# Chapter 5

# Algebraic Eigenvalue Problems

Eigenvalue analysis is an important practice in many fields of engineering or physics. We also have seen in this course that eigenvalue analysis play an important role in the performance of many numerical algorithms. A general eigenvalue problem is stated as follows:

**Definition 5.0.1** *Given $n \times n$ matrices $A$ and $B$, find numbers $\lambda$ such that the equation*

$$Ax = \lambda Bx \tag{5.1}$$

*is satisfied for some nontrivial vector $x \neq 0$.*

If $B$ is invertible, then (5.1) can be reduced to

$$Cx = \lambda x. \tag{5.2}$$

Even if both $A$ and $B$ are real-valued, it is likely that $\lambda$ and $x$ are complex-valued. Finding the solution of eigensystems is a fairly complicated procedure. It is at least as difficult as finding the roots of polynomials. Therefore, any numerical method for solving eigenvalue problems is expected to be iterative in nature. Algorithms for solving eigenvalue problems include the power method, subspace iteration, the QR algorithm, the Jacobi method, the Arnoldi method and the Lanczos algorithm.

Some major references in this field are given at the end of this note.

## 5.1 Perturbation Theory

Given $A = (a_{ij}) \in R^{n \times n}$, define $r_i := \sum_{j \neq i}^{n} |a_{ij}|$ for all $i$. Let $C_i := \{x \in C| \ |x - a_{ii}| \leq r_i\}$. Then

**Theorem 5.1.1** *(Gerschgorin) Every eigenvalue $\lambda$ of $A$ belongs to one of the circles $C_i$. Moreover, if $m$ of the circles form a connected set $S$ disjoint from the remaining $n-m$ circles, then $S$ contains exactly $m$ eigenvalues.*

(pf): Let $\lambda$ be a eigenvalue of $A$ and let $x$ be the corresponding eigenvector, i.e. $Ax = \lambda x$. Let $k$ be the index for which $\|x\|_\infty = |x_k|$. Then $\Sigma a_{kj}x_j = \lambda x_k$ implies $(\lambda - a_{kk})x_k = \sum_{j\neq k} a_{kj}x_j$. It follows that $|\lambda - a_{kk}| \leq r_k$, i.e., $\lambda \in C_k$.

We know that eigenvalues are roots of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$ whereas coefficients of $p(\lambda)$ are polynomials in $a_{ij}$. We also know that the zeros of a polynomial depend continuously on its coefficients. So we conclude that the eigenvalues of a matrix depend continuously on its elements. Consider $A(\epsilon) := D + \epsilon(A - D)$ where $D$ is the diagonal of $A$ and $0 \leq \epsilon \leq 1$. When $\epsilon = 0$, all circles degenerate into points since $A(0) = D$. As $\epsilon$ increases, all circles expand. Note $A(1) = A$. As $\epsilon$ is changed, each eigenvalue $\lambda(\epsilon)$ is varied continuously in the complex plane and mark out a path from $\lambda(0)$ to $\lambda(1)$. Circles that are disjoint for $\epsilon = \epsilon_0$ are necessarily disjoint for all $\epsilon < \epsilon_0$. By the continuity of $\lambda_i(\epsilon)$ in $\epsilon$, eigenvalues cannot go from a connected set to another disjoint set. The theorem is proved. $\oplus$

**Theorem 5.1.2** *(Bauer-Fike) If $\mu$ is an eigenvalue of $A + E \in \mathbf{C^{n\times n}}$ and $P^{-1}AP = D = \ diag\ \{\lambda_1, \ldots, \lambda_n\}$. Then*

$$\min |\lambda - \mu| \leq \kappa(P)\|E\| \tag{5.3}$$

*where $\|\cdot\|$ denotes any induced norm.*

(pf): Consider $(A + E)x = \mu x$. Then $(\mu I - A)x = Ex$, $(\mu I - D)P^{-1}x = (P^{-1}EP)P^{-1}x$. Assuming $\mu \neq \lambda_i$, then it follows that $\|P^{-1}x\| \leq \|(\mu I - D)^{-1}P^{-1}EP\|\|P^{-1}x\|$. If an induced norm is used, then $\|\mu I - D)^{-1}\| = \max \frac{1}{|\mu-\lambda|}$. The theorem is proved. $\oplus$

**Example.** Let $A = \begin{bmatrix} 101, & -90 \\ 110, & -98 \end{bmatrix}, E = \begin{bmatrix} -\epsilon, & -\epsilon \\ 0, & 0 \end{bmatrix}$. The eigenvalues of $A$ are $\{1, 2\}$. The eigenvalues of $A + E$ are $\{\frac{3-\epsilon\pm\sqrt{1-838\epsilon+\epsilon^2}}{2}\}$. When $\epsilon = 0.001$, the eigenvalues are approximately $\{1.298, 1.701\}$. Thus $\min|\lambda - \mu| \approx 0.298$. This example shows that a small perturbation $E$ can lead to relative large perturbation in the eigenvalues of $A$.

**Remark.** When $A$ is a normal matrix, i.e., when $AA^* = A^*A$ (This class of matrices include symmetric matrices, orthogonal matrices, hermitian matrices, etc.), it is known that $P$ may be chosen to be a unitary matrix (i.e., $PP^* = I$). If we use the $L_2$-norm, then $\|P\|_2 = \|P^{-1}\|_2 = 1$. In this case, we have $\min\|\lambda - \mu| \leq \|E\|_2$. This implies that eigenvalue problems for normal matrices are well-conditioned.

**Remark.** When talking about the perturbation of eigenvectors, some cautions should be taken.

(1) Eigenvectors are unique up to multiplicative constants. Continuity should be discussed only under the assumption that all vectors are normalized in the same way.

(2) There are difficulties associated with multiply eigenvalues. For example, the matrix $\begin{bmatrix} 1+\epsilon, 0 \\ 0, 1-\epsilon \end{bmatrix}$ has eigenvectors $[1,0]^T$ and $[0,1]^T$ whereas the matrix $I$ has eigenvectors everywhere in $\mathbf{R^2}$.

**Examples.** (1) Consider the matrix

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

which is already in the Jordan Canonical form. It is not diagonalizable and has eigenvalues $\lambda = 1$ as an $n$-ford eigenvalues. Consider a slightly perturbed matrix

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

The eigenvalues of $A + E$ are roots of $p(\lambda) = (1-\lambda)^n - (-1)^n\epsilon = 0$. So $1 - \lambda_i = -(\epsilon)_i^{1/n}$ where $(\epsilon)_i^{1/n}$ is the $i$-th of the $n$-th roots of $\epsilon$. Note that $|\lambda_i(\epsilon) - \lambda_i| = |\epsilon^{1/n}|$. With $n = 10$ and $\epsilon = 10^{-10}$, we find that $|\lambda_i(\epsilon) - \lambda_i| = .1$. Note also that the matrix $A$ has only one eigenvector whereas $A + E$ has a complete set of eigenvectors.

(2) Consider the matrix $\begin{bmatrix} 2, & -10^{10} \\ 0, & 2 \end{bmatrix}$. $A$ is not diagonalizable. The exact eigenvalue of $A$ are $\{2, 2\}$ Suppose that $\lambda = 1$ is an approximate eigenvalue with eigenvector $x = [1, 10^{-10}[^T$. Then we find the residue $r := Ax - \lambda x = [0, 10^{-10}]^T$. Obviously, the residue is misleading.

## 5.2 Power Method and Its Variants

Given a matrix $A \in \mathbf{C^{n\times n}}$ and $x^{(0)} \in \mathbf{C^n}$. We define an iteration $x^{(k+1)} := Ax^{(k)}$ for $k = 0, 1, \ldots$ What will happen to the sequence $\{x^{(k)}\}$. If $A$ is convergent, then $x^{(k)} = A^k x^{(0)} \to 0$. If $A$ is not convergent, then $x^{(k)}$ may grow unboundly. Either case is undesirable. So we modify the process as follows.

**Algorithm 5.2.1** *(Power Method)*
*Given* $x^{(0)} \in \mathbf{C^n}$ *arbitrary,*
*For* $k = 1, 2, \ldots,$ *do*

$$\begin{aligned} w^{(k)} &:= Ax^{(k-1)} \\ x^{(k)} &:= \frac{w^{(k)}}{\|w^{(k)}\|_\infty} \end{aligned}$$

There is nothing special about doing a sup-norm normalization. In fact, any kind of normalization will be fine. We now examine the convergence properties of the power method.

For illustration purpose, we shall assume $A$ is diagonalizable with eigenvalues $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Let the corresponding eigenvectors be denoted by $x_1, \ldots x_n$. Detail discussion for other cases can be found in [1]. Suppose $x^{(0)} = \sum_{i=1}^{n} \alpha_i x_i$. Then

$$\begin{aligned} Ax^{(0)} &= \sum_{i=1}^{n} \alpha_i \lambda_i x_i \\ A^k x^{(0)} &= \sum_{i=1}^{n} \alpha_i \lambda_i^k x_i = \lambda_1^k \left( \alpha_1 x_1 + \sum_{i=2}^{n} \alpha_i \left( \frac{\lambda_i}{\lambda_1} \right)^k x_i \right). \end{aligned}$$

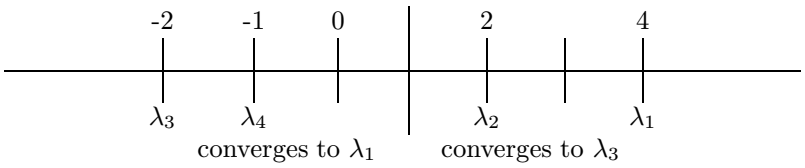
Assume $\alpha_1 \neq 0$ (This is guaranteed if $x^{(0)}$ is selected randomly). We observe

$$\begin{aligned} x^{(k)} &= \frac{Ax^{(k-1)}}{\|Ax^{(k-1)}\|_\infty} = \frac{A\left(\frac{w^{(k-1)}}{\|w^{(k-1)}\|_\infty}\right)}{\|A\left(\frac{w^{(k-1)}}{\|w^{(k-1)}\|_\infty}\right)\|_\infty} \\ &= \frac{A^2 x^{(k-2)}}{\|A^2 x^{(k-2)}\|_\infty} = \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|_\infty}. \end{aligned}$$

It is clear that as $k \to \infty$, the vector $A^k x^{(0)}$ *behaves like* $\alpha_1 \lambda_1^k x_1$ in the sense that contribution from $x_2, \ldots x_n$ becomes less and less significant. Normalization makes $x^{(k)} \to \frac{\alpha_1 \lambda_1^k}{|\alpha_1 \lambda_1^k|} \frac{x_1}{\|x_1\|_\infty}$. That is, $x^{(k)}$ converges to an eigenvector associated with the eigenvalue $\lambda_1$. Also, $w^{(k+1)} = Ax^{(k)} \to \lambda_1 x^{(k)}$. So $\frac{(Ax^{(k)})j}{x_j^{(k)}} \to \lambda_1$.

**Remark.** It is clear that the rate of convergence of power method depends on the ratio $\frac{\lambda_2}{\lambda_1}$.

**Remark.** The eigenvalues of the matrix $A - bI$ for a scalar $b$ are $\lambda_i - b$ if and only if $\lambda_i$ are eigenvalues of $A$. With this shift in mind, we can work on the matrix $A - bI$ instead of $A$ with the hope that the ratio of the first two dominant eigenvalues $\lambda_i - b$ will become smaller. This is called a shifted power method. It is not hard to find that the application is very limited. For example, assume all eigenvalues are real and are distributed as follows: Then with all choices of




$b$ the shifted power method will converge to either $\lambda_1$ or $\lambda_3$, but not any other eigenvalues. The idea is plausible, however, in the following setting.

Suppose $A$ is nonsingular, then $A^{-1}$ has eigenvalues $\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$. We apply the power method to $A^{-1}$, i.e., we define

**Algorithm 5.2.2** *(Inverse Power Method)*
*Given* $x^{(0)} \in \mathbf{C^n}$ *arbitrary, For* $k = 1, 2, \ldots,$ *do*

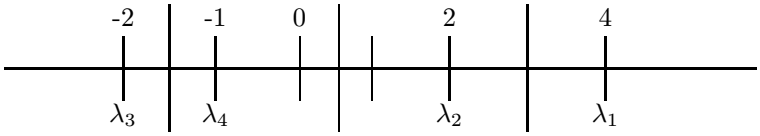
$$\begin{aligned} Aw^{(k)} &:= x^{(k-1)} \\ x^{(k)} &:= \frac{w^{(k)}}{\|w^{(k)}\|_\infty}. \end{aligned}$$

If we assume $|\lambda_1| \geq \ldots \geq |\lambda_{n-1}| > |\lambda_n|$, then $\{x^{(k)}\}$ converges to an eigenvector associated with $\lambda_n$ at the rate depending on $|\frac{\lambda_{n-1}}{\lambda_n}|$. Now we incorporate with the shift idea to obtain the following algorithm.

**Algorithm 5.2.3** *(Shifted Inverse Power Method)*
*Given* $x^{(0)} \in \mathbf{C^n}$ *arbitrary,*
*For* $k = 1, 2, \ldots,$ *do*

$$\begin{aligned} (A - bI)w^{(k)} &:= x^{(k-1)} \\ x^{(k)} &:= \frac{w^{(k)}}{\|w^{(k)}\|_\infty} \end{aligned}$$

The eigenvalues of $(A - bI)^{-1}$ are $\frac{1}{\lambda_1 - b}, \ldots, \frac{1}{\lambda_n - b}$. So whenever $b$ is chosen close enough to the eigenvalue $\lambda_i$, the sequence $\{x^{(k)}\}$ by the shifted inverse power method converges to an eigenvector associate with the eigenvalue $\frac{1}{\lambda_i - b}$.
**Example.** Suppose eigenvalues are distributed as follows: Then the longer




vertical bars separates the regions of $b$ by which the shifted inverse power method will be able to find, respectively, different eigenvalues.
**Remark.** There are several ways to choose the shift $b$:

(1) If some estimate of $\lambda_i$ has been found, we may use it for $b$.

(2) We may generate $b^{(0)}$ at random. Then define

$$b^{(k+1)} := \frac{(w^{(k)})^* A w^{(k)}}{(w^{(k)})^* w^{(k)}} \tag{5.4}$$

successively. This is a variable-shift inverse power method, and is known as the Rayleigh Quotient Iteration. It can be shown that the rate of convergence is cubic [2].

**Remark.** In the inverse power method, we have to solve a linear system

$$(A - bI)w^{(k)} = x^{(k-1)}. \tag{5.5}$$

When $b$ is close to some eigenvalue $\lambda_i$, the matrix $A - bI$ is nearly singular. We expect that small errors in $x^{(k-1)}$ would lead to large errors in the computation of $w^{(k)}$. Fortunately, this error mainly occurs in the direction of $x^{(k)}$. Thus after normalization the method is still useful in practice. (See, [2], for a justification.)

## 5.3 The OR algorithm

The basic idea behind the QR algorithm is from a sequence of matrices $\{A_i\}$ that are isospectral to the original matrix $A_1 := A$ and that converge to a special form from which the eigenvalues can readily be known. For simplicity, we shall restrict our discussion for real matrices in the sequel.

**Algorithm 5.3.1** *(Basic QR algorithm)*
*Given $A \in \mathbf{R^{n\times n}}$, define $A_1 := A$.*
*For $k = 1, 2, \ldots,$ do*
*Calculate the QR decomposition $A_k = Q_k R_k$,*
*Define $A_{k+1} := R_k Q_k$.*

**Remark.** We observe that $A_{k+1} = R_k Q_k = Q_k^T A_k Q_k$. So $A_{k+1}$ is orthogonally similar to $A_k$. By induction, the sequence $\{A_k\}$ is isospectral to $A$.

**Definition 5.3.1** *We say matrix $A \in \mathbf{R^{n\times n}}$ is upper Hessenberg if and only if $A_{ij} = 0$ for all $i - 2 \geq j$.*

The QR algorithm is expensive because of the large number of computations that must be performed at each step. To save this overhead, the matrix $A$ usually is simplified first by orthogonal similarity transformations to an upper Hessenberg matrix. This can be accomplished as follows:

Given $A$, let $U_2$ be the unitary matrix $U_2 = \begin{bmatrix} 1, & 0, & \ldots, & 0 \\ 0, & & & \\ & & \tilde{U}_2 & \\ 0, & & & \end{bmatrix}$ where $\tilde{U}_2$ is formed by the Householder transformation for the column vector $[a_{21}, \ldots a_{n1}]^T$.

Thus $U_2 A = \begin{bmatrix} \text{unchanged} & & & \\ x & x & \ldots & x \\ 0 & x & & x \\ \vdots & \vdots & & \vdots \\ 0 & x & \ldots & x \end{bmatrix}$. It is clear that $U_2 A U_2^T$ does not change the first column of $U_2 A$. Continuing this procedure, $A$ is transformed by orthogonal similarity transformations into an upper Hessenberg matrix.

**Definition 5.3.2** *An orthogonal matrix of the form*

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & 0 & \sin\theta & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -\sin\theta & 0 & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \longrightarrow k-th\ row \\ \\ \longrightarrow j-th\ row \end{matrix}$$

*is called a rotation matrix in the $(j,k)$-plane and is denoted as $\Omega_{jk}$.*

**Example.** Consider the $2\times 2$ matrix $A = \begin{bmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{bmatrix}$. Then

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{bmatrix} = \begin{bmatrix} ca_{11}+sa_{21}, & ca_{12}+sa_{22} \\ -sa_{11}+ca_{21}, & -sa_{12}+ca_{22} \end{bmatrix}.$$

Thus if we choose $c := \frac{a_{11}}{\sqrt{a_{11}^2+a_{21}^2}}$ and $s = \frac{a_{21}}{\sqrt{a_{11}^2+a_{21}^2}}$, then the $(2,1)$-position of the product becomes zero. The matrix $\begin{bmatrix} c & s \\ -s & c \end{bmatrix}$ amounts to the rotation of the coordinate axes by an angle determined by $\tan\theta = \frac{s}{c} = \frac{a_{21}}{a_{11}}$. This idea has been incorporated into the so called Given's method for solving eigenvalue problems:

Suppose $A$ is an upper Hessenberg matrix. Then the QR decomposition of $A$ can be calculated by a sequence of plane rotations. This is demonstrated as follows:

$$A_1 = \begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix} \longrightarrow \Omega_{12}A_1 = A_2 = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix}$$

$$\longrightarrow \Omega_{23}A_2 \quad = A_3 = \begin{bmatrix} x & x & x & x \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

$$\longrightarrow \Omega_{34}A_3 \quad = A_4 = \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix} = R.$$

Working with upper Hessenberg matrices, the QR decomposition will only take $O(n^2)$ flops, while with general matrices, the QR decomposition requires $O(n^3)$ flops.

**Remark.** It is important to note that if $A_1 = A$ is upper Hessenberg, then the orthogonal matrix $Q_1$ in the QR decomposition of $A_1 = Q_1R_1$ is also upper

Hessenberg (Prove this!). Therefore, $A_2 = R_1 Q_1$ maintains to be upper Hessenberg. It follows that the upper Hessenberg structure is maintained throughout the QR algorithm.

We have seen that

$$A_{k+1} = Q_k^T A_k Q_k = Q_k^T \dots Q_1^T A Q_1 \dots Q_k.$$

Let us call

$$\begin{aligned} P_k : &= Q_1 \dots Q_k && (5.6) \\ U_k : &= R_k \dots R_1. && (5.7) \end{aligned}$$

Note that $P_k$ is still orthogonal and $U_k$ upper triangular. Moreover,

$$\begin{aligned} P_k U_k &= (Q_1 \dots Q_k)(R_k \dots R_1) = (Q_1 \dots Q_{k-1}) A_k (R_{k-1} \dots R_1) \\ &= (Q_1 \dots Q_{k-1})(Q_{k-1}^T \dots Q_1^T A Q_1 \dots Q_{k-1})(R_{k-1} \dots R_1) \\ &= A P_{k-1} U_{k-1} = \dots = A^k. \quad (5.8) \end{aligned}$$

In other words, we have found that $P_k U_k$ is the QR decomposition of the matrix $A^k$

**Theorem 5.3.1** *Suppose $A \in \mathbf{R^{n \times n}}$ and suppose $|\lambda_1| > \dots > |\lambda_n| > 0$. Then the sequence $\{A_k\}$ converges to an upper triangular matrix.*

(pf): From the assumption, we know all eigenvalues of $A$ are real and distanct. So $A$ is diagonalizable. Let $P$ be the nonsingular matrix such that $P^{-1}AP = D := \text{ diag } \{\lambda_1, \dots, \lambda_n\}$. Then $A^k = PD^kP^{-1}$. Let $P^{-1} = LU$ be the $LU$ decomposition of $P^{-1}$ with value 1 along the diagonal of $L$ (We assume this LU decomposition exists with some prior pivoting if necessary). Then $A^k = PD^kLU = P(D^kLD^{-k})D^kU$. Consider $D^kLD^{-k}$. The $(i,j)$-component is $(\frac{\lambda_i}{\lambda_j})^k l_{ij}$. Since $L$ is lower triangular, only the lower triangular i.e., $i \geq j$, portion needs to be considered.

Let $D^kLD^{-k} := I + E_k$. By the ordering of $\lambda_i$, we see that $E_k \longrightarrow 0$ as $k \longrightarrow \infty$. Let $P := QR$. Then

$$\begin{aligned} A^k &= QR(D^kLD^{-k})D^kU \\ &= QR(I + E_k)D^kU = Q(I + RE_kR^{-1})RD^kU \\ &= Q(\tilde{Q}_k\tilde{R}_k)RD^kU = (Q\tilde{Q}_k)(\tilde{R}_kRD^kU). \quad (5.9) \end{aligned}$$

where $\tilde{Q}_k\tilde{R}_k$ is the QR decomposition of $I + RE_kR^{-1}$.

We recall the issue of uniqueness of the QR decomposition of a matrix. Suppose a nonsingular matrix $B$ has two QR decompositions, say $B = QR = \tilde{Q}\tilde{R}$. Then $Q^T\tilde{Q} = R\tilde{R}^{-1}$. But $Q^T\tilde{Q}$ is orthogonal whereas $R\tilde{R}^{-1}$ is upper triangular. An orthogonal and upper triangular matrix is diagonal with absolute value 1 along its diagonal. If we call $R\tilde{R}^{-1} := D$, then $\tilde{Q} = QD$ and $\tilde{R} = D^{-1}R$. So the QR decomposition of $B$ is unique if the diagonal elements of $R$ are required to be positive.

We have already known $A^k = P_k U_k$. Thus $A^k$ has two QR decompositions. It follows that

$$\begin{aligned} P_k &= Q\tilde{Q}_k\tilde{D}_k \\ U_k &= \tilde{D}_k^{-1}\tilde{R}_k R D^K U = \tilde{D}_k\tilde{R}_k R D^k U \end{aligned}$$

where $\tilde{D}_k$ is a diagonal matrix with elements either 1 or -1. Now

$$\begin{aligned} A_{k+1} &= P_k^T A P_k = (Q\tilde{Q}_k\tilde{D}_k)^T A(Q\tilde{Q}_k\tilde{D}_k) = \tilde{D}_k\tilde{Q}_k^T(Q^T AQ)(\tilde{Q}_k\tilde{D}_k) \\ &= \tilde{D}_k\tilde{Q}_k^T(RP^{-1}APR^{-1})(\tilde{Q}_k\tilde{D}_k) \\ &= \tilde{D}_k\tilde{Q}_k^T(RDR^{-1})(\tilde{Q}_k\tilde{D}_k). \end{aligned} \tag{5.10}$$

Note that $RDR^{-1}$ is an upper triangular matrix with diagonal elements $\lambda_1, \ldots, \lambda_n$. Observe also that $I + RE_kR^{-1}$ converges to $I$ as $k \longrightarrow \infty$ since $E_k \longrightarrow 0$. So $\tilde{R}_k = \tilde{Q}_k(I + RE_kR^{-1}) \longrightarrow \tilde{Q}_k^T$ as $k \longrightarrow \infty$. Since $\tilde{R}_k$ is upper triangular with positive diagonal elements, $\tilde{R}_k$ and $\tilde{Q}_k \longrightarrow I$ as $k \longrightarrow \infty$. Thus we see from (5.10) that $A_{k+!} \longrightarrow$ an upper triangular matrix with $\lambda_1, \ldots, \lambda_n$ along the diagonal. $\oplus$

**Remark.** From the above proof, we see that the rate of convergence depends on the ration $\max_j |\frac{\lambda_{j+1}}{\lambda_j}|$. It is, therefore , desired to use the shift technique to accelerate the convergence.

**Algorithm 5.3.2** *(QR algorithm with origin shift)(Explicit shift)*
*Given $A \in \mathbf{R^{n\times n}}$ , define $A_1 := A$.*
*For $k = 1, 2, \ldots,$ do*
*Select a shift factor $c_k$;*
*Calculate the QR decomposition,*

$$A_k - c_k I = Q_k R_k; \tag{5.11}$$

*Define*

$$A_{k+1} := R_k Q_k + c_k I. \tag{5.12}$$

**Remark.** We note that $R_k = Q_k^T(A_k - c_k I)$. So $A_{k+1} = Q_k^T(A_k - c_k I)Q_k + c_k I = Q_k^T A_k Q_k$. That is, $A_{k+1}$ and $A_k$ are orthogonally similar.

**Definition 5.3.3** *Suppose $A_k$ is upper Hessenberg. The shift factor $c_k$ may be determine from the eigenvalues, say $\mu_k$ or $\nu_k$, of the bottom $2 \times 2$ submatrix of $A_k$. If both $\mu_k$ and $\nu_k$ are real, we take $c_k$ to be $\mu_k$ or $\nu_k$ according to whether $|\mu_k - a_{nn}^{(k)}|$ or $|\nu_k - a_{nn}^{(k)}|$ is smaller. If $\mu_k = \overline{\nu}_k$, then we choose $c_k = \Re\nu_k$. Such a choice of shift factor is known as the Wilkinson shift.*

**Theorem 5.3.2** *Given any matrix $A$ and the relationship $B = Q^T AQ$. Suppose $B$ is upper Hessenberg with positive and diagonal elements and suppose $Q$ is orthogonal. Then matrices $Q$ and $B$ are uniquely determined from the first column of $Q$.*

(pf): This is homework problem.

**Remark.** Let $A$ be any given matrix and let $B = Q^T AQ$ be an upper Hessenberg matrix. Let $P^T$ be an orthogonal matrix whose first column agrees with that of $Q$. Suppose $PAP^T$ is reduced to upper Hessenberg, that is, suppose $U_1, \ldots, U_{n-2}$ are orthogonal matrices such that

$$\tilde{B} := U_{n-2} \ldots U_1 (PAP^T) U_1 \ldots U_{n-2}^T$$

becomes an upper Hessenberg matrix (Recall how this is done!). Let $\tilde{Q} = P^T U_1^T \ldots U_{n-2}^T$. Then we have $\tilde{Q}^T A \tilde{Q} = \tilde{B}$. Note that the first column of $\tilde{Q}$ is the same as that of $Q$. From Theorem 5.3.1, we conclude that $\tilde{B} = B$.

Suppose $A_k$ is upper Hessenberg. There is an easier way to compute $A_{k+1} = Q_k^T A_k Q_k$ from $A_k$ based on the above theorem and remark: Since $A_k$ is upper Hessenberg, the first column of $Q_k$ is known, i.e., it is the column $[c, -s, 0, \ldots, 0]^T$ with

$$c := \frac{a_{11}^{(k)} - c_k}{\sqrt{(a_{11}^{(k)} - c_k)^2 + a_{21}^{(k)}}}$$

$$s := \frac{a_{21}^{(k)}}{\sqrt{(a_{11}^{(k)} - c_k)^2 + a_{21}^{(k)}}}.$$

So the first column of $P^T$ is known and $P^T$ can be chosen to be the matrix

$$P^T := \begin{bmatrix} c & s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It is easy to see that $PA_kP^T$ is at most a matrix of the form

$$PA_kP^T = \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix}$$

So we can make a rotation in the (2.3)-plane to eliminate the nonzero element at the $(3,1)$-position. In doing so, the $(4,2)$-position becomes nonzero. We continue to chase the nonzero elements by $n-2$ rotations to produce $A_{k+1}$.

Suppose we have applied one QR step with shift $c_1$ to obtain $A_2 = Q_1^T A_1 Q_1$ and other QR step with shift $c_2$ to obtain $A_3 = Q_2^T A_2 Q_2 = Q_2^T Q_1^T A_1 Q_1 Q_2$. Then we may as well using the first column of $Q_1 Q_2$ to compute $A_3$ from $A_1$ by using Theorem 5.3.1 and the above remark. We recall that $A_1 - c_1 I = Q_1 R_1$, $A_2 = R_1 Q_1 + c_1 I$, $A_2 - c_2 I = Q_2 R_2$ and $A_3 = R_2 Q_2 + c_2 I$. It follows that $T_2 = (A_3 - c_2 I) Q_2^T = (Q_2^T Q_1^T A_1 Q_1 Q_2 - c_2 I) Q_2^T = Q_2^T Q_1^T (A_1 - c_2 I) Q_1$. Thus $R_2 R_1 = Q_2^T Q_1^T (A_1 - c_2 I) Q_1 R_1$, or equivalently $Q_1 Q_2 R_2 R_1 = (A_1 - c_2 I)(A - c_1 I) = A^2 - (c_0 + c_1) A + c_0 c_1 I$. That is, the first column of $Q_1 Q_2$ is seem to be the normalized first column of the matrix $A^2 - (c_0 + c_1) A + c_0 c_1 I$. This is implicit double shift.

## 5.4 Exercises

1. Let $p$ be a chosen integer satisfying $1 \leq p \leq n$. Given an $n \times p$ matrix $Q_0$ with orthonormal columns, consider the iteration

$$\begin{aligned} Z_k &= AQ_{k-1} \\ Q_k R_k &= Z_k \ (QR \text{ factorization}) \end{aligned}$$

   for $k = 1, 2, \ldots$. Explain why this iteration can usually be used to compute the $p$ largest eigenvalues of $A$ in absolute value. How then should you modify the iteration when the $p$ smallest eigenvalues of $A$ in absolute value are needed.

2. Prove that the upper Hessenberg structure is maintained throughout the QR algorithm. That is, if $A_k = Q_k R_k$ is an upper Hessenberg matrix, show that $A_{k+1} = R_k Q_k$ is also upper Hessenberg.

3. Prove Theorem 5.3.2.

4. Compute a $QR$ step with the matrix

$$A = \begin{bmatrix} 2 & \epsilon \\ \epsilon & 1 \end{bmatrix}$$

   (a) without shift;

   (b) with shift $k = 1$.

5. Assume $A = diag(\lambda_1, \lambda_2)$ with $\lambda_1 > \lambda_2$. Assume

$$x_k = \begin{pmatrix} c_k \\ s_k \end{pmatrix}$$

   with $c_k^2 + s_k^2 = 1$ is an approximate eigenvector.

   (a) Compute the Rayleigh quotient of $x_k$.

   (b) Compute, by pen and paper, the next Raleigh quotient iteration $x_{k+1}$.

   (c) Comparing $x_k$ and $x_{k+1}$, give an explanation why the convergence of the Rayleigh quotient iteration is ultimately cubic.

# Bibliography


[1] D. K. Faddeev and V. N. Faddeva, Computational Methods of Linear Algebra, W. H. Freeman and Co., San Francisco, 1963.

[2] B. N. Parlett, The Symmetric Eigenvalue Problems, Prentice-Hall, Englewood Cliffs, N.J., 1980.

[3] B. T. Smith et al, Matrix Eigensystem Routines, EISPACK Guide, Lecture notes in computer science, Vol. 6, Berlin, Springer-Verlag, 1974.

[4] G.W. Stewart and Ji-guang Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.

[5] D. S. Watkins, Understanding the QR algorithm, SIAM Review, 24(1982), 427-440.

[6] J. H. Wilkinson and C. Reinsch, Linear Algebra, Handbook for Automatic Computation, Vol. II, ed. F. L. Bauer, Berlin, Springer-Verlag, 1971.

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