R}^{m\times n}$ denotes the set of all $m \times n$ real matrices. Both $A(i, j)$ and A_{ij} refer to the (i, j)th element of matrix A. On the other hand, $A = [a_{ij}]$ signifies that the (i, j)th element of A is a_{ij}. Given $A \in C^{m\times n}$, we let A^T denote the transpose of A and we let A^* denote the conjugate transpose of A. If A is square and invertible, then A^{-1} denotes the inverse of A. The $k \times k$ identity matrix is denoted as I_k. In circumstances where the dimension is unambiguous, the subscript of a matrix is dropped. For example, if confusion seems unlikely, we may write I instead of I_k. The unit vectors representing the coordinate system of dimension n is denoted as e_i, for all $i = 1, 2, \ldots, n$. Mathematically, $e_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T$, where the number 1 is the i-th element of e_i.

The notation $\operatorname{diag}([d_1, d_2, \cdots, d_n])$ represents the diagonal matrix with d_1, d_2, …, d_n on the diagonal in that order. Similarly, if $x = [x_1, x_2, \ldots, x_n]^T$ (or $x = [x_1, x_2, \ldots, x_n]$) is a vector, then the diagonal matrix with x on the diagonal is defined as $\operatorname{diag}(x) = \operatorname{diag}([x_1, x_2, \ldots, x_n])$. The notation Dg applied on an $n \times n$ matrix A is defined as $\operatorname{Dg}(A) = \operatorname{diag}([A(1,1), A(2,2), \ldots, A(n,n)])$. The zero scalar, vectors and matrices are all represented by 0 for simplicity. The distinction should be clear in context. The Kronecker delta function is denoted as δ_{ij} and $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$. The character λ is used exclusively to represent an eigenvalue.

For a scalar a, $|a|$ represents the absolute value of a and $\overline{a}$ denote the complex conjugate of a. For a matrix A, $|A|$ is the matrix whose elements are the absolute value of the corresponding elements of A. In other words, $(|A|)_{ij} = |A_{ij}|$. The Euclidean norm of a vector v in C^n or R^n is denoted by $\|v\|_2$. We reserve the notation $\|\!|\cdot|\!\|$ for matrix norms. For example, $\|\!|\cdot|\!\|_p$ is the induced p-norm of a square matrix for $p \geq 1$ defined as $\|\!|A|\!\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$. As a known fact, square matrix p-norms are sub-multiplicative; that is, if $A, B \in \mathbb{C}^{n \times n}$, then $\|\!|AB|\!\|_p \leq \|\!|A|\!\|_p \|\!|B|\!\|_p$. Specifically, the matrix 1-norm of a matrix $A \in \mathbb{C}^{n \times n}$ is

$$\|\!|A|\!\|_1 = \max_{j=1,\cdots,n} \sum_{i=1}^{n} |A(i,j)|. \tag{1.1}$$

The Frobenius norm of a matrix $A \in \mathbb{C}^{n \times n}$ is defined as

$$\|\!|A|\!\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |A(i,j)|^2}. \tag{1.2}$$

The limit of a sequence of matrices is a matrix consisted of the component-wise limits of the matrix elements. For example, if $A_k,\ k = 1, 2, \ldots$ are $n \times n$ complex matrices, then $\lim_{k\to\infty} A_k = 0$ means that $\lim_{k\to\infty} A_k(i, j) = 0$, for all $i, j = 1, 2, \ldots, n$.

The direct sum of matrices $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ is defined as $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, which is an $(m + n) \times (m + n)$ matrix.

Let $a, b \in \mathbb{C}^{n\times 1}$, then the inner product between a and b denoted by $\langle a, b\rangle$ is b^*a.

The projection of a vector a on the vector subspace W with respect to $\langle , \rangle$ is denoted as $\mathbb{P}_W(a)$.

Chapter 2

Useful Lemmas and Theorems

Lemma 2.1. *Suppose $A \in \mathbb{C}^{n\times n}$. If A is unitary and upper triangular, then A is diagonal.*

Proof. Let $A = [a_{ij}]$. Since A is unitary,

$$\begin{cases} \sum_{j=1}^{n} |a_{ij}|^2 = 1, \forall i = 1, \dots, n \\ \sum_{i=1}^{n} |a_{ij}|^2 = 1, \forall j = 1, \dots, n. \end{cases} \tag{2.1}$$

We prove by induction. The lemma is obviously true when $n = 1$. Assume that the lemma is true for $n = k$, $k \geq 1$. For $n = k + 1$, write

$$A = \begin{bmatrix} a_{11} & c \\ 0 & A_k \end{bmatrix}, \tag{2.2}$$

where A_k is $k \times k$ and c is $1 \times k$. Since A is unitary,

$$A^* A = I_{k+1}. \tag{2.3}$$

Plugging (2.2) into (2.3) we obtain

$$\begin{bmatrix} a_{11}^* & 0 \\ c^* & A_k^* \end{bmatrix} \begin{bmatrix} a_{11} & c \\ 0 & A_k \end{bmatrix} = \begin{bmatrix} |a_{11}|^2 & a_{11} c \\ a_{11} c^* & c^* c + A_k^* A_k \end{bmatrix} = I_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & I_k \end{bmatrix}.$$

From the above, we infer that $|a_{11}|^2 = 1$ and $a_{11} c = 0$. Thus, $a_{11} \neq 0$. Consequently c must be 0. This means that A_k must be unitary and upper triangular. By induction hypothesis, A_k is diagonal. We conclude that A is diagonal. □

Corollary 2.1. *Suppose $A \in \mathbb{C}^{n\times n}$. If A is unitary and upper triangular with positive diagonal elements, then $A = I_n$.*

Proof. By Lemma 2.1, A is diagonal. Since A is also unitary, we must have $|a_{ii}|^2 = 1$ for each i. But $a_{ii} > 0$ for each i. Therefore $a_{ii} = 1$ for each i. So A must be I_n. □

Lemma 2.2. *If $A \in \mathbb{C}^{n\times n}$ is upper triangular and invertible, then A^{-1} is upper triangular.*

Proof. Let $A = [a_1, a_2, \ldots, a_n]$ and $B = A^{-1} = [b_1^T, b_2^T, \ldots, b_n^T]^T$, where a_i, $i = 1, \ldots, n$ are the columns of A, and b_i, $i = 1, \ldots, n$, are the rows of B. Since A is upper triangular and invertible, $a_{ij} = 0$, for all $i > j$ and $a_{ii} \neq 0$ for all $i = 1, \ldots, n$. From $BA = I$, we get $b_i a_j = \delta_{ij}$. We use induction to show that $b_{ij} = 0$ when $i > j$.

For $j = 1$ and $i > 1$, $(BA)_{i1} = b_{i1}a_{11}$. Since $a_{11} \neq 0$, we must have $b_{i1} = 0$, for all $i > 1$. Now suppose that there exists a positive integer $1 \leq k < n$ such that for all $j \leq k$, $b_{ij} = 0$ for all $n \geq i > j \geq 1$. For $j = k+1$, and $i > k+1$, $(BA)_{ij} = (BA)_{i,k+1} = b_{i1}a_{1,k+1} + b_{i2}a_{2,k+1} + \cdots + b_{i,k+1}a_{k+1,k+1} = b_{i,k+1}a_{k+1,k+1} = 0$. But $a_{k+1,k+1} \neq 0$, so $b_{i,k+1} = 0$, for all $i > k+1$. □

Corollary 2.2. *If $A \in \mathbb{C}^{n\times n}$ is upper triangular with positive diagonal elements, then A^{-1} is also upper triangular with positive diagonal elements.*

Proof. Let $B = [b_{ij}] = A^{-1}$ where $A = [a_{ij}]$. By Lemma 2.2, B is upper triangular; moreover, for each i, $1 \leq i \leq n$, we have $(BA)_{ii} = b_{ii}a_{ii} = 1$. Therefore, $b_{ii} = 1/a_{ii} > 0$. □

Corollary 2.3. *If $A \in \mathbb{C}^{n\times n}$ is lower triangular and invertible, then A^{-1} is lower triangular.*

Proof. Since A is lower triangular, A^T is upper triangular. Thus $\left(A^T\right)^{-1}$ is upper triangular. But $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$. Hence A^{-1} is lower triangular. □

Lemma 2.3. *If R_1 and R_2 are both upper triangular matrices, then R_1R_2 is upper triangular.*

Proof. Let $R_1 = [a_{ij}]$ and $R_2 = [b_{ij}]$. Then for $i > j$, $(R_1R_2)_{ij} = \sum_{t=1}^{n} a_{it}b_{tj} = 0$ because it is always true that either $i > t$ or $t > j$. □

Definition 1. *A matrix* $B \in \mathbb{C}^{n\times n}$ *is said to be similar to* $A \in \mathbb{C}^{n\times n}$ *if there exists an invertible matrix* $S \in \mathbb{C}^{n\times n}$ *such that* $B = S^{-1}AS$. S *is called the similarity matrix of the similarity transformation* $A \to S^{-1}AS$.

Definition 2. *If the matrix* $A \in \mathbb{C}^{n\times n}$ *is similar to a diagonal matrix, then* A *is said to be diagonalizable. Furthermore, if the similarity matrix is unitary, then* A *is said to be unitarily diagonalizable.*

Lemma 2.4. *If* $A \in \mathbb{C}^{n\times n}$, *then* A *is diagonalizable if and only if there is a set of* n *linearly independent vectors, each of which is an eigenvector of* A.

Proof. Please refer to page 46 of [1] for the proof of this Lemma. □

Definition 3. *The matrix* $A \in \mathbb{C}^{n\times n}$ *is said to be normal if* $A^*A = AA^*$.

Lemma 2.5. *If* $A \in \mathbb{C}^{n\times n}$ *is normal, then* A *is unitarily diagonalizable.*

Proof. This is a result of Theorem 2.5.4 of [1]. □

Corollary 2.4. *Suppose* $A \in \mathbb{C}^{n\times n}$ *is normal and has* n *distinct eigenvalues. Let* $A = XDX^{-1}$ *be any diagonalization of* A. *Then the columns of* X *are mutually orthogonal.*

Proof. Since A is normal, by Lemma 2.5, A is unitarily diagonalizable. Thus there exists a unitary matrix U and a diagonal matrix D_1 such that $A = UD_1U^{-1}$. This is equivalent to $AU = UD_1$, or

$$A[u_1, u_2, \ldots, u_n] = [u_1, u_2, \ldots, u_n]\,\mathrm{diag}\left([d_{11}, d_{22}, \ldots, d_{nn}]\right), \tag{2.4}$$

where u_i is the i-th column vector of U and d_{ii} is the i-th diagonal element of D_1, for each $i = 1, 2, \ldots, n$. Since A has n distinct eigenvalues, $d_{ii} \neq d_{jj}$, for each $i \neq j$. Equation 2.4 means that u_i is the corresponding eigenvector of the eigenvalue d_{ii} of A. Furthermore, u_i, for all $i = 1, 2, \ldots, n$ are mutually orthogonal since U is unitary.

For any matrix X such that $A = XDX^{-1}$, all the columns of X have to be eigenvectors of A corresponding to different eigenvalues of A listed on the diagonal of D. Without loss of generality, we assume that $D = D_1$. Let the i-th column of X be x_i. Then x_i is an eigenvector of A corresponding to the eigenvalue d_{ii}. Thus we have $x_i = c_i u_i$, where c_i is a non-zero constant for each i. Since u_i, for all $i = 1, 2 \ldots, n$ are mutually orthogonal, x_i, for all $i = 1, 2, \ldots, n$ are also mutually orthogonal because

$$\langle x_i, x_j \rangle = \langle c_i u_i, c_j u_j \rangle = c_i \bar{c}_j \langle u_i, u_j \rangle = c_i \bar{c}_j \delta_{ij}.$$

□

Definition 4. *A function* $\|\cdot\| : \mathbb{C}^{n\times n} \to \boldsymbol{R}$ *is called a matrix norm if for all* $A, B \in \mathbb{C}^{n\times n}$ *it satisfies the following axioms:*

(1) $\|A\| \geq 0$,

(2) $\|A\| = 0$ *if and only if* $A = 0$,

(3) $\|cA\| = |c| \, \|A\|$ *for all complex scalars* c,

(4) $\|A + B\| \leq \|A\| + \|B\|$,

(5) $\|AB\| \leq \|A\| \|B\|$.

Lemma 2.6. *The Frobenius norm* $\|\cdot\|_F$ *is a matrix norm.*

Proof. Please refer to page 291 of [1]. □

Lemma 2.7. *If* $U \in \mathbb{C}^{n\times n}$ *is unitary, and* $A \in \mathbb{C}^{n\times n}$, *then* $\|UA\|_F = \|A\|_F$. *In other words, the Frobenius norm is invariant under unitary multiplication.*

Proof. By definition (1.2), it is easy to see that

$$\|A\|_F^2 = \operatorname{trace} A^* A.$$

Thus,

$$\begin{aligned}
\|UA\|_F^2 &= \operatorname{trace}(UA)^*(UA) \\
&= \operatorname{trace} A^*U^*UA \\
&= \operatorname{trace} A^*A \\
&= \|A\|_F^2.
\end{aligned}$$

□

Lemma 2.8. *Suppose $A_k \in \mathbb{C}^{n\times n}$, $k = 1, 2, \ldots$, is a sequence of matrices, then the following two conditions are equivalent,*

1. $\lim_{k\to\infty} \|A_k\|_F = 0$;
2. $\lim_{k\to\infty} A_k = 0$.

Proof. Assume that $\lim_{k\to\infty} \|A_k\|_F = 0$, then $\forall i, j \in \{1, 2, \ldots, n\}$,

$$\begin{aligned}
\lim_{k\to\infty} |A_k(i,j)| &\le \lim_{k\to\infty} \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |A_k(i,j)|^2} \\
&= \lim_{k\to\infty} \|A_k\|_F \\
&= 0.
\end{aligned}$$

Therefore, $\lim_{k\to\infty} A_k = 0$.

Now assuming that $\lim_{k\to\infty} A_k = 0$, we have

$$\begin{aligned}
\lim_{k\to\infty} \|A_k\|_F^2 &= \lim_{k\to\infty} \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |A_k(i,j)|^2\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\lim_{k\to\infty} |A_k(i,j)|^2\right) \\
&= 0.
\end{aligned}$$

□

Lemma 2.9. *Let t and n be positive integers so that $t \le n$. Suppose $Q_1, Q_2, Q_3, \ldots$ is a sequence of $n \times n$ unitary matrices. Let*

$$Q_k = \begin{bmatrix} Q_{11}^{(k)} & Q_{12}^{(k)} \\ Q_{21}^{(k)} & Q_{22}^{(k)} \end{bmatrix}, \tag{2.5}$$

where each $Q_{11}^{(k)}$ is a $t \times t$ matrix and each $Q_{22}^{(k)}$ is an $(n-t) \times (n-t)$ matrix. If $\lim_{k\to\infty} \| Q_{12}^{(k)} \|_F = 0$, then $\lim_{k\to\infty} \| Q_{21}^{(k)} \|_F = 0$.

Proof. Note that for each k we have

$$(n-t) \;=\; \|Q_{21}^{(k)}\|_F^2 + \|Q_{22}^{(k)}\|_F^2 \;=\; \|Q_{12}^{(k)}\|_F^2 + \|Q_{22}^{(k)}\|_F^2.$$

Thus, $\|Q_{21}^{(k)}\|_F^2 = \|Q_{12}^{(k)}\|_F^2$ for each positive integer k. That, $\lim_{k\to\infty} \|Q_{12}^{(k)}\|_F^2 = 0$, then implies that $\lim_{k\to\infty} \|Q_{21}^{(k)}\|_F^2 = 0$. □

Corollary 2.5. *Suppose $Q_1, Q_2, Q_3, \ldots$ is a sequence of $n \times n$ unitary matrices such that all the elements above the diagonal have limit zero, i.e., $\lim_{k\to\infty} Q_k(i,j) = 0$ for all $1 \le i < j \le n$. Then all the elements below the diagonal also have limit zero, i.e., $\lim_{k\to\infty} Q_k(i,j) = 0$ for all $1 \le j < i \le n$.*

Proof. For any i and j such that $1 \le j < i \le n$, let $Q_k = \begin{bmatrix} Q_{11}^{(k)} & Q_{12}^{(k)} \\ Q_{21}^{(k)} & Q_{22}^{(k)} \end{bmatrix}$, for all $k = 1, 2, \ldots$, where each $Q_{11}^{(k)}$ is a $j \times j$ matrix and each $Q_{22}^{(k)}$ is an $(n-j) \times (n-j)$ matrix. Then $Q_k(i,j)$ belongs to $Q_{21}^{(k)}$. By Lemma 2.9, since $Q_{12}^{(k)}$ has limit zero, $Q_{21}^{(k)}$ also has limit zero, which means that $Q_k(i,j)$ has limit zero for all $1 \le j < i \le n$. □

Lemma 2.10. *If $A_k \in \mathbb{C}^{n\times n}$, for $k = 1, 2, \ldots$ and $\lim_{k\to\infty} A_k = 0$, then for any $B, C \in \mathbb{C}^{n\times n}$, $\lim_{k\to\infty} B A_k C = 0$.*

Proof. Since $\lim_{k\to\infty} A_k = 0$, $\lim_{k\to\infty} A_k(i,j) = 0$, for all $i, j = 1, 2, \ldots, n$. Thus by the definition of matrix 1-norm (1.1),

$$\lim_{k\to\infty} \|A_k\|_1 = 0.$$

Using the sub-multiplicative property of the matrix 1-norm, we have

$$\|BA_kC\|_1 \le \|B\|_1\|A_k\|_1\|C\|_1.$$

Taking the limit as k goes to infinity on both sides, we get that $\lim_{k\to\infty}\|BA_kC\|_1 = 0$. □

Lemma 2.11. *Suppose $A_1, A_2, A_3, \ldots$ and $B_1, B_2, B_3, \ldots$ are sequences of $n \times n$ matrices such that $\lim_{k\to\infty} A_k = \tilde{A}$ and $\lim_{k\to\infty} B_k = \tilde{B}$. Then $\lim_{k\to\infty} A_kB_k = \tilde{A}\tilde{B}$.*

Proof. We use Frobenius norm to prove this lemma. First, $\|B_k\|_F = \|(B_k - \tilde{B}) + \tilde{B}\|_F \le \|(B_k - \tilde{B})\|_F + \|\tilde{B}\|_F$. Thus, $\lim_{k\to\infty}\|B_k\|_F = \|\tilde{B}\|_F$. Since the Frobenius norms of B_k converge, they are bounded. In other words, there exists a positive number M, such that $\|B_k\|_F < M$ for all $k = 1, 2, \ldots$.

Now we have

$$\begin{aligned}
\|A_kB_k - \tilde{A}\tilde{B}\|_F &= \|A_kB_k - \tilde{A}B_k + \tilde{A}B_k - \tilde{A}\tilde{B}\|_F \\
&\le \|(A_k - \tilde{A})B_k\|_F + \|\tilde{A}(B_k - \tilde{B})\|_F \\
&\le \|(A_k - \tilde{A})\|_F\|B_k\|_F + \|\tilde{A}\|_F\|B_k - \tilde{B}\|_F \\
&\le M\|(A_k - \tilde{A})\|_F + \|\tilde{A}\|_F\|B_k - \tilde{B}\|_F.
\end{aligned}$$

Thus, $\lim_{k\to\infty}\|A_kB_k - \tilde{A}\tilde{B}\|_F = 0$. According to Lemma 2.8, we have $\lim_{k\to\infty}(A_kB_k - \tilde{A}\tilde{B}) = 0$. Therefore, $\lim_{k\to\infty} A_kB_k = \tilde{A}\tilde{B}$. □

Lemma 2.12. *Suppose $Q_1, Q_2, \ldots$ is a sequence of $n \times n$ unitary matrices, and $R_1, R_2, \ldots$ is a sequence of $n \times n$ upper triangular matrices each of which has positive diagonal. If $\lim_{k\to\infty} Q_kR_k = I$, then $\lim_{k\to\infty} Q_k = I$, and $\lim_{k\to\infty} R_k = I$.*

Proof. Let $Q_k = [q_{ij}^{(k)}]$, and let $R_k = [r_{ij}^{(k)}]$ for each k. Note that $\lim_{k\to\infty} \|Q_k R_k - I\|_F = 0$. Moreover, because each R_k is upper triangular we have

$$\|Q_k R_k - I\|_F^2 \;=\; \|Q_k(R_k - Q_k^*)\|_F^2 \;=\; \|R_k - Q_k^*\|_F^2 \;\geq\; \sum_{i=2}^{n}\sum_{j=1}^{i-1} |q_{ji}^{(k)}|^2.$$

where the second equality follows by Lemma 2.7. Since $\lim_{k\to\infty} \|Q_k R_k - I\|_F^2 = 0$, we must have $\lim_{k\to\infty} q_{ji}^{(k)} = 0$ whenever $i > j$. This means that the upper triangular part excluding the diagonal of the Q_k has limit zero. By Corollary 2.5, the lower triangular part of Q_k excluding the diagonal also has limit zero. Then for each i, $1 \leq i \leq n$, we must have $\lim_{k\to\infty} |q_{ii}^{(k)}| = 1$, simply because each row and column of each Q_k has 2-norm equal to 1.

We now consider R_k. Since $\lim_{k\to\infty} \|R_k - Q_k^*\|_F^2 = 0$, and

$$\|R_k - Q_k^*\|_F^2 \geq \sum_{i\neq j} |r_{ij}^{(k)} - \bar{q}_{ji}^{(k)}|^2,$$

we have $\lim_{k\to\infty} \sum_{i\neq j} |r_{ij}^{(k)} - \bar{q}_{ji}^{(k)}|^2 = 0$. Thus, $\lim_{k\to\infty} |r_{ij}^{(k)} - \bar{q}_{ji}^{(k)}| = 0$ when $i \neq j$. But, we have $0 \leq |r_{ij}^{(k)}| \leq |r_{ij}^{(k)} - \bar{q}_{ji}^{(k)}| + |\bar{q}_{ji}^{(k)}|$ by the triangle inequality. Moreover, $\lim_{k\to\infty} q_{ji}^{(k)} = 0$ when $i \neq j$. Therefore we must have $\lim_{k\to\infty} r_{ij}^{(k)} = 0$ for each $i \neq j$.

Since $\lim_{k\to\infty} Q_k R_k = I$, we have $\lim_{k\to\infty} (Q_k R_k)_{ii} = 1$ for each i, $1 \leq i \leq n$. But,

$$(Q_k R_k)_{ii} \;=\; \sum_{t=1}^{n} q_{it}^{(k)} r_{ti}^{(k)} \;=\; q_{ii}^{(k)} r_{ii}^{(k)} + \sum_{t\neq i} q_{it}^{(k)} r_{ti}^{(k)}.$$

Since the off-diagonal parts of the Q_k and R_k both tend to 0 as k goes to infinity, we have

$$\lim_{k\to\infty} (Q_k R_k)_{ii} \;=\; \lim_{k\to\infty} q_{ii}^{(k)} r_{ii}^{(k)} \;=\; 1.$$

But, we have already shown that $\lim_{k\to\infty} |q_{ii}^{(k)}| = 1$ for each i. This and the fact that $\lim_{k\to\infty} q_{ii}^{(k)} r_{ii}^{(k)} = 1$ together imply that $\lim_{k\to\infty} |r_{ii}^{(k)}| = 1$ for each i. But, each of the R_k has positive diagonal. Therefore, $\lim_{k\to\infty} |r_{ii}^{(k)}| = \lim_{k\to\infty} r_{ii}^{(k)} = 1$. We have now shown that $\lim_{k\to\infty} R_k = I$. The final step is to show that $\lim_{k\to\infty} q_{ii}^{(k)} = 1$ for each i. This follows immediately from the fact that both $\lim_{k\to\infty} q_{ii}^{(k)} r_{ii}^{(k)} = 1$ and $\lim_{k\to\infty} r_{ii}^{(k)} = 1$. $\square$

Lemma 2.13. *Let $B_1, B_2, \ldots$ be a sequence of $n \times n$ matrices whose elements are bounded. Let $Q_1, Q_2, \ldots$ be a sequence of $n \times n$ unitary matrices such that $\lim_{k\to\infty} Q_k = I_n$. Let $A_k = Q_k^* B_k Q_k$, for all $k = 1, 2, \ldots$. Then $lim_{k\to\infty} (A_k - B_k) = 0$.*

Proof. Since the elements of B_k, $k = 1, 2, \ldots$ are bounded, there exists an $M > 0$, such that $|B_k(i,j)| \leq M$, for all $k = 1, 2, \ldots$, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$. Let $Q_k = [q_1^{(k)}, q_2^{(k)}, \ldots, q_n^{(k)}]$, where $q_i^{(k)}$ is the i-th column of Q_k, for all $i = 1, 2, \ldots, n$. Since $\lim_{k\to\infty} Q_k = I_n$, we have

$$\lim_{k\to\infty} q_i^{(k)} = e_i.$$

If we let $\delta_i^{(k)} = q_i^{(k)} - e_i$, for all $k = 1, 2, \ldots$ and $i = 1, 2, \ldots, n$, then $\lim_{k\to\infty} \delta_i^{(k)} = 0$. Thus,

$$\begin{aligned}
|A_k(i,j) - B_k(i,j)| &= \left| \left(q_i^{(k)}\right)^T B_k \left(q_j^{(k)}\right) - B_k(i,j) \right| \\
&= \left| \left(e_i + \delta_i^{(k)}\right)^T B_k (e_j + \delta_j^{(k)}) - B_k(i,j) \right| \\
&\leq \left| \left(\delta_i^{(k)}\right)^T B_k e_j \right| + \left| \left(\delta_i^{(k)}\right)^T B_k \delta_j^{(k)} \right| + \left| e_i^T B_k \delta_j^{(k)} \right| \\
&\leq \left| \left(\delta_i^{(k)}\right)^T \right| |B_k| \, e_j + \left| \left(\delta_i^{(k)}\right)^T \right| |B_k| \left| \delta_j^{(k)} \right| + e_i^T |B_k| \left| \delta_j^{(k)} \right| \\
&\leq M \left(\left| \left(\delta_i^{(k)}\right)^T \right| e_j + \left| \left(\delta_i^{(k)}\right)^T \right| \Xi \left| \delta_j^{(k)} \right| + e_i^T \left| \delta_j^{(k)} \right| \right),
\end{aligned}$$

where Ξ is the $n \times n$ matrix of which all elements are 1. Since

$$\lim_{k\to\infty} M \left(\left| \left(\delta_i^{(k)}\right)^T \right| e_j + \left| \left(\delta_i^{(k)}\right)^T \right| \Xi \left| \delta_j^{(k)} \right| + e_i^T \left| \delta_j^{(k)} \right| \right) = 0,$$

we get

$$\lim_{k\to\infty} (A_k(i,j) - B_k(i,j)) = 0.$$

□

Lemma 2.14. *If $A \in \mathbb{C}^{n\times n}$ is invertible, then A can be uniquely factorized as $A = QR$, where $Q \in \mathbb{C}^{n\times n}$ is unitary and $R \in \mathbb{C}^{n\times n}$ is upper triangular with positive diagonal elements.*

Proof. The proof is essentially a description of the Gram-Schmidt orthogonalization process. We give a sketch of that procedure. Denote $A = [a_1, a_2, \dots, a_n]$, where a_i, $i = 1, \dots, n$ are columns of A. Since A is invertible, $\{a_1, a_2, \dots, a_n\}$ are linearly independent. Let $\beta_1 = a_1/\|a_1\|_2$, or $a_1 = \|a_1\|_2\beta_1$. Then let $\beta_2 = (a_2 - \mathbb{P}_{W_1}(a_2))/\|a_2 - \mathbb{P}_{W_1}(a_2)\|_2$, where $W_1 = \text{span}\{\beta_1\}$. Note since $\{a_1, a_2, \dots, a_n\}$ are linearly independent, $a_2 - \mathbb{P}_{W_1}(a_2) \neq 0$. One gets $a_2 = \mathbb{P}_{W_1}(a_2) + \|a_2 - \mathbb{P}_{W_1}(a_2)\|_2\beta_2$. Notice that $\mathbb{P}_{W_1}(a_2) = \langle a_2, \beta_1\rangle\beta_1$. So $a_2 = \langle a_2, \beta_1\rangle\beta_1 + \|a_2 - \mathbb{P}_{W_1}(a_2)\|_2\beta_2$. In general, one gets $\beta_{i+1} = (a_{i+1} - \mathbb{P}_{W_i}(a_{i+1}))/\|a_{i+1} - \mathbb{P}_{W_i}(a_{i+1})\|_2$, where $W_i = \text{span}\{\beta_1, \beta_2, \dots, \beta_i\}$. Let $r_{11} = \|a_{11}\|_2$, $r_{ii} = \|a_i - \mathbb{P}_{W_{i-1}}(a_i)\|_2$, $r_{ij} = 0$, when $j < i$ and

$$r_{ij} = \langle a_j, \beta_i\rangle, mbox for all j > i. \tag{2.6}$$

Then

$$A = [a_1, a_2, \dots, a_n] = [\beta_1, \beta_2, \dots, \beta_n]\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ & r_{22} & \dots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}.$$

Let $Q = [\beta_1, \beta_2, \dots, \beta_n]$ and $R = [r_{ij}]$. We get $A = QR$, where Q is unitary by construction and R is invertible because $r_{ii} > 0$, $i = 1, 2, \dots, n$.

Now suppose that there exist two different factorizations with Q_1, Q_2, R_1 and R_2, where Q_1 and Q_2 are unitary and R_1 and R_2 are upper triangular with positive diagonal elements, such that $A = Q_1R_1 = Q_2R_2$. Then $Q_2^{-1}Q_1 = R_2^{-1}R_1$. So $R_2^{-1}R_1$ is upper triangular, unitary and has positive diagonal elements. By Corollary 2.1, $R_2^{-1}R_1 = I$.

Thus $R_1 = R_2$. Also, $Q_2^{-1}Q_1 = I$ leads to the conclusion that $Q_1 = Q_2$. This proves the uniqueness of the factorization.

□

Lemma 2.15. *If L is an $n \times n$ lower triangular matrix with unit diagonal, U is an upper triangular matrix, and P_1 and P_2 are permutation matrices such that $L = P_1UP_2$, then $P_2 = P_1^T$.*

Proof. Let $l_{i,j} = L(i,j)$, $u_{i,j} = U(i,j)$, $\alpha_{i,j} = P_1(i,j)$ and $\beta_{i,j} = P_2(i,j)$, for all $i, j \in \{1, 2, \ldots, n\}$. Elementwise,

$$l_{i,j} = [\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n}]U \begin{bmatrix} \beta_{1,j} \\ \beta_{2,j} \\ \vdots \\ \beta_{n,j} \end{bmatrix}$$

$$= u_{t_i,s_j},$$

where $t_i, s_j \in \{1, 2, \ldots, n\}$ are the indices of element 1 in the i-th row of P_1 and the j-th column of P_2, respectively, for all $i, j \in \{1, 2, \ldots, n\}$. Note that s and t are permutations on $\{1, 2, \ldots, n\}$. We focus on the cases when $i = j$, i.e., $1 = l_{i,i} = u_{t_i,s_i}$. Immediately we get that $t_i \leq s_i$ since U is upper triangular. Since this has to be true for all $i \in \{1, 2, \ldots, n\}$, we claim that $s_i = t_i$ for each $i \in \{1, 2, \ldots, n\}$. Indeed, Let i_1 be chosen such that $t_{i_1} = n$. Then $s_{i_1} = n$. Otherwise the element in U corresponding to the position (i_1, i_1) would be 0. Now choose i_2 such that $t_{i_2} = n - 1$. Then, s_{i_2} cannot be n because s is a permutation and $s_{i_1} = n$. Therefore, $s_{i_2} = n - 1$. Proceed for $t_{i_k} = n - k$, $k = 3, 4, \ldots, n - 1$, we get that $t_i = s_i$, for all $i \in \{1, 2, \ldots, n\}$. Remembering the definitions of t_i and s_i, we get that the t_i-th row of P_1 and the t_i-th column of P_2 are identical vectors, except the transposition, for all $t_i = 1, 2, \ldots, n$. In other words, $P_1(t_i, j) = P_2(j, t_i)$, for all $j = 1, 2, \ldots, n$. Thus, $P_2 = P_1^T$. □

Lemma 2.16. *If $A \in \mathbb{C}^{n\times n}$ and A is invertible, then A can be factored as $A = LPU$ where L is a unit lower triangular matrix, P is a permutation matrix, and U is an invertible upper triangular matrix. Furthermore, P is unique.*

Proof. Let $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Start from the first column of A, and find the first non-zero element (such an element must exist since A is invertible). Suppose this element is $a_{i,1}$, where $1 \le i \le n$. Let

$$L_1 = \begin{bmatrix} I_i & 0 \\ C & I_{n-i} \end{bmatrix},$$

where

$$C = \begin{bmatrix} 0 & \dots & 0 & c_1 \\ 0 & \dots & 0 & c_2 \\ & \vdots & & \\ 0 & \dots & 0 & c_{n-i} \end{bmatrix}$$

and $c_t = -\dfrac{a_{i+t,1}}{a_{i,1}}$. Then the first column of L_1A are all zeros except the i-th element, which remains $a_{i,1}$. This is essentially the Gaussian elimination applied on the first column of A.

Let

$$U_1 = \begin{bmatrix} a_{i,1}^{-1} & b_1 & b_2 & \dots & b_{n-1} \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & \vdots & & \ddots & \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

where $b_t = -\dfrac{a_{i,t+1}}{a_{i,1}}$, for all $t = 1, 2, \dots, n-1$. Then the i-th row of L_1AU_1 are all zero except the first element, which is normalized to 1. Note that the only change to the first column of L_1A by right multiplying U_1 is that its i-th element is normalized to 1. This in essence is the Gaussian elimination applied on the i-th row of L_1A.

Assume that the $a'_{j,2}$ is the first non-zero element in the second column of L_1AU_1. Let

$$L_2 = \begin{bmatrix} I_j & 0 \\ G & I_{n-j} \end{bmatrix},$$

where

$$G = \begin{bmatrix} 0 & \dots & 0 & g_1 \\ 0 & \dots & 0 & g_2 \\ & \vdots & & \\ 0 & \dots & 0 & g_{n-j} \end{bmatrix}$$

and $g_t = -\frac{a_{j+t,2}}{a_{j,2}}$. Then the second column of $L_2L_1AU_1$ are all zeros except the j-th element, which remains $a_{j,2}$. This is essentially the Gaussian elimination applied on the second column of L_1AU_1.

Let

$$U_2 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{j,2}^{-1} & h_1 & \dots & h_{n-2} \\ 0 & 0 & 1 & \dots & 0 \\ & \vdots & & \ddots & \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

where $h_t = -\frac{a_{j,t+2}}{a_{j,2}}$, for all $t = 1, 2, \dots, n-2$. Then the j-th row of $L_2L_1AU_1U_2$ are all zero except the first element, which is normalized to 1. This in essence is the Gaussian elimination applied on the j-th row of $L_2L_1AU_1$.

Repeat the above process for another $(n-2)$ times. Since A is invertible, we will get a permutation matrix. Mathematically, we express the process as

$$L_n L_{n-1} \cdots L_2 L_1 A U_1 U_2 \cdots U_n = P,$$

where L_k and U_k for any $k = 1, 2, \ldots, n$ are the unit lower triangular matrix and the upper triangular matrix used in the k-th iteration and P is the final permutation matrix. Let $L = L_n L_{n-1} \cdots L_1$ and $U = U_1 U_2 \cdots U_n$. Then L is unit lower triangular and U is upper triangular. Thus we have $LAU = P$. Multiplying L^{-1} on the left and U^{-1} on the right gives $A = L^{-1} P U^{-1}$. By Lemma 2.2 and Corollary 2.3, L^{-1} is unit lower triangular and U^{-1} is upper triangular. The existence of the LPU decomposition is proved.

Now, since A is invertible, U is invertible, which means all the diagonal elements of U are non-zero. Suppose there exist permutation matrices P_1, P_2 and lower and upper triangular matrices L_1, L_2 and U_1, U_2, such that $A = L_1 P_1 U_1 = L_2 P_2 U_2$. Then, $P_1 = L_1^{-1} L_2 P_2 U_2 U_1^{-1} = L P_2 U$, where $L = L_1^{-1} L_2$ is unit lower-triangular and $U_2 U_1^{-1}$ is upper triangular.

The condition $P_1 = L P_2 U$ is equivalent to $L = P_1 U^{-1} P_2^T = P_1 V P_2^T$, where $V = U^{-1}$. Since U is upper triangular, V is upper triangular. By Lemma 2.15, we get that $P_2^T = P_1^T$, i.e., $P_2 = P_1$.

This decomposition is called modified Bruhat decomposition [8].

□

Definition 5. *A multiset of cardinality n is a collection of n members where multiple presence of the same element is allowed and is counted as multiple members.*

A multiset is like a set, whose members are not ordered, but some or all of its members could be the same element. For example, the collection $\{1, 2, 1, 3\}$ is a multiset of cardinality 4. Furthermore, it is the same multiset as $\{2, 1, 1, 3\}$.

Definition 6. *The eigenvalues of $A \in \mathbb{C}^{n\times n}$, counting multiplicity, compose a multiset. We call it the eigenvalue multiset of A.*

Definition 7. *Let* $\Phi = \{\phi_1, \phi_2, \ldots, \phi_n\}$ *and* $\Psi = \{\psi_1, \psi_2, \ldots, \psi_n\}$ *be two multisets with complex elements. Then we define the distance* $d(\Phi, \Psi)$ *between multisets* Φ *and* Ψ *as*

$$d(\Phi, \Psi) = \min_{\sigma} \max_{i=1,\ldots,n} \left|\Phi_i - \Psi_{\sigma(i)}\right|. \tag{2.7}$$

where the minimum is taken over all permutations σ *of* $1, 2, \ldots, n$.

Theorem 2.1. *Let* $n \geq 1$ *and let*

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad a_n \neq 0$$

be a polynomial with complex coefficients. Then, for every $\epsilon > 0$*, there is a* $\delta > 0$ *such that for any polynomial*

$$q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$$

satisfying $b_n \neq 0$ *and*

$$\max_{0 \leq i \leq n} |a_i - b_i| < \delta,$$

we have

$$d(\Lambda, M) \leq \epsilon,$$

where multiset $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$ *contains all of the zeros of* $p(x)$*, multiset* $M = \{\mu_1, \ldots, \mu_n\}$ *contains all of the zeros of* $q(x)$*, both counting multiplicities.*

Proof. Please see [1] and [2] for the proof of this theorem. □

Theorem 2.2. *Suppose* $n \geq 1$ *and* $A, B \in \mathbb{C}^{n\times n}$*. Let* $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ *and* $\mu = \{\mu_1, \mu_2, \ldots, \mu_n\}$ *be the eigenvalue multisets of* A *and* B *respectively. Then for every* $\epsilon > 0$, *there exists a* $\delta > 0$*, such that if* $\max_{i,j=1,2,\ldots,n} |A(i,j) - B(i,j)| < \delta$*, then*

$$d(\lambda, \mu) \leq \epsilon. \tag{2.8}$$

Proof. The eigenvalues of A and B are the zeros of the corresponding characteristic polynomials $p_A(x) = \det(\lambda I - A)$ and $p_B(x) = \det(\lambda I - B)$. Let $a = \{a_n, a_{n-1}, \ldots, a_1, a_0\}$ and $b = \{b_n, b_{n-1}, \ldots, b_1, b_0\}$ be the coefficient vectors for p_A and p_B, i.e.,

$$p_A(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

and

$$p_B(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0.$$

According to Theorem 2.1, there is a $\delta' > 0$, so that if $\|a - b\|_\infty < \delta'$, then $d(\lambda, \mu) \leq \epsilon$.

Since a_i and b_i are polynomial functions of elements of A and B, they are continuous functions of elements of A and B. Hence there exists a $\delta > 0$, such that if $\max_{i,j=1,2,\ldots,n} |A(i,j) - B(i,j)| < \delta$, $\|a - b\|_\infty < \delta'$. Combining the above arguments, we finish the proof. □

Remark 1. *Theorem 2.2 illustrates the continuous dependence of the eigenvalues of a matrix on its elements.*

Definition 8. *Let* $\{\Phi_k$*:* Φ_k *is a multiset of* n *complex elements,* $k = 1, 2, \ldots\}$*. Let* Ψ *be a multiset of* n *complex elements. If* $\lim_{k\to\infty} d(\Phi_k, \Psi) = 0$*, we say that the sequence of multisets* $\{\Phi_k\}$ *converges to multiset* Ψ*.*

Lemma 2.17. *Suppose* $\{A_k \in \mathbb{C}^{n\times n} : k = 1, 2, \ldots\}$ *and* $B \in \mathbb{C}^{n\times n}$*. Let the eigenvalue multiset of* A_k *be* $\Gamma^{(k)} = \{\gamma_1^{(k)}, \gamma_2^{(k)}, \ldots, \gamma_n^{(k)}\}$*. Similarly, let the eigenvalue multiset of* B *be* $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$*. If* $\lim_{k\to\infty} A_k = B$*, then* $\{\Gamma^{(k)}\}$ *converge to* Λ*.*

Proof. The proof follows directly from Theorem 2.2. □

Chapter 3

Convergence of the QR Algorithm

If $A \in \mathbb{C}^{n\times n}$ has n distinct eigenvalues $\{\lambda_i : i = 1, \ldots, n\}$ where $|\lambda_i| > |\lambda_{i+1}|$ for each i such that $1 \leq i \leq n-1$ and $|\lambda_n| > 0$. Then A has n linearly independent eigenvectors. By Lemma 2.4, there exists an X such that $A = XDX^{-1}$, where $D = \operatorname{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n])$. Furthermore, by Lemma 2.16, X^{-1} can be factorized as $X^{-1} = LPU$, where L is a unit lower triangular matrix, P is a unique permutation matrix, and U is an invertible upper triangular matrix.

Theorem 3.1. *Suppose $A \in \mathbb{C}^{n\times n}$ has n distinct eigenvalues $\{\lambda_i : i = 1, \ldots, n\}$ where $|\lambda_i| > |\lambda_{i+1}|$ for each i such that $1 \leq i \leq n-1$ and $|\lambda_n| > 0$. There exists an invertible matrix X such that $A = XDX^{-1}$, where $D = \operatorname{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n])$. Let $X^{-1} = LPU$, where L is unit lower-triangular, P is a unique permutation matrix and U is upper triangular. Let $A = Q_1R_1$ be the unique QR factorization of A with R_1 having positive diagonal elements. Let $A_2 = R_1Q_1 = Q_2R_2$, where Q_2R_2 is the unique QR factorization of A_2. Repeat the above process so that for $k \geq 2$, $A_k = R_{k-1}Q_{k-1} = Q_kR_k$, where Q_kR_k is the unique QR factorization of A_k with R_k having positive diagonal elements. Then $\operatorname{Dg}(A_k)$ converges to $P^T \operatorname{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n])\, P$. Furthermore, as k goes to infinity, the elements in the lower triangular part of A_k go to zero and the elements in the upper triangular part of A_k converge in magnitude.*

Proof. By description, $A_k = R_{k-1}Q_{k-1} = Q_{k-1}^*Q_{k-1}R_{k-1}Q_{k-1} = Q_{k-1}^*A_{k-1}Q_{k-1}$, which means that A_k is similar to A_{k-1}, for each $k > 1$. As a result of this similarity, the matrices A_k all have the same characteristic polynomial and, hence, the same eigenvalues.

For each positive integer k, let

$$P_k = Q_1 Q_2 \cdots Q_k \tag{3.1}$$

and

$$U_k = R_k R_{k-1} \cdots R_1. \tag{3.2}$$

Note that by Lemma 2.3, U_k is upper triangular with positive diagonal elements. Also notice that

$$Q_{k-1} A_k = Q_{k-1} R_{k-1} Q_{k-1} = A_{k-1} Q_{k-1}. \tag{3.3}$$

Using (3.3), we compute the product $P_k U_k$ as

$$\begin{aligned} P_k U_k &= Q_1 Q_2 \cdots Q_k R_k \cdots R_1 \\ &= Q_1 \cdots Q_{k-1} A_k U_{k-1} \\ &= Q_1 \cdots Q_{k-2} A_{k-1} Q_{k-1} U_{k-1} \\ &= Q_1 \cdots Q_{k-3} A_{k-2} Q_{k-2} Q_{k-1} U_{k-1} \\ &= \cdots \\ &= A Q_1 Q_2 \cdots Q_{k-2} Q_{k-1} U_{k-1} \\ &= A P_{k-1} U_{k-1}. \end{aligned} \tag{3.4}$$

Since (3.4) is true for each $k \geq 2$, we have

$$\begin{aligned} P_k U_k &= A P_{k-1} U_{k-1} \\ &= A(A P_{k-2} U_{k-2}) \\ &= \cdots \\ &= A^{k-1} P_1 U_1 \\ &= A^k. \end{aligned}$$

Since $Q_k R_k$ is the unique QR-factorization of A_k as guaranteed by Lemma 2.14, $P_k U_k$ is the unique QR-factorization of A^k with U_k having positive diagonal elements.

Let $XP = QR$ be the unique QR factorization of XP such that R has positive diagonal elements. Thus, $X = QRP^T$. Now we have

$$\begin{aligned} A^k = X D^k X^{-1} = QRP^T D^k LPU \\ &= QR\left(P^T (D^k L D^{-k}) P\right) P^T D^k P U \\ &= QR\left(P^T (D^k L D^{-k}) P\right) D_p^k U, \end{aligned} \tag{3.5}$$

where $D_p = P^T D P$. Since D is an invertible diagonal matrix, $D^k L D^{-k}$ is still a unit lower-triangular matrix. Thus, when $1 \leq i < j \leq n$, $\left(D^k L D^{-k}\right)_{ij} = 0$. When $1 \leq j < i \leq n$, $\left(D^k L D^{-k}\right)_{ij} = l_{ij} \lambda_i^k / \lambda_j^k$ goes to zero as $k \to \infty$ since $|\lambda_i| < |\lambda_j|$. Since the diagonal elements of $P^T(D^k L D^{-k})P$ are the diagonal elements of $D^k L D^{-k}$, only rearranged by P, we have $\lim_{k\to\infty} D^k L D^{-k} = I$ and

$$\lim_{k\to\infty} P^T D^k L D^{-k} P = I.$$

Write $P^T D^k L D^{-k} P = I + E_k$, where $\lim_{k\to\infty} E_k = 0$. Plugging into Equation (3.5), we get $A^k = QR(I + E_k) D_p^k U = Q(I + R E_k R^{-1}) R D_p^k U$. By Lemma 2.10, $\lim_{k\to\infty} R E_k R^{-1} = 0$ and so $\lim_{k\to\infty}(I + R E_k R^{-1}) = I$. For each k, let $\tilde{Q}_k \tilde{R}_k$ be the unique QR factorization of $I + RE_kR^{-1}$ such that R has positive diagonal elements. Notice that by Lemma 2.12, $\lim_{k\to\infty} \tilde{Q}_k = \lim_{k\to\infty} \tilde{R}_k = I$. We get

$$A^k = (Q\tilde{Q}_k)(\tilde{R}_k R D_p^k U).$$

Now we focus on the diagonal elements of $\tilde{R}_k R D_p^k U$. Note that the diagonal elements of D_p are those of D rearranged by permutation P. Let $\lambda_{p,i} = D_p(i,i)$. Also let $U = [u_{ij}]$. The i-th diagonal element of $\tilde{R}_k R D_p^k U$ can be written as $(\tilde{R}_k)_{ii}(R)_{ii}\lambda_{p,i}^k u_{ii}$. By construction, $(\tilde{R}_k)_{ii} > 0$ and $(R)_{ii} > 0$ for each $i = 1, \ldots, n$. Let

$$\Lambda_k = \operatorname{diag}\left(\left[\frac{\overline{\lambda_{p,1}^k u_{11}}}{\left|\lambda_{p,1}^k\right| |u_{11}|}, \frac{\overline{\lambda_{p,2}^k u_{22}}}{\left|\lambda_{p,2}^k\right| |u_{22}|}, \ldots, \frac{\overline{\lambda_{p,n}^k u_{nn}}}{\left|\lambda_{p,n}^k\right| |u_{nn}|}\right]\right). \tag{3.6}$$

Note that Λ_k is a unitary diagonal matrix. Moreover, $\Lambda_k \tilde{R}_k R D_p^k U$ has positive diagonal elements, for its i-th diagonal element is $(\tilde{R}_k)_{ii}(R)_{ii} |\lambda_{p,i}|^k |u_{ii}|$. Since $Q\tilde{Q}_k\Lambda_k^*$ is unitary,

$$A^k = (Q\tilde{Q}_k\Lambda_k^*)(\Lambda_k \tilde{R}_k R D_p^k U) \tag{3.7}$$

is the unique QR factorization of A^k. But we have already shown that P_kU_k is the unique QR factorization of A^k for each k; therefore,

$$P_k = Q\tilde{Q}_k\Lambda_k^*, \tag{3.8}$$

and

$$U_k = \Lambda_k \tilde{R}_k R D_p^k U. \tag{3.9}$$

From (3.1) and (3.8),

$$\begin{aligned} Q_k = P_{k-1}^* P_k &= (Q\tilde{Q}_{k-1}\Lambda_{k-1}^*)^*(Q\tilde{Q}_k\Lambda_k^*) \\ &= \Lambda_{k-1}\tilde{Q}_{k-1}^*\tilde{Q}_k\Lambda_k^* \\ &= \Lambda_{k-1}\Lambda_k^*(\Lambda_k\tilde{Q}_{k-1}^*\tilde{Q}_k\Lambda_k^*). \end{aligned}$$

By Lemma 2.11, $(\tilde{Q}_{k-1}^*\tilde{Q}_k) \to I$, as $k \to \infty$ since each of the $\tilde{Q}_k$ converges to I. Now,

$$\begin{aligned} \|(\Lambda_k\tilde{Q}_{k-1}^*\tilde{Q}_k\Lambda_k^*) - I\|_F &= \left\|\left(\Lambda_k\left(\tilde{Q}_{k-1}^*\tilde{Q}_k - I\right)\Lambda_k^*\right)\right\|_F \\ &= \|\tilde{Q}_{k-1}^*\tilde{Q}_k - I\|_F. \end{aligned}$$

Thus, $\lim_{k\to\infty}\|(\Lambda_k\tilde{Q}_{k-1}^*\tilde{Q}_k\Lambda_k^*) - I\|_F = 0$. By Lemma 2.8, we get

$$\lim_{k\to\infty}(\Lambda_k\tilde{Q}_{k-1}^*\tilde{Q}_k\Lambda_k^*) = I.$$

Furthermore,

$$\begin{aligned} (\Lambda_{k-1}\Lambda_k^*)_{ii} &= \frac{\overline{\lambda_{p,i}^{k-1}u_{ii}}}{\left|\lambda_{p,i}^{k-1}\right||u_{ii}|}\frac{\lambda_{p,i}^{k}u_{ii}}{\left|\lambda_{p,i}^{k}\right||u_{ii}|} \\ &= \frac{\lambda_{p,i}}{|\lambda_{p,i}|}. \end{aligned}$$

Thus,

$$\lim_{k\to\infty} Q_k = \operatorname{diag}\left(\left[\frac{\lambda_{p,1}}{|\lambda_{p,1}|}, \frac{\lambda_{p,2}}{|\lambda_{p,2}|}, \ldots, \frac{\lambda_{p,n}}{|\lambda_{p,n}|}\right]\right). \tag{3.10}$$

In other words, Q_k converge to a unitary diagonal matrix.

Similarly, plugging (3.9) into $R_k = U_k U_{k-1}^{-1}$, we get

$$R_k = (\Lambda_k\tilde{R}_k R D_p^k U)(\Lambda_{k-1}\tilde{R}_{k-1}RD_p^{k-1}U)^{-1} = \Lambda_k\tilde{R}_k R D_p R^{-1}\tilde{R}_{k-1}^{-1}\Lambda_{k-1}^{-1}.$$

Again, focusing on the i-th diagonal elements, we have

$$(R_k)_{ii} = \frac{\overline{\lambda_{p,i}^k u_{ii}}}{|\lambda_{p,i}|^k}(\tilde{R}_k)_{ii} R_{ii} \lambda_{p,i} R_{ii}^{-1} (\tilde{R}_{k-1})_{ii}^{-1} \left(\frac{\overline{\lambda_{p,i}^{k-1} u_{ii}}}{|\lambda_{p,i}|^{k-1}} \right)^{-1}$$
$$= (\tilde{R}_k)_{ii} (\tilde{R}_{k-1})_{ii}^{-1} |\lambda_{p,i}|,$$

Hence,

$$\lim_{k\to\infty} (R_k)_{ii} = \lim_{k\to\infty} (\tilde{R}_k)_{ii} |\lambda_{p,i}| / (\tilde{R}_{k-1})_{ii} = |\lambda_{p,i}|. \tag{3.11}$$

This shows that the diagonal elements of the matrices R_k, $k = 1, 2, \ldots$, converge to the magnitudes of the eigenvalues of A.

Note that

$$\begin{aligned} A_{k+1} = P_k^* A P_k &= \left(Q \tilde{Q}_k \Lambda_k^* \right)^* A \left(Q \tilde{Q}_k \Lambda_k^* \right) \\ &= \Lambda_k \tilde{Q}_k^* Q^* X D X^{-1} Q \tilde{Q}_k \Lambda_k^* \\ &= \Lambda_k \tilde{Q}_k^* Q^* Q R P^T D P R^{-1} Q^* Q \tilde{Q}_k \Lambda_k^* \\ &= \Lambda_k \tilde{Q}_k^* R D_p R^{-1} \tilde{Q}_k \Lambda_k^* \\ &= \Lambda_k \tilde{Q}_k^* \Lambda_k^* (\Lambda_k R D_p R^{-1} \Lambda_k^*) \Lambda_k \tilde{Q}_k \Lambda_k^* \\ &= \hat{Q}_k (\Lambda_k R D_p R^{-1} \Lambda_k^*) \hat{Q}_k^*, \end{aligned} \tag{3.12}$$

where $\hat{Q}_k = \Lambda_k \tilde{Q}_k^* \Lambda_k^*$. Since $\lim_{k\to\infty} \tilde{Q}_k = I$ and $|(\Lambda_k)_{ii}| = 1$, $\lim_{k\to\infty} \hat{Q}_k = I$. Obviously the elements in the various $\Lambda_k R D_p R^{-1} \Lambda_k^*$ are uniformly bounded, by Lemma 2.13 , we have

$$\lim_{k\to\infty} \left(A_{k+1} - \Lambda_k R D_p R^{-1} \Lambda_k^* \right) = 0. \tag{3.13}$$

Hence, as k goes to infinity, the elements in the lower triangular part of A_{k+1} have limit zero, that is, $\lim_{k\to\infty} A_{k+1}(i,j) = 0$ for $1 \le j < i \le n$; and on the diagonal,

$$\lim_{k\to\infty} A_{k+1}(i,i) = (\Lambda_k R D_p R^{-1} \Lambda_k^*)_{ii} = \lambda_{p,i}, \tag{3.14}$$

which proves the convergence of the QR algorithm on the lower triangular and the diagonal.

Let $RD_pR^{-1} = W$. Then $W = [w_{ij}]$ is an upper triangular matrix. That is, $w_{ij} = 0$ when $1 \le j < i \le n$. Also let $\lambda_{p,i} = |\lambda_{p,i}|\, e^{\theta_i}$ and $u_{ii} = |u_{ii}|\, e^{\alpha_i}$, for each $i = 1, \ldots, n$. Thus

$$\begin{aligned}(\Lambda_k R D_p R^{-1} \Lambda_k^*)_{st} &= \frac{\overline{\lambda_{p,s}^k u_{ss}}}{|\lambda_{p,s}^k|\, |u_{ss}|} w_{st} \frac{\lambda_{p,t}^k u_{tt}}{|\lambda_{p,t}^k|\, |u_{tt}|} \\ &= w_{st} e^{jk(\theta_t - \theta_s) + \alpha_t - \alpha_s}..\end{aligned}$$

From Equation (3.13), for $s < t$,

$$\lim_{k\to\infty} \left((A_{k+1})_{st} - w_{st} e^{jk(\theta_t - \theta_s) + \alpha_t - \alpha_s} \right) = 0. \tag{3.15}$$

So $\lim_{k\to\infty} (A_{k+1})_{st}$ may not exist for $1 \le s < t \le n$. However, from Equation (3.15), we conclude that $\lim_{k\to\infty} |(A_{k+1})_{st}|$ exists and is equal to $|w_{st}|$.

Notice that if X^{-1} has LU decomposition, then $P = I$. In this case $\mathrm{Dg}\left(A^k\right)$ converges to $\mathrm{diag}\left([\lambda_1, \lambda_2, \ldots, \lambda_n]\right)$.

□

Corollary 3.1. *Given that A satisfies all the assumptions described in Theorem 3.1, suppose that A is also normal. Then not only the convergence of the QR algorithm described in Theorem 3.1 holds, but also the elements in the upper triangular part of A_k off the diagonal converge to zero.*

Proof. Since A is normal and has n distinct eigenvalues, by Lemma 2.4, for any diagonalization of A in the form of $A = XDX^{-1}$, the columns of X are mutually orthogonal. Hence $(XP)^*(XP) = P^T(X^*X)P$ is diagonal. This means that the columns of XP are mutually orthogonal as well. Recall that $XP = QR$ is the unique QR factorization of XP where R has positive diagonal elements. According to the Gram-Schmidt orthogonalization process described in the proof of Lemma 2.14, any element in the upper triangular part of R off the diagonal is described by Equation (2.6). The mutual orthogonality of column vectors of XP then means $r_{ij} = 0$, for all $j > i$. So R is diagonal.

From Equation 3.13, we have

$$\begin{aligned}\lim_{k\to\infty} |A_{k+1}| &= \lim_{k\to\infty} \left|\Lambda_k R D_p R^{-1} \Lambda_k^*\right| \\ &= \lim_{k\to\infty} \left|R D_p R^{-1}\right| \\ &= \left|R D_p R^{-1}\right|. \end{aligned} \tag{3.16}$$

Because R is diagonal, RD_pR^{-1} is diagonal, which is equivalent as saying that the upper triangular part of RD_pR^{-1} is zero. Hence

$$\lim_{k\to\infty} |A_{k+1}(i,j)| = 0, \text{ for all } 1 \le i < j \le n,$$

which means that

$$\lim_{k\to\infty} A_{k+1}(i,j) = 0, \text{ for all } 1 \le i < j \le n. \tag{3.17}$$

□

Chapter 4

Generalization: Equal-magnitude eigenvalues

In the proof given in Chapters 3, we assumed that the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of our $n \times n$ matrix A satisfy $|\lambda_{i+1}| > |\lambda_i|$, for each $i \in \{1, 2, \ldots, n-1\}$. In this section, we relax that assumption a bit and assume only that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n| \,.$$

In Chapter 3, in the expression $A = XDX^{-1}$ where D is a diagonal matrix with eigenvalues of A on the diagonal, no constraints on X^{-1} were imposed to guarantee convergence. However, in this chapter, extra constraints have to be assumed to guarantee convergence in general, as stated in the following theorem.

Theorem 4.1. *Suppose $A \in \mathbb{C}^{n \times n}$. Let $\{\lambda_i : i = 1, \ldots, n\}$ be the eigenvalues of A, counting multiplicity. Suppose $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n| > 0$. Let $D = \operatorname{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n])$. Suppose there exists invertible $X \in \mathbb{C}^{n \times n}$ such that $A = XDX^{-1}$. Let $X^{-1} = LPU$ be the modified Bruhat decomposition of X^{-1}, where L is unit lower-triangular, P is a unique permutation matrix and U is upper triangular. Assume that*

$$P = P_r \oplus P_{n-r}, \tag{4.1}$$

where P_r and P_{n-r} are permutation matrices of sizes $r \times r$ and $(n-r) \times (n-r)$. Let $\lambda_{p,i} = (P^T D P)_{i,i}$. Then in the QR algorithm iteration, the sequence of eigenvalue multisets of the top-left $r \times r$ blocks converge to $\{\lambda_{p,1}, \lambda_{p,2}, \ldots, \lambda_{p,r}\}$ and the sequence of eigenvalue multisets of the bottorm-right $(n-r) \times (n-r)$ blocks converge to $\{\lambda_{p,r+1}, \ldots, \lambda_{p,n}\}$.

Proof. We define P_k and U_k as in Chapter 3. See (3.1) and (3.2). Thus $A^k = P_k U_k$ is the unique QR-factorization of A^k where U_k has positive diagonal elements.

Let $XP = QR$ be the unique QR factorization of XP such that R has positive diagonal elements. Thus $X = QRP^T$. Similar to equation (3.5), we have

$$A^k = Q(RP^T D^k L D^{-k} P R^{-1}) R D_p^k U.$$

where $D_p = P^T D P$. Let

$$H_k = RP^T D^k L D^{-k} P R^{-1} \tag{4.2}$$

and let $H_k = \tilde{Q}_k \tilde{R}_k$ be the unique QR factorization of H_k such that $\tilde{R}_k$ has positive diagonal elements. Let

$$\tilde{Q}_k = \begin{bmatrix} \tilde{Q}_k^r & \tilde{Q}_k^{12} \\ \tilde{Q}_k^{21} & \tilde{Q}_k^{n-r} \end{bmatrix},$$
$$\tilde{R}_k = \begin{bmatrix} \tilde{R}_k^r & \tilde{R}_k^{12} \\ 0 & \tilde{R}_k^{n-r} \end{bmatrix},$$

and

$$H_k = \begin{bmatrix} H_k^r & H_k^{12} \\ H_k^{21} & H_k^{n-r} \end{bmatrix},$$

where H_k^r, $\tilde{Q}_k^r$ and $\tilde{R}_k^r$ are of size $r \times r$, H_k^{n-r}, $\tilde{Q}_k^{n-r}$ and $\tilde{R}_k^{n-r}$ are of size $(n-r) \times (n-r)$. Since $\tilde{R}_k$ is upper triangular, its inverse $\tilde{R}_k^{-1}$ is also upper triangular according to Lemma

2.2. Given $\tilde{R}_k$ in the above block form, we can easily show that

$$\tilde{R}_k^{-1} = \begin{bmatrix} (\tilde{R}_k^r)^{-1} & W_k \\ 0 & (\tilde{R}_k^{n-r})^{-1} \end{bmatrix} \tag{4.3}$$

where $W_k = -\left(\tilde{R}_k^r\right)^{-1} \tilde{R}_k^{12} \left(\tilde{R}_k^{n-r}\right)^{-1}$.

Let $F_k = D^k L D^{-k}$ and write F_k in block form as

$$F_k = \begin{bmatrix} F_{11}^{(k)} & 0 \\ F_{21}^{(k)} & F_{22}^{(k)} \end{bmatrix}, \tag{4.4}$$

where $F_{11}^{(k)}$ is of size $r \times r$ and $F_{22}^{(k)}$ is of size $(n-r) \times (n-r)$. Since $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n|$, we have $\lim_{k\to\infty} F_k(i,j) = \lim_{k\to\infty} l_{ij}\lambda_i^k/\lambda_j^k = 0$, for all $i \geq r+1 > r \geq j$. That is to say $\lim_{k\to\infty} F_{21}^{(k)} = 0$. Now let

$$P^T D^k L D^{-k} P = \begin{bmatrix} Z_{11}^{(k)} & Z_{12}^{(k)} \\ Z_{21}^{(k)} & Z_{22}^{(k)} \end{bmatrix}, \tag{4.5}$$

where $Z_{11}^{(k)}$ are of size $r \times r$ and $Z_{22}^{(k)}$ are of size $(n-r) \times (n-r)$. Plugging (4.1) and (4.4) into (4.5) and expand the multiplications, we get $Z_{21}^{(k)} = P_{n-r}^T F_{21}^{(k)} P_r$. Taking the limit as k goes to infinity, we get $\lim_{k\to\infty} Z_{21}^{(k)} = \lim_{k\to\infty} P_{n-r}^T F_{21}^{(k)} P_r = 0$. Furthermore, let

$$R_k = \begin{bmatrix} R_{11}^{(k)} & R_{12}^{(k)} \\ 0 & R_{22}^{(k)} \end{bmatrix}$$

where $R_{11}^{(k)}$ are of size $r \times r$ and $R_{22}^{(k)}$ are of size $(n-r) \times (n-r)$. Then we can write

$$R_k^{-1} = \begin{bmatrix} \left(R_{11}^{(k)}\right)^{-1} & R'_{12} \\ 0 & \left(R_{22}^{(k)}\right)^{-1} \end{bmatrix},$$

where $R'_{12} = -\left(R_{11}^{(k)}\right)^{-1} R_{12}^{(k)} \left(R_{22}^{(k)}\right)^{-1}$. From (4.2), we get

$$\lim_{k\to\infty} H_k^{21} = \lim_{k\to\infty} R_{22}^{(k)} Z_{21}^{(k)} \left(R_{11}^{(k)}\right)^{-1} = 0. \tag{4.6}$$

However, here $\lim_{k\to\infty} H_k = I$ may not be true.

Let

$$\Lambda_k = \operatorname{diag}\left(\left[\frac{\overline{\lambda_{p,1}^k u_{11}}}{\left|\lambda_{p,1^k}\right| |u_{11}|}, \frac{\overline{\lambda_{p,2}^k u_{22}}}{\left|\lambda_{p,2}^k\right| |u_{22}|}, \ldots, \frac{\overline{\lambda_{p,n}^k u_{nn}}}{\left|\lambda_{p,n}^k\right| |u_{nn}|}\right]\right),$$

and we get the following equation (in the same form as equation (3.7)),

$$A^k = (Q\tilde{Q}_k\Lambda_k^*)(\Lambda_k\tilde{R}_k R D_p^k U). \tag{4.7}$$

Again, since $Q\tilde{Q}_k\Lambda_k^*$ is unitary and $\Lambda_k\tilde{R}_k R D_p^k U$ is upper triangular with positive diagonal elements, we recognize that (4.7) is the unique QR factorization of A^k. Thus,

$$P_k = Q\tilde{Q}_k\Lambda_k^*, \tag{4.8}$$

and

$$U_k = \Lambda_k\tilde{R}_k R D_p^k U. \tag{4.9}$$

From $H_k = \tilde{Q}_k \tilde{R}_k$, we get $\tilde{R}_k^{-1} = H_k^{-1} \tilde{Q}_k$. Applying the Frobenius norm,

$$\begin{aligned}
\|\tilde{R}_k^{-1}\|_F &= \|H_k^{-1}\tilde{Q}_k\|_F \\
&= \|H_k^{-1}\|_F, \qquad \text{by Lemma 2.7} \\
&= \|\left(RP^T D^k L D^{-k} P R^{-1}\right)^{-1}\|_F \\
&= \|RP^T D^k L^{-1} D^{-k} P R^{-1}\|_F \\
&\leq \|RP^T\|_F \|D^k L^{-1} D^{-k}\|_F \|PR^{-1}\|_F \\
&= C_{R,P} \|D^k L^{-1} D^{-k}\|_F \\
&\leq C_{R,P} \|L^{-1}\|_F
\end{aligned} \tag{4.10}$$

where $C_{R,P} = \|RP^T\|_F \|PR^{-1}\|_F$ is a constant. The last inequality holds because

$$\left| \left(D^k L^{-1} D^{-k}\right)_{i,j} \right| \leq \left|L^{-1}(i,j)\right|, \text{ for all } i, j = 1, 2, \ldots, n.$$

On account of (4.10), we have,

$$\|(\tilde{R}_k^r)^{-1}\|_F \leq \|\tilde{R}_k^{-1}\|_F \leq C_{R,P} \|L^{-1}\|_F. \tag{4.11}$$

Thus,

$$\|\tilde{Q}_k^{21}\|_F \leq \|H_k^{21}\|_F \|(\tilde{R}_k^r)^{-1}\|_F \leq C_{R,P} \|L^{-1}\|_F \|H_k^{21}\|_F \to 0, \text{ as } k \to \infty \tag{4.12}$$

since $\|H_k^{21}\|_F \to 0$ by (4.6). Therefore, $\lim_{k\to\infty} \tilde{Q}_k^{21} = 0$.

Furthermore, by Lemma 2.9, we get $\lim_{k\to\infty} \tilde{Q}_k^{12} = 0$. Consequently, $\lim_{k\to\infty} \tilde{Q}_k^r \tilde{Q}_k^{r*} = I_r$ and $\lim_{k\to\infty} \tilde{Q}_k^{n-r} \tilde{Q}_k^{n-r*} = I_{n-r}$. In conclusion, we showed that

$$\lim_{k\to\infty} \left(\tilde{Q}_k - \begin{bmatrix} \tilde{Q}_k^r & 0 \\ 0 & \tilde{Q}_k^{n-r} \end{bmatrix} \right) = 0.$$

Notice that the same computation in (3.12) applies to the factorization shown in (4.7). We get

$$A_{k+1} = \hat{Q}_k(\Lambda_k R D_p R^{-1} \Lambda_k^*)\hat{Q}_k^*$$

where $\hat{Q}_k = \Lambda_k \tilde{Q}_k^* \Lambda_k^*$. If we denote $\Lambda_k = \begin{bmatrix} \Lambda_k^r & 0 \\ 0 & \Lambda_k^{n-r} \end{bmatrix}$ and $\hat{Q}_k = \begin{bmatrix} \hat{Q}_k^r & \hat{Q}_k^{12} \\ \hat{Q}_k^{21} & \hat{Q}_k^{n-r} \end{bmatrix}$, then

$$\begin{aligned} \hat{Q}_k &= \begin{bmatrix} \Lambda_k^r & 0 \\ 0 & \Lambda_k^{n-r} \end{bmatrix} \begin{bmatrix} \tilde{Q}_k^r & \tilde{Q}_k^{12} \\ \tilde{Q}_k^{21} & \tilde{Q}_k^{n-r} \end{bmatrix} \begin{bmatrix} \Lambda_k^{r*} & 0 \\ 0 & \Lambda_k^{n-r*} \end{bmatrix} \\ &= \begin{bmatrix} \Lambda_k^r \tilde{Q}_k^r \Lambda_k^{r*} & \Lambda_k^r \tilde{Q}_k^{12} \Lambda_k^{n-r*} \\ \Lambda_k^{n-r} \tilde{Q}_k^{21} \Lambda_k^{r*} & \Lambda_k^{n-r} \tilde{Q}_k^{n-r} \Lambda_k^{n-r*} \end{bmatrix}. \end{aligned}$$

From above, we get $\hat{Q}_k^{12} = \Lambda_k^r \tilde{Q}_k^{12} \Lambda_k^{n-r*} \to 0$ and $\hat{Q}_k^{21} = \Lambda_k^{n-r} \tilde{Q}_k^{21} \Lambda_k^{r*} \to 0$, as $k \to \infty$. Moreover,

$$\lim_{k \to \infty} \hat{Q}_k^r = I_r, \tag{4.13}$$

and

$$\lim_{k \to \infty} \hat{Q}_k^{n-r} = I_{n-r}. \tag{4.14}$$

Let

$$J_k = \Lambda_k R D_p R^{-1} \Lambda_k^*. \tag{4.15}$$

Then J_k is upper triangular with $J_k(i,i) = \lambda_{p,i}$, $i = 1, \ldots, n$. If we write $J_k = \begin{bmatrix} J_k^r & J_k^{12} \\ 0 & J_k^{n-r} \end{bmatrix}$, then,

$$
\begin{aligned}
& A_{k+1} \\
&= \hat{Q}_k J_k \hat{Q}_k^* \\
&= \begin{bmatrix} \hat{Q}_k^r & \hat{Q}_k^{12} \\ \hat{Q}_k^{21} & \hat{Q}_k^{n-r} \end{bmatrix} \begin{bmatrix} J_k^r & J_k^{12} \\ 0 & J_k^{n-r} \end{bmatrix} \begin{bmatrix} \hat{Q}_k^{r*} & \hat{Q}_k^{21*} \\ \hat{Q}_k^{12*} & \hat{Q}_k^{n-r*} \end{bmatrix} \\
&= \begin{bmatrix} \hat{Q}_k^r J_k^r \hat{Q}_k^{r*} + (\hat{Q}_k^r J_k^{12} + \hat{Q}_k^{12} J_k^{n-r}) \hat{Q}_k^{12*} & \hat{Q}_k^r J_k^r \hat{Q}_k^{21*} + (\hat{Q}_k^r J_k^{12} + \hat{Q}_k^{12} J_k^{n-r}) \hat{Q}_k^{n-r*} \\ \hat{Q}_k^{21} J_k^r \hat{Q}_k^{r*} + (\hat{Q}_k^{21} J_k^{12} + \hat{Q}_k^{n-r} J_k^{n-r}) \hat{Q}_k^{12*} & \hat{Q}_k^{21} (J_k^r \hat{Q}_k^{21*} + J_k^{12} \hat{Q}_k^{n-r*}) + \hat{Q}_k^{n-r} J_k^{n-r} \hat{Q}_k^{n-r*} \end{bmatrix}.
\end{aligned}
$$

If we write $A_{k+1} = \begin{bmatrix} A_{k+1}^r & A_{k+1}^{12} \\ A_{k+1}^{21} & A_{k+1}^{n-r} \end{bmatrix}$, then as $k \to \infty$, for the lower left block A_{k+1}^{21}, we have

$$
\begin{aligned}
\lim_{k\to\infty} A_{k+1}^{21} &= \lim_{k\to\infty} \hat{Q}_k^{21} J_k^r \hat{Q}_k^{r*} + (\hat{Q}_k^{21} J_k^{12} + \hat{Q}_k^{n-r} J_k^{n-r}) \hat{Q}_k^{12*} \\
&= 0
\end{aligned}
$$

for the following reasons. By (4.15), J_k is uniformly bounded, hence J_k^r, J_k^{12} and J_k^{n-r} are all uniformly bounded. Furthermore, both $\hat{Q}_k^{r*}$ and $\hat{Q}_k^{n-r}$ are uniformly bounded because they are part of a unitary matrix. Finally both $\hat{Q}_k^{21}$ and $\hat{Q}_k^{12}$ go to zero as k goes to infinity.

For the top right block A_{k+1}^{12}, we have

$$
\begin{aligned}
& \lim_{k\to\infty} \left(A_{k+1}^{12} - \hat{Q}_k^r J_k^{12} \hat{Q}_k^{n-r*} \right) \\
&= \lim_{k\to\infty} \left(\left(\hat{Q}_k^r J_k^r \hat{Q}_k^{21*} + (\hat{Q}_k^r J_k^{12} + \hat{Q}_k^{12} J_k^{n-r}) \hat{Q}_k^{n-r*} \right) - \hat{Q}_k^r J_k^{12} \hat{Q}_k^{n-r*} \right) \\
&= 0, \qquad (4.16)
\end{aligned}
$$

for the same reasons stated above for A^{21}_{k+1}. Furthermore, by (4.13), (4.14) and (4.15), we get

$$\begin{aligned}
\lim_{k\to\infty}\left|A^{12}_{k+1}\right| &= \lim_{k\to\infty}\left|\hat{Q}^{r}_{k}J^{12}_{k}\hat{Q}^{n-r*}_{k}\right| \\
&= \lim_{k\to\infty}\left|J^{12}_{k}\right| \\
&= \lim_{k\to\infty}\left|\left(\Lambda_k R D_p R^{-1}\Lambda^*_k\right)^{12}\right| \\
&= \lim_{k\to\infty}\left|\left(R D_p R^{-1}\right)^{12}\right|,
\end{aligned} \tag{4.17}$$

where $\left(\Lambda_k R D_p R^{-1}\Lambda^*_k\right)^{12}$ and $\left(R D_p R^{-1}\right)^{12}$ represent the top right $r \times (n-r)$ blocks of $\Lambda_k R D_p R^{-1}\Lambda^*_k$ and $R D_p R^{-1}$, respectively. Equation 4.17 shows that the elements of A^{12}_{k+1} converge in magnitude to those of the corresponding top right block of $R D_p R^{-1}$, which is a fixed matrix. Note that this is a similar result of (3.15), except that (3.15) applies to all the upper triangular elements of A_{k+1} off the diagonal while (4.17) only applies to the upper right block off the diagonal blocks of A_{k+1}.

We also want to point out that a similar result to Corollary 3.1 exists for the equal-magnitude eigenvalue case. If all the eigenvalues of A are distinct and A is normal, then by the same reasoning presented in Corollary 3.1, R is diagonal. Hence $\left(R D_p R^{-1}\right)^{12} = 0$. Thus, $\lim_{k\to\infty}\left|A^{12}_{k+1}\right| = 0$, which means that

$$\lim_{k\to\infty} A^{12}_{k+1} = 0. \tag{4.18}$$

For the two diagonal blocks, we have

$$\lim_{k\to\infty}\left(A^{r}_{k+1} - \hat{Q}^{r}_{k}J^{r}_{k}\hat{Q}^{r*}_{k}\right) = \lim_{k\to\infty}\left((\hat{Q}^{r}_{k}J^{12}_{k} + \hat{Q}^{12}_{k}J^{n-r}_{k})\hat{Q}^{12*}_{k}\right) = 0 \tag{4.19}$$

and

$$\lim_{k\to\infty}\left(A^{n-r}_{k+1} - \hat{Q}^{n-r}_{k}J^{n-r}_{k}\hat{Q}^{n-r*}_{k}\right) = \lim_{k\to\infty}\left(\hat{Q}^{21}_{k}J^{r}_{k}\hat{Q}^{21*}_{k} + \hat{Q}^{21}_{k}J^{12}_{k}\hat{Q}^{n-r*}_{k}\right) = 0. \tag{4.20}$$

Remember that in Chapter 3, the whole lower triangular block excluding the diagonal goes to zero as k goes to infinity. Here for the equal eigenvalue magnitude case, under the assumption of (4.1), only the lower left block of size $(n-r) \times r$ goes to zero as k goes to infinity. The two diagonal blocks A^r_{k+1} and A^{n-r}_{k+1} does not converge in a conventional sense. However, using Lemma 2.17, we conclude that the sequence of eigenvalue multisets of $\{A^r_k : k = 1, 2, \ldots,\}$ converge to the multiset $\{\lambda_{p,1}, \lambda_{p,2}, \ldots, \lambda_{p,r}\}$ and the sequence of eigenvalue multisets of $\{A^{n-r}_k : k = 1, 2, \ldots,\}$ converge to the multiset $\{\lambda_{p,r+1}, \ldots, \lambda_{p,n}\}$.

Note that in (4.1), if $P_r = I_r$ and $P_{n-r} = I_{n-r}$, then $P = I$. This means that X^{-1} has LU decomposition $X^{-1} = LU$.

□

Chapter 5

Experiments

To show the convergence of the QR algorithm that we have proved in Chapters 3 and 4, we have designed several experiments which we performed using MATLAB (The Mathworks, Natick, MA). The symbols used in this chapter will be the same symbols used in Chapters 3 and 4.

We choose the dimension of the matrices to be $n = 5$. First, we construct a random unit lower triangular matrix L and a random upper triangular matrix U. To guarantee numerical stability, we constrained the 2-norm condition numbers of both L and U to be no more than 100.

To illustrate Theorem 3.1, we choose the magnitude of the eigenvalues of the the matrix A to be $\lambda_i = 2n+1-i$, for all $i = 1, 2, ...n$. Their phases are generated randomly in the range of 0 to 2π. Using these eigenvalues, we form the diagonal matrix $D = \operatorname{diag}([\lambda_1, \lambda_2, \ldots, \lambda_n])$.

In the first experiment, we form the matrix X^{-1} by letting $X^{-1} = LU$. The matrix A is formed by $A = X \Lambda X^{-1}$. Then we performed the QR algorithm on A for 800 iterations. The results are shown in Figures 5.1, 5.2 and 5.3.

Figure 5.1 shows the convergence of the lower triangular part of A_k off the diagonal. The curve represents the evolution of the maximum absolute value of all the lower triangular off diagonal elements of A_k with the iterations.

Figure 5.2 shows the convergence of the diagonal elements of A_k in the complex plane. The trajectory of the diagonal elements were plotted with the iterations. The triangles represent the starting point of each diagonal element (Note that the diagonal elements in the first several iterations tend to be far from the final converging value. In order to show more detail, we chose to start the plot from the 7-th iteration). The circles represent the

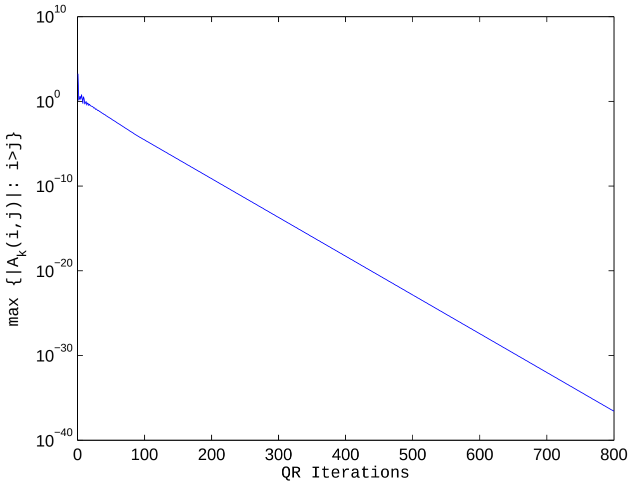

Figure 5.1: Lower triangular (off diagonal) part of A_k converge to zero

ending points of the trajectories. The "+" signs mark the real eigenvalues of A. From this figure, we can see that the diagonal elements of A_k converge to the eigenvalues of A.

The evolution of four randomly selected upper triangular elements of A_k are shown in Fig. 5.3. On the top row, the magnitudes of these elements are shown against iterations. On the bottom row, their trajectories in the complex plane are plotted. Again triangles and circles represent beginning and ending points of the trajectories. One can see that these upper triangular elements of A_k converge in magnitude (top row) but do not converge in value (bottom row).

In the second experiment, we generate a random permutation matrix P. We relax the constraint such that $X^{-1} = LPU$. The results are shown in Figures 5.4, 5.5 and 5.6. These figures show similar convergence results of the lower triangular, diagonal and upper triangular parts of A_k. Compared to the Figs. 5.1, 5.2 and 5.3, one can see that there are some more oscillations presented in the LPU case than the LP case. Also, notice that in Fig. 5.5, the trajectories of $A_k(1,1)$ and $A_k(2,2)$ trade places with each other during the

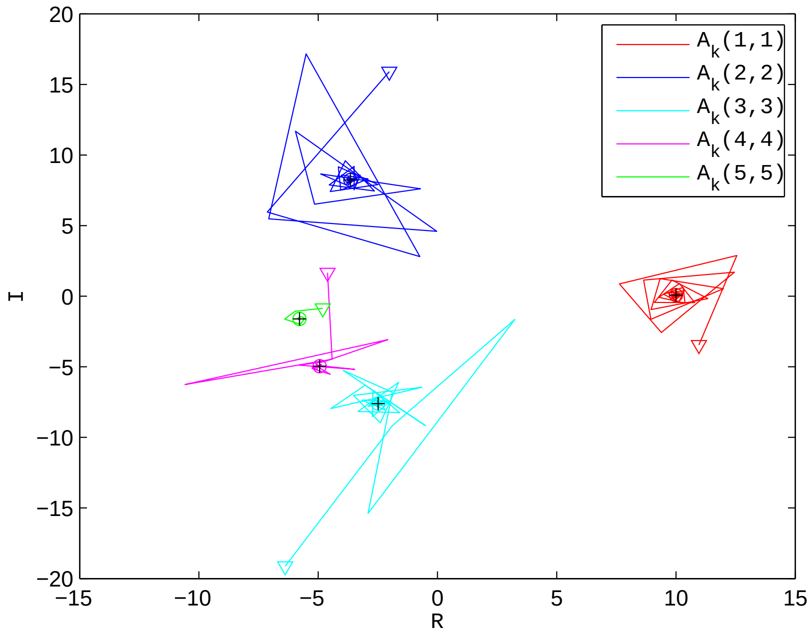

Figure 5.2: Diagonal elements of A_k converge to eigenvalues of A

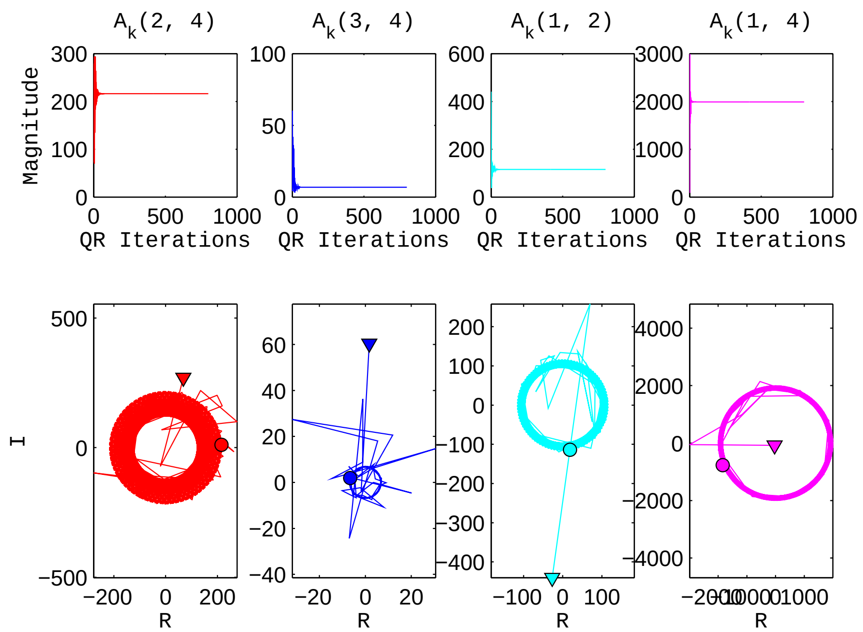

Figure 5.3: Upper triangular (off diagonal) part of A_k converge in magnitude

iterations. This actually reflects the permutation matrix P. In this case,

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

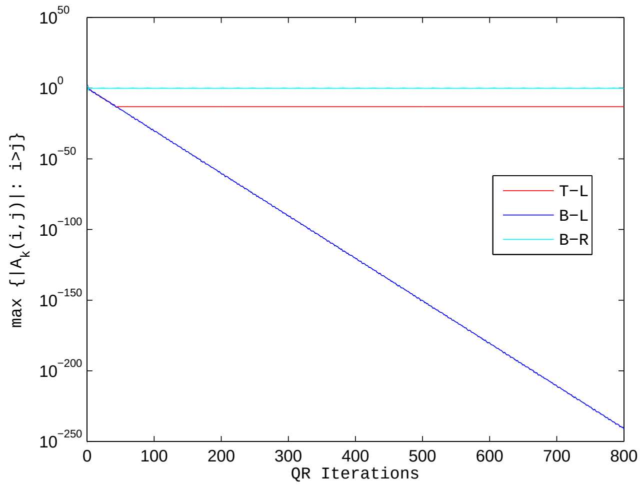

Figure 5.4: Lower triangular (off diagonal) part of A converge to zero

The above experiments validate the QR algorithm that we presented in Chapter 3.

Then we changed the eigenvalues to validate Theorem 4.1 presented in Chapter 4. The new eigenvalues are divided into two groups, the first group of 2 eigenvalues λ_1 and λ_2 have the same magnitude of 2, but with random phases. The second group of 3 eigenvalues λ_3, λ_4 and λ_5 have the same magnitude of 1, again with random phases. The diagonal matrix D is then formed by $D = \text{diag}([\lambda_1, \lambda_2, \ldots, \lambda_5])$.

We constructed two random permutation matrices P_r and P_{n-r} of sizes 2 and 3 respectively. We let $P = P_r \oplus P_{n-r}$ and X^{-1} is constructed as $X^{-1} = LPU$. Then A is formed by $A = XDX^{-1}$. The QR algorithm iteration results are shown in Figures 5.7, 5.9 and 5.10.

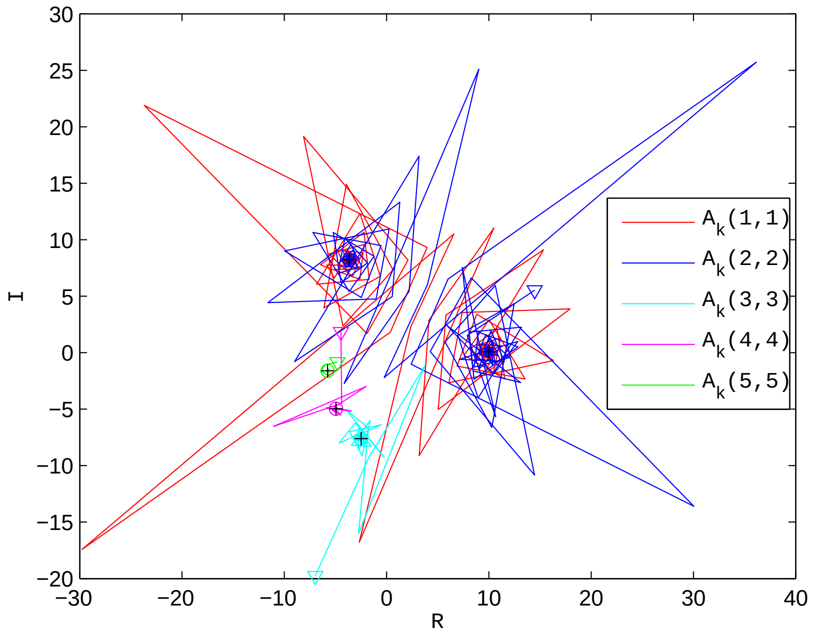

Figure 5.5: Diagonal elements of A_k converge to eigenvalues of A

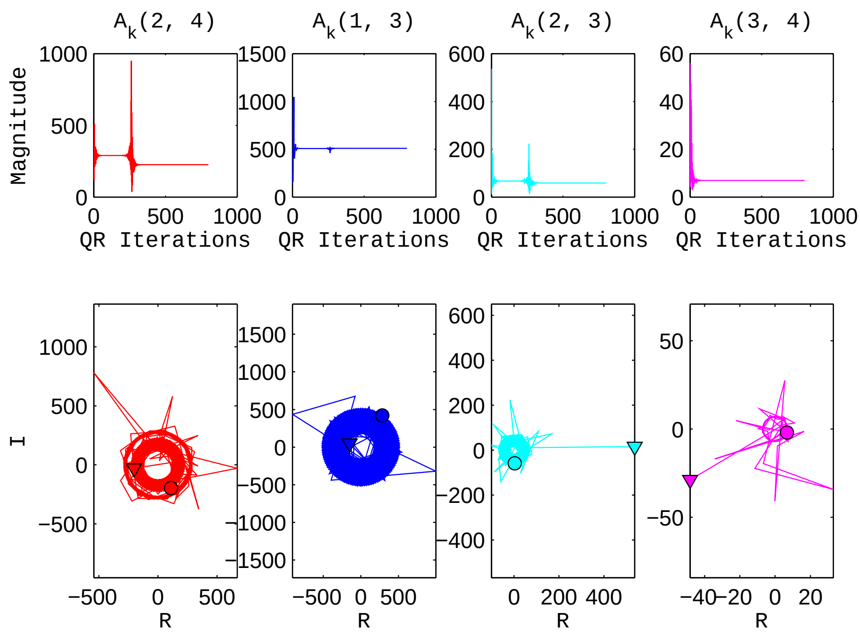

Figure 5.6: Upper triangular (off diagonal) part of A converge in magnitude

In Fig. 5.7, the convergence of several blocks of the lower triangular part of A_k off the diagonal are shown. Here, T-L is the set whose only member is $A_k(2,1)$; B-L is the set containing $A_k(i,j)$, $i = 3,4,5$ and $j = 1,2$; B-R is the set that contains $A_k(4,3)$, $A_k(5,3)$ and $A_k(5,4)$. A schematic illustration of the different blocks are shown in Fig. 5.8. From Fig. 5.7, we can see that the B-L block converge to zero while the other two blocks do not.

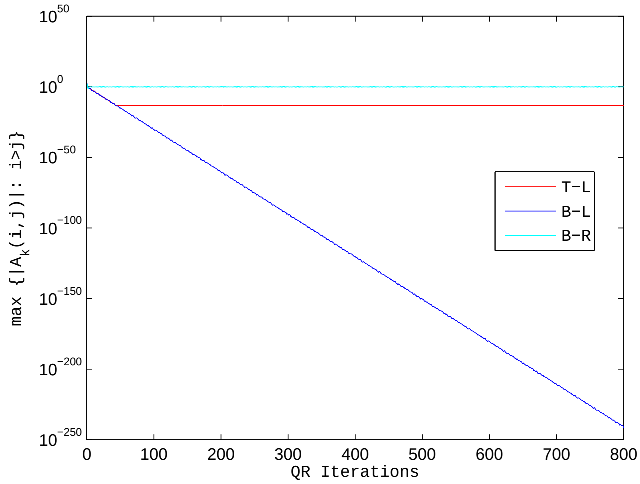

Figure 5.7: Lower triangular (off diagonal) part of A converge to zero

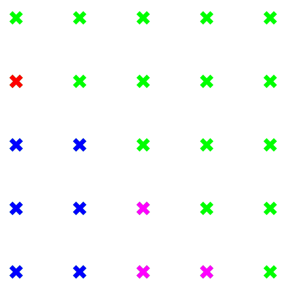

Figure 5.8: Schematic illustration of the lower triangular off diagonal blocks of A_k

Figure 5.9 shows the non-convergence of the diagonal elements of A_k. However, the eigenvalue multisets of the two blocks (top-left 2×2 block and bottom-right 3×3 block) of A_k converge to the eigenvalue multisets of the corresponding blocks of A, as shown in Fig. 5.10.

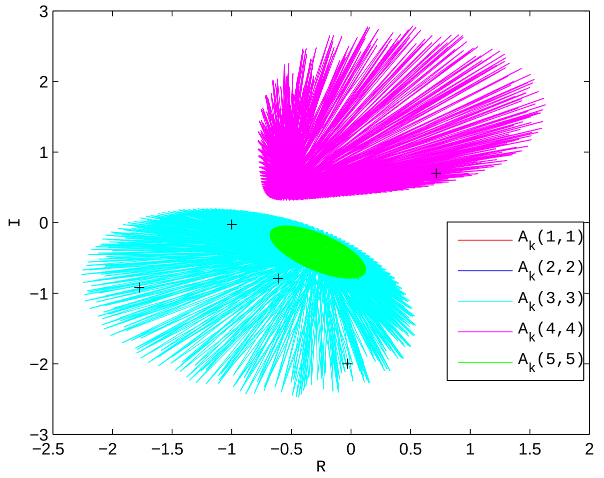

Figure 5.9: Diagonal elements of A_k do NOT converge to eigenvalues of A

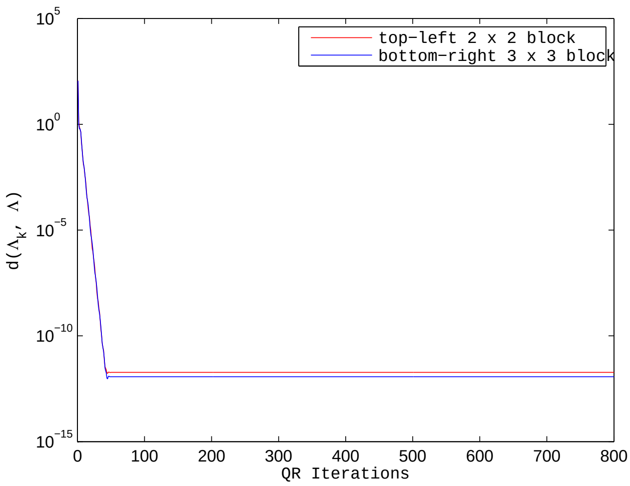

Figure 5.10: Convergence of eigenvalue multisets of diagonal blocks of A_k to those of A

Although we do not explore QR algorithms with shift [7, 9, 10, 11], we are aware of these advanced algorithms. The most simple shift algorithm for QR is the single shift QR algorithm. As shown above, when some eigenvalues of A share the same magnitude, then the diagonal elements of A_k would not converge to these eigenvalues. In the extreme case, if all the eigenvalues of A have the same magnitude, then the QR algorithm fails totally. Shift algorithms are proposed to solve this problem. Here we provide one example illustrating the single shift QR algorithm.

Let

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then all 5 eigenvalues of A have the same magnitude 1. In fact, the 5 eigenvalues of A are the fifth complex roots of 1 listed as follows, $1, \cos\left(\frac{2\pi}{5}\right) + i\sin\left(\frac{2\pi}{5}\right), \cos\left(\frac{4\pi}{5}\right) + i\sin\left(\frac{4\pi}{5}\right), \cos\left(\frac{6\pi}{5}\right) + i\sin\left(\frac{6\pi}{5}\right), \cos\left(\frac{8\pi}{5}\right) + i\sin\left(\frac{8\pi}{5}\right)$. The original QR algorithm fails to converge at all for this matrix. However, if we let

$$A_s = A + I,$$

then the eigenvalues of A_s are those of A plus 1. Indeed, if we let λ be an eigenvalue of A and x be a corresponding eigenvector of A, then

$$A_s x = (A + I)x = \lambda x + x = (\lambda + 1)x.$$

The eigenvalues of A_s are now

$$[\lambda_{s,1},\dots,\lambda_{s,5}] = \Big[2, 1+\cos\left(\frac{2\pi}{5}\right)+i\sin\left(\frac{2\pi}{5}\right),\\ 1+\cos\left(\frac{4\pi}{5}\right)+i\sin\left(\frac{4\pi}{5}\right),\\ 1+\cos\left(\frac{6\pi}{5}\right)+i\sin\left(\frac{6\pi}{5}\right),\\ 1+\cos\left(\frac{8\pi}{5}\right)+i\sin\left(\frac{8\pi}{5}\right)\Big].$$

Obviously they do not all have the same magnitude. But still $|\lambda_{s,2}| = |\lambda_{s,5}|$ and $|\lambda_{s,3}| = |\lambda_{s,4}|$. So after shift, we have 3 multisets of eigenvalues composed respectively of $\lambda_{s,1}$, $\lambda_{s,2}$ and $\lambda_{s,5}$, and $\lambda_{s,3}$ and $\lambda_{s,4}$. According to our proof in Chapter 4, the QR iterations should produce convergence at $\lambda_{s,1}$ on the diagonal. It should also produce two 2×2 block matrices along the diagonal that converge in terms of eigenvalue multisets.

We implemented the QR algorithm on A_s for 200 iterations. The iteration result A_{200} is shown below,

$$A_{200} = \begin{bmatrix} 2.0000 & -1.7000\times10^{-16} & -3.2641\times10^{-16} & -3.1697\times10^{-16} & -1.5618\times10^{-17} \\ -2.4792\times10^{-19} & 1.3090 & 9.5106\times10^{-1} & -1.0435\times10^{-16} & 4.6808\times10^{-17} \\ 7.6303\times10^{-19} & -9.5106\times10^{-1} & 1.3090 & 2.3168\times10^{-16} & 2.0064\times10^{-16} \\ 1.5620\times10^{-101} & -1.9469\times10^{-83} & 6.3259\times10^{-84} & 1.9098\times10^{-1} & -5.8779\times10^{-1} \\ -1.1349\times10^{-101} & 1.4145\times10^{-83} & -4.5961\times10^{-84} & 5.8779\times10^{-1} & 1.9098\times10^{-1} \end{bmatrix}. \tag{5.1}$$

As seen in (5.1), $A_k(1,1)$ converges to $\lambda_{s,1} = 2$. Also we see that there are two diagonal blocks that do not converge to zero. All other elements under these diagonal blocks converge to zero.

Also notice that the all the elements above the diagonal blocks also converge to zero. This is not by accident. In this case A is normal because $A^*A = AA^* = I$. By Equation 4.18, all the elements above the diagonal blocks converge to zero.

The diagonal block of $\begin{bmatrix} 1.3090 & 9.5106\times10^{-1} \\ -9.5106\times10^{-1} & 1.3090 \end{bmatrix}$ has two eigenvalues: $1.3090 + 0.9511i$ and $1.3090-0.9511i$, which are approximately equal to $\lambda_{s,2}$ and $\lambda_{s,5}$. The eigenvalues of the diagonal block of $\begin{bmatrix} 1.9098 \times 10^{-1} & -5.8779 \times 10^{-1} \\ 5.8779 \times 10^{-1} & 1.9098 \times 10^{-1} \end{bmatrix}$ are $0.1910 + 0.5878i$ and $0.1910 + 0.5878i$, which are approximately equal to $\lambda_{s,3}$ and $\lambda_{s,4}$. This result not only validates our proof in Chapter 4, but also illustrates that the single shift is effective as to enable the convergence to one eigenvalue of A.

Figure 5.11 shows the QR iterations of $A + (2 + i) * I$. In each iteration, the constant $(2 + i)$ is subtracted from the diagonal elements of A_k before they are plotted. One can see that all 5 eigenvalues converge with this shift. This is because the complex shift resulted in all 5 eigenvalues having 5 different magnitudes.

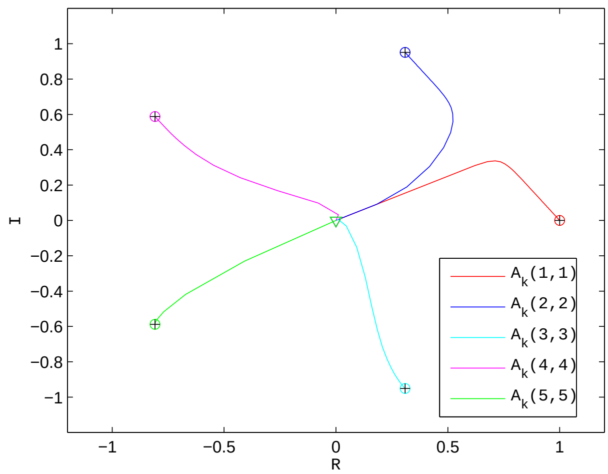

Figure 5.11: Convergence of QR: A is shifted by $(2 + i)$ in the complex plane

Bibliography

[1] Roger A. Horn and Charles R. Johnson, "Matrix Analysis," Cambridge University Press, 1990

[2] Ludwig Elsner, "On the Variation of the Spectra of Matrices," Linear Algebra and Its Applications 47 (1982), 127-138

[3] J.G.F. Francis, "The QR Transformation Part I," Computer Journal 4 (1961), 135-148,

[4] J.G.F. Francis, "The QR Transformation Part II," Computer Journal 4 (1962), 332-345

[5] V. N. Kublanovskaya, "On Some Algorithms for the Solution of the Complete Eigenvalue Problems," USSR Computational Mathematics and Mathematical Physics 1 (1961), 637-657

[6] J.H. Wilkinson, "Convergence of the LR, QR and Related Algorithms," Computer Journal 8 (1965), 77-84

[7] J.H. Wilkinson, "The Algebraic Eigenvalue Problem," Oxford University Press, 1988

[8] E. E. Tyrtyshnikov, "Matrix Bruhat Decompositions with a Remark on the QR (GR) Algorithm", Linear Algebra and its Applications 250 (1997), 61-68

[9] K. Braman, R. Byers, R. Mathias, "The Multishift QR Algorithm. Part I: Maintaining Well-Focused Shifts and Level 3 Performance," SIAM J. Matrix Anal. Appl. 23 (2001), 929-947

[10] K. Braman, R. Byers, R. Mathias, "The Multishift QR Algorithm. Part II: Aggressive Early Deflation," SIAM J. Matrix Anal. Appl. 23 (2001), 948-973

[11] D.S. Watkins, "Forward Stability and Transmission of Shifts in the QR Algorithm," SIAM J. Matrix Anal. Appl. 16 (1995), 469-487

[12] D.S. Watkins, "Understanding the QR Algorithm," SIAM Rev. 24 (1982), 427-440

[13] B.N. Parlett, "The QR Algorithm," Computing in Science and Engg. 2 (2000), 38-42

[14] H. Huang, T. Tam, "On the QR iterations of real matrices," Linear Algebra and its Applications. 408 (2005), 161-176