Classical Numerical Analysis, Chapter 08

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

Chapter 08, Part 2 of 2
Eigenvalue Problems

Power Iteration Methods

Definition (power iteration)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *and* $\mathbf{q} \in \mathbb{C}^n_\star$ *is given. The* **power method** *is an algorithm that generates a sequence of vectors* $\{\mathbf{v}_k\}_{k=0}^\infty \subset \mathbb{C}^n_\star$ *according to the following recursive formula:*

$$\mathbf{v}_0 = \frac{\mathbf{q}}{\|\mathbf{q}\|_2}.$$

For $k \geq 1$,

$$\mathbf{q}_k = \mathsf{A}\mathbf{v}_{k-1},$$

and, provided $\mathbf{q}_k \neq \mathbf{0}$,

$$\mathbf{v}_k = \frac{\mathbf{q}_k}{\|\mathbf{q}_k\|_2}.$$

If $\mathbf{q}_k = \mathbf{0}$, *the algorithm terminates at step* k.

The name of the method becomes clear once we realize that

$$\mathbf{v}_k = \frac{1}{\|\mathbf{q}_k\|_2}\mathbf{q}_k = \frac{1}{\|\mathbf{q}_k\|_2}\mathsf{A}\mathbf{v}_{k-1} = \cdots = c_k \mathsf{A}^k \mathbf{v}_0.$$

Theorem (convergence of the power iteration)

Let $\mathsf{A} \in \mathbb{C}^{n \times n}$ *be Hermitian and* $\mathbf{q} \in \mathbb{C}^n_\star$ *be given. Suppose that the spectrum* $\sigma(\mathsf{A}) = \{\lambda_i\}_{i=1}^n$ *has the ordering*

$$0 \leq |\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_{n-1}| < |\lambda_n|.$$

Assume that $\{\mathbf{w}_i\}_{i=1}^n$ *is the associated basis of orthonormal eigenvectors of* A*. If* $(\mathbf{q}, \mathbf{w}_n)_2 \neq 0$*, where* $\mathbf{w}_n$ *is the eigenvector corresponding to the dominant eigenvalue* λ_n*, then, for some constant* $C > 0$*,*

$$\|\mathbf{v}_k - s_k \mathbf{w}_n\|_2 \leq C \left| \frac{\lambda_{n-1}}{\lambda_n} \right|^k, \qquad |R(\mathbf{v}_k) - \lambda_n| \leq 2|\lambda_n| C^2 \left| \frac{\lambda_{n-1}}{\lambda_n} \right|^{2k},$$

when k *is sufficiently large. Here* s_k *is a sequence of modulus one complex numbers.*

The proof is on the next few slides.

Proof.

The vector $\mathbf{v}_0$ is obtained by normalizing $\mathbf{q}$. Let us then expand $\mathbf{v}_0$ in the basis of eigenvectors: there exist unique constants $\alpha_j \in \mathbb{C}$ such that

$$\mathbf{v}_0 = \sum_{j=1}^{n} \alpha_j \mathbf{w}_j.$$

From our assumptions $\alpha_n \neq 0$ and $\sum_{j=1}^{n} |\alpha_j|^2 = 1$. Now

$$\mathbf{q}_1 = \mathsf{A}\mathbf{v}_0 = \sum_{j=1}^{n} \lambda_j \alpha_j \mathbf{w}_j.$$

From this it follows that

$$\mathbf{v}_1 = \frac{1}{\sqrt{\sum_{j=1}^{n} |\alpha_j|^2 \lambda_j^2}} \sum_{j=1}^{n} \lambda_j \alpha_j \mathbf{w}_j$$

Proof, Cont.

Proceeding this way we observe that

$$
\begin{aligned}
\mathbf{v}_k &= \frac{1}{\sqrt{\sum_{j=1}^{n} |\alpha_j|^2 \lambda_j^{2k}}} \sum_{j=1}^{n} \lambda_j^k \alpha_j \mathbf{w}_j \\
&= \frac{1}{\sqrt{\sum_{j=1}^{n} |\alpha_j|^2 \lambda_j^{2k}}} \left(\lambda_n^k \alpha_n \mathbf{w}_n + \sum_{j=1}^{n-1} \lambda_j^k \alpha_j \mathbf{w}_j \right) \\
&= \frac{\alpha_n \lambda_n^k}{|\alpha_n \lambda_n^k|} \frac{1}{\sqrt{1 + \sum_{j=1}^{n-1} \left| \frac{\alpha_j}{\alpha_n} \right|^2 \left(\frac{\lambda_j}{\lambda_n} \right)^{2k}}} \left(\mathbf{w}_n + \sum_{j=1}^{n-1} \frac{\alpha_j}{\alpha_n} \left(\frac{\lambda_j}{\lambda_n} \right)^k \mathbf{w}_j \right).
\end{aligned}
$$

Define $s_k = \frac{\alpha_n \lambda_n^k}{|\alpha_n \lambda_n^k|}$. Given the assumptions on $\sigma(A)$, we observe that, for $j = 1, \ldots, n-1$, that

$$
\left(\frac{\lambda_j}{\lambda_n} \right)^k \to 0 \qquad k \to \infty.
$$

Proof, Cont.

It is clear that

$$\frac{1}{s_k}\mathbf{v}_k - \mathbf{w}_n \to \mathbf{0} \qquad k \to \infty.$$

We leave it to the reader as an exercise to prove the estimate

$$\|\mathbf{v}_k - s_k\mathbf{w}_n\|_2 \le C\left|\frac{\lambda_{n-1}}{\lambda_n}\right|^k,$$

for some $C > 0$, provided k is sufficiently large. If this holds, the Rayleigh quotient estimate theorem yields the second order convergence rate of the eigenvalue approximations. □

Proposition (shifts)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *and* $\mu \in \mathbb{C}$ *has the the property that* $\mu \notin \sigma(\mathsf{A})$. *Then* $(\mathsf{A} - \mu\mathsf{I})^{-1}$ *exists, and* $(\lambda, \mathbf{w})$ *is an eigenpair of* A *iff* $\left((\lambda - \mu)^{-1}, \mathbf{w}\right)$ *is an eigenpair of* $(\mathsf{A} - \mu\mathsf{I})^{-1}$.

Proof.

The matrix $\mathsf{A} - \mu\mathsf{I}$ is singular iff $\mu \in \sigma(\mathsf{A})$. Thus, if $\mu \notin \sigma(\mathsf{A})$, $\mathsf{A} - \mu\mathsf{I}$ is invertible. Next, suppose that $(\lambda, \mathbf{w})$ is an eigenpair of A. Then, if $\mu \notin \sigma(\mathsf{A})$,

$$\mathsf{A}\mathbf{w} = \lambda\mathbf{w},$$

which implies

$$(\mathsf{A} - \mu\mathsf{I})\,\mathbf{w} = (\lambda - \mu)\mathbf{w},$$

which implies

$$(\mathsf{A} - \mu\mathsf{I})^{-1}\,\mathbf{w} = (\lambda - \mu)^{-1}\mathbf{w}.$$

The other direction uses the same reasoning. □

Inverse Iteration Method

If μ is close to $\lambda_r \in \sigma(\mathsf{A})$, then $(\lambda_r - \mu)^{-1}$ is much larger in modulus than $(\lambda_j - \mu)^{-1}$, for any other eigenvalue. Thus, to approximate λ_r we can apply power iterations to the matrix $(\mathsf{A} - \mu\mathsf{I})^{-1}$. This is the idea of the inverse iteration method.

Definition (inverse iteration method)

Suppose that $\mathsf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{q} \in \mathbb{C}^n_\star$, *and* $\mu \in \mathbb{C}$ *are given. Assume that* $\mathsf{A} - \mu \mathsf{I}$ *is invertible. The inverse iteration method is an algorithm that generates a sequence of vectors* $\{\mathbf{v}_k\}_{k=0}^\infty \in \mathbb{C}^n_\star$ *according to the following recursive formula:*

$$\mathbf{v}_0 = \frac{\mathbf{q}}{\|\mathbf{q}\|_2}.$$

For $k \geq 1$

$$\mathbf{q}_k = (\mathsf{A} - \mu \mathsf{I})^{-1} \mathbf{v}_{k-1},$$

and, provided $\mathbf{q}_k \neq \mathbf{0}$,

$$\mathbf{v}_k = \frac{\mathbf{q}_k}{\|\mathbf{q}_k\|_2}.$$

If $\mathbf{q}_k = \mathbf{0}$, *the algorithm terminates at step* k.

Theorem (convergence of the inverse iteration)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is Hermitian with spectrum* $\sigma(\mathsf{A}) = \{\lambda_i\}_{i=1}^n \subset \mathbb{R}$. *Let* $\{\mathbf{w}_i\}_{i=1}^n$ *be an associated orthonormal basis of eigenvectors of* A. *Suppose that* $\mu \in \mathbb{C}$ *is given with the property that* $\mu \notin \sigma(\mathsf{A})$. *Define*

$$\lambda_r = \operatorname*{argmin}_{j=1}^{n} |\lambda_j - \mu|, \quad \lambda_s = \operatorname*{argmin}_{\substack{j=1\\ j\neq r}}^{n} |\lambda_j - \mu|$$

and assume that

$$|\lambda_r - \mu| < |\lambda_s - \mu| \leq |\lambda_j - \mu|,$$

for all $j \in \{1, \ldots, n\} \setminus \{r, s\}$. *If* $\mathbf{q} \in \mathbb{C}^n_\star$ *is given with the property that* $(\mathbf{q}, \mathbf{w}_r)_2 \neq 0$, *then the inverse iterations converges. Moreover, we have the following convergence estimates: there is some* $C > 0$ *such that*

$$\|\mathbf{v}_k - s_k \mathbf{w}_r\|_2 \leq C \left|\frac{\mu - \lambda_r}{\mu - \lambda_s}\right|^k, \quad |R(\mathbf{v}_k) - \lambda_r| \leq 2\rho(\mathsf{A}) C^2 \left|\frac{\mu - \lambda_r}{\mu - \lambda_s}\right|^{2k},$$

provided k *is sufficiently large. As before,* s_k *is a sequence of modulus one complex numbers.*

Theorem (deflation)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is Hermitian, and its spectrum is denoted* $\sigma(\mathsf{A}) = \{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{R}$. *Let* $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ *be an orthonormal basis of eigenvectors of* A, *with* $\mathsf{A}\mathbf{w}_k = \lambda_k \mathbf{w}_k$, *for* $k = 1, \ldots, n$. *Assume that the eigenvalues of* A *are ordered as follows:*

$$0 \leq |\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_{n-2}| < |\lambda_{n-1}| < |\lambda_n|.$$

Assume that λ_n *is known, and that* $\mathbf{v}_0 \in \mathbb{C}^n_\star$ *is such that* $\mathbf{w}_{n-1}^{\mathsf{H}}\mathbf{v}_0 \neq 0$. *Then the sequence,*

$$\mathbf{v}_{k+1} = \mathsf{A}\mathbf{v}_k, \qquad \mathbf{y}_{k+1} = \mathbf{v}_{k+1} - \lambda_n \mathbf{v}_k,$$

satisfies

$$\frac{\mathbf{y}_{k+1}^{\mathsf{H}}\mathbf{y}_k}{\mathbf{y}_k^{\mathsf{H}}\mathbf{y}_k} \to \lambda_{n-1}, \qquad k \to \infty.$$

The proof is on the next few slides.

Proof.

There exist unique constants $\alpha_j \in \mathbb{C}$, $j = 1, 2, \ldots, n$, such that $\mathbf{v}_0 = \sum_{j=1}^{n} \alpha_j \mathbf{w}_j$. By assumption $\alpha_{n-1} \neq 0$. We observe that $\mathbf{v}_k = \mathsf{A}^k \mathbf{v}_0$, and, consequently,

$$\mathbf{y}_k = \mathsf{A}^k \mathbf{v}_0 - \lambda_n \mathsf{A}^{k-1} \mathbf{v}_0 = \mathsf{A}^{k-1} \left(\mathsf{A} - \lambda_n \mathsf{I}_n\right) \mathbf{v}_0 = \sum_{j=1}^{n-1} \alpha_j \lambda_j^{k-1} \left(\lambda_j - \lambda_n\right) \mathbf{w}_j .$$

Therefore,

$$\begin{aligned}
\frac{\mathbf{y}_{k+1}^{\mathsf{H}} \mathbf{y}_k}{\mathbf{y}_k^{\mathsf{H}} \mathbf{y}_k} &= \frac{\sum_{j=1}^{n-1} |\alpha_j|^2 \lambda_j^{2k-1} \left(\lambda_j - \lambda_n\right)^2}{\sum_{j=1}^{n-1} |\alpha_j|^2 \lambda_j^{2(k-1)} \left(\lambda_j - \lambda_n\right)^2} \\
&= \frac{\sum_{j=1}^{n-1} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \lambda_j \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1} - \lambda_n}\right)^2}{\sum_{j=1}^{n-1} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1} - \lambda_n}\right)^2} \\
&= \frac{\lambda_{n-1} + \sum_{j=1}^{n-2} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \lambda_j \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1} - \lambda_n}\right)^2}{1 + \sum_{j=1}^{n-2} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1} - \lambda_n}\right)^2} .
\end{aligned}$$

Proof, Cont.

Notice that, in sums within the expression

$$\frac{\mathbf{y}_{k+1}^{\mathrm{H}}\mathbf{y}_k}{\mathbf{y}_k^{\mathrm{H}}\mathbf{y}_k} = \frac{\lambda_{n-1} + \sum_{j=1}^{n-2} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \lambda_j \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1}-\lambda_n}\right)^2}{1 + \sum_{j=1}^{n-2} \left|\frac{\alpha_j}{\alpha_{n-1}}\right|^2 \left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} \left(\frac{\lambda_j - \lambda_n}{\lambda_{n-1}-\lambda_n}\right)^2}$$

we have the terms

$$\left(\frac{\lambda_j}{\lambda_{n-1}}\right)^{2(k-1)} .$$

It follows that

$$\frac{\mathbf{y}_{k+1}^{\mathrm{H}}\mathbf{y}_k}{\mathbf{y}_k^{\mathrm{H}}\mathbf{y}_k} \to \lambda_{n-1},$$

as $k \to \infty$, since $\left(\frac{\lambda_j}{\lambda_{n-1}}\right)^2 < 1$, for each $j = 1, 2, \ldots, n-2$. $\square$

Reduction to Hessenberg Form

We have seen that it is impossible, in general, to reduce a square matrix to a triangular or diagonal matrix through similarity transformations in a finite number of steps (via a direct method).

We will now see that it is possible to reduce a matrix to a so–called *Hessenberg matrix*, one that has zeros below the first sub–diagonal.

Definition (Hessenberg matrix)

The square matrix $\mathsf{A} = [a_{i,j}] \in \mathbb{C}^{n \times n}$ *is called a* **lower Hessenberg matrix** *or is said to have* **lower Hessenberg form** *iff* $a_{i,j} = 0$ *for all* $1 \le i, j \le n$ *satisfying* $i \ge j + 2$. A *is said to be* **upper Hessenberg** *iff* $\mathsf{A}^{\intercal}$ *is lower Hessenberg.*

Hessenberg Via Similarity Transformation

We want to come up with a similarity transformation that effects the following change, where by $\times$ we denote nonzero entries of the result:

$$
A \to Q^H A Q = \begin{bmatrix} \times & \times & \times & \cdots & \times \\ \times & \times & \times & \ddots & \vdots \\ 0 & \times & \times & \ddots & \times \\ \vdots & \ddots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & \times & \times \end{bmatrix}.
$$

We now explain how this is achieved.

Hessenberg Via Similarity Transformation

Suppose that H_{n-1} is the Householder matrix that leaves the first row of A unchanged and introduces zeros below the second row of the first column:

$$
A \to H_{n-1}^{H} A = \begin{bmatrix} \times & \times & \times & \cdots & \times \\ \otimes & \otimes & \otimes & \cdots & \otimes \\ 0 & \otimes & \otimes & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \otimes \\ 0 & \otimes & \otimes & \cdots & \otimes \end{bmatrix},
$$

where as before, by $\times$ we denote entries of the matrix A that did not change and by $\otimes$ those that did change.

Hessenberg Via Similarity Transformation

When we calculate $\mathsf{H}_{n-1}^{\mathsf{H}}\mathsf{A}\mathsf{H}_{n-1}$ it leaves the first column unchanged, so that we obtain

$$\mathsf{A} \to \mathsf{H}_{n-1}^{\mathsf{H}}\mathsf{A} = \begin{bmatrix} \times & \times & \times & \cdots & \times \\ \otimes & \otimes & \otimes & \cdots & \otimes \\ 0 & \otimes & \otimes & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \otimes \\ 0 & \otimes & \otimes & \cdots & \otimes \end{bmatrix} \to \mathsf{H}_{n-1}^{\mathsf{H}}\mathsf{A}\mathsf{H}_{n-1} \begin{bmatrix} \times & \otimes & \otimes & \cdots & \otimes \\ \otimes & \oplus & \oplus & \cdots & \oplus \\ 0 & \oplus & \oplus & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \oplus \\ 0 & \oplus & \oplus & \cdots & \oplus \end{bmatrix}.$$

The symbol $\oplus$ indicates that the entry has been changed twice. Repeating this idea we can reduce any matrix to Hessenberg form.

If $\mathsf{A} \in \mathbb{C}^{n\times n}$ is Hermitian, then the Hessenberg transformation results in a tridiagonal Hermitian matrix similar to A. We present a result on the existence of this transformation for the Hermitian case. We leave it to the reader to prove the existence in the more general setting.

Theorem (existence)

Suppose that $A \in \mathbb{C}^{n\times n}$ *is Hermitian. There exists a unitary matrix* $Q \in \mathbb{C}^{n\times n}$, *the product of* $n-2$ *Householder matrices*

$$Q = H_{n-1} \cdots H_2,$$

where $H_k \in \mathbb{C}^{n\times n}$ *is a Householder matrix, such that*

$$Q^H A Q = T,$$

where $T \in \mathbb{C}^{n\times n}$ *is a Hermitian tridiagonal matrix.*

Proof.

The proof can be found in the book. $\square$

The QR Method

Definition (QR iteration method)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$. The **QR iteration method** is a recursive algorithm for computing the sequence $\{\mathsf{A}_k\}_{k=0}^{\infty} \subset \mathbb{C}^{n\times n}$ according to the following rules:

1. Set $\mathsf{A}_0 = \mathsf{A}$.
2. For $k = 0, 1, 2, \ldots$, given A_k,
    1. Compute the factorization
    $$\mathsf{Q}_{k+1}\mathsf{R}_{k+1} = \mathsf{A}_k,$$
    where $\mathsf{Q}_{k+1} \in \mathbb{C}^{n\times n}$ is unitary and $\mathsf{R}_{k+1} \in \mathbb{C}^{n\times n}$ is upper triangular.
    2. Define the next iterate as
    $$\mathsf{A}_{k+1} = \mathsf{R}_{k+1}\mathsf{Q}_{k+1}.$$

Recall that every square matrix has a (not necessarily unique) *Schur decomposition*

$$A = Q^H U Q,$$

where Q is unitary and U is upper triangular. We will show that, under suitable assumptions, the QR iteration method computes the factor U in a Schur decomposition of A as the limit of the sequence $\{A_k\}_{k=0}^{\infty}$. To minimize the cost of the QR iteration, first one converts the matrix A to Hessenberg form. This preserves the spectrum, and the QR iteration will preserve this form, among other properties.

Proposition (invariance)

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ and suppose that $\{\mathsf{A}_k\}_{k=0}^{\infty}$ is computed according to the QR iteration method. Then we have:

1. *The matrix A_k is similar to A_{k-1}.*
2. *If A_{k-1} is Hermitian, so is A_k.*
3. *If A_{k-1} is Hessenberg, so is A_k.*
4. *If A_{k-1} is tridiagonal, so is A_k.*

Proof.

Since $\mathsf{R}_k = \mathsf{Q}_k^{\mathsf{H}}\mathsf{A}_{k-1}$ and $\mathsf{A}_k = \mathsf{R}_k\mathsf{Q}_k$, it follows that

$$\mathsf{A}_k = \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{\mathsf{H}}\mathsf{A}_{k-1}\mathsf{Q}_k.$$

Since Q_k is unitary, this shows that A_k is similar to A_{k-1}, and that A_k is Hermitian if A_{k-1} is Hermitian as well.

The last two statements are left for the reader. $\square$

Example

Consider the symmetric, tridiagonal matrix

$$A = \begin{bmatrix} 9 & 17 & 0 & 0 & 0 \\ 17 & 3 & 18 & 0 & 0 \\ 0 & 18 & 20 & 2 & 0 \\ 0 & 0 & 2 & 1 & 8 \\ 0 & 0 & 0 & 8 & 16 \end{bmatrix}.$$

Example (Cont.)

Observe that the QR factorization, $QR = A$, has factors

$$Q = \begin{bmatrix} -0.467888 & 0.533293 & 0.700794 & -0.008036 & -0.074181 \\ -0.883788 & -0.282332 & -0.371009 & 0.004254 & 0.039272 \\ 0 & 0.797425 & -0.600026 & 0.006880 & 0.063515 \\ 0 & 0 & -0.105874 & -0.107091 & -0.988596 \\ 0 & 0 & 0 & -0.994184 & 0.107696 \end{bmatrix}$$

and

$$R = \begin{bmatrix} -19.235384 & -10.605455 & -15.908182 & 0 & 0 \\ 0 & 22.572645 & 10.866537 & 1.594851 & 0 \\ 0 & 0 & -18.890423 & -1.305926 & -0.846990 \\ 0 & 0 & 0 & -8.046802 & -16.763666 \\ 0 & 0 & 0 & 0 & -6.185635 \end{bmatrix},$$

to six decimal digits of precision. Observe that we have not demanded that R has positive diagonal components in this factorization, though that can be easily remedied. In any case, note the placement of the zeros in the factor matrices.

Example (Cont.)

Recall that, in the QR algorithm, $A_1 = RQ$, and we obtain

$$A_1 = \begin{bmatrix} 18.372973 & -19.949431 & 0 & 0 & 0 \\ -19.949431 & 2.292279 & -15.063703 & 0 & 0 \\ 0 & -15.063703 & 11.473011 & 0.851945 & 0 \\ 0 & 0 & 0.851945 & 17.527905 & 6.149659 \\ 0 & 0 & 0 & 6.149659 & -0.666167 \end{bmatrix},$$

again showing six decimal digits of precision. The structure of A is preserved in this first step of QR algorithm, namely, A_1 is symmetric and tridiagonal, as predicted by our theory.

Theorem (convergence)

Suppose that $\mathsf{A} \in \mathbb{C}^{n \times n}$ *is invertible and all its eigenvalues are distinct in modulus, that is*

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0.$$

Let $\mathsf{P} \in \mathbb{C}^{n \times n}$ *be an invertible matrix such that*

$$\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^{-1},$$

where $\mathsf{D} = \mathbf{diag}(\lambda_1, \lambda_2 \ldots, \lambda_n)$*. If* P^{-1} *has an LU factorization — i.e., there exists a unit lower triangular matrix* $\mathsf{L} \in \mathbb{C}^{n \times n}$ *and an upper triangular matrix* $\mathsf{U} \in \mathbb{C}^{n \times n}$*, such that* $\mathsf{P}^{-1} = \mathsf{L}\mathsf{U}$ *— then the sequence of matrices* $\{\mathsf{A}_k\}_{k=1}^{\infty}$ *produced by the QR iteration method is such that*

$$\lim_{k \to \infty} [\mathsf{A}_k]_{i,i} = \lambda_i, \quad 1 \le i \le n,$$

and

$$\lim_{k \to \infty} [\mathsf{A}_k]_{i,j} = 0, \quad 1 \le j < i \le n.$$

Theorem (convergence II)

Suppose that the QR iteration method is applied to a Hermitian matrix $\mathsf{A} \in \mathbb{C}^{n\times n}$*, whose eigenvalues satisfy*

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

and whose corresponding unitary eigenvector matrix Q *has all nonsingular leading principal sub–matrices. Then, as* $k \to \infty$*, the matrices* A_k *converge linearly, with constant*

$$\max_j \left| \frac{\lambda_{j+1}}{\lambda_j} \right|,$$

to $\Lambda = \mathsf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *and* Q_k *converges, with the same rate, to* Q*.*

Theorem (convergence III)

Suppose that $\mathsf{A} \in \mathbb{C}^{n\times n}$ *is HPD. Then the sequence of matrices* $\{\mathsf{A}_k\}_{k=1}^{\infty}$ *produced by the QR iteration converges to a diagonal matrix* $\mathsf{D} = \mathbf{diag}(\lambda_n, \ldots, \lambda_1)$, *where*

$$0 < \lambda_1 \leq \cdots \leq \lambda_n,$$

are the eigenvalues of A.