Chapter 5

Eigenvalues and Eigenvectors

Core Topics

(i) The Characteristic Equation (5.2)

(ii) Basic Power Method (5.3)

(iii) Inverse Power Method (5.4)

(iv) Shifted Power Method (5.5)

(v) QR Factorisation and Iteration Method (5.6)

(vi) Use of Matlab’s built-in functions for determining eigenvalues and eigenvectors (5.7)

Background:

For a given square matrix , the number is an eigenvalue of the matrix if:

$$[a][u] = \lambda[u] \tag{5.1}$$

The vector is a column vector called the eigenvector associated with the eigenvalue . It should be noted that there are usually more than one eigenvalue and eigenvector. In fact for an matrix there are eigenvalues and an infinite number of eigenvectors.

In general:

$$Lu = \lambda u \tag{5.2}$$

where is some mathematical operator

Background:

Equation (5.2) is a general statement of an eigenvalue problem where is the eigenvalue associated with the operator
 is the eigenvector or eigenfunction corresponding to the eigenvalue and the operator

For example consider:

$$\frac{d^2 y}{dx^2} = k^2 y \qquad (5.3)$$

Here is the second derivative w.r.t. is the eigenfunction and is the eigenvalue associated with .
Here however, we will concern ourselves with matrix operations only.

Background:

How to find the eigenvalues and eigenvectors:

Equation (5.1) can be rewritten:

$$[a - \lambda I][u] = 0 \qquad (5.4)$$

If has an inverse then . If on the other hand has no inverse then a non-trivial solution for is possible. Another way of stating this problem is to use Cramer's Rule. has no inverse (is singular) if:

$$\det[a - \lambda I] = 0 \qquad (5.5)$$

This equation is known as the characteristic equation for .

Background:

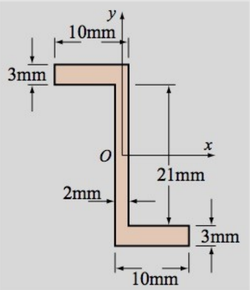
The characteristic equation for yields a polynomial equation in whose roots are the eigenvalues. Once the eigenvalues are known then the eigenvectors, , can be determined by substituting the the eigenvalues, one at a time, into Equation (5.1) and solving for

This process is relatively straightforward for small matrices but for large matrices we need to use numerical methods such as the power method or the QR factorisation method.

Example:

Example 5-1: Principal moments of inertia.

Determine the principal moments of inertia and the orientation of the principal axes of inertia for the cross-sectional area shown.



The moment of inertia I_x, I_y, and the product of inertia I_{xy} are:

$I_x = 10228.5\ \text{mm}^4$, $I_y = 1307.34\ \text{mm}^4$, and $I_{xy} = -2880\ \text{mm}^4$

SOLUTION

In matrix form, the two-dimensional moment of inertia matrix is given by:

$$I_{Iner} = \begin{bmatrix} I_x & -I_{xy} \\ -I_{xy} & I_y \end{bmatrix} = \begin{bmatrix} 10228.5 & 2880 \\ 2880 & 1307.34 \end{bmatrix} \tag{5.6}$$

The principal moments of inertia and the orientation of the principal axes of inertia can be calculated by solving the following eigenvalue problem:

$$[I_{Iner}][u] = \lambda[u] \tag{5.7}$$

where the eigenvalues λ are the principal moments of inertia and the associated eigenvectors $[u]$ are unit vectors in the direction of the principal axes of inertia. The eigenvalues are determined by calculating the determinant in Eq. (5.5):

$$det[I_{Iner} - \lambda I\,] = 0 \tag{5.8}$$

$$det\begin{bmatrix} (10228.5-\lambda) & 2880 \\ 2880 & (1307.34-\lambda) \end{bmatrix} = 0 \tag{5.9}$$

The polynomial equation for λ is:

$$(10228.5-\lambda)(1307.34-\lambda) - 2880^2 = 0 \quad \text{or} \quad \lambda^2 - 11535.84\lambda + 5077727.19 = 0 \tag{5.10}$$

The solutions of the quadratic polynomial equation are the eigenvalues $\lambda_1 = 11077.46\ \text{mm}^4$ and $\lambda_2 = 458.38\ \text{mm}^4$, which are the principal moments of inertia.

The eigenvectors that correspond to each eigenvalue are calculated by substituting the eigenvalues in Eq. (5.7). For the first eigenvector $u^{(1)}$:

$$\begin{bmatrix} 10228.5 & 2880 \\ 2880 & 1307.34 \end{bmatrix}\begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix} = 11077.46\begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -848.96 & 2880 \\ 2880 & -9770.12 \end{bmatrix}\begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5.11}$$

The two equations in Eqs. (5.11) give $u_2^{(1)} = 0.29478u_1^{(1)}$. By using the additional condition that the eigenvector in this problem is a unit vector, $(u_1^{(1)})^2 + (u_2^{(1)})^2 = 1$, the eigenvector associated with the first eigenvalue, $\lambda_1 = 11077$, is determined to be $u^{(1)} = 0.95919i + 0.28275j$.

For the second eigenvector $u^{(2)}$:

$$\begin{bmatrix} 10228.5 & 2880 \\ 2880 & 1307.34 \end{bmatrix}\begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \end{bmatrix} = 458.38\begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 9770.12 & 2880 \\ 2880 & 848.96 \end{bmatrix}\begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5.12}$$

The two equations in Eqs. (5.12) give $u_2^{(2)} = -3.3924u_1^{(2)}$. By using the additional condition that the eigenvector is a unit vector, $(u_1^{(2)})^2 + (u_2^{(2)})^2 = 1$, the eigenvector associated with the second eigenvalue, $\lambda_2 = 458.38$, is determined to be $u^{(1)} = -0.28275i + 0.95919j$.

The Basic Power Method:

This method is an iterative procedure for determining the largest real eigenvalue and corresponding eigenvector of a matrix. It is recommended that you study this proof over that given in the text.

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$, not necessarily distinct, that satisfy the relations $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$. The eigenvalue λ_1, which is largest in magnitude, is known as the *dominant* eigenvalue of the matrix A. Furthermore, assume that the associated eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3, \ldots, \mathbf{v}_n$ are linearly independent, and therefore form a basis for $\mathbf{R}^n$. It should be noted at this point that not all matrices have eigenvalues and eigenvectors which satisfy the conditions we've assumed here. At the end of the section and in the exercises, we will explore what happens when these conditions are violated.

Let $\mathbf{x}^{(0)}$ be a nonzero element of $\mathbf{R}^n$. Since the eigenvectors of A form a basis for $\mathbf{R}^n$, it follows that $\mathbf{x}^{(0)}$ can be written as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3, \ldots, \mathbf{v}_n$; that is, there exist constants α_1, α_2, $\alpha_3, \ldots, \alpha_n$ such that

$$\mathbf{x}^{(0)} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \cdots + \alpha_n \mathbf{v}_n.$$

The Basic Power Method:

Next, construct the sequence of vectors $\{\mathbf{x}^{(m)}\}$ according to the rule $\mathbf{x}^{(m)} = A\mathbf{x}^{(m-1)}$ for $m \geq 1$. By direct calculation we find

$$\begin{aligned}
\mathbf{x}^{(1)} &= A\mathbf{x}^{(0)} = \alpha_1(A\mathbf{v}_1) + \alpha_2(A\mathbf{v}_2) + \alpha_3(A\mathbf{v}_3) + \cdots + \alpha_n(A\mathbf{v}_n) \\
&= \alpha_1(\lambda_1\mathbf{v}_1) + \alpha_2(\lambda_2\mathbf{v}_2) + \alpha_3(\lambda_3\mathbf{v}_3) + \cdots + \alpha_n(\lambda_n\mathbf{v}_n), \\
\mathbf{x}^{(2)} &= A\mathbf{x}^{(1)} = A^2\mathbf{x}^{(0)} \\
&= \alpha_1(A^2\mathbf{v}_1) + \alpha_2(A^2\mathbf{v}_2) + \alpha_3(A^2\mathbf{v}_3) + \cdots + \alpha_n(A^2\mathbf{v}_n) \\
&= \alpha_1(\lambda_1^2\mathbf{v}_1) + \alpha_2(\lambda_2^2\mathbf{v}_2) + \alpha_3(\lambda_3^2\mathbf{v}_3) + \cdots + \alpha_n(\lambda_n^2\mathbf{v}_n)
\end{aligned}$$

and, in general,

$$\begin{aligned}
\mathbf{x}^{(m)} &= A\mathbf{x}^{(m-1)} = \cdots = A^m\mathbf{x}^{(0)} \\
&= \alpha_1(A^m\mathbf{v}_1) + \alpha_2(A^m\mathbf{v}_2) + \alpha_3(A^m\mathbf{v}_3) + \cdots + \alpha_n(A^m\mathbf{v}_n) \\
&= \alpha_1(\lambda_1^m\mathbf{v}_1) + \alpha_2(\lambda_2^m\mathbf{v}_2) + \alpha_3(\lambda_3^m\mathbf{v}_3) + \cdots + \alpha_n(\lambda_n^m\mathbf{v}_n).
\end{aligned}$$

In deriving these expressions we have made repeated use of the relation $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$, which follows from the fact that $\mathbf{v}_j$ is an eigenvector associated with the eigenvalue λ_j.

The Basic Power Method:

Factoring λ_1^m from the right-hand side of the equation for $\mathbf{x}^{(m)}$ gives

$$\mathbf{x}^{(m)} = \lambda_1^m \left[\alpha_1 \mathbf{v}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^m \mathbf{v}_2 + \alpha_3 \left(\frac{\lambda_3}{\lambda_1}\right)^m \mathbf{v}_3 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^m \mathbf{v}_n\right]. \quad (1)$$

By assumption, $|\lambda_j/\lambda_1| < 1$ for each j, so $|\lambda_j/\lambda_1|^m \to 0$ as $m \to \infty$. It therefore follows that

$$\lim_{m\to\infty} \frac{\mathbf{x}^{(m)}}{\lambda_1^m} = \alpha_1 \mathbf{v}_1.$$

Since any nonzero constant times an eigenvector is still an eigenvector associated with the same eigenvalue, we see that the scaled sequence $\{\mathbf{x}^{(m)}/\lambda_1^m\}$ converges to an eigenvector associated with the dominant eigenvalue provided $\alpha_1 \neq 0$. Furthermore, convergence toward the eigenvector is linear with asymptotic error constant $|\lambda_2/\lambda_1|$.

The Basic Power Method:

An approximation for the dominant eigenvalue of A can be obtained from the sequence $\{\mathbf{x}^{(m)}\}$ as follows. Let i be an index for which $x_i^{(m-1)} \neq 0$, and consider the ratio of the ith element from the vector $\mathbf{x}^{(m)}$ to the ith element from $\mathbf{x}^{(m-1)}$. By equation (1),

$$\frac{x_i^{(m)}}{x_i^{(m-1)}} = \frac{\lambda_1^m \alpha_1 v_{1,i}\left[1 + O\left((\lambda_2/\lambda_1)^m\right)\right]}{\lambda_1^{m-1} \alpha_1 v_{1,i}\left[1 + O\left((\lambda_2/\lambda_1)^{m-1}\right)\right]} = \lambda_1\left[1 + O\left((\lambda_2/\lambda_1)^{m-1}\right)\right],$$

provided $v_{1,i} \neq 0$, where $v_{1,i}$ denotes the ith element from the vector $\mathbf{v}_1$. Hence, the ratio $x_i^{(m)}/x_i^{(m-1)}$ converges toward the dominant eigenvalue, and the convergence is linear with asymptotic rate constant $|\lambda_2/\lambda_1|$.

The Basic Power Method:

Algorithm for the Power Method:

1. Choose a column vector of length . The vector can be any non-zero vector.

2. Multiply this vector by . This gives a column vector .

3. Normalise this column vector.

4. Go back to 2. with

5. Continue with 2. – 4. until a desired accuracy is reached for both the eigenvalue and eigenvector.

Example:

Example 5-2: Using the power method to determine the largest eigenvalue of a matrix.

Determine the largest eigenvalue of the following matrix:

$$\begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \tag{5.21}$$

Use the power method and start with the vector $x = [1, 1, 1]^T$.

SOLUTION

Starting with $i = 1$, $x_1 = [1, 1, 1]^T$. With the power method, the vector $[x]_2$ is first calculated by $[x]_2 = [a][x]_1$ (Step 2) and is then normalized (Step 3):

$$[x]_2 = [a][x]_1 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \\ 2 \end{bmatrix} = 7 \begin{bmatrix} 0.5714 \\ 1 \\ 0.2857 \end{bmatrix} \tag{5.22}$$

For $i = 2$, the normalized vector $[x]_2$ (without the multiplicative factor) is multiplied by $[a]$. This results in $[x]_3$, which is then normalized:

$$[x]_3 = [a][x]_2 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.5714 \\ 1 \\ 0.2857 \end{bmatrix} = \begin{bmatrix} 3.7143 \\ 7.1429 \\ 4 \end{bmatrix} = 7.1429 \begin{bmatrix} 0.52 \\ 1 \\ 0.56 \end{bmatrix} \tag{5.23}$$

The next three iterations are:

$$i = 3: \qquad [x]_4 = [a][x]_3 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.52 \\ 1 \\ 0.56 \end{bmatrix} = \begin{bmatrix} 2.96 \\ 7.52 \\ 2.8 \end{bmatrix} = 7.52 \begin{bmatrix} 0.3936 \\ 1 \\ 0.3723 \end{bmatrix} \tag{5.24}$$

$$i = 4: \qquad [x]_5 = [a][x]_4 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3936 \\ 1 \\ 0.3723 \end{bmatrix} = \begin{bmatrix} 2.8298 \\ 7.5851 \\ 3.2979 \end{bmatrix} = 7.5851 \begin{bmatrix} 0.3731 \\ 1 \\ 0.4348 \end{bmatrix} \tag{5.25}$$

$$i = 5: \qquad [x]_6 = [a][x]_5 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3731 \\ 1 \\ 0.4348 \end{bmatrix} = \begin{bmatrix} 2.6227 \\ 7.6886 \\ 3.0070 \end{bmatrix} = 7.6886 \begin{bmatrix} 0.3411 \\ 1 \\ 0.3911 \end{bmatrix} \tag{5.26}$$

After three more iterations, the results are:

$$i = 8 \qquad [x]_9 = [a][x]_8 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3272 \\ 1 \\ 0.3946 \end{bmatrix} = \begin{bmatrix} 2.5197 \\ 7.7401 \\ 3.0760 \end{bmatrix} = 7.7401 \begin{bmatrix} 0.3255 \\ 1 \\ 0.3974 \end{bmatrix} \tag{5.27}$$

The results show that the differences between the vector $[x_i]$ and the normalized vector $[x_{i+1}]$ are getting smaller. The value of the multiplicative factor (7.7401) is an estimate of the largest eigenvalue. As shown in Section 5.5, a value of 7.7504 is obtained for the eigenvalue by MATLAB's built-in function `eig`.

The Inverse Power Method:

This method is used to find the smallest eigenvalue. By applying the power method to . The eigenvalues of are the reciprocals of the eigenvalues of This can be seen as follows:

Then:

$$[a]^{-1}[a][x] = [a]^{-1}\lambda[x] = \lambda[a]^{-1}[x] \qquad (5.28)$$

Hence:

$$[x] = \lambda[a]^{-1}[x] \text{ or } [a]^{-1}[x] = \frac{[x]}{\lambda} \qquad (5.29)$$

This shows that is an eigenvalue of

The Inverse Power Method:

Thus the Inverse Power Method can be used to find the largest eigenvalue of which corresponds to the smallest eigenvalue of (because the eigenvalues of are reciprocals of the eigenvalues of).

The procedure is the same as before. The starting vector is multiplied by to give , which is then normalised and multiplied by again:

$$[x]_{i+1} = [a]^{-1}[x]_i \qquad (5.30)$$

se:

$$[a][x]_{i+1} = [x]_i \qquad (5.31)$$

omposition.

The Shifted Power Method:

Given then if is the largest (or smallest) eigenvalue obtained using the Power Method (or Inverse Power Method) then the eigenvalues of a new shifted matrix formed by are

. The reason for this is as follows:

$$[a - \lambda_1 I][x] = \alpha[x] \qquad (5.32)$$

where the are the eigenvalues of . But so:

$$(\lambda - \lambda_1)[x] = \alpha[x] \qquad (5.33)$$

where .

The Shifted Power Method:

Procedure:

1. Determine the largest eigenvalue of using the Power   Method.

2. Now determine the largest eigenvalue of the shifted matrix
 (using the Power Method). Then from which can be determined.

3. Next determine the largest eigenvalue of the shifted matrix
. Then from which can be determined.

4. Continue in this fashion until all the eigenvalues of are determined.
This is done in a total of steps where is an matrix.

The Shifted Power Method:

The Shifted Power Method is an inefficient process. A preferred method for obtaining all eigenvalues of a matrix is the QR Factorisation Method. This is described next.

QR Factorisation:

Definitions:

1. Two square matrices and are similar if:

where is an invertible matrix. and have the same   eigenvalues.

2. A matrix is orthogonal if:

3. A matrix is symmetric if:

4. The eigenvalues of an upper triangular matrix are its diagonal elements.

QR Factorisation:

QR Factorisation is used to find all the eigenvalues of a matrix. We begin with the matrix whose eigenvalues are to be determined.

Let:

where is an orthogonal matrix and is upper triangular. It will be shown later how to find and .

We now define:

QR Factorisation:

But

So we can write:

This means that and are similar – thus having the same eigenvalues. This completes the first iteration in the QR factorisation procedure.

QR Factorisation:

Now write:

Again is orthogonal and is upper triangular.

Then define:

QR Factorisation:

Again

Then:

This means that and are similar – thus having the same eigenvalues (which are the same as the eigenvalues of). This completes the second iteration in the QR factorisation procedure.

QR Factorisation:

The iterations continue until the sequence of matrices generated results in an upper triangular matrix whose eigenvalues are its diagonal elements which are the same as the eigenvalues of .

QR Factorisation:

We now obtain and - that is and for the iteration.

To do this we use what is known as the Householder matrix which has the form:

$$[H] = [I] - \frac{2}{[v]^T[v]}[v][v]^T \qquad (5.38)$$

where is the identity matrix and is an element column vector given by:

$$[v] = [c] + \|c\|_2[e] \qquad (5.39)$$

$$\|c\|_2 = \sqrt{c_1^2 + c_2^2 + c_3^2 + \ldots + c_n^2} \qquad (5.40)$$

QR Factorisation:

The Householder matrix is both symmetric and orthogonal. This means:

Hence:

yields a matrix that is similar to .

QR Factorisation:

Factoring the matrix into an orthogonal matrix and an upper triangular matrix such that is done in steps.

Step 1:

We choose to be the first column of

$$[c] = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{bmatrix} \qquad (5.41)$$

The vector is defined as:

$$[e] = \begin{bmatrix} \pm 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad (5.42)$$

The first element of is +1 if the first element of is positive. Otherwise it is negative.

QR Factorisation:

Now we have enough information to calculate . The matrix is then factored into where:

and

The matrix is orthogonal because is orthogonal and is a matrix with zeros as the elements in the first column below the diagonal.

$$\begin{bmatrix} R_{11}^{(1)} & R_{12}^{(1)} & R_{13}^{(1)} & R_{14}^{(1)} & R_{15}^{(1)} \\ 0 & R_{22}^{(1)} & R_{23}^{(1)} & R_{24}^{(1)} & R_{25}^{(1)} \\ 0 & R_{32}^{(1)} & R_{33}^{(1)} & R_{34}^{(1)} & R_{35}^{(1)} \\ 0 & R_{42}^{(1)} & R_{43}^{(1)} & R_{44}^{(1)} & R_{45}^{(1)} \\ 0 & R_{52}^{(1)} & R_{53}^{(1)} & R_{54}^{(1)} & R_{55}^{(1)} \end{bmatrix}$$

QR Factorisation:

Step 2:

The vector is defined as the second column of the matrix with its first element set to zero:

$$[c] = \begin{bmatrix} 0 \\ R_{22}^{(1)} \\ R_{32}^{(1)} \\ \dots \\ R_{n2}^{(1)} \end{bmatrix} \qquad (5.45)$$

The vector is:

$$[e] = \begin{bmatrix} 0 \\ \pm 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} \qquad (5.46)$$

The second element of is is +1 if the second element of is positive. Otherwise it is negative.

QR Factorisation:

Now a new Householder matrix, , is constructed. Now we factor the matrix into where:

and

The matrix is orthogonal because is orthogonal and is a matrix with zeros as the elements in the second column below the diagonal.

$$\begin{bmatrix} R_{11}^{(2)} & R_{12}^{(2)} & R_{13}^{(2)} & R_{14}^{(2)} & R_{15}^{(2)} \\ 0 & R_{22}^{(2)} & R_{23}^{(2)} & R_{24}^{(2)} & R_{25}^{(2)} \\ 0 & 0 & R_{33}^{(2)} & R_{34}^{(2)} & R_{35}^{(2)} \\ 0 & 0 & R_{43}^{(2)} & R_{44}^{(2)} & R_{45}^{(2)} \\ 0 & 0 & R_{53}^{(2)} & R_{54}^{(2)} & R_{55}^{(2)} \end{bmatrix}$$

QR Factorisation:

Step 3:

Moving to the third column of , the vectors and are defined as:

$$[c] = \begin{bmatrix} 0 \\ 0 \\ R_{33}^{(2)} \\ R_{34}^{(2)} \\ \cdots \\ R_{n3}^{(2)} \end{bmatrix} \qquad (5.49)$$

$$[e] = \begin{bmatrix} 0 \\ 0 \\ \pm 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad (5.50)$$

where the sign rule applies as before.

is then:

$$\begin{bmatrix} R_{11}^{(3)} & R_{12}^{(3)} & R_{13}^{(3)} & R_{14}^{(3)} & R_{15}^{(3)} \\ 0 & R_{22}^{(3)} & R_{23}^{(3)} & R_{24}^{(3)} & R_{25}^{(3)} \\ 0 & 0 & R_{33}^{(3)} & R_{34}^{(3)} & R_{35}^{(3)} \\ 0 & 0 & 0 & R_{44}^{(3)} & R_{45}^{(3)} \\ 0 & 0 & 0 & R_{54}^{(3)} & R_{55}^{(3)} \end{bmatrix}$$

QR Factorisation:

We continue in this latter fashion until is upper triangular. Then we have:

This completes the factorisation of

QR Factorisation:

Procedure:

1. Factor into an orthogonal matrix and an upper triangular matrix . This is done in steps using a Householder matrix such that:

2. Calculate:

3. Repeat 1. and 2. until is reached. At this point is upper triangular and so its eigenvalues, which are the same as those of , appear on its diagonal.

Example:

Example 5-3: QR factorization of a matrix.

Factor the following matrix $[a]$ into an orthogonal matrix $[Q]$ and an upper triangular matrix $[R]$:

$$[a] = \begin{bmatrix} 6 & -7 & 2 \\ 4 & -5 & 2 \\ 1 & -1 & 1 \end{bmatrix} \tag{5.56}$$

SOLUTION

The solution follows the steps listed in pages 176–179. Since the matrix $[a]$ is (3×3), the factorization requires only two steps.

Step 1: The vector $[c]$ is defined as the first column of the matrix $[a]$:

$$[c] = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}$$

The vector $[e]$ is defined as the following three-element column vector:

$$[e] = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Using Eq. (5.40), the Euclidean norm, $\|c\|_2$, of $[c]$ is:

$$\|c\|_2 = \sqrt{c_1^2 + c_2^2 + c_3^2} = \sqrt{6^2 + 4^2 + 1^2} = 7.2801$$

Using Eq. (5.39), the vector $[v]$ is:

$$[v] = [c] + \|c\|_2[e] = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix} + 7.2801 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 13.2801 \\ 4 \\ 1 \end{bmatrix}$$

Next, the products $[v]^T[v]$ and $[v][v]^T$ are calculated:

$$[v]^T[v] = \begin{bmatrix} 13.2801 & 4 & 1 \end{bmatrix} \begin{bmatrix} 13.2801 \\ 4 \\ 1 \end{bmatrix} = 193.3611$$

$$[v][v]^T = \begin{bmatrix} 13.2801 \\ 4 \\ 1 \end{bmatrix} \begin{bmatrix} 13.2801 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 176.3611 & 53.1204 & 13.2801 \\ 53.1204 & 16 & 4 \\ 13.2801 & 4 & 1 \end{bmatrix}$$

The Householder matrix $[H]^{(1)}$ is then:

$$[H]^{(1)} = [I] - \frac{2}{[v]^T[v]}[v][v]^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{193.3611} \begin{bmatrix} 176.3611 & 53.1204 & 13.2801 \\ 53.1204 & 16 & 4 \\ 13.2801 & 4 & 1 \end{bmatrix} = \begin{bmatrix} -0.8242 & -0.5494 & -0.1374 \\ -0.5494 & 0.8345 & -0.0414 \\ -0.1374 & -0.0414 & 0.9897 \end{bmatrix}$$

Once the Housholder matrix $[H]^{(1)}$ is constructed, $[a]$ can be factored into $[Q]^{(1)}[R]^{(1)}$, where:

$$[Q]^{(1)} = [H]^{(1)} = \begin{bmatrix} -0.8242 & -0.5494 & -0.1374 \\ -0.5494 & 0.8345 & -0.0414 \\ -0.1374 & -0.0414 & 0.9897 \end{bmatrix}$$

and

$$[R]^{(1)} = [H]^{(1)}[a] = \begin{bmatrix} -0.8242 & -0.5494 & -0.1374 \\ -0.5494 & 0.8345 & -0.0414 \\ -0.1374 & -0.0414 & 0.9897 \end{bmatrix} \begin{bmatrix} 6 & -7 & 2 \\ 4 & -5 & 2 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -7.2801 & 8.6537 & -2.8846 \\ 0 & -0.2851 & 0.5288 \\ 0 & 0.1787 & 0.6322 \end{bmatrix}$$

This completes the first step.

Step 2: The vector $[c]$, which has three elements, is now defined as:

$$[c] = \begin{bmatrix} 0 \\ R_{22}^{(1)} \\ R_{32}^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ -0.2851 \\ 0.1787 \end{bmatrix}$$

The vector $[e]$ is defined as the following three-element column vector:

$$[e] = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

Example:

Using Eq. (5.40), the Euclidean norm, $\|c\|_2$, of $[c]$ is:

$$\|c\|_2 = \sqrt{c_1^2 + c_2^2 + c_3^2} = \sqrt{0^2 + (-0.2851)^2 + 0.1787^2} = 0.3365$$

Using Eq. (5.39), the vector $[v]$ is:

$$[v] = [c] + \|c\|_2[e] = \begin{bmatrix} 0 \\ -0.2851 \\ 0.1787 \end{bmatrix} + 0.3365\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.6215 \\ 0.1787 \end{bmatrix}$$

Next, the products $[v]^T[v]$ and $[v][v]^T$ are calculated:

$$[v]^T[v] = \begin{bmatrix} 0 & -0.6215 & 0.1787 \end{bmatrix}\begin{bmatrix} 0 \\ -0.6215 \\ 0.1787 \end{bmatrix} = 0.4183$$

$$[v][v]^T = \begin{bmatrix} 0 \\ -0.6215 \\ 0.1787 \end{bmatrix}\begin{bmatrix} 0 & -0.6215 & 0.1787 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.3864 & -0.1111 \\ 0 & 0.1111 & 0.0319 \end{bmatrix}$$

The Householder matrix $[H]^{(2)}$ is then:

$$[H]^{(2)} = [I] - \frac{2}{[v]^T[v]}[v][v]^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{0.4183}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.3864 & -0.1111 \\ 0 & 0.1111 & 0.0319 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.8474 & 0.5311 \\ 0 & 0.5311 & 0.8473 \end{bmatrix}$$

Once the Housholder matrix $[H]^{(2)}$ is constructed, $[a]$ can be factored into $[Q]^{(2)}[R]^{(2)}$, where:

$$[Q]^{(2)} = [Q]^{(1)}[H]^{(2)} = \begin{bmatrix} -0.8242 & -0.5494 & -0.1374 \\ -0.5494 & 0.8345 & -0.0414 \\ -0.1374 & -0.0414 & 0.9897 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.8474 & 0.5311 \\ 0 & 0.5311 & 0.8473 \end{bmatrix} = \begin{bmatrix} -0.8242 & 0.3927 & -0.4082 \\ -0.5494 & -0.7291 & 0.4082 \\ -0.1374 & 0.5607 & 0.8166 \end{bmatrix}$$

and

$$[R]^{(2)} = [H]^{(2)}[R]^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.8474 & 0.5311 \\ 0 & 0.5311 & 0.8473 \end{bmatrix}\begin{bmatrix} -7.2801 & 8.6537 & -2.8846 \\ 0 & -0.2851 & 0.5288 \\ 0 & 0.1787 & 0.6322 \end{bmatrix} = \begin{bmatrix} -7.2801 & 8.6537 & -2.8846 \\ 0 & 0.3365 & -0.1123 \\ 0 & 0 & 0.8165 \end{bmatrix}$$

This completes the factorization, which means that:

$$[a] = [Q]^{(2)}[R]^{(2)} \quad \text{or} \quad \begin{bmatrix} 6 & -7 & 2 \\ 4 & -5 & 2 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -0.8242 & 0.3927 & -0.4082 \\ -0.5494 & -0.7291 & 0.4082 \\ -0.1374 & 0.5607 & 0.8166 \end{bmatrix}\begin{bmatrix} -7.2801 & 8.6537 & -2.8846 \\ 0 & 0.3365 & -0.1123 \\ 0 & 0 & 0.8165 \end{bmatrix}$$

The results can be verified by using MATLAB. First, it is verified that the matrix $[Q]^{(2)}$ is orthogonal. This is done by calculating the inverse of $[Q]^{(2)}$ with MATLAB's built-in function, `inv`, and verifying that it is equal to the transpose of $[Q]^{(2)}$. Then the multiplication $[Q]^{(2)}[R]^{(2)}$ is done with MATLAB, and the result is compared with $[a]$.

```
>> Q2=[-0.8242 0.3927 -0.4082; -0.5494 -0.7291 0.4084; -0.1374 0.5607 0.8166];
>> R2=[-7.2801 8.6537 -2.8846; 0 0.3365 -0.1123; 0 0 0.8165];
>> invQ2=inv(Q2)
invQ2 =
  -0.8242   -0.5494   -0.1372
   0.3924   -0.7290    0.5607
  -0.4081    0.4081    0.8165
>> a=Q2*R2
a =
   6.0003   -7.0002    2.0001
   3.9997   -4.9997    2.0001
   1.0003   -1.0003    1.0001
```

The results (other than errors due to rounding) verify the factorization.

Example:

Example 5-4: Calculating eigenvalues using the QR factorization and iteration method.

The three-dimensional state of stress at a point inside a loaded structure is given by:

$$\sigma_{ij} = \begin{bmatrix} 45 & 30 & -25 \\ 30 & -24 & 68 \\ -25 & 68 & 80 \end{bmatrix} \text{ MPa}$$

Determine the principal stresses at this point by determining the eigenvalues of the stress matrix, using the QR factorization method.

SOLUTION

The problem is solved with MATLAB. First, a user-defined function named `QRFactorization` is written. Then, the function is used in a MATLAB program written in a script file for determining the eigenvalues using the QR factorization and iteration method.

The user-defined MATLAB function `QRFactorization`, which is listed below, uses the Householder matrix construct in the procedure that is described in pages 176–179 to calculate the QR factorization of a square matrix.

Program 5-1: User-defined function, QR factorization of a matrix.

```
function [Q R] = QRFactorization(R)
% The function factors a matrix [A] into an orthogonal matrix [Q] and an
upper-triangular matrix [R].
% Input variables:
% A  The (square) matrix to be factored.
% Output variables:
% Q  Orthogonal matrix.
% R  Upper-triangular matrix.

nmatrix=size(R);
n=nmatrix(1);
I=eye(n);
Q=I;
for j=1:n-1
    c=R(:,j);                          Define the vector [c].
    c(1:j-1)=0;
    e(1:n,1)=0;
     if c(j) > 0
        e(j)=1;                        Define the vector [e].
     else
        e(j)=-1;
     end
     clength=sqrt(c'*c);               Eq. (5.40).
     v=c+clength*e;                    Generate the vector [v], Eq. (5.39).
     H=I-2/(v'*v)*v*v';                Construct the Householder matrix [H], Eq. (5.38).
     Q=Q*H;
     R=H*R;
end
```

The determination of the eigenvalues follows the procedure in the algorithm.

```
A=[45 30 -25; 30 -24 68; -25 68 80]
for i=1:40
   [q R]=QRFactorization(A);
   A=R*q;
end
A
e=diag(A)
```

The program repeats the QR factorization 40 times and then displays (in the Command Window) the last matrix $[A]$ that is obtained. The diagonal elements of the matrix are the eigenvalues of the original matrix $[A]$.

Example:

```
A =
    45    30   -25
    30   -24    68
   -25    68    80
A =
  114.9545    0.0000    0.0000
    0.0000  -70.1526   -1.5563
    0.0000   -1.5563   56.1981
e =
  114.9545
  -70.1526
   56.1981
```

The results show that after 40 iterations, the matrix $[A]$ is nearly upper triangular. Actually, in this case, QR factorization results in a diagonal matrix because the original matrix $[\sigma]$ is symmetric.

Recommended Problems:

Problems to be solved by hand (do at least 2):

5.3, 5.7, 5.9

Problems to be programmed in MATLAB (do at least 1):

5.10, 5.12

Problems in Science and Engineering:

5.18

You should pick out problems that you find interesting/challenging and do these too.