Eigenvalues and Eigenvectors of a Matrix

Jeff Chak Fu WONG

Department of Mathematics
The Chinese University of Hong Kong

2526 EDB HSMMC Math Modelling Advanced Workshop

1. Eignevalues and Eigenvectors
2. Diagonalization

Eignevalues and Eigenvectors

In this lecture, every matrix considered is a square matrix.

- Let A be an $n \times n$ matrix.
- Then, as we recalled from the concept of Matrix Transformation and Linear Transformation, the function $L : \mathbb{R}^n \to \mathbb{R}^n$ defined by $L(\mathbf{x}) = \mathbf{Ax}$, for $\mathbf{x}$ in $\mathbb{R}^n$, a linear transformation.
- A question of considerable importance in great many applied problems is the determination of vectors $\mathbf{x}$,
    - if they are any,
    - such $\mathbf{x}$ and $A\mathbf{x}$ are parallel (see Figure 1).

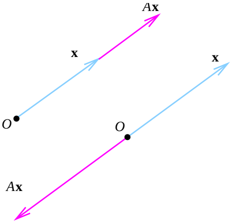

Figure 1: $\mathbf{x}$ is an eigenvector of A. $A\mathbf{x}$ is in same or opposite direction as $\mathbf{x}$, if $\lambda \neq 0$

Such questions arise in all applications involving vibrations; they arise in machine learning, AI, differential equations, and others.

In this note we shall formulate this problem precisely; we also define some pertinent terminology.

In the next section we solve this problem for symmetric matrices and briefly discuss the situation in general case.

Definition 1.

- Let A be an $n \times n$ matrix.
- The number λ is called an **eigenvalue of** A if there exists a ***nonzero*** vector $\mathbf{x}$ in $\mathbb{R}^n$ such that

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{1}$$

- Every nonzero vector $\mathbf{x}$ satisfying Equation (1) is called an **eigenvector of** A **associated with the eigenvalue** λ.
- The word *eigenvalue* is a hybrid one (*eigen* in German means "proper").
- Eigenvalues are also called **proper values**, **characteristic values** and **latent values**; and eigenvectors are also called **proper vectors**, and so on, accordingly.

Note that $\mathbf{x} = \mathbf{0}$ always satisfies Equation (1), but $\mathbf{0}$ is not an eigenvector, since we insist that an eigenvector be a nonzero vector.
Remark: In the preceding definition, the number λ can be real or complex and the vector $\mathbf{x}$ can have real or complex components.

Example 2.

- If A is the identity matrix I_n, then the only eigenvalue is $\lambda = 1$;
- every nonzero vector in $\mathbb{R}^n$ is an eigenvector of A associated with the eigenvalue $\lambda = 1$:

$$I_n\mathbf{x} = 1\mathbf{x}.$$

Example 3.

Let

$$A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

Then

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

so that

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is an eigenvector of A associated with the eigenvalue $\lambda_1 = \frac{1}{2}$. Also

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

so that

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

is an eigenvector of A associated with the eigenvalue $\lambda_2 = -\frac{1}{2}$.

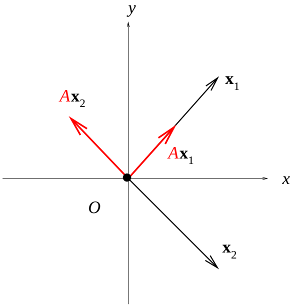

Figure 2:

Figure 2 shows that $\mathbf{x}_1$ and $A\mathbf{x}_1$ are parallel, and $\mathbf{x}_2$ and $A\mathbf{x}_2$ are parallel also.

This illustrates the fact that if $\mathbf{x}$ is an eigenvector of A, then $\mathbf{x}$ and $A\mathbf{x}$ are parallel.

Let λ be an eigenvalue of A with corresponding eigenvector $\mathbf{x}$.

We show $\mathbf{x}$ and $A\mathbf{x}$ for cases

1. $\lambda > 1$, see Figure 3
2. $0 < \lambda < 1$, see Figure 4 and
3. $\lambda < 0$, see Figure 5.

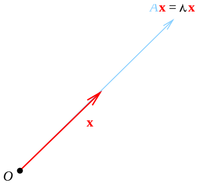

Figure 3: $\lambda > 1$

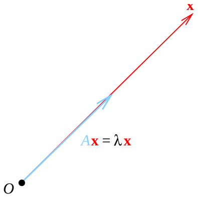

Figure 4: $0 < \lambda < 1$

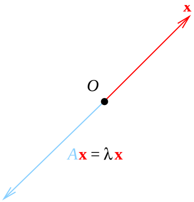

Figure 5: $\lambda < 0$

An eigenvalue of λ can have associated with it many different eigenvectors. If fact,

- if $\mathbf{x}$ is an eigenvector of A associated with λ (i.e., $A\mathbf{x} = \lambda\mathbf{x}$) and r is any nonzero real number,
- then

$$A(r\mathbf{x}) = r(A\mathbf{x}) = r(\lambda\mathbf{x}) = \lambda(r\mathbf{x}).$$

Thus $r\mathbf{x}$ is also an eigenvector of A associated with λ.

Remark 1.

Note that two eigenvectors associated with the same eigenvalue need not be in the same direction.

They must only be parallel.

Thus, in Example 3, it can be verified easily that

$$\mathbf{x}_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

is another eigenvector associated with the eigenvalue $\lambda_1 = \frac{1}{2}$.

Example 4.

Let

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

Then

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

so that $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of A associated with the eigenvalue $\lambda_1 = 0$. Also

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is an eigenvector of A associated with the eigenvalue $\lambda_2 = 1$ (verify).

Example 4 points out the fact that although the zero vector, by definition, cannot be an eigenvector, the number zero can be an eigenvalue.

Computing Eigenvalues and Eigenvectors

Thus far we have found the eigenvalues and associated eigenvectors of a given matrix by

- *inspection*,
- *geometric arguments* or
- *very simple algebraic approaches*.

In the following example, we compute the eigenvalues and associated eigenvectors of a matrix by a somewhat more systematic method.

Example 5.

- Let

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}.$$

  we wish to find the eigenvalues of A and their associated eigenvectors.

- Thus we wish to find all real numbers λ and all nonzero vectors

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

satisfying Equation (1), that is, $A\mathbf{x} = \lambda\mathbf{x}$

$$\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- Equation (2) becomes

$$\begin{aligned} x_1 + x_2 &= \lambda x_1 \\ -2x_1 + 4x_2 &= \lambda x_2 \end{aligned}$$

or

$$\begin{aligned} (\lambda - 1)x_1 - x_2 &= 0 \\ 2x_1 + (\lambda - 4)x_2 &= 0. \end{aligned}$$

- Equation (3) is a homogeneous system of two equations in two unknowns.
- We notice that the homogeneous system in (3) has a nontrivial solution if and only if the determinant of its coefficient matrix is zero; that is, if and only if

$$\begin{vmatrix} \lambda - 1 & -1 \\ 2 & \lambda - 4 \end{vmatrix} = 0.$$

- This means that

$$(\lambda - 1)(\lambda - 4) + 2 = 0,$$

or

$$\lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda - 2).$$

- Hence $\lambda_1 = 2$ and $\lambda_2 = 3$ are the eigenvalues of A.
- To find all eigenvectors of A associated with $\lambda_1 = 2$, we form the linear system

$$A\mathbf{x} = 2\mathbf{x},$$

or

$$\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- This gives

$$\begin{aligned} x_1 + x_2 &= 2x_1 \\ -2x_1 + 4x_2 &= 2x_2 \end{aligned}$$

or

$$\begin{aligned} x_1 - x_2 &= 0 \\ 2x_1 - 2x_2 &= 0. \end{aligned}$$

- Note that we could have obtained this last homogeneous system by merely substituting $\lambda = 2$ in (3). All solutions to this last system are given by

$$x_1 = x_2$$

$$x_2 = \text{any real number } r.$$

- Hence all eigenvectors associated with the eigenvalue $\lambda_1 = 2$ are given by $\begin{bmatrix} r \\ r \end{bmatrix}$, r any nonzero real numbers. In particular,

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

  is an eigenvector associated with $\lambda_1 = 2$.

- Similarly, for $\lambda_2 = 3$ we obtain from (3)

$$\begin{aligned} (3-1)x_1 - x_2 &= 0 \\ 2x_1 + (3-4)x_2 &= 0 \end{aligned}$$

  or

$$\begin{aligned} 2x_1 - x_2 &= 0 \\ 2x_1 - x_2 &= 0. \end{aligned}$$

- All solutions to this last homogeneous system are given by

$$x_1 = \frac{1}{2}x_2$$

$$x_2 = \text{any real number } r.$$

- Hence all eigenvectors associated with the eigenvalue $\lambda_2 = 3$ are given by $\begin{bmatrix} \frac{1}{2}r \\ r \end{bmatrix}$, r any nonzero real number.
- In particular,

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is an eigenvector associated with the eigenvalue $\lambda_2 = 3$.

In Examples 2, 3 and 4 we found eigenvalues and eigenvectors by in inspection, whereas in Example 5 we proceed in a more systematic fashion.

We use the procedure of Example 5 as our standard method, as follows.

Definition 6.

- Let $A = [a_{ij}]$ be an $n \times n$ matrix.
- The determinant

$$f(\lambda) = \det(\lambda I_n - A) = \left| \begin{array}{cccc} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{array} \right]$$

  is called the **characteristic polynomial** of A.

- The equation

$$f(\lambda) = \det(\lambda I_n - A) = 0$$

  is called the **characteristic equation of** A.

Example 7.

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix}.$$

The characteristic polynomial of A is

$$f(\lambda) = \det(\lambda I_3 - A) = \begin{vmatrix} \lambda - 1 & -2 & 1 \\ -1 & \lambda - 0 & -1 \\ -4 & 4 & \lambda - 5 \end{vmatrix} = \lambda^3 - 6\lambda^2 + 11\lambda - 6.$$

Eigenvalues of an Upper Triangular Matrix

Example 8.

Find the eigenvalues of the upper triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}.$$

Solution

Recalling that the determinant of a triangular matrix is the product of the entries on the main diagonal, we obtain

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & -a_{14} \\ 0 & \lambda - a_{22} & -a_{23} & -a_{24} \\ 0 & 0 & \lambda - a_{33} & -a_{34} \\ 0 & 0 & 0 & \lambda - a_{44} \end{vmatrix} \\ &= (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) \end{aligned}$$

Thus, the characteristic equation is

$$(\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) = 0$$

and the eigenvalues are

Eigenvalues of an Lower Triangular Matrix

Example 9.

Find the eigenvalues of the lower triangular matrix

$$A = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -1 & \frac{2}{3} & 0 \\ 5 & -8 & -\frac{1}{4} \end{bmatrix}.$$

Solution The eigenvalues of the lower triangular matrix are $\lambda = \frac{1}{2}$, $\lambda = \frac{2}{3}$ and $\lambda = -\frac{1}{4}$

Recall from the properties of determinant that

each term in the expansion of the determinant of an $n \times n$ matrix is a product of n elements of the matrix, containing

- exactly one element from each row and
- exactly one element from each column.

Thus, if we expand $f(\lambda) = \det(\lambda I_n - A)$, we obtain a polynomial of degree n.

-
  - A polynomial of degree n with real coefficients has n roots (counting repeats), some of which may be complex numbers.
  - The expression involving λ^n in the characteristic polynomials of A comes from the product

    $$(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

    so the coefficient of λ^n is 1.
-
  - We can then write

    $$f(\lambda) = \det(\lambda I_n - A) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n$$

  - If we let $\lambda = 0$ in $\det(\lambda I_n - A)$ as well as in the expression on the right, then we get $\det(-A) = c_n$, which shows that the constant term c_n is $(-1)^n \det(A)$.

The next theorem establishes a relationship between the eigenvalues and the invertibility of a matrix.

Theorem 10.

An $n \times n$ matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof:

- Assume that A is an $n \times n$ matrix and observe first that $\lambda = 0$ is a solution of the characteristic equation
- We can then write

  $$\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n = 0$$

  if and only if the constant term c_n is zero.
- Thus it suffices to prove that A is invertible if and only if $c_n \neq 0$.
- But

  $$\det(\lambda I_n - A) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n$$

  or, on setting $\lambda = 0$,

  $$\det(-A) = c_n \quad \text{or} \quad (-1)^n \det(A) = c_n$$
- It follows from the last equation that $\det(A) = 0$ if and only if $c_n = 0$, and this in turn implies that A is invertible if and only if $c_n \neq 0$.

We now connect the characteristic polynomial of a matrix with its eigenvalues in the following theorem.

Theorem 11.

The eigenvalues of A are the roots of the characteristic polynomial of A.

Proof:

- Let λ be an eigenvalue of A with associated eigenvector $\mathbf{x}$.
- Then $A\mathbf{x} = \lambda\mathbf{x}$, which can be rewritten as

  $$A\mathbf{x} = (\lambda I_n)\mathbf{x}$$

  or

  $$(\lambda I_n - A)\mathbf{x} = \mathbf{0}, \tag{2}$$

  a homogeneous system of n equations in n unknowns.
- This system has a nontrivial solution if and only if the determinant of its coefficient matrix vanishes, that is, if and only if $\det(\lambda I_n - A) = 0$.

- Conversely,
    - if λ is a root of the characteristic polynomial of A, then $\det(\lambda I_n - A) = 0$,
    - so the homogeneous system has a nontrivial solution $\mathbf{x}$.
- Hence λ is an eigenvalue of A. $\blacksquare$

Thus, to find the eigenvalues of a given matrix A, we must find the roots of its characteristic polynomial $f(\lambda)$.

There are many methods for finding approximations to the roots of a polynomial, some of them more effective than others; indeed, many computer programs are available to find the roots of a polynomial.

Two results that are sometimes useful in this connection are as follows: The product of all the roots of the polynomial

$$f(\lambda) = \det(\lambda I_n - A) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n$$

is $(-1)^n a_n$, and if $a_1, a_2, \cdots, a_n$ are integers, then $f(\lambda)$ cannot have a rational root that is not already an integer. Thus, as possible rational roots of $f(\lambda)$, one need only try the integer factors of a_n.

- Of course, $f(\lambda)$ might well have irrational roots or complex roots.
- However, to minimize the computational effort and as a convenience to the reader, many of the characteristic polynomials to be considered in the rest of this lecture note have only integer roots, and each of these roots is a factor of the constant term of the characteristic polynomial of A.
- The corresponding eigenvectors are obtained by substituting the value of λ in Equation (2) and solving the resulting homogeneous system.

Example 12.

Consider the matrix of Example 7. The characteristic polynomial is

$$f(\lambda) = \lambda^3 - 6\lambda^2 + 11\lambda - 6.$$

The possible integer roots of $f(\lambda)$ are $\pm 1, \pm 2, \pm 3$ and ± 6. By substituting these values in $f(\lambda)$, we find $f(1) = 0$, so $\lambda = 1$ is a root of $f(\lambda)$. Hence $(\lambda - 1)$ is a factor of $f(\lambda)$. Dividing $f(\lambda)$ by $(\lambda - 1)$, we obtain

$$f(\lambda) = (\lambda - 1)(\lambda^2 - 5\lambda + 6).$$

Factoring $\lambda^2 - 5\lambda + 6$, we have

$$f(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

The eigenvalues of A are then

$$\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3.$$

To find an eigenvector $\mathbf{x}_1$ associated with $\lambda_1 = 1$, we form the linear system

$$(1I_3 - A)\mathbf{x} = \mathbf{0},$$

$$\begin{bmatrix} 1-1 & -2 & 1 \\ -1 & 1 & -1 \\ -4 & 4 & 1-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & -2 & 1 \\ -1 & 1 & -1 \\ -4 & 4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to \left[\begin{array}{rrr|r} 1 & 0 & 1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

A solution is

$$\begin{bmatrix} -\frac{1}{2}r \\ \frac{1}{2}r \\ r \end{bmatrix}$$

for any real number r. Thus for $r = 2$,

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_1 = 1$.

To find an eigenvector associated with $\lambda_2 = 2$, we form the linear system

$$(2I_3 - A)\mathbf{x} = \mathbf{0},$$

$$\begin{bmatrix} 2-1 & -2 & 1 \\ -1 & 2 & -1 \\ -4 & 4 & 2-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ -4 & 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 0 & 1/2 & 0 \\ 0 & 1 & -1/4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

A solution is

$$\begin{bmatrix} -\frac{1}{2}r \\ \frac{1}{4}r \\ r \end{bmatrix}$$

for any real number r. Thus for $r = 4$,

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_2 = 2$.

To find an eigenvector associated with $\lambda_3 = 3$, we form the linear system

$$(3I_3 - A)\mathbf{x} = \mathbf{0},$$

and find that a solution is

$$\begin{bmatrix} -\frac{1}{4}r \\ \frac{1}{4}r \\ r \end{bmatrix}$$

for any real number r. Thus for $r = 4$,

$$\mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_3 = 3$.

Example 13.

Compute the eigenvalues and associated eigenvectors of

$$A = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & -1 \\ 0 & 1 & 3 \end{bmatrix}.$$

Solution: The characteristic polynomial of A is

$$\begin{aligned} p(\lambda) = \det(\lambda I_3 - A) &= \begin{vmatrix} \lambda - 0 & 0 & -3 \\ -1 & \lambda - 0 & 1 \\ 0 & -1 & \lambda - 3. \end{vmatrix} \\ &= \lambda^3 - 3\lambda^2 + \lambda - 3 \end{aligned}$$

(verify). We find that $\lambda = 3$ is a root of $p(\lambda)$. Dividing $p(\lambda)$ by $(\lambda - 3)$, we obtain $p(\lambda) = (\lambda - 3)(\lambda^2 + 1)$. The eigenvalues of A are then

$$\lambda_1 = 3, \quad \lambda_2 = i, \quad \lambda_3 = -i.$$

To obtain an eigenvector $\mathbf{x}_1$ associated with $\lambda_1 = 3$, we substitute $\lambda = 3$ in (5), obtaining

$$\begin{bmatrix} 3-0 & 0 & -3 \\ -1 & 3-0 & 1 \\ 0 & -1 & 3-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

we find that the vector $\begin{bmatrix} r \\ 0 \\ r \end{bmatrix}$ is a solution for any number r (verify). Letting $r = 1$, we conclude that

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_1 = 3$.

To obtain an eigenvector $\mathbf{x}_2$ associated with $\lambda_2 = i$, we substitute $\lambda = i$ in Equation (2), obtaining

$$\begin{bmatrix} i-0 & 0 & -3 \\ -1 & i-0 & 1 \\ 0 & -1 & i-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

we find that the vector $\begin{bmatrix} (-3i)r \\ (-3+i)r \\ r \end{bmatrix}$ is a solution for any number r (verify).
Letting $r = 1$, we conclude that

$$\mathbf{x}_2 = \begin{bmatrix} -3i \\ -3+i \\ 1 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_2 = i$.

Similarly we find that

$$\mathbf{x}_3 = \begin{bmatrix} 3i \\ -3 - i \\ 1 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_3 = -i$.

The procedure for finding the eigenvalues and associated eigenvectors of a matrix is as follows:

Step 1: Determine the roots of the characteristic polynomial $f(\lambda) = \det(\lambda I_n - A)$. These are eigenvalues of A.

Step 2: For each eigenvalue λ, find all the nontrivial solutions to the homogeneous system $(\lambda I_n - A)\mathbf{x} = \mathbf{0}$. These are the eigenvectors of A associated with the eigenvalue λ.

- Of course, the characteristic polynomial of a matrix may have some complex roots and it may even have no real roots.
- However, in the important case of symmetric matrices, all the roots of the characteristic polynomial are real (Exercise).
- Eigenvalues and eigenvectors satisfy many important and interesting properties. For example, if A is an upper (lower) triangular matrix, or a diagonal matrix, then the eigenvalues of A are the elements on the main diagonal of A (Exercise).
- The set S consisting of all eigenvectors of A associated with λ_j as well as the zero vector is a subspace of $\mathbb{R}^n$ called the **eigenspace associated with** λ_j.

Theorem 14.

Let A be $n \times n$ matrix and λ an eigenvalue of A. The set of all eigenvectors corresponding to λ, together with the zero vector, is a subspace of $\mathbb{R}^n$. This subspace is called the **eigenspace** of λ.

Proof Let V be the set of all eigenvectors corresponding to λ, together with the zero vector. In order to show that V is a subspace, we have to show that it is closed under vector addition and scalar multiplication.
Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two vectors in V and c be a scalar. Then $A\mathbf{x}_1 = \lambda\mathbf{x}_1$ and $A\mathbf{x}_2 = \lambda\mathbf{x}_2$. Hence,

$$A\mathbf{x}_1 + A\mathbf{x}_2 = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2$$
$$A(\mathbf{x}_1 + \mathbf{x}_2) = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$$

Thus $\mathbf{x}_1 + \mathbf{x}_2$ is an eigenvector corresponding to λ. V is closed under addition.
Further, since $A\mathbf{x}_1 = \lambda\mathbf{x}_1$,

$$cA\mathbf{x}_1 = c\lambda\mathbf{x}_1$$
$$A(c\mathbf{x}_1) = \lambda(c\mathbf{x}_1)$$

Therefore $c\mathbf{x}_1$ is an eigenvector corresponding to λ. V is closed under scalar multiplication.
Thus V is a subspace of $\mathbb{R}^n$.

Example 15.

Find the eigenvalues and corresponding eigenspaces of the matrix

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

Solution: The matrix $\lambda I_3 - A$ is obtained by subtracting λ from the diagonal elements of A. Thus

$$\lambda I_3 - A = \begin{bmatrix} \lambda - 5 & -4 & -2 \\ -4 & \lambda - 5 & -2 \\ -2 & -2 & \lambda - 2 \end{bmatrix}.$$

The characteristic polynomial of A is $\det(\lambda I_3 - A)$. Using row and column operations to simplify determinants, we get

$$\begin{aligned}
\det(\lambda I_3 - A) &= \begin{vmatrix} \lambda-5 & -4 & -2 \\ -4 & \lambda-5 & -2 \\ -2 & -2 & \lambda-2 \end{vmatrix} = \begin{vmatrix} \lambda-1 & -\lambda+1 & 0 \\ -4 & \lambda-5 & -2 \\ -2 & -2 & \lambda-2 \end{vmatrix} \\
&= \begin{vmatrix} \lambda-1 & 0 & 0 \\ -4 & \lambda-9 & -2 \\ -2 & -4 & \lambda-2 \end{vmatrix} \\
&= (\lambda-1)[(\lambda-9)(\lambda-2)-8] \\
&= (\lambda-1)[\lambda^2-11\lambda+10] \\
&= (\lambda-1)(\lambda-1)(\lambda-10) \\
&= (\lambda-1)^2(\lambda-10)
\end{aligned}$$

We now solve the characteristic equation of A:

$$\begin{aligned}
(\lambda-1)^2(\lambda-10) &= 0 \\
\lambda &= 10 \text{ or } 1
\end{aligned}$$

The eigenvalues of A are 10 and 1.

The corresponding eigenvectors are found by using these values of λ in the equation $\det(\lambda I_3 - A)\mathbf{x} = \mathbf{0}$.

$\lambda = 10$: We get

$$\text{Characteristic equation :} \qquad \det(10I_3 - A)\mathbf{x} = \mathbf{0}$$

$$\text{Homogeneous system :} \qquad (10I_3 - A)\mathbf{x} = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The solutions to this system of equations are $x_1 = 2r$, $x_2 = 2r$, and $x_3 = r$, where r is a scalar. Thus the eigenspace of $\lambda = 10$ is the one-dimensional space of vectors of the form

$$r \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \qquad \text{where} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

$\lambda = 1$: We get

$$\text{Characteristic equation :} \qquad \det(1I_3 - A)\mathbf{x} = \mathbf{0}$$

$$\text{Homogeneous system :} \qquad (1I_3 - A)\mathbf{x} = \begin{bmatrix} -4 & -4 & -2 \\ -4 & -4 & -2 \\ -2 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The solutions to this system of equations are $x_1 = -s - t$, $x_2 = s$, and $x_3 = 2t$, where s and t are scalars. Thus the eigenspace of $\lambda = 1$ is the space of vectors of the form

$$\begin{bmatrix} -s - t \\ s \\ 2t \end{bmatrix}, \qquad \text{where} \begin{bmatrix} 1 & 1 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Separating the parameters s and t, we can write

$$\begin{bmatrix} -s - t \\ s \\ 2t \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

Thus the eigenspace of $\lambda = 1$ is a two-dimensional subspace of $\mathbb{R}^2$ with basis

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}$$

If an eigenvalue occurs as k times repeated root of the characteristic equations, we say that it is of **multiplicity** k.

Thus $\lambda = 10$ has multiplicity 1, which $\lambda = 1$ has multiplicity 2 in this example.

Remark 2.

It must be pointed out that the method for finding the eigenvalues of a linear transformation or matrix by obtaining the roots of the characteristic polynomial is not practical for $n > 4$, since it involves evaluating a determinant.

Efficient numerical methods for finding eigenvalues and eigenvectors are studied in numerical analysis.

Warning:

When finding the eigenvalues and associated eigenvectors of a matrix A, do not make the common mistake of the first transforming A to reduced row echelon form B and then finding the eigenvalues and eigenvectors of B.

To see quickly how this approach fails, consider the matrix A defined in Example 5.

- Its eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 3$.
- Since A is a nonsingular matrix, when we transform it to reduced row echelon form B, we have $B = I_2$.
- The eigenvalues of I_2 are $\lambda_1 = 1$ and $\lambda_2 = 1$.

Diagonalization

- In this section we know how to find the eigenvalues and associated eigenvectors of a given matrix A by finding the eigenvalues and eigenvectors of a related matrix B that has the same eigenvalues and eigenvectors as A.
- The matrix B has the helpful property that its eigenvalues are easily obtained. Thus, we will have found the eigenvalues of A.
- In the following section, this approach will shed much light on the eigenvalue-eigenvector problem. For convenience, we only work with matrices all of those entries and eigenvalues are real numbers.

Similar Matrices

Definition 16.

A matrix B is said to be **similar** to a matrix A if there is a nonsingular (or an invertible) matrix P such that

$$B = P^{-1}AP.$$

Example 17.

Let

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

be the matrix of Example 5. Let

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Then

$$P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

and

$$B = P^{-1}AP = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Thus B is similar to A.

Note that the following elementary properties hold for similarity:

1. A is similar to A.
2. If B is similar to A, then A is similar to B.
3. If A is similar to B and B is similar to C, then A is similar to C.

By property 2 we replace the statements "A is similar to B" and "B is similar to A" by A and B are similar.

Definition 18.

We shall say that the matrix A is **diagonalizable** if it is similar to a diagonal matrix. In this case we also say that A can be **diagonalized**.

Example 19.

If A and B are as in Example 17, then A is diagonalizable, since it is similar to B.

Theorem 20.

Similar matrices have the same eigenvalues.

Proof:

- Let A and B be similar matrices.
- Then $B = P^{-1}AP$ for some nonsingular matrix P.
- We prove that A and B have the same characteristic polynomial, $f_A(\lambda)$ and $f_B(\lambda)$, respectively.
- We have

$$\begin{aligned}
f_B(\lambda) &= \det(\lambda I_n - B) = \det(\lambda I_n - P^{-1}AP) \\
&= \det(P^{-1}\lambda I_n P - P^{-1}AP) = \det(P^{-1}(\lambda I_n - A)P) \\
&= \det(P^{-1})\det(\lambda I_n - A)\det(P) \\
&= \det(\lambda I_n - A)\det(P^{-1})\det(P) \\
&= \det(\lambda I_n - A)\det(P^{-1}P) \\
&= \det(\lambda I_n - A)\det(I) \\
&= \det(\lambda I_n - A) \\
&= f_A(\lambda). \qquad (3)
\end{aligned}$$

- Since $f_A(\lambda) = f_B(\lambda)$, it follows that A and B have the same eigenvalues.

Note that the eigenvalues of a diagonal matrix are the entries on its main diagonal (Exercise). The following theorem establishes when a matrix is diagonalizable.

Theorem 21.

An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

Proof:
The matrix P whose columns consists of n linearly independent eigenvectors can be used in a similarity transformation $P^{-1}AP$ to give a diagonal matrix D. The diagonal elements of D will be the eigenvalues of A.

1.

- Let A have eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, (which need not be distinct), with corresponding linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$.
- Let P be the matrix having $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ as column vectors.

$$P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{array} \right].$$

- Since $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$, $A\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$, $\cdots$, $A\mathbf{x}_n = \lambda_n \mathbf{x}_n$, matrix multiplication in terms of columns gives

$$\begin{aligned} AP &= A \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{array} \right] \\ &= \left[\begin{array}{cccc} A\mathbf{x}_1 & A\mathbf{x}_2 & \ldots & A\mathbf{x}_n \end{array} \right] \\ &= \left[\begin{array}{cccc} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \ldots & \lambda_n\mathbf{x}_n \end{array} \right] \\ &= \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{array} \right] \left[\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{array} \right] \\ &= P \left[\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{array} \right] \end{aligned}$$

- Since the columns of P are linearly independent, P is nonsingular.
- Therefore,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

- Therefore, if an $n \times n$ matrix A has n linearly independent eigenvectors, these eigenvectors can be used as the columns of a matrix P that diagonalizes A.
- The diagonal matrix has the eigenvalues of A as diagonal elements.

2. The converse is proved by retracting the above steps.
    - Commence with the assumption that P is a matrix $\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix}$ that diagonalizes A.
    - Thus, there exist scalars $\lambda_1, \lambda_2, \cdots, \lambda_n$ such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

    - Retracting the above steps, we arrive at the conclusion that

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, A\mathbf{x}_2 = \lambda_2\mathbf{x}_2, \cdots, A\mathbf{x}_n = \lambda_n\mathbf{x}_n$$

    - The $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are eigenvectors of A.
    - Since P is nonsingular, these vectors (column vectors of P) are linearly independent.
    - Thus if an $n \times n$ matrix A is diagonalizable, it has n linearly independent eigenvectors.

Remark 3.

- If A is diagonalizable matrix, then $P^{-1}AP = D$, where D is a diagonal matrix.
- It follows from the proof of Theorem 21 that the diagonal elements of D are the eigenvalues of A.
- Moreover, P is a matrix whose columns are, respectively, n linearly independent eigenvectors of A.
- Observe also that in Theorem 21, the order of the columns of P determines the order of the diagonal entries in D.

If an $n \times n$ matrix A has fewer than n linearly independent eigenvectors, we say that A is *defective*. It follows from Theorem 21 that a defective matrix is not diagonalizable.

Example 22.

Let A be as in Example 19. The eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 3$ (see Example 5). The corresponding eigenvectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

are linearly independent. Hence A is diagonalizable. Here

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Thus, as in Example 2,

$$P^{-1}AP = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

On the other hand, if we let $\lambda_1 = 3$ and $\lambda_2 = 2$, then

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Then

$$P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad P^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}.$$

Hence

$$P^{-1}AP = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}.$$

- If A is similar to a diagonal matrix D under the transformation $P^{-1}AP$, then it can be shown that $A^k = PD^kP^{-1}$.
- This result can be used to compute A^k.
- Let us derive this result and then apply it.

$$D^k = (P^{-1}AP)^k = \underbrace{(P^{-1}A\textcolor{red}{P})(\textcolor{red}{P^{-1}}A\textcolor{red}{P})\cdots(\textcolor{red}{P^{-1}}A\textcolor{red}{P})(\textcolor{red}{P^{-1}}AP)}_{k\ \text{times}} = P^{-1}A^kP$$

- This leads to

$$A^k = PD^kP^{-1}.$$

Example 23.

Compute A^9 for the following matrix A.

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}.$$

Solution A is the matrix of Example 22. Use the values of P and D from that example. We get

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^9 = \begin{bmatrix} 2^9 & 0 \\ 0 & 3^9 \end{bmatrix} = \begin{bmatrix} 512 & 0 \\ 0 & 19683 \end{bmatrix}$$

The transformation now gives

$$\begin{aligned} A^9 &= PD^9P^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 512 & 0 \\ 0 & 19683 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -18659 & 19171 \\ -38342 & 138854 \end{bmatrix} \end{aligned}$$

Example 24.

Let

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

- Using the eigenvalue-eigenvector system, we have

$$\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

- The characteristic equation becomes

$$\begin{vmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2 = 0.$$

- The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 1$.
- Now

$$(\lambda I_2 - A)\mathbf{x} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- Eigenvectors associated with λ_1 and λ_2 are vectors of the form

$$\begin{bmatrix} r \\ 0 \end{bmatrix},$$

  where r is any nonzero real number.
- Since A does not have two linearly independent eigenvectors, we conclude that A is not diagonalizable.

Example 25.

Show that the following matrix A is not diagonalizable.

$$A = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}.$$

Solution

- Let us compute the eigenvalues and corresponding eigenvectors of A.
- We get

$$\lambda I_2 - A = \begin{bmatrix} \lambda - 5 & 3 \\ -3 & \lambda + 1 \end{bmatrix}$$

- The characteristic equation is

$$\det(\lambda I_2 - A) = \det\left(\begin{bmatrix} \lambda - 5 & 3 \\ -3 & \lambda + 1 \end{bmatrix}\right) = 0$$

$$(\lambda - 5)(\lambda + 1) + 9 = 0$$

$$(\lambda - 2)(\lambda - 2) = 0$$

- There is a single repeated eigenvalue, $\lambda = 2$.
- We now find the corresponding eigenvectors.
- $(2I_2 - A)\mathbf{x} = \mathbf{0}$ gives

$$\begin{bmatrix} -3 & 3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- This gives $3x_1 - 3x_2 = 0$.
- Thus $x_1 = r$, $x_2 = r$.
- The eigenvectors are nonzero vectors of the form

$$r \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

  The eigenspace is a one-dimensional space.
- A is 2×2 matrix, but it does not have two linearly independent eigenvectors. Thus A is not diagonalizable.

The following is a useful theorem because it identifies a large class of matrices that can be diagonalized.

Theorem 26.

If the roots of the characteristic polynomial of an $n \times n$ matrix A are all distinct (i.e., different from each other), then A is diagonalizable.

Proof:
Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the distinct eigenvalues of A and let $S = \{\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots\ \mathbf{x}_n\}$ be a set of associated eigenvectors.
We wish to show that S is linearly independent.

- Suppose that S is linearly dependent.
- Then the vector space lecture note implies that some vectors $\mathbf{x}_j$ is a linear combination of the preceding vectors in S.
- We can assume that $S_1 = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{j-1}\}$ is linearly independent, for otherwise one of the vectors in S_1 is a linear combination of the preceding ones, and we can choose a new set S_2, and so on.
- We thus have that S_1 is linearly independent and that

  $$\mathbf{x}_j = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_{j-1}\mathbf{x}_{j-1}, \tag{4}$$

  where $c_1, c_2, \ldots, c_{j-1}$ are scalars.
- Pre-multiplying both sides of Equation (4) by A, we obtain

  $$\begin{aligned} A\mathbf{x}_j &= A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_{j-1}\mathbf{x}_{j-1}) \\ &= c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 + \ldots + c_{j-1}A\mathbf{x}_{j-1}. \end{aligned} \tag{5}$$

- Since $\lambda_1, \lambda_2, \ldots, \lambda_j$ are eigenvalues of A and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j$, its associated eigenvectors, we know that $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$ for $i = 1, 2, \ldots, j$.
- Substituting in Equation (5), we have

$$\lambda_j \mathbf{x}_j = c_1 \lambda_1 \mathbf{x}_1 + c_2 \lambda_2 \mathbf{x}_2 + \ldots + c_{j-1} \lambda_{j-1} \mathbf{x}_{j-1}, \tag{6}$$

- Multiplying Equation (4) by λ_j we obtain

$$\lambda_j \mathbf{x}_j = \lambda_j c_1 \mathbf{x}_1 + \lambda_j c_2 \mathbf{x}_2 + \ldots + \lambda_j c_{j-1} \mathbf{x}_{j-1}, \tag{7}$$

- Substituting in Equation (5), we have

$$\lambda_j \mathbf{x}_j = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \ldots + c_{j-1}\lambda_{j-1}\mathbf{x}_{j-1}, \tag{8}$$

- Multiplying Equation (4) by λ_j we obtain

$$\lambda_j \mathbf{x}_j = \lambda_j c_1\mathbf{x}_1 + \lambda_j c_2\mathbf{x}_2 + \ldots + \lambda_j c_{j-1}\mathbf{x}_{j-1}, \tag{9}$$

- Subtracting Equations (9) from (8) we have

  $$\mathbf{0} = \lambda_j\mathbf{x}_j - \lambda_j\mathbf{x}_j = c_1(\lambda_1 - \lambda_j)\mathbf{x}_1 + c_2(\lambda_2 - \lambda_j)\mathbf{x}_2 + \ldots + c_{j-1}(\lambda_{j-1} - \lambda_j)\mathbf{x}_{j-1}.$$

- Since S_1 is linearly independent, we must have

$$c_1(\lambda_1 - \lambda_j) = 0, c_2(\lambda_2 - \lambda_j) = 0, \ldots, c_{j-1}(\lambda_{j-1} - \lambda_j) = 0$$

- Now

  $$\lambda_1 - \lambda_j \neq 0, \lambda_2 - \lambda_j \neq 0, \ldots, \lambda_{j-1} - \lambda_j \neq 0$$

  (because the λ's are distinct), which implies that

  $$c_1 = c_2 = \ldots = c_{j-1} = 0.$$

- From (4) we conclude that $\mathbf{x}_j = \mathbf{0}$, which is impossible if $\mathbf{x}_j$ is an eigenvector.
- Hence S is linearly independent, and from Theorem 21 it follows that A is diagonalizable.

Remark 4.

In the proof of Theorem 26 we have actually established the following somewhat stronger result:

Let A be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be k distinct eigenvalues of A with associated eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$. Then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ are linearly independent.

- If the roots of the characteristic polynomial of A are not all distinct, then A may or may not be diagonalizable.
- The characteristic polynomial of A can be written as the product of n factors, each of the form $\lambda - \lambda_j$, where λ_j is a root of the characteristic polynomial of A.
- Thus the characteristic polynomial can be written as

  $$(\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_r)^{k_r},$$

  where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct eigenvalues of A, and $k_1, k_2, \cdots, k_r$ are integers whose sum is n. The integer k_i is called the multiplicity of λ_i.
- Thus in Example 24, λ_1 is an eigenvalue of

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

  of multiplicity 2.

- It can be shown that A can be diagonalized if and only if for each eigenvalue λ_j of multiplicity k_j linearly independent eigenvectors.
- This means that the solution space of linear system $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$ has dimension k_j.
- It can also be shown that if λ_j is an eigenvalue of A of multiplicity k_j, then we can never find more than k_j linearly independent eigenvectors associated with λ_j.

We consider the following examples.

Example 27.

Let

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

- The characteristic polynomial of A is $f(\lambda) = \lambda(\lambda - 1)^2$, so the eigenvalues of A are $\lambda_1 = 0, \lambda_2 = 1$ and $\lambda_3 = 1$.
- Thus $\lambda_2 = 1$ is an eigenvalue of multiplicity 2.
- We now consider the eigenvectors associated with the eigenvalues $\lambda_2 = \lambda_3 = 1$.
- They are obtained by solving the linear system $(1I_3 - A)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Example 28.

Let

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

- The characteristic polynomial of A is $f(\lambda) = \lambda(\lambda - 1)^2$, so the eigenvalues of A are $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 1$; $\lambda_2 = 1$ is again an eigenvalue of multiplicity 2.
- Now we consider the solution space of $(1I_3 - A)\mathbf{x} = \mathbf{0}$, that is, of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- A solution is any vector of the form

$$\begin{bmatrix} 0 \\ r \\ s \end{bmatrix}$$

  for any real numbers r and s.
- Thus we can take as eigenvectors $\mathbf{x}_2$ and $\mathbf{x}_3$ the vectors

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

- A solution is any vector of the form

$$\begin{bmatrix} 0 \\ r \\ s \end{bmatrix}$$

  for any real numbers r and s.

- Thus we can take as eigenvectors $\mathbf{x}_2$ and $\mathbf{x}_3$ the vectors

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

- Now we look for an eigenvector associated with $\lambda_1 = 0$.
- We have to solve the homogeneous system $(0I_3 - A)\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- A solution is any vector of the form

$$\begin{bmatrix} t \\ 0 \\ -t \end{bmatrix}$$

  for any real number t.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

  is an eigenvector associated with $\lambda_1 = 0$.
- Since $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent, A can be diagonalized.

Example 29.

Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Solution

There are four main steps to implement the description of Theorem 21.

Step 1: Find the eigenvalues of A. The characteristic equation turns out to involve a cubic polynomial that can be factored:

$$0 = \det(\lambda I_3 - A) = \lambda^3 + 3\lambda^2 - 4 = (\lambda - 1)(\lambda + 2)^2$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$.

Step 2: Find three linearly eigenvectors of A. This is the crucial step. If it fails, A cannot be diagonalized.

1. For $\lambda = 1$:
   We have to solve the homogeneous system $(1I_3 - A)\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} 0 & -3 & -3 \\ 3 & 6 & 3 \\ -3 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

   Now,

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

2. For $\lambda = -2$:
   We have to solve the homogeneous system $(-2I_3 - A)\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} -3 & -3 & -3 \\ 3 & 3 & 3 \\ -3 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

   Now,

$$\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

   Thus $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is a linearly independent set (verify).

Step 3: Construct P from the vectors in Step 2. Using the order chosen in **Step 2** form

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Step 4: Construct D from the corresponding eigenvalues. The order of the eigenvalues matches the order chosen for P.

Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors to $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Let us check that P and D really work. To avoid confusion, we simply verify that $AP = PD$. This is equivalent to $A = PDP^{-1}$ when P is invertible. We compute

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

Example 30.

Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

Solution

The characteristic equation of A turns out be exactly the same as that in the last example:

$$0 = \det(\lambda I_3 - A) = (\lambda - 1)(\lambda + 2)^2$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$. However, when we look for eigenvectors, we find that each eigenspace is only one-dimensional.

1 For $\lambda = 1$:
We have to solve the homogeneous system $(1I_3 - A)\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} -1 & -4 & -3 \\ 4 & 7 & 3 \\ -3 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Now,

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

2 For $\lambda = -2$:
We have to solve the homogeneous system $(-2I_3 - A)\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} -4 & -4 & -3 \\ 4 & 4 & 3 \\ -3 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Now,

$$\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

There are no other eigenvalues, and every eigenvector of A is multiple of either $\mathbf{u}_1$ and $\mathbf{u}_2$. Hence it is impossible to construct a basis of R^3 using eigenvectors of A.
By Theorem 21, A is not diagonalizable.

Thus an $n \times n$ matrix will fail to be diagonalizable only if it does not have n linearly independent eigenvectors.

The procedure for diagonalizing a matrix A is as follows

Step 1 Form the characteristic polynomial $f(\lambda) = \det(\lambda I_n - A)$ of A.

Step 2 Find the roots of the characteristic polynomial of A.

Step 3 For each eigenvalue λ_j of A of multiplicity k_j, find a basis for the solution space $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$ (the eigenspace associated with λ_j). If the dimension of the eigenspace is less than k_j, then A is not diagonalizable. We thus determine n linearly independent eigenvectors of A. Then we solve the problem of finding a basis for the solution space of a homogeneous system.

Step 4 Let P be the matrix whose columns are the n linearly independent eigenvectors determined in Step 3. Then $P^{-1}AP = D$, a diagonal matrix whose diagonal elements are the eigenvalues of A that corresponds to the columns of P.

Diagonalization of Symmetric Matrices

In this lecture we consider the diagonalization of symmetric matrices (an $n \times n$ matrix A with real entries that $A = A^T$).

We restrict our attention to this case because symmetric matrices arise in many applied problems.

As an example of such a problem, consider the task of *identifying the conic* represented by the equation

$$2x^2 + 2xy + 2y^2 = 9,$$

which can be written in matrix form as

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 9$$

Observe that the matrix used here is a symmetric matrix.

We shall merely remark here that the solution calls for the determination of the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

The x- and y-axes are then rotated to a new set of axes, which lie along the eigenvectors of the matrix. In the new set of axes, the given conic can be identified readily.

Theorem 31.

All the roots of the characteristic polynomial of a symmetric matrix are real numbers.

Theorem 32.

If A is a symmetric matrix, then eigenvectors that are associated with distinct eigenvalues of A are orthogonal.

Example 33.

Given the symmetric matrix

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix},$$

we find that the characteristic polynomial of A is (verify)

$$f(\lambda) = (\lambda + 2)(\lambda - 4)(\lambda + 1),$$

so the eigenvalues of A are

$$\lambda_1 = -2,\ \lambda_2 = 4, \quad \lambda_3 = -1.$$

Then we can find the associated eigenvectors by solving the homogeneous system $(\lambda_j I_3 - A)\mathbf{x} = \mathbf{0}, j = 1, 2, 3$ and obtain the respective eigenvectors (verify)

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

It is easy to check that $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is an orthogonal set of vectors in $\mathbb{R}^3$.

Thus A is diagonalizable and is similar to

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

- We call that
    - if A can be diagonalized,
    - then there exists a nonsingular matrix P such that $P^{-1}AP$ diagonal.
- Moreover, the columns of P are eigenvectors of A.
- Now,
    - if the eigenvectors of A form <u>an orthogonal set</u> S, as happens when A is symmetric and the eigenvalues of A are distinct,
    - then since any scalar multiple of an eigenvector of A is also an eigenvector of A, we can normalize S to obtain <u>an orthonormal set</u> $T = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ of eigenvectors of A.

- The jth column of P is the eigenvector λ_j, and we now examine what type of matrix P must be.
- Write P as
$$P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array} \right].$$
Then
$$P^T = \left[\begin{array}{c} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{array} \right],$$
where $\mathbf{x}_i^T$, $1 \leq i \leq n$, is the transpose of the $n \times 1$ matrix (or vector) $\mathbf{x}_i$.
- We find that the i, jth entry in $P^T P$ is $\mathbf{x}_i \cdot \mathbf{x}_j$. Since
$$\mathbf{x}_i \cdot \mathbf{x}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$
then $P^T P = I_n$. Thus $P^T = P^{-1}$.
- Such matrices are important enough to have a special name.

Orthogonal Matrix

Example 34.

Let

$$A = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}.$$

It is easy to check that $A^T A = I_3$. Hence A is an orthogonal matrix.

Example 35.

Let A be the matrix in Example 33. We already know that the set of eigenvectors

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is orthogonal. If we normalize these vectors, we find that

$$T = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}, \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix} \right\}$$

is an orthonormal set of vectors. The matrix P such that $P^{-1}AP$ is diagonal is the matrix whose columns are the vectors in T. Thus

$$P = \begin{bmatrix} 0 & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

We leave the student to verify that P is an orthogonal matrix and that

$$P^{-1}AP = P^{T}AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Theorem 36.

The $n \times n$ matrix A is orthogonal if and only if the columns (and rows) of A form an orthonormal set of vectors in $\mathbb{R}^n$.

If A is an orthogonal matrix, then it is easy to show that $\det(A) = \pm 1$.

We now turn to the general situation for a symmetric matrix; even if A has eigenvalues whose multiplicities are greater than one, it turns out that we can still diagonalize A. We omit the proof of the theorem. For a proof, see J. M. Ortege, Matrix Theory, a second course, New York: Plenum Press, 1987.

Theorem 37.

- If A is symmetric $n \times n$ matrix, then there exists an orthogonal matrix P such that $P^{-1}AP = D$, a diagonal matrix.
- The eigenvalues of A lie on the main diagonal of D.

Thus, not only is a symmetric matrix always diagonalizable, but it is diagonalizable by means of an orthogonal matrix. In such a case, we say that A is **orthogonally diagonalizable**.

- It can be shown that
  - if a symmetric matrix A has an eigenvalue λ_j of multiplicity k_j,
  - then the solution space of the linear system $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$ (the eigenspace of λ_j) has dimension k_j.

  This means that there exist k_j linearly independent eigenvectors of A associated with the eigenvalue λ_j.
-
  - By the Gram-Schmidt process, we can construct an orthonormal basis for the solution space.
  - Thus we obtain a set of k_j orthonormal eigenvectors associated with the eigenvalue λ_j.
-
  - Since eigenvectors associated with distinct eigenvalues are orthogonal, if we form the set of all eigenvectors, we get an orthonormal set.
  - Hence the matrix P whose columns are the eigenvectors is orthogonal.

The procedure for diagonalizing a symmetric matrix A by an orthogonal matrix P is as follows:

Step 1 Form the characteristic polynomial $f(\lambda) = \det(\lambda I_n - A)$.

Step 2 Find the roots of the characteristic polynomial of A. These will all be real.

Step 3 For each eigenvalue λ_j of A of multiplicity k_j, find a basis of k_j eigenvectors for the the solution space $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$
(the eigenspace associated with λ_j).

Step 4 For each eigenspace, transform the basis obtained in **Step 3** to an orthonormal basis by the Gram-Schmidt process. The totality of all these orthonormal bases determines an orthonormal set of n linearly independent eigenvectors of A.

Step 5 Let P be the matrix whose columns are the n linearly independent eigenvectors determined in **Step 4**. Then P is an orthogonal matrix and $P^{-1}AP = P^T AP = D$, a diagonal matrix whose diagonal elements are the eigenvalues of A that correspond to the columns of P.

Example 38.

Let

$$A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}.$$

The characteristic polynomial of A is (verify)

$$f(\lambda) = (\lambda + 2)^2(\lambda - 4),$$

so the eigenvalues are

$$\lambda_1 = -2, \quad \lambda_2 = -2, \quad \lambda_3 = 4.$$

That is, -2 is an eigenvalue whose multiplicity is 2. Next, to find the eigenvectors associated with λ_1 and λ_2, we solve the homogeneous linear system $(-2I_3 - A)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{10}$$

A basis for the solution space of Equation (10) consists of the eigenvectors (verify)

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Now $\mathbf{x}_1$ and $\mathbf{x}_2$ are not orthogonal, since $\mathbf{x}_1 \cdot \mathbf{x}_2 \neq 0$. We can use the Gram-Schmidt process to obtain an orthonormal basis for the solution space of Equation (10) (the eigenspace of $\lambda_1 = -2$) as follows. Let

$$\mathbf{y}_1 = \mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

and

$$\mathbf{y}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\mathbf{y}_1 \cdot \mathbf{y}_1}\mathbf{y}_1 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}.$$

Let

$$\mathbf{y}_2^* = 2\mathbf{y}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

The set $\{\mathbf{y}_1, \mathbf{y}_2^*\}$ is an orthogonal set of eigenvectors. Normalizing these eigenvectors we obtain

$$\mathbf{z}_1 = \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{z}_2 = \frac{1}{\|\mathbf{y}_2^*\|}\mathbf{y}_2^* = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}.$$

The set $\{\mathbf{z}_1, \mathbf{z}_2\}$ is an orthonormal basis of eigenvectors of A for the solution space of Equation (10).

Now we find a basis for the solution space of $(4I_3 - A)\mathbf{x} = \mathbf{0}$,

$$\begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{11}$$

to consist of (verify)

$$\mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Normalizing this vector, we have the eigenvector

$$\mathbf{z}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

as an orthonormal basis for the solution space of Equation (11).

Since eigenvectors associated with distinct eigenvalues are orthogonal, we observe that $\mathbf{z}_3$ is orthogonal to both $\mathbf{z}_1$ and $\mathbf{z}_2$.

Thus the set $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$ is an orthonormal basis for $\mathbb{R}^3$ consisting of eigenvectors of A.

The matrix P is the matrix whose jth column is $\mathbf{z}_j$:

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

We leave the student to verify that

$$P^{-1}AP = P^T AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Example 39.

Let

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{bmatrix}.$$

The characteristic polynomial of A is (verify)

$$f(\lambda) = (\lambda + 1)^2 (\lambda - 3)^2,$$

so the eigenvalues of A are

$$\lambda_1 = -1, \;\; \lambda_2 = -1, \;\; \lambda_3 = 3, \;\; \lambda_4 = 3.$$

We find (verify) that a basis for the solution space of

$$(-1 I_3 - A)\mathbf{x} = \mathbf{0} \tag{12}$$

consists of eigenvectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix},$$

which are orthogonal. Normalizing these eigenvectors, we obtain

$$\mathbf{z}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{z}_4 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

as an orthonormal basis of eigenvectors for the solution space of Equation (**??**). Since eigenvectors associated with distinct eigenvalues are orthogonal, we conclude that

$$\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4\}$$

is an orthonormal basis for $\mathbb{R}^4$ consisting of eigenvectors of A. The matrix P is the matrix whose jth column is $\mathbf{z}_j$:

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

- Suppose now that A is an $n \times n$ matrix that is orthogonally diagonlizable.
- Thus we have an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix D.
- Then $P^{-1}AP = D$, or $A = PDP^{-1}$.
- Since $P^{-1} = P^T$, we can write $A = PDP^T$.
- Then
  $$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$
  ($D = D^T$ since D is a diagonal matrix). Thus A is symmetric.

This result, together with Theorem 21, yields the following theorem.

Theorem 40.

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is symmetric.