Chapter 6

NUMERICAL CALCULATION OF EIGENVALUES

6.1 STABILITY OF THE NUMERICAL EIGENVALUE PROBLEM

In this chapter, this problem is considered.

> Given $A_{n \times n}$ with real entries, find numerical approximations to the eigenvalues and eigenvectors of A.

We outline two methods of solution for this problem in Secs. 6.2 and 6.3. However, there is a warning to be heeded.

> If A and B are almost equal, the eigenvalues of A and B need **not** be almost equal.

Another way of saying this is as follows:

> Small changes in A do **not** necessarily lead to small changes in the eigenvalues of A.

These are important points, since the computer stored form of A may not be exactly equal to A.

To illustrate, consider the matrices

$$A = \begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1000 \\ 0.001 & 1 \end{pmatrix}$$

Matrix A has eigenvalues 1 and 1, and B has eigenvalues 0 and 2. Even though A differs from B by only 0.001 in the row 2, column 1 entry, the eigenvalues differ by 1. That is, a change of only 0.001 in one entry of A led to a 100 percent change in the eigenvalue.

For a more extreme example, consider matrices

$$Z = \begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & 1000 \\ -0.001 & 1 \end{pmatrix}$$

Matrix C differs from A by only 0.001 in the second row and first column. The eigenvalues of A are 1 and 1, but C has no real eigenvalues at all since the characteristic polynomial is $2 - 2\lambda + \lambda^2$.

What these examples tell us is that when error enters into the calculation of the entries of a matrix A and we later try to calculate the eigenvalues of A, we must view the results carefully. But if A is symmetric, small changes in A will generally not lead to large changes in the eigenvalues. So in applications involving symmetric matrices, numerical methods are generally quite successful in computing the needed eigenvalues.[1]

To make a concrete statement, we define the **Frobenius norm** of a matrix A as

$$\|A\|_F = \sqrt{\sum_{1 \le i,j \le n} |a_{ij}|^2}$$

STABILITY THEOREM. Let A be an $n \times n$ matrix, and let E be an $n \times n$ "error matrix." Suppose that A and E are real and symmetric, and set $\hat{A} = A + E$ (that is, $\hat{A}$ is the "error version" of A). Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A and $\hat{\lambda}_1, \dots, \hat{\lambda}_n$ be the eigenvalues of $\hat{A}$. Then

$$(\lambda_1 - \hat{\lambda}_1)^2 + (\lambda_2 - \hat{\lambda}_2)^2 + \cdots + (\lambda_n - \hat{\lambda}_n)^2 \le \|E\|_F^2$$

---

[1] There exist certain matrices which can be used to test numerical methods. If a method performs well on these matrices, confidence in the method increases. See, for example, **A Collection of Matrices for Testing Computational Algorithms** by R. Gregory and D.L. Karney. Wiley, NY, 1969.

STABILITY COROLLARY. With the same hypotheses of the stability theorem,

$$|\lambda_k - \hat{\lambda}_k| \le \|E\|_F$$

for $k = 1, 2, \ldots, n$. This means that the process of finding $\lambda_1, \ldots, \lambda_n$ is **stable**: Small errors in A ($\|E\|_F$ small) lead to small errors in determination of $\lambda_1, \ldots, \lambda_n(|\lambda_k - \hat{\lambda}_k|$ small).

We cannot prove the theorem, but the corollary is obtained by noting that by the theorem

$$(\lambda_k - \hat{\lambda}_k)^2 \le (\lambda_1 - \hat{\lambda}_1)^2 + \cdots + (\lambda_k - \hat{\lambda}_k)^2 + \cdots + (\lambda_n - \hat{\lambda}_n)^2 \le \|E\|_F^2$$

Thus

$$\sqrt{(\lambda_k - \hat{\lambda}_k)^2} \le \|E\|_F$$

This corollary tells us that if $\|E\|_F$ is less than or equal to a small number ε then all the computed eigenvalues $\hat{\lambda}_k$ of $A + E$ will be within ε of the true eigenvalues λ_k of A. That is, small error in real symmetric A leads to small absolute error in a calculation of the eigenvalues. Our initial example in this section does not involve a symmetric matrix.

Example 1. Suppose we need to calculate the eigenvalues of

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 7 \\ 2 & 7 & 5 \end{pmatrix}$$

but because of errors in calculation of the entries of A, the computer works with

$$\hat{A} = \begin{pmatrix} 1.01 & -1.05 & 2.1 \\ -1.05 & 1.97 & 7.1 \\ 2.1 & 7.1 & 4.9 \end{pmatrix}$$

What will be the maximum difference in the calculated eigenvalues and the eigenvalues of A?

Solution We have $\hat{A} = A + E$, so $E = \hat{A} - A$. Therefore

$$E = \begin{pmatrix} 0.01 & -0.05 & 0.1 \\ -0.05 & -0.03 & 0.1 \\ 0.1 & 0.1 & -0.1 \end{pmatrix}$$

Now

$$\|E\|_F = \sqrt{(0.01)^2 + 2(-0.05)^2 + 4(0.1)^2 + (0.03)^2 + (-0.1)^2}$$
$$= 0.23664\cdots$$

so we are guaranteed that

$$|\lambda_k - \hat{\lambda}_k| \le 0.23664$$

The maximum error is 0.23664.

Example 2. Suppose that we want to calculate the eigenvalues of

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

in a computer which keeps only six significant digits. What is a bound for the error in calculating the eigenvalues of A, assuming an error of 0.000001 in each entry of A?

Solution Suppose that

$$E = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$$

where $\varepsilon = 0.000001$. Then by the stability corollary

$$|\lambda - \hat{\lambda}_k| \le \sqrt{9 \cdot (0.000001)^2} = 0.000003$$

Of course, this estimate neglects errors introduced by calculations used in a particular method of computing eigenvalues.

The main point has been made: In general, small errors in A need **not** lead to small errors in computation of $\lambda_1, \ldots, \lambda_n$; however, if A is symmetric, with symmetrically distributed error, then small errors in A **do** lead to small absolute errors in determination of $\lambda_1, \ldots, \lambda_n$.

Finally, even if the absolute error is small

$$|\lambda_k - \hat{\lambda}_k| \le \varepsilon$$

where ε is small, the **relative** error could be large. For instance, if $\lambda_k \doteq \varepsilon/2$, then

$$\frac{|\lambda_k - \hat{\lambda}_k|}{|\lambda_k|} \doteq 200\%$$

PROBLEMS 6.1

1. Consider the matrices

$$
H_{10\times10} = \begin{pmatrix}
1 & & & & & \mathbf{0} \\
1 & 1 & & & & \\
0 & 1 & 1 & & & \\
\vdots & \ddots & \ddots & \ddots & & \\
0 & \cdots & 0 & 1 & 1
\end{pmatrix} \quad \text{and} \quad H + E
$$

where

$$
E_{10\times10} = \begin{pmatrix}
0 & 0 & \cdots & 0 & \frac{1}{2^{10}} \\
0 & 0 & \cdots & 0 & 0 \\
\hdotsfor{5} \\
0 & 0 & \cdots & 0 & 0
\end{pmatrix}
$$

(a) What are the eigenvalues of H?

(b) Show that $\lambda = \frac{1}{2}$ is an eigenvalue of $H + E$.

(c) Show that $\|E\|_F = 1/2^{10}$.

(d) Why does the stability corollary not apply to H and $H + E$?

2. Calculate the Frobenius norm for the following matrices.

(a) $D = \begin{pmatrix}
d_1 & & & & \mathbf{0} \\
& d_2 & & & \\
& & \ddots & & \\
\mathbf{0} & & & & d_n
\end{pmatrix}$ (b) I (c) 0 (zero matrix)

(d) $D + I$

3. Regarding Prob. 2, which is larger?

$$
\|D + I\|_F \quad \text{or} \quad \|D\|_F + \|I\|_F
$$

4. Let A be an $n \times n$ matrix. If $\|A\|_F = 0$, must A be the zero matrix?

5. Let $A = (a_{ij})_{n\times n}$. Define the 1 norm of A by

$$
\|A\|_1 = \sum_{1 \le i,j \le n} |a_{ij}|
$$

Let

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix}$$

Calculate $\|A\|_1$ and $\|A\|_F$. Which norm is larger?

6. Suppose the eigenvalues of an $n \times n$ symmetric matrix A are to be computed. Because of a data entry error, every entry of A has 0.0001 added to it. What is the error bound for $|\lambda_k - \hat{\lambda}_k|$, as given in the stability corollary? How does the error bound change as n increases? What can you say about the stability of the eigenvalue problem for large versus small matrices?

6.2   POWER METHOD

The problem we are considering is this: Given an $n \times n$ real matrix A, find numerical approximations to the eigenvalues and eigenvectors of A. This **numerical eigenproblem** is difficult to solve in general. In many applications, A may be symmetric, or tridiagonal or have some other special form or property. Consequently, most numerical methods are designed for special matrices.

The **power method**, the subject of this section, can be used when

1. $A_{n \times n}$ has n linearly independent eigenvectors   (6.2.1)
2. The eigenvalues can be ordered in magnitude as

$$|\lambda_1| \underset{\uparrow}{>} |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \qquad (6.2.2)$$

<div align="center">Note the strict inequality</div>

When this ordering can be done, λ_1 is called the **dominant** eigenvalue of A.

Example 1. Let A have eigenvalues 2, 5, 0, -7, and -2. Does A have a dominant eigenvalue? If so, which is dominant?

Solution Since

$$|-7| > |5| > |2| \geq |-2| > |0|$$

A has a dominant eigenvalue of $\lambda_1 = -7$.

Now suppose A satisfies condition (6.2.1) and (6.2.2). What must be true about A? If X_0 is any vector, it can be written as

$$X_0 = c_1 V_1 + c_2 V_2 + \cdots + c_n V_n$$

where $\{V_1, \ldots, V_n\}$ is the set of n linearly independent eigenvectors. Thus

$$AX_0 = c_1\lambda_1 V_1 + c_2\lambda_2 V_2 + \cdots + c_n\lambda_n V_n$$
$$A^2 X_0 = c_1\lambda_1{}^2 V_1 + c_2\lambda_2{}^2 V_2 + \cdots + c_n\lambda_n{}^2 V_n$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$A^m x_0 = c_1\lambda_1{}^m V_1 + c_2\lambda_2{}^m V_2 + \cdots + c_n\lambda_n{}^m V_n$$

And if we divide the last equation by $\lambda_1{}^m$, we have

$$\frac{1}{\lambda_1{}^m} A^m X_0 = c_1 V_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^m V_1 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^m c_n V_n \qquad (6.2.3)$$

As m gets larger and larger, the terms $(\lambda_1/\lambda_1)^m, \ldots, (\lambda_n/\lambda_1)^m$ all get closer and closer to zero (remember: $|\lambda_2/\lambda_1| < 1, \ldots, |\lambda_n/\lambda_1| < 1$). Therefore, for large m

$$\frac{1}{\lambda_1{}^m} A^m X_0 \doteq c_1 V_1 \qquad (6.2.4)$$

As long as $c_1 \neq 0$, this last equation can give us an approximation to λ_1. (To guarantee that $c_1 \neq 0$, X_0 should not be orthogonal to V_1.) To obtain the approximation, note that in addition to (6.2.4) we would have, since $m + 1 > m$, that

$$\frac{1}{\lambda_1{}^{m+1}} A^{m+1} X_0 \doteq c_1 V_1 \qquad (6.2.5)$$

Now taking the dot product of both sides of (6.2.4) and (6.2.5) with any Y which is not orthogonal to V_1, we have

$$\frac{1}{\lambda_1{}^m} (A^m X_0 \cdot Y) \doteq c_1 V_1 \cdot Y$$

$$\frac{1}{\lambda_1^{m+1}} (A^{m+1} X_0 \cdot Y) \doteq c_1 V_1 \cdot Y$$

and so

$$\frac{1}{\lambda_1^{m+1}} (A^{m+1} X_0 \cdot Y) \doteq \frac{1}{\lambda_1{}^m} (A^m X_0 \cdot Y) \neq 0$$

Finally, by dividing we have

$$\frac{A^{m+1} X_0 \cdot Y}{A^m X_0 \cdot Y} \doteq \frac{\lambda_1^{m+1}}{\lambda_1{}^m} = \lambda_1$$

The powers of A give the power method its name.

Power Method

Suppose that A has n linearly independent eigenvectors and has a dominant eigenvalue λ_1 with eigenvector V_1. Let Y be any vector not orthogonal to V_1. Then for large m, if X_0 is not orthogonal to V_1,

$$\frac{A^{m+1}X_0 \cdot Y}{A^m X_0 \cdot Y} \doteq \lambda_1 \qquad (6.2.6)$$

Clearly the concepts of linear independence and basis were essential to showing that the power method works. Once again, the theory of linear algebra supplies the foundation for a numerical method.

Example 2. Use the power method to estimate the largest eigenvalue of

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}$$

Solution Let

$$X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

We compute $AX_0, A^2 X_0$, and so on.[2]

$$AX_0 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \qquad\qquad A^4 X_0 = A\begin{pmatrix} 34 \\ 36 \end{pmatrix} = \begin{pmatrix} 142 \\ 140 \end{pmatrix}$$

$$A^2 X_0 = A\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \end{pmatrix} \qquad A^5 X_0 = A\begin{pmatrix} 142 \\ 140 \end{pmatrix} = \begin{pmatrix} 562 \\ 564 \end{pmatrix}$$

$$A^3 X_0 = A\begin{pmatrix} 10 \\ 8 \end{pmatrix} = \begin{pmatrix} 34 \\ 36 \end{pmatrix} \qquad A^6 X_0 = A\begin{pmatrix} 562 \\ 564 \end{pmatrix} = \begin{pmatrix} 2254 \\ 2252 \end{pmatrix}$$

Now let

$$Y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

---

[2]Note that the components of the computed vectors are getting larger. To overcome this, one can modify the power method to the power method with scaling, which is outlined in Prob. 6.

In Eq. (6.2.6) with $m = 5$:

$$\lambda_1 \doteq \frac{A^6 X_0 \cdot Y}{A^5 X_0 \cdot Y} = \frac{\begin{pmatrix} 2254 \\ 2252 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 562 \\ 564 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \frac{2254}{562} = 4.0106 \cdots$$

Checking by the methods of Chap. 5, we find

$$\left(4, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$$

is the dominant eigenpair; the approximation to λ_1 has absolute error $=$ $0.0106 \cdots$, which is less than 0.3 percent relative error.

In Example 2 the vectors $A^m X_0$ appear to be almost parallel to the eigenvector which corresponds to λ_1. In fact, this will always be the case. To see this, first we note that for any eigenpair (λ, V) of A we have

$$AV = \lambda V$$

So if we take the dot product of both sides with V, we see that

$$\frac{AV \cdot V}{V \cdot V} = \lambda \tag{6.2.7}$$

We can choose Y in Eq. (6.2.6) as

$$Y = A^m X_0$$

to find

$$\lambda_1 \doteq \frac{A^{m+1} X_0 \cdot A^m X_0}{A^m X_0 \cdot A^m X_0} \doteq \frac{A(A^m X_0) \cdot A^m X_0}{A^m X_0 \cdot A^m X_0}$$

The last expression matches Eq. (6.2.7) with $V = A^m X_0$. Therefore, it is reasonable to believe that $A^m X_0$ is (**approximately**) **an eigenvector corresponding to λ_1**.

Example 3. Use the power method to calculate an approximation to the dominant eigenpair for

$$A = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix}$$

Solution Let

$$X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Then

$$A(X_0) = \begin{pmatrix} -7 \\ 8 \end{pmatrix}$$

$$A^2(X_0) = \begin{pmatrix} 65 \\ -64 \end{pmatrix}$$

$$A^3(X_0) = \begin{pmatrix} -583 \\ 584 \end{pmatrix}$$

We stop here because the vectors $A^m X_0$ already appear to be approaching a multiple of

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Now

$$\lambda_1 \doteq \frac{A^3 X_0 \cdot A^2 X_0}{A^2 X_0 \cdot A^2 X_0} = \frac{-583 \cdot 65 - 584 \cdot 64}{65 \cdot 65 + 64 \cdot 64} = -9.05$$

So a dominant eigenpair is (approximately)

$$\left(-9.05, \frac{1}{583} \begin{pmatrix} -583 \\ 584 \end{pmatrix} \right) = \left(-9.05, \begin{pmatrix} -1 \\ 1.002 \end{pmatrix} \right)$$

Since the actual dominant eigenpair is

$$\left(-9, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right)$$

the method has worked well in this example.

Although the power method has worked well in these examples, we must say something about cases in which the power method may fail. There are basically three such cases:

1. Using the power method when A is not diagonalizable. Recall that A has n linearly independent eigenvectors **if and only if** A is diagonalizable. Of course, it is not easy to tell by just looking at A whether it is diagonalizable.

2. Using the power method when A does not have a dominant eigenvalue, or when the dominant eigenvalue is such that

$$|\lambda_1| > |\lambda_2| \quad \text{but} \quad |\lambda_1| \doteq |\lambda_2|$$

(Then $|\lambda_1/\lambda_2|$ is barely less than 1, and high powers of $|\lambda_1/\lambda_2|$ do not tend to zero quickly.) Again, it is not easy to determine whether A has this defect by just looking at A.

3. If the entries of A contain significant error. Powers A^m of A will have significant roundoff error in their entries.

Here is a rule of thumb for using the power method:

1. Try it, and if the numbers

$$\frac{A^{m+1}X_0 \cdot A^m X_0}{A^m X_0 \cdot A^m X_0}$$

approach a single number λ_1, then stop and go to step 2.

2. Check whether $(\lambda_1, A^m X_0)$ is an eigenpair by checking whether

$$A(A^m X_0) \doteq \lambda_1 (A^m X_0)$$

3. If step 2 checks, accept
$$(\lambda_1, A^m X_0)$$
as a dominant eigenpair.

Example 4. Check the answer to Example 3.

Solution The proposed eigenpair is

$$\left(-9.05, \left(\begin{array}{c} -1 \\ 1.005 \end{array} \right) \right)$$

To check, we calculate

$$A \left(\begin{array}{c} -1 \\ 1.005 \end{array} \right) \quad \text{and} \quad -9.05 \left(\begin{array}{c} -1 \\ 1.005 \end{array} \right)$$

We have

$$A \begin{pmatrix} -1 \\ 1.005 \end{pmatrix} = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1.005 \end{pmatrix} = \begin{pmatrix} 9.01 \\ -9.005 \end{pmatrix}$$

and

$$-9.05 \begin{pmatrix} -1 \\ 1.005 \end{pmatrix} = \begin{pmatrix} 9.05 \\ -9.09525 \end{pmatrix}$$

Thus

$$A \begin{pmatrix} -1 \\ 1.005 \end{pmatrix} - (-9.05) \begin{pmatrix} -. \\ 1.005 \end{pmatrix} = \begin{pmatrix} -0.04 \\ 0.09025 \end{pmatrix} \doteq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so

$$A \begin{pmatrix} -1 \\ 1.005 \end{pmatrix} \doteq -9.05 \begin{pmatrix} -1 \\ 1.005 \end{pmatrix}$$

and the answer checks.

To illustrate possible failure of the power method, we show an example.

Example 5. Try the power method on

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

with

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad X_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Explain the results.

Solution Let

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Then we have

$$AX_0 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$A^2 X_0 = A \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$A^3 X_0 = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$A^4 X_0 = A \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\vdots$$

We see that

$$A^{2n}X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad A^{2n+1}X_0 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

so that $A^m X_0$ is not becoming parallel to any vector. Also

$$\frac{A^{2n+1}X_0 \cdot A^{2n}X_0}{A^{2n}X_0 \cdot A^{2n}X_0} = \frac{\begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \frac{1}{2}$$

and

$$\frac{A^{2n}X_0 \cdot A^{2n-1}X_0}{A^{2n-1}X_0 \cdot A^{2n-1}X_0} = \frac{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}}{\begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}} = \frac{1}{5}$$

So we have no approximation to the eigenvalues. The power method has failed when

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

When

$$X_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

We find

$$AX_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A^2 X_0 = A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$A^3 X_0 = A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A^4 X_0 = A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\vdots$$

and for the same reasons as before, the power method fails. Also in this case the approximations to λ_1 oscillate between $-\frac{1}{2}$ and -1.

An explanation for the failure is that A has no dominant eigenvalue. In fact, the eigenvalues are 1 and -1, which have the same magnitude.

When to Stop in the Power Method   We would like to have a rule about when to stop in using the power method. Usually we would stop when $|\lambda_1^{\text{calc}} - \lambda_1^{\text{actual}}|$ is small. However, in most realistic problems we do not know $\lambda_1^{\text{actual}}$, so we can only estimate $|\lambda_1^{\text{calc}} - \lambda_1^{\text{actual}}|$. This can be done for symmetric A.

Theorem 6.2.1. *Let A be real symmetric with dominant eigenvalue λ_1. Then if $\lambda_1^{\text{calc}} = (AX \cdot X)/(X \cdot X)$, where $X = A^m X_0$ as in the power method, we have*

$$|\lambda_1^{\text{calc}} - \lambda_1^{\text{actual}}| \leq \sqrt{\frac{AX \cdot AX}{X \cdot X} - (\lambda_1^{\text{calc}})^2}$$

Example 6. Apply the power method to the symmetric matrix

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} \quad \text{with} \quad X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Estimate the error in using

$$\frac{A^6 X_0 \cdot A^5 X_0}{A^5 X_0 \cdot A^5 X_0}$$

for λ. (Eigenvalues are 4 and 9.)

Solution   With

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

we find

$$AX_0 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

$$A^2X_0 = A\begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 42 \end{pmatrix}$$

$$A^3X_0 = A\begin{pmatrix} 3 \\ 42 \end{pmatrix} = \begin{pmatrix} -69 \\ 330 \end{pmatrix}$$

$$A^4X_0 = A\begin{pmatrix} -69 \\ 330 \end{pmatrix} = \begin{pmatrix} -1005 \\ 2778 \end{pmatrix}$$

$$A^5X_0 = A\begin{pmatrix} -1005 \\ 2778 \end{pmatrix} = \begin{pmatrix} -10,581 \\ 24,234 \end{pmatrix}$$

$$A^6X_0 = A\begin{pmatrix} -10,581 \\ 24,234 \end{pmatrix} = \begin{pmatrix} -101,373 \\ 215,034 \end{pmatrix} = \frac{1}{101,373}\begin{pmatrix} -1 \\ 2.1212157 \end{pmatrix}$$

Now using A^6X_0, we have

$$\lambda_1 \doteq \frac{A^6X_0 \cdot A^5X_0}{A^5X_0 \cdot A^5X_0} \doteq 8.9865$$

The error estimate is

$$|8.9865 - \lambda_1^{\text{actual}}| \leq \sqrt{\frac{A^6X_0 \cdot A^6X_0}{A^5X_0 \cdot A^5X_0} - (8.9865)^2}$$

$$\doteq \sqrt{\frac{565.16106 \times 10^8}{6.9924435 \times 10^8} - (8.9865)^2}$$

$$= \sqrt{80.824544 - 80.7587182}$$

$$= \sqrt{0.067362}$$

$$\doteq 0.26$$

From the error estimate we know our error is at most 0.26. In practice, we would not stop here because the percentage error could be

$$\frac{\text{Maximum absolute error}}{\text{Calculated value}} = \frac{0.26}{8.9865} \doteq 2.9\%$$

which is fairly large. Ordinarily, we would continue calculations until the error estimate were smaller than some preset tolerance.

We should make two observations concerning Example 6. First, the error estimate is **pessimistic**. The estimate is 0.26, but 8.9865 is only 0.0135 from the actual dominant eigenvalue of 9. That is, the approximation is much closer to the actual eigenvalue than the error estimate predicts. Because of this, another check for accuracy (although not totally reliable) is to calculate the relative error

$$E_{n+1} = \frac{|\lambda_1^{\text{calc at step } n} - \lambda_1^{\text{calc at step } n+1}|}{|\lambda_1^{\text{calc at step } n+1}|} \tag{6.2.8}$$

and stop if E is small. In fact, this second check must be used for non-symmetric A since the estimate in Theorem 6.2.1 is for symmetric matrices only.

Second, the components of the approximate eigenvectors get quite large. To overcome this problem, we can scale the approximate eigenvectors at each stage by multiplying the approximate eigenvector by the reciprocal of the largest (in absolute value) component of the approximate eigenvector, then use the scaled vector in the next step (see Prob. 6). If we repeat Example 6 with this scaling process, we obtain

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Step 1: $\qquad AX_0 = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \xrightarrow{\text{Scale}} \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} = W_1$

Step 2: $\qquad AW_1 = \begin{pmatrix} 0.5 \\ 7 \end{pmatrix} \xrightarrow{\text{Scale}} \begin{pmatrix} \frac{1}{14} \\ 1 \end{pmatrix} = W_2$

Step 3: $\qquad AW_2 = \begin{pmatrix} -1.64\cdots \\ 7.85\cdots \end{pmatrix} \xrightarrow{\text{Scale}} \begin{pmatrix} -0.209\cdots \\ 1 \end{pmatrix} = W_3$

$$\vdots$$

Step 10: $\qquad \xrightarrow{\text{Scale}} \begin{pmatrix} -0.4988\cdots \\ 1 \end{pmatrix} = W_{10}$

Step 11: $\qquad AW_{10} = \begin{pmatrix} -4.49\cdots \\ 8.997\cdots \end{pmatrix} \qquad\qquad \begin{pmatrix} -0.4994\cdots \\ 1 \end{pmatrix}$

The approximate eigenvalue is (see Prob. 6)

$$\frac{AW_{10} \cdot W_{10}}{W_{10} \cdot W_{10}} \doteq 9.002$$

The approximate eigenvector as calculated above is close to the actual (any multiple of)

$$\begin{pmatrix} -0.5 \\ 1 \end{pmatrix}$$

Example 7. Calculate E_{n+1} [from Eq. (6.2.8)] in Example 2, for $n = 3, 4, 5$.

Solution We have

$$\lambda_1^{\text{step } 3} = 3.829268\cdots \qquad \lambda_1^{\text{step } 5} = 3.9926566\cdots$$
$$\lambda_1^{\text{step } 4} = 4.0244698\cdots \qquad \lambda_1^{\text{step } 6} = 4.0017604\cdots$$

Therefore

$$E_4 = \frac{|\lambda_1^{\text{step } 3} - \lambda_1^{\text{step } 4}|}{|\lambda_1^{\text{step } 4}|} \doteq 0.04858037 \doteq 4.9\%$$

$$E_5 = \frac{|\lambda_1^{\text{step } 4} - \lambda_1^{\text{step } 5}|}{|\lambda_1^{\text{step } 5}|} \doteq 0.0079679 \doteq 0.8\%$$

$$E_6 = \frac{|\lambda_1^{\text{step } 5} - \lambda_1^{\text{step } 6}|}{|\lambda_1^{\text{step } 6}|} \doteq 0.0022749 \doteq 0.023\%$$

If we had decided to accept $\lambda_1^{\text{step } n}$ as the approximate eigenvalue when $E_{n+1} < 0.005$, then we would accept 4.0017604 as λ_1. But if we had initially decided to accept $\lambda_1^{\text{step } n}$ when $E_{n+1} < 0.01$, then we would accept 3.9926566.

Finding Nondominant Eigenvalues Once the dominant eigenpair (λ_1, V_1) of A is computed, we may wish to compute λ_2. (Recall that $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$.) If A is symmetric, it can be proved that if $U_1 = V_1/|V_1|$, then

$$\mathcal{A} = A - \lambda_1 U_1 U_1^T$$

has eigenvalues $0, \lambda_2, \lambda_3, \ldots, \lambda_n$ and the eigenvectors of $\mathcal{A}$ are eigenvectors of A. Therefore, to find λ_2, we could apply the power method to $\mathcal{A}$. However, a **warning** is in order. Since λ_1 is not exact, some error will be introduced in the power method applied to $\mathcal{A}$. The application of the power method to $\mathcal{A}$ to find λ_2 is called the **method of deflation**.

Example 8. Apply the method of deflation to the matrix of Example 6 to find λ_2. Assume that $\lambda_1 = 9$ and

$$V_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Solution We form $\mathcal{A}$:

$$\mathcal{A} = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} - 9 \begin{pmatrix} \dfrac{1}{\sqrt{5}} \\ \dfrac{-2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \dfrac{1}{\sqrt{5}} & \dfrac{-2}{\sqrt{5}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{16}{5} & \frac{8}{5} \\ \frac{8}{5} & \frac{4}{5} \end{pmatrix} = \frac{4}{5} \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$$

Now, applying the power method to $\mathcal{A}$ with

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

we have

$$\mathcal{A}X_0 = \frac{4}{5} \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$

$$\mathcal{A}^2 X_0 = \frac{4}{5} \mathcal{A} \begin{pmatrix} 6 \\ 3 \end{pmatrix} = \frac{16}{25} \begin{pmatrix} 30 \\ 15 \end{pmatrix}$$

$$\mathcal{A}^3 X_0 = \frac{16}{25} \mathcal{A} \begin{pmatrix} 30 \\ 15 \end{pmatrix} = \frac{64}{125} \begin{pmatrix} 150 \\ 75 \end{pmatrix}$$

Clearly, the vectors generated by the power method are all multiples of

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Also

$$\lambda_2^{\text{step } 2} = \frac{\dfrac{16}{25} \begin{pmatrix} 30 \\ 15 \end{pmatrix} \cdot \dfrac{4}{5} \begin{pmatrix} 6 \\ 3 \end{pmatrix}}{\dfrac{4}{5} \begin{pmatrix} 6 \\ 3 \end{pmatrix} \cdot \dfrac{4}{5} \begin{pmatrix} 6 \\ 3 \end{pmatrix}} = 4$$

$$\lambda_2^{\text{step } 3} = \frac{\dfrac{65}{125} \begin{pmatrix} 150 \\ 75 \end{pmatrix} \cdot \dfrac{16}{25} \begin{pmatrix} 30 \\ 15 \end{pmatrix}}{\dfrac{16}{25} \begin{pmatrix} 30 \\ 15 \end{pmatrix} \cdot \dfrac{16}{25} \begin{pmatrix} 30 \\ 15 \end{pmatrix}} = 4$$

Therefore, the first (and only, since A is 2×2) nondominant eigenvalue for A is 4, with eigenvector

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

In fact, this is exactly correct.

Example 9. Rework Example 8 with λ_1 and V_1 **as calculated in Example 6**. Compare the results with Example 8.

Solution We have $\lambda_1 = 8.987$ and

$$V_1 = \begin{pmatrix} -1 \\ 2.12 \end{pmatrix}$$

Now

$$U_1 = \begin{pmatrix} -1 \\ 2.12 \end{pmatrix} \Big/ \left| \begin{pmatrix} -1 \\ 2.12 \end{pmatrix} \right| = \frac{1}{2.344} \begin{pmatrix} -1 \\ 2.12 \end{pmatrix} = \begin{pmatrix} -0.427 \\ 0.904 \end{pmatrix}$$

so

$$
\begin{aligned}
\mathcal{A} &= \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} - 8.987 \begin{pmatrix} -0.427 \\ 0.904 \end{pmatrix} (-0.427 \quad 0.904) \\
&= \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} - 8.987 \begin{pmatrix} 0.182 & -0.386 \\ -0.386 & 0.817 \end{pmatrix} \\
&= \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} - \begin{pmatrix} 1.636 & -3.469 \\ -3.469 & 7.342 \end{pmatrix} = \begin{pmatrix} 3.364 & 1.469 \\ 1.469 & 0.658 \end{pmatrix}
\end{aligned}
$$

Now we apply the power method with

$$X_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

to find

$$\mathcal{A}X_0 = \begin{pmatrix} 4.833 \\ 2.127 \end{pmatrix}$$

$$\mathcal{A}^2 X_0 = \mathcal{A} \begin{pmatrix} 4.833 \\ 2.127 \end{pmatrix} = \begin{pmatrix} 19.383 \\ 8.5 \end{pmatrix} = \frac{1}{8.5} \begin{pmatrix} 2.28 \\ 1 \end{pmatrix}$$

$$\mathcal{A}^3 X_0 = \mathcal{A} \begin{pmatrix} 19.383 \\ 8.5 \end{pmatrix} = \begin{pmatrix} 77.691 \\ 34.066 \end{pmatrix} = \frac{1}{34.066} \begin{pmatrix} 2.28 \\ 1 \end{pmatrix}$$

$$\mathcal{A}^4 X_0 = \mathcal{A} \begin{pmatrix} 77.691 \\ 34.066 \end{pmatrix} = \begin{pmatrix} 311.4 \\ 136.54 \end{pmatrix} = \frac{1}{136.54} \begin{pmatrix} 2.28 \\ 1 \end{pmatrix}$$

Since the vectors generated seem to be fixed at multiples of

$$\begin{pmatrix} 2.28 \\ 1 \end{pmatrix}$$

we know that

$$\lambda_2^{\text{step 3}} = \frac{\mathcal{A}^3 X_0 \cdot \mathcal{A}^2 X_0}{\mathcal{A}^2 X_0 \cdot \mathcal{A}^2 X_0} = 4.00813$$

$$\lambda_2^{\text{step 4}} = \frac{\mathcal{A}^4 X_0 \cdot \mathcal{A}^3 X_0}{\mathcal{A}^3 X_0 \cdot \mathcal{A}^3 X_0} = 4.00817$$

The relative error in the eigenvalue is very small, and we will accept

$$\left(4.008, \begin{pmatrix} 2.28 \\ 1 \end{pmatrix} \right)$$

as the first nondominant eigenpair of A. However, the eigenvector is inaccurate because

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

is correct. The results are less accurate than in Example 8 because we started with inaccurate λ_1 and V_1.

Example 10. Rework Example 9, using $\lambda_1 = 9.002$ and

$$V_1 = \begin{pmatrix} -0.4994 \\ 1 \end{pmatrix}$$

as found in the calculations following Example 6.

Solution This time,

$$U_1 = \begin{pmatrix} -0.4467 \\ 0.8946 \end{pmatrix} \quad \text{and} \quad \mathcal{A} = \begin{pmatrix} 3.203 & 1.597 \\ 1.597 & 0.795 \end{pmatrix}$$

After four steps, beginning with

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

we find $\lambda_2 \doteq 3.997$ and

$$V_2 \doteq \begin{pmatrix} 1 \\ 0.498 \end{pmatrix}$$

The calculated approximations are more accurate this time because we started with more accurate values for λ_1 and V_1.

For an $n \times n$ matrix (symmetric), to find λ_3 by deflation, we proceed as follows. After finding (λ_2, V_2), we form

$$\mathcal{B} = \mathcal{A} - \lambda_2 U_2 U_2{}^T$$

where $U_2 = V_2/|V_2|$. Then $\mathcal{B}$ will have eigenvalues $0, 0, \lambda_3, \lambda_4, \ldots, \lambda_n$, and the eigenvalues of $\mathcal{B}$ will be eigenvectors of A. The power method applied to $\mathcal{B}$ will yield λ_3 as the dominant eigenvalue of $\mathcal{B}$. To find $\lambda_4, \lambda_5, \ldots$, continue the procedure.

In general, the deflation method becomes more inaccurate as we calculate more eigenvalues, because error is introduced in each eigenvalue and eigenvector and this error accumulates as the process continues. Luckily, in many applications only the dominant eigenvalue of A is needed.

PROBLEMS 6.2

In Probs. 1 to 5, use the power method to calculate approximations to the dominant eigenpair (if a dominant eigenpair exists). If the method does not work, give a reason.

1. $\begin{pmatrix} 1 & 5 \\ 5 & 6 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 4 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

3. $\begin{pmatrix} 2 & 3 \\ -2 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}$

5. $\begin{pmatrix} 3 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix}$

6. **The power method with scaling**. From the examples in this chapter we saw vectors with large components generated by the power method. To avoid this problem, we can at each step multiply the vector

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{by} \quad \frac{1}{\max\{|x_1|, |x_2|, \ldots, |x_n|\}}$$

This is called the scaling of X. For example, the scaling of

$$\begin{pmatrix} 7 \\ 5 \end{pmatrix} \quad \text{is} \quad \frac{1}{5} \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{5}{7} \end{pmatrix}$$

and the scaling of

$$\begin{pmatrix} 3 \\ -6 \end{pmatrix} \quad \text{is} \quad \frac{1}{6} \begin{pmatrix} 3 \\ -6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -1 \end{pmatrix}$$

The power method with scaling proceeds as follows: Choose X_0.
Step 1. Calculate AX_0. Let $V_1 =$ scaled version of AX_0.
Step 2. Calculate AV_1. Let $V_2 =$ scaled version of AX_0.
Step 3. Calculate AV_2. Let $V_3 =$ scaled version of AX_0.
Continue in this way. We then have at step m:

$$\lambda_1 \doteq \frac{A_{m-1} \cdot V_{m-1}}{V_{m-1} \cdot V_{m-1}}$$

and V_m is an approximate eigenvector.
Use the power method with scaling on Probs. 1, 2, and 5.

7. If a matrix has complex eigenvalues, can the power method as described in this section work?

8. Use the relative error E_{n+1} from Eq. (6.2.8) to estimate the error in the computed dominant eigenvalue in Probs. 1, 2, and 5.

9. Use deflation to fluid the nondominant eigenvalues for Probs. 1, 2, and 5.

6.3 QR METHOD

The basis of the QR method for calculating the eigenvalues of A is the fact that an $n \times n$ real matrix can be written as

$$A = QR \quad (QR \text{ factorization of } A)$$

where Q is orthogonal and R is upper triangular. The method is efficient for the calculation of all eigenvalues of a matrix.

The construction of Q and R proceeds as follows. Matrices $P_1, P_2, \ldots, P_{n-1}$ are constructed so that $P_{n-1}P_{n-2} \cdots P_2 P_1 A = R$ is upper triangular. These matrices can be chosen as orthogonal matrices and are called **householder matrices**. Since the P's are orthogonal, the stability of the eigenvalue problem will not be worsened (this is proved in numerical analysis texts). If we let

$$Q^T = P_{n-1}P_{n-2} \cdots P_2 P_1$$

then we have $Q^T A = R$ and

$$QQ^T A = QR$$
$$IA = QR$$
$$A = QR$$

We discuss the construction of the P's presently. First, we state how the QR factorization of A is used to find eigenvalues of A. We define sequences of matrices $A_1, A_2, \ldots, A_m, \ldots, Q_1, Q_2, \ldots, Q_m, \ldots$, and $R_1, R_2, \ldots, R_m, \ldots$ by this process:

Step 1. Set $A_1 = A, Q_1 = Q$, and $R_1 = R$.

Step 2. First set $A_2 = R_1 Q_1$; then factor A_2 as $A_2 = Q_2 R_2$ (QR factorization of A_2).

Step 3. First, set $A_3 = R_2 Q_2$; then factor A_3 as $A_3 = Q_3 R_3$ (QR factorization of A_3).

Step 4. Set $A_m = R_{m-1} Q_{m-1}$; then factor A_m as $A_m = Q_m R_m$ (QR factorization of A_m).

At the kth step, a matrix A_k is found, first, by using Q_{k-1} and R_{k-1} from the previous step; second, A_k is factored into $Q_k R_k$. Thus a QR factorization takes place at each step. Matrix A_m will tend toward a triangular or nearly triangular form. Thus the eigenvalues of A_m will be easy to calculate. The importance is that if the eigenvalues can be ordered as $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$, then the following is true:

> As m increases the eigenvalues of A_m approach the eigenvalues of A.

The proof of this fact is well beyond the scope of this book. Before applying the QR algorithm to some examples, we discuss the QR factorization of a matrix A.

The idea in QR factorization is to first find P_1 which, when multiplied on the left of A, will produce zeros below a_{11}. That is, we want

$$P_1 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & \cdots & \tilde{a}_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 0 & \tilde{a}_{n2} & \cdots & \tilde{a}-nn \end{pmatrix}$$

After this is done, we find P_2 which will produce

$$P_2 P_1 A = P_2 \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \mathbf{0} & \tilde{a}_{22} & \cdots & \tilde{a}_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{0} & \tilde{a}_{n2} & \cdots & \tilde{a}_{nn} \end{pmatrix} = \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} & \hat{a}_{13} & \cdots & \hat{a}_{1n} \\ \mathbf{0} & \hat{a}_{22} & \hat{a}_{23} & \cdots & \hat{a}_{2n} \\ \mathbf{0} & \mathbf{0} & \hat{a}_{33} & \cdots & \hat{a}_{3n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{0} & \mathbf{0} & \hat{a}_{n3} & \cdots & \hat{a}_{nn} \end{pmatrix}$$

The process is continued until we have

$$P_{n-1} P_{n-2} \cdots P_2 P_1 A = R = \begin{pmatrix} \\ \mathbf{0's} \end{pmatrix}$$

The problem is to find the P matrices. It turns out that the matrices P_k can be chosen as orthogonal matrices. In fact, the construction proceeds as follows.

To construct P_k:

1. Pull column k out of the matrix $P_{k-1}P_{k-2}\cdots P_2P_1A$ (just A if $k=0$):

$$\begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{kk} \\ a_{k+1,k} \\ \vdots \\ a_{nk} \end{pmatrix}$$

2. Normalize this column vector, and call the new vector

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{k-1} \\ d_k \\ \vdots \\ d_n \end{pmatrix}$$

3. Set $D = \pm\sqrt{d_k{}^2 + \cdots + d_n{}^2}$ (choose $+$ if $d_k \le 0$).

4. Set $v_1 = v_2 = \cdots = v_{k-1} = 0$. Also set

$$v_k = \sqrt{\frac{1}{2}\left(1 - \frac{d_k}{D}\right)} \quad p = -Dv_k$$

and

$$v_j = \frac{d_j}{2p} \quad \text{for} \quad j = k+1, k+2, \ldots, n$$

5. Write

$$V = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ v_k \\ v_{k+1} \\ \vdots \\ v_n \end{pmatrix}$$

Note that $|V| = 1$.

6. For the matrix

$$P_k \equiv I - 2VV^T$$

The matrix P_k will work for finding a QR factorization of A. These matrices, because of their form, are called **householder matrices**. It can be shown that householder matrices are orthogonal.

Definition 6.3.1. A householder matrix is any matrix of the form $H = I - 2VV^T$, where $|V| = 1$.

Theorem 6.3.1. *Householder matrices are orthogonal.*

Proof. We show that $H^T H = I$. By definition, then, H would be orthogonal. First we note that H is symmetric:

$$H^T = (I - 2VV^T)^T = I^T - 2V^{T^T}V^T = I - 2VV^T = H$$

Now

$$
\begin{aligned}
H^T H = H^2 &= (I - 2VV^T)^2 = I - 4VV^T + 4(VV^T)(VV^T) \\
&= I - 4VV^T + 4V(V^TV)V^T = I - 4VV^T + 4V|V|^2V^T \\
&= I - 4VV^T + 4VV^T = I
\end{aligned}
$$

$\square$

Example 1. Find a QR factorization of

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

keeping four digits to the right of the decimal point.

Solution The first column normalized is

$$\begin{pmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix} = \begin{pmatrix} 0.8165 \\ 0.4082 \\ 0.4082 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

The "diagonal" element is 0.8165. We want zeros below it. First we calculate D:

$$D = -\sqrt{(0.8165)^2 + (0.4082)^2 + (0.4082)^2} = -1$$

The minus sign was chosen since $0.8165 > 0$. Now we set

$$v_1 = \sqrt{\frac{1}{2}\left(1 - \frac{0.8165}{-1}\right)} = 0.9530 \qquad P = -Dv_1 = 0.9530$$

$$v_2 = \frac{d_2}{2p} = \frac{0.4082}{2(0.9530)} = 0.214 \qquad v_3 = \frac{d_3}{2p} = 0.214$$

Thus

$$P_1 = I - 2\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}(v_1 v_2 v_3) = I - 2\begin{pmatrix} 0.9082 & 0.204 & 0.204 \\ 0.204 & 0.0458 & 0.0458 \\ 0.204 & 0.0458 & 0.0458 \end{pmatrix}$$

$$= \begin{pmatrix} -0.8164 & -0.408 & -0.408 \\ 0.408 & 0.9084 & -0.0916 \\ 0.408 & -0.0916 & 0.9084 \end{pmatrix}$$

Multiply A by P_1; we obtain

$$P_1 A = \begin{pmatrix} -2.4488 & -2.04 & -2.04 \\ 0^* & 1.3172 & 0.3172 \\ 0^* & 0.3172 & 1.3172 \end{pmatrix}$$

To construct P_2, we look at column 2 of $P_1 A$:

$$\begin{pmatrix} -2.04 \\ 1.3172 \\ 0.3172 \end{pmatrix}$$

Normalized, this is

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} -.833 \\ 0.5379 \\ 0.1295 \end{pmatrix}$$

In this case $D = \sqrt{(0.5379)^2 + (0.1295)^2} = -0.5533$. The minus sign is chosen since $0.5379 > 0$. We set

$$v_1 = 0 \quad \text{and} \quad v_2 = \sqrt{\frac{1}{2}\left(1 - \frac{d_2}{D}\right)} = \sqrt{\frac{1}{2}\left[1 - \frac{0.5379}{(-0.5533)}\right]} = 0.993$$

So

---

*Actually we obtain 0.0008. Because of rounding, we called this zero.

$$p = 0.5533 v_2 = 0.5494$$

Therefore,

$$v_3 = \frac{d_3}{2p} = \frac{0.1295}{2(0.5494)} = 0.1179$$

and we have

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.993 \\ 0.1179 \end{pmatrix}$$

$$P_2 = I - 2 \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} (v_1 v_2 v_3) = I - 2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.986 & 0.1171 \\ 0 & 0.1171 & 0.0139 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.972 & -0.2342 \\ 0 & 0.2342 & 0.972 \end{pmatrix}$$

Finally,

$$P_2 P_1 A = P_2 \begin{pmatrix} -2.4488 & -2.04 & -2.04 \\ 0 & 1.3172 & 0.3172 \\ 0 & -0.3172 & 1.3172 \end{pmatrix}$$

$$\begin{pmatrix} -2.4488 & -2.04 & -.204 \\ 0 & -1.355 & -0.6168 \\ 0 & 0^* & 1.206 \end{pmatrix} = R$$

Therefore,

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = QR = (P_2 P_1)^T R$$

From the example just calculated, we see that finding the QR factorization for a 3×3 matrix is tedious by hand. A computer, of course, is necessary to find QR factorizations and, therefore, to use the QR method for finding eigenvalues.

For a 2×2 matrix, only one householder matrix must be found, so we consider the QR factorization for a general 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

The first column, normalized, is

---

*Actually 0.0001. Because of rounding we call this zero.

$$\begin{pmatrix} \dfrac{a}{\sqrt{a^2+c^2}} \\ \dfrac{c}{\sqrt{a^2+c^2}} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

Now $D = \pm 1$, so we can write $D = -\text{sign } a$, where sign $a = 1$ if $a \geq 0$ and sign $a = -1$ if $a < 0$.

$$v_1 = -\text{sign } a \sqrt{\frac{1}{2}\left(1 + \frac{a}{\sqrt{a^2+c^2}}\right)}$$

where sign $a = 1$ if $a \geq 0$ and sign $a = -1$ if $a < 0$. Since $p = -Dv_1$, we have $p = (\text{sign } a)v_1$. For v_2 we have

$$v_2 = \frac{d_v}{2v_1} = \frac{-(\text{sign } a)c}{\sqrt{2}\sqrt{a^2+c^2}\sqrt{1+a\sqrt{a^2+c^2}}}$$

Therefore,

$$P_1 = I - 2\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}(v_1 v_2) = \begin{pmatrix} 1 - 2v_1{}^2 & -2v_1 v_2 \\ -2v_1 v_2 & 1 - 2v_2{}^2 \end{pmatrix}$$

$$= \frac{1}{\sqrt{a^2+c^2}}\begin{pmatrix} -a & -c \\ -c & a \end{pmatrix}$$

snf

$$P_1 A = \frac{1}{\sqrt{a^2+c^2}}\begin{pmatrix} -a & -c \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\sqrt{a^2+c^2}}\begin{pmatrix} -(a^2+c^2) & -ab-cd \\ 0 & ad-bc \end{pmatrix}$$

Since P_1 is symmetric, $Q = P_1$.

Example 2. Find a QR factorization of

$$\begin{pmatrix} 3 & 7 \\ 4 & 4 \end{pmatrix}$$

Solution Using the formulas as above, we find that

$$\begin{pmatrix} 3 & 7 \\ 4 & 4 \end{pmatrix} = \begin{pmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} -5 & -\frac{37}{5} \\ 0 & -\frac{16}{5} \end{pmatrix}$$

Using the formulas for the 2×2 case, we now calculate the eigenvalues of

$$\begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$$

by the QR method.

Example 3. Use the QR method to calculate the eigenvalues of

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$$

(The true eigenvalues are 4 and 9.)

Solution We use the formulas for the 2×2 case each time we need a QR factorization. The calculated matrices are listed below (rounded); after step 3, only A_m is listed.

Step 1: $A_1 = A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$ $Q_1 = Q = \begin{pmatrix} 0.928 & 0.371 \\ 0.371 & 0.928 \end{pmatrix}$

$\quad\quad\quad R_1 = R = \begin{pmatrix} -5.385 & 4.828 \\ 0 & 6.685 \end{pmatrix}$

Step 2: $A_2 = R_1 Q_1 = \begin{pmatrix} 6.793 & -2.482 \\ -2.482 & 6.207 \end{pmatrix}$ $Q_2 = \begin{pmatrix} -0.939 & -0.343 \\ -0.343 & 0.939 \end{pmatrix}$

$\quad\quad\quad R_2 = \begin{pmatrix} -7.233 & -4.462 \\ 0 & 4.977 \end{pmatrix}$

Step 3: $A_3 = R_2 Q_2 = \begin{pmatrix} 8.324 & -1.708 \\ -1.708 & 4.675 \end{pmatrix}$ $Q_3 = \begin{pmatrix} -0.979 & 0.201 \\ 0.201 & 0.979 \end{pmatrix}$

Step 4: $A = \begin{pmatrix} 8.850 & 0.852 \\ 0.852 & 4.149 \end{pmatrix}$

Step 5: $A_5 = \begin{pmatrix} 8.969 & -0.387 \\ -0.387 & 4.030 \end{pmatrix}$

Step 6: $A_6 = \begin{pmatrix} 8.993 & 0.173 \\ 0.173 & 4.006 \end{pmatrix}$

$\quad\quad\quad \vdots$

Step 12: $A_{12} = \begin{pmatrix} 8.9999996 & 0.00134 \\ 0.00134 & 4.000018 \end{pmatrix}$

Approximate eigenvalues are on the diagonal.

In Example 3, A_m appeared to be converging to a diagonal matrix; of course, the diagonal elements are the approximate eigenvalues. This illustrates the following important result.

Theorem 6.3.2. *Let A be a real $n \times n$ matrix with eigenvalues satisfying*

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

Then matrices A_m in the QR method will converge to an upper triangular matrix with diagonal entries $\{\lambda_k\}$, $k = 1, 2, \ldots, n$. If A is symmetric, matrices A_m converge to a diagonal matrix with the eigenvalues on the diagonal.

If the hypotheses of Theorem 6.3.2 are not satisfied by A, the QR method may fail. If the difference in the magnitudes of the eigenvalues is small, convergence of the QR method can be slow.

Example 4. Applying the QR method to attempt calculation of eigenvalues of

$$\begin{pmatrix} 3 & 3 \\ 0.33333 & 5 \end{pmatrix}$$

Note that the true eigenvalues are 4.001 and 3.999.

Solution After computation we find

$$A_{10} = \begin{pmatrix} 4.7058 & -3.1764 \\ 0.1568 & 3.2941 \end{pmatrix}$$

$$A_{20} = \begin{pmatrix} 4.2582 & -3.3132 \\ 0.0201 & 3.7418 \end{pmatrix}$$

$$A_{30} = \begin{pmatrix} 4.1571 & -3.3259 \\ 0.0074 & 3.8428 \end{pmatrix}$$

We see that as the theorem guarantees, A_m is converging to an upper triangular matrix and the diagonal elements are heading in the right direction; however, the convergence is very slow. The slowness of convergence is due to the fact that $|\lambda_1| \doteq |\lambda_2|$.

As we have seen, the convergence of the QR method can be slow: This cosets money because of the computer time used. There exist methods for accelerating convergence of the QR method; these are covered in advanced numerical analysis texts.

Finally, after we find the eigenvalues of A, the corresponding eigenvectors can be found by solving $(\lambda I - A)X = 0$, subject to some side condition such as $|X| = 1$

PROBLEMS 6.3

In Probs. 1 to 5, find QR factorizations for the given matrices

1. $\begin{pmatrix} 1 & 5 \\ 5 & 6 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 4 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

3. $\begin{pmatrix} 2 & 3 \\ -2 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}$

5. $\begin{pmatrix} 3 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix}$

6. Use the QR method to find approximations to the eigenvalues of the matrices in Probs. 1, 3, and 4.

7. The matrix
$$\begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix}$$
has complex eigenvalues. Use the QR method to attempt to calculate the eigenvalues. What happens?

8. Do as in Prob. 7 with
$$A = \begin{pmatrix} 1 & 4 \\ -4 & 1 \end{pmatrix}$$

SUMMARY

The **power method** and QR **algorithm** are two methods for numerical calculation of eigenvalues of real matrices. The **stability** of a numerical eigenvalue problem depends on the matrix under consideration. If the matrix is symmetric with symmetrically distributed error, then the calculated eigenvalues will approximate the actual eigenvalues, provided the eigenvalues are all simple. Otherwise, the numerical methods may fail to find all eigenvalues.

Space has not permitted the presentation of all the numerical methods of linear algebra. To cover all these methods requires a fairly intensive course in numerical analysis. But the issues we have briefly addressed in the sections on numerical methods are important issues: error propagation, approximation, iteration, roundoff error, ill-conditioning, and theoretical underpinnings of numerical methods.

The next chapter is similar in flavor. To understand the methods discussed requires a course in optimization. Not having the luxury of that much time, we are content to present the bare bones of the simplex algorithm for linear programming problems and hope that it whets the reader's appetite for a more in-depth experience.

ADDITIONAL PROBLEMS

Shown in Probs. 1 through 5 are physical systems and a matrix associated with the differential equations for the systems. Compute the largest eigenvalue.

1. Coupled pendula.
$$\begin{pmatrix} \dfrac{g}{L} + \dfrac{k}{m} & \dfrac{-k}{m} \\ \dfrac{-k}{m} & \dfrac{g}{L} + \dfrac{k}{m} \end{pmatrix}$$
   Use $g = -32, k = 0.1, L = 2$, and $m = 1$.

2. Masses in a grove.
$$\left(\frac{k}{m} \right) \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

3. An electric circuit

$$\begin{pmatrix} 0 & \dfrac{-R}{L} \\ \dfrac{1}{RC} & \dfrac{-1}{RC} \end{pmatrix}$$

Use $R = 0.5, L = 1$, and $C = 3$.

4. Trolleys connected with springs.

$$\begin{pmatrix} \dfrac{k_1 + k_2}{m_1} & \dfrac{-k_2}{m_1} & 0 \\ \dfrac{-k_2}{m_2} & \dfrac{k_2 + k_3}{m_2} & \dfrac{-k_3}{m_2} \\ 0 & \dfrac{-k_3}{m_3} & \dfrac{-k_3}{m_3} \end{pmatrix}$$

Use $k_i = 1$ and $m_i = 2$

5. Solute diffusion.

$$\begin{pmatrix} -(k_1 + k_3) & k_1 & k_3 \\ k_1 & -(k_1 + k_2) & k_2 \\ k_3 & k_2 & -(k_2 + k_3) \end{pmatrix}$$

Use $k_1 = 2k_2$ and $k_1 = 3k_3$.

6. Gershgorin's theorem states for an $n \times n$ matrix A that any eigenvalue must lie in one of the circles in the complex plane

$$\left\{ z \text{ such that } |z - a_{kk}| \leq \sum_{\substack{j=1, \\ j \neq k}}^{n} |a_{kj}| \right\}$$

Thus for the matrix

$$A = \begin{pmatrix} i & 0.1 & 0.5 \\ 0.3 & 2 & 0.1 \\ i/2 & -1 & -3 \end{pmatrix}$$

we have the circles as shown in Fig. AP6.6. Gershgorin's theorem gives us an estimate as to where to look for the eigenvalues, real or complex. Apply Gershgorin's theorem to the matrices in Probs. 1 through 5.

7. Use Gershgorin's theorem to show that

$$A = \begin{pmatrix} i & 0.4 & 0.1i \\ 0.4 & 1+i & 0 \\ 0.1i & 0 & -3i \end{pmatrix}$$

has no real eigenvalues. Note that A is symmetric, but not real symmetric.

8. Use Gershgorin's theorem to show that

$$A = \begin{pmatrix} -4 & 1 & 1 \\ 0 & -6 & 2 \\ 1 & 0 & -2 \end{pmatrix}$$

has no eigenvalues with positive real part.