

MA2001

MULTI-VARIABLE CALCULUS AND LINEAR ALGEBRA

Chapter 1

Eigenvalues and Eigenvectors

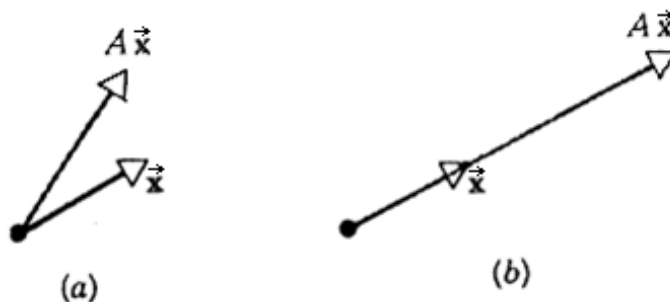
Dr. Emily Chan

Chapter 1: Eigenvalues and Eigenvectors

- 1 Introduction**
- 2 Determination of Eigenvalues and Eigenvectors**
- 3 Eigenvalue Results**
 - 3.1 Diagonalization**
- 4 Quadratic Forms**
 - 4.1 Introduction**
 - 4.2 Definiteness**

1 Introduction

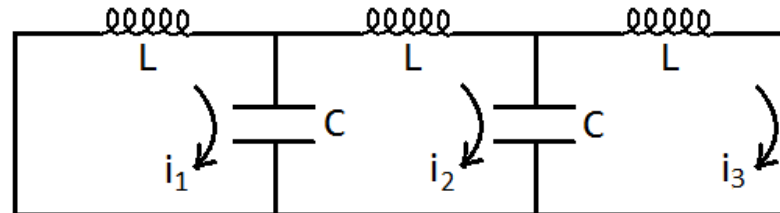
If A is a 3×3 matrix and $\vec{x}$ is a vector in $\mathbb{R}^3$, then there is usually no general geometric relationship between the vector $\vec{x}$ and the vector $A\vec{x}$, see figure (a). However, there are often certain nonzero vectors $\vec{x}$ such that $\vec{x}$ and $A\vec{x}$ are scalar multiples of one another, that is, $A\vec{x} = \lambda\vec{x}$ for some real number λ , see figure (b). Such vectors arise naturally in the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, mechanical stress, economics, and geometry.



Let A be an $n \times n$ square matrix and consider the equation $A\vec{x} = \lambda\vec{x}$. A value of λ for which $A\vec{x} = \lambda\vec{x}$ has a non-trivial solution $\vec{x} \neq \vec{0}$ is called an **eigenvalue** (or characteristic value) of A and the corresponding solution $\vec{x} \neq \vec{0}$ is an **eigenvector** (or characteristic vector). The set of all eigenvalues of A is known as the spectrum of A .

Application (for your reference) – Electric Circuit

Consider the electrical circuit



The currents flowing in the loops satisfy

$$L \frac{di_1}{dt} + \frac{1}{C} \int (i_1 - i_2) dt = 0 \Rightarrow LCi_1'' + i_1 - i_2 = 0$$

$$L \frac{di_2}{dt} + \frac{1}{C} \int (i_2 - i_3) dt + \frac{1}{C} \int (i_2 - i_1) dt = 0 \Rightarrow LCi_2'' + 2i_2 - i_3 - i_1 = 0$$

$$L \frac{di_3}{dt} + \frac{1}{C} \int (i_3 - i_2) dt = 0 \Rightarrow LCi_3'' + i_3 - i_2 = 0$$

In matrix form,

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix} = -LC \begin{pmatrix} i_1'' \\ i_2'' \\ i_3'' \end{pmatrix}.$$

If we look for solutions which are sinusoidal and all have the same frequency, that is,

$$i_1 = A_1 \sin(\omega t + \theta_1), \quad i_2 = A_2 \sin(\omega t + \theta_2), \quad i_3 = A_3 \sin(\omega t + \theta_3),$$

then $i_1'' = -\omega^2 i_1$, $i_2'' = -\omega^2 i_2$ and $i_3'' = -\omega^2 i_3$ and the problem becomes

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix} = \omega^2 LC \begin{pmatrix} i_1 \\ i_2 \\ i_3 \end{pmatrix},$$

which is an eigenvalue problem with $\lambda = \omega^2 LC$. □

2 Determination of Eigenvalues and Eigenvectors

$$A\vec{x} = \lambda\vec{x} \Leftrightarrow A\vec{x} = \lambda I\vec{x} \Leftrightarrow A\vec{x} - \lambda I\vec{x} = \vec{0} \Leftrightarrow (A - \lambda I)\vec{x} = \vec{0}.$$

For this equation $(A - \lambda I)\vec{x} = \vec{0}$ to have **non-trivial solutions** $\vec{x} \neq \vec{0}$, we must have $(A - \lambda I)^{-1}$ **not existing**, otherwise

$$(A - \lambda I)\vec{x} = \vec{0} \Rightarrow (A - \lambda I)^{-1}(A - \lambda I)\vec{x} = (A - \lambda I)^{-1}\vec{0} \Rightarrow I\vec{x} = \vec{0} \Rightarrow \vec{x} = \vec{0}$$

which is the unique (**trivial**) **solution**.

Hence,

$$\begin{aligned} \lambda \text{ is an } \mathbf{eigenvalue} &\Leftrightarrow (A - \lambda I)^{-1} \text{ does not exist} \\ &\Leftrightarrow \boxed{\det(A - \lambda I) = 0}. \end{aligned}$$

Consider $\det(A - \lambda I)$. After expansion, $\det(A - \lambda I)$ is a polynomial of degree n in λ , known as the **characteristic polynomial** of A , and $\det(A - \lambda I) = 0$ is the **characteristic equation** of A .

Clearly, if repeated roots are also counted, an $n \times n$ square matrix A has n eigenvalues including real and complex eigenvalues.

Procedure for finding eigenvalues and their corresponding eigenvectors of a square matrix A

Step 1: Compute $\det(A - \lambda I)$.

Step 2: Set $\boxed{\det(A - \lambda I) = 0}$ and solve for λ ($\leftarrow$ **eigenvalues** of A).

Step 3: For each eigenvalue λ , solve the homogeneous system $\boxed{(A - \lambda I)\vec{x} = \vec{0}}$ to find the non-trivial solutions $\vec{x}$ ($\leftarrow$ **eigenvectors** of A corresponding to eigenvalue λ).

Remark: In step 3, for each eigenvalue λ , the system $(A - \lambda I)\vec{x} = \vec{0}$ must have infinitely many solutions. The number of free variables in the solution depends on the number of zero rows in the row echelon form. This tells us the number of linearly independent eigenvectors corresponding to that particular eigenvalue λ .

Example 1

Find the eigenvalues of $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ and the corresponding eigenvectors.

Solution

$$A - \lambda I = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \lambda & 4 \\ 2 & 3 - \lambda \end{pmatrix}.$$

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 4 \\ 2 & 3 - \lambda \end{vmatrix} = (1 - \lambda)(3 - \lambda) - 8 = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1).$$

Setting $\boxed{\det(A - \lambda I) = 0}$, we get $\lambda = 5$ or $\lambda = -1$.

$\therefore$ The **eigenvalues** of A are

$$\lambda = 5 \text{ and } \lambda = -1.$$

To find the corresponding eigenvectors $\vec{x}$, we solve the system

$$\boxed{(A - \lambda I)\vec{x} = \vec{0}}$$

for each eigenvalue.

For $\lambda = 5$, we solve $(A - 5I)\vec{x} = \vec{0}$, i.e.

$$\begin{pmatrix} -4 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} -4 & 4 & 0 \\ 2 & -2 & 0 \end{array} \right) \xrightarrow{R_2 + \frac{1}{2}R_1} \left(\begin{array}{cc|c} -4 & 4 & 0 \\ 0 & 0 & 0 \end{array} \right) \xrightarrow{-\frac{1}{4}R_1} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

This is equivalent to the equation

$$x_1 - x_2 = 0 \Rightarrow x_1 = x_2.$$

Let $x_2 = k$, then $x_1 = x_2 = k$.

$\therefore$ The corresponding **eigenvectors** of A for $\lambda = 5$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ where } k \neq 0.$$

(Note: One of the eigenvectors of A for $\lambda = 5$ is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (when putting $k = 1$).)

For $\lambda = -1$, we solve $(A + I)\vec{x} = \vec{0}$, i.e.

$$\begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} 2 & 4 & 0 \\ 2 & 4 & 0 \end{array} \right) \xrightarrow{R_2 - R_1} \left(\begin{array}{cc|c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array} \right) \xrightarrow{\frac{1}{2} \times R_1} \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

This is equivalent to the equation

$$x_1 + 2x_2 = 0.$$

Let $x_2 = k$, then $x_1 = -2x_2 = -2k$.

$\therefore$ The corresponding **eigenvectors** of A for $\lambda = -1$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \text{ where } k \neq 0.$$

(Note: Putting $k = 1$, then $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ is one of the corresponding eigenvectors of A for $\lambda = -1$.) □

Example 2

Find the eigenvalues of $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 1 - \lambda & -1 & 0 \\ -1 & 2 - \lambda & -1 \\ 0 & -1 & 1 - \lambda \end{vmatrix} \\ &= (1 - \lambda) \begin{vmatrix} 2 - \lambda & -1 \\ -1 & 1 - \lambda \end{vmatrix} - (-1) \begin{vmatrix} -1 & -1 \\ 0 & 1 - \lambda \end{vmatrix} \\ &= (1 - \lambda)[(2 - \lambda)(1 - \lambda) - (-1)(-1)] + [(-1)(1 - \lambda) - (-1)(0)] \\ &= (1 - \lambda)[(2 - \lambda)(1 - \lambda) - 1 - 1] \\ &= (1 - \lambda)(\lambda^2 - 3\lambda) \\ &= \lambda(1 - \lambda)(\lambda - 3) \end{aligned}$$

Setting $|A - \lambda I| = 0$, we get the **eigenvalues**

$$\lambda = 0, 1, 3.$$

Next, we find the corresponding eigenvectors for $\lambda = 0, 1$ and 3 .

For $\lambda = 0$, we solve $(A - 0I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Using Gaussian Elimination,

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array} \right) \xrightarrow{R_2+R_1} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array} \right) \xrightarrow{R_3+R_2} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

This is equivalent to the system

$$\begin{cases} x_1 - x_2 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

Let $x_3 = k$, then $x_2 = x_3 = k$, $x_1 = x_2 = k$.

$\therefore$ The eigenvectors of A corresponding to $\lambda = 0$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ where } k \neq 0.$$

Note: One of the eigenvectors of A for $\lambda = 0$ is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. (Take $k = 1$)

For $\lambda = 1$, we solve $(A - I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Using Gaussian Elimination,

$$\left(\begin{array}{ccc|c} 0 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{array} \right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \right) \xrightarrow{R_3 - R_2} \left(\begin{array}{ccc|c} -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

This is equivalent to the system

$$\begin{cases} -x_1 + x_2 - x_3 = 0 \\ -x_2 = 0 \end{cases}$$

Let $x_3 = k$, then $x_2 = 0$, $x_1 = x_2 - x_3 = -k$.

$\therefore$ The eigenvectors of A corresponding to $\lambda = 1$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ where } k \neq 0.$$

Note: One of the eigenvectors of A for $\lambda = 1$ is $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. (Take $k = 1$)

For $\lambda = 3$, we solve $(A - 3I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Using Gaussian Elimination,

$$\begin{pmatrix} -2 & -1 & 0 & | & 0 \\ -1 & -1 & -1 & | & 0 \\ 0 & -1 & -2 & | & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} -1 & -1 & -1 & | & 0 \\ -2 & -1 & 0 & | & 0 \\ 0 & -1 & -2 & | & 0 \end{pmatrix} \xrightarrow{R_2 - 2R_1} \begin{pmatrix} -1 & -1 & -1 & | & 0 \\ 0 & 1 & 2 & | & 0 \\ 0 & -1 & -2 & | & 0 \end{pmatrix} \\ \xrightarrow{\begin{matrix} (-1)R_1 \\ R_3 + R_2 \end{matrix}} \begin{pmatrix} 1 & 1 & 1 & | & 0 \\ 0 & 1 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

This is equivalent to the system $\begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 + 2x_3 = 0 \end{cases}$.

Let $x_3 = k$, then $x_2 = -2x_3 = -2k$, $x_1 = -x_2 - x_3 = -(-2k) - k = k$.

$\therefore$ The eigenvectors of A corresponding to $\lambda = 3$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \text{where } k \neq 0.$$

Note: One of the eigenvectors of A for $\lambda = 3$ is $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. (Take $k = 1$) □

Notice that the eigenvalues of a matrix could be complex numbers.

Example 3 (Complex eigenvalues)

Find the eigenvalues of $A = \begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution

$$|A - \lambda I| = \begin{vmatrix} 1 - \lambda & -1 \\ 4 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 + 4 = (\lambda - 1)^2 + 4.$$

$$\text{Set } |A - \lambda I| = 0 \Rightarrow (\lambda - 1)^2 + 4 = 0$$

$$\Rightarrow (\lambda - 1)^2 = -4$$

$$\Rightarrow \lambda - 1 = \pm\sqrt{-4} = \pm 2i$$

$$\Rightarrow \lambda = 1 \pm 2i$$

$\therefore$ The eigenvalues of A are $\lambda = 1 \pm 2i$.

For $\lambda = 1 + 2i$, we solve $[A - (1 + 2i)I]\vec{x} = \vec{0}$, i.e.

$$\begin{pmatrix} -2i & -1 \\ 4 & -2i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} -2i & -1 & 0 \\ 4 & -2i & 0 \end{array} \right) \xrightarrow{\frac{1}{-2i}R_1} \left(\begin{array}{cc|c} 1 & \frac{1}{2i} & 0 \\ 4 & -2i & 0 \end{array} \right) \xrightarrow[\frac{1}{2i} = \frac{i}{2i^2} = -\frac{i}{2}]{=} \left(\begin{array}{cc|c} 1 & -\frac{i}{2} & 0 \\ 4 & -2i & 0 \end{array} \right) \xrightarrow{R_2 - 4R_1} \left(\begin{array}{cc|c} 1 & -\frac{i}{2} & 0 \\ 0 & 0 & 0 \end{array} \right)$$

This corresponds to the equation

$$x_1 - \frac{i}{2}x_2 = 0.$$

Let $x_2 = k$, then $x_1 = \frac{i}{2}x_2 = \frac{i}{2}k$.

$\therefore$ The eigenvectors of A corresponding to the eigenvalue $\lambda = 1 + 2i$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}, \quad k \neq 0.$$

Note: One of the eigenvectors corresponding to $\lambda = 1 + 2i$ is $\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}$.

Discussion:

Let's take a look at the results of Example 3:

- $\lambda = 1 + 2i$ is a complex eigenvalue of A with corresponding eigenvector $\begin{pmatrix} \frac{i}{2} \\ 1 \end{pmatrix}$.
- $\lambda = 1 - 2i$ is a complex eigenvalue of A with corresponding eigenvector $\begin{pmatrix} -\frac{i}{2} \\ 1 \end{pmatrix}$.

It seems that if λ is a complex eigenvalue of A with corresponding eigenvector $\vec{x}$, then $\bar{\lambda}$ (the complex conjugate of λ) is also a complex eigenvalue of A with a corresponding eigenvector which is the complex conjugate of $\vec{x}$.

Question: Is this always true for any square matrix A ?

Example 4 (Repeated eigenvalue)

Find the eigenvalues of $A = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}$ and the corresponding eigenvectors.

Solution

$$|A - \lambda I| = \begin{vmatrix} 5 - \lambda & 3 \\ -3 & -1 - \lambda \end{vmatrix} = (5 - \lambda)(-1 - \lambda) - (-9) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$$

Setting $|A - \lambda I| = 0 \Rightarrow (\lambda - 2)^2 = 0 \Rightarrow \lambda = 2, 2$ (i.e. $\lambda = 2$ is the only eigenvalue.)

(We say that A has eigenvalue 2 of **multiplicity 2**.)

For $\lambda = 2$, solve $(A - 2I)\vec{x} = \vec{0}$, i.e.

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

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} 3 & 3 & 0 \\ -3 & -3 & 0 \end{array} \right) \xrightarrow{R_2 + R_1} \left(\begin{array}{cc|c} 3 & 3 & 0 \\ 0 & 0 & 0 \end{array} \right) \xrightarrow{\frac{1}{3} \times R_1} \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

This corresponds to the equation $x_1 + x_2 = 0$.

Let $x_2 = k$, then $x_1 = -x_2 = -k$.

∴ The **eigenvectors** of A corresponding to the eigenvalue $\lambda = 2$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad k \neq 0.$$

Note: One of the corresponding eigenvectors for $\lambda = 2$ is $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Any other eigenvector corresponding to $\lambda = 2$ are not linearly independent with $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. □

It is also possible to have a repeated eigenvalue which does lead to more than one **linearly independent** eigenvector. (See the following example.)

Example 5 (Repeated eigenvalue)

Find the eigenvalues of $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and the corresponding eigenvectors.

Solution

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

Setting $|A - \lambda I| = 0 \Rightarrow (2 - \lambda)^2 = 0 \Rightarrow \lambda = 2, 2$.

For $\lambda = 2$, solve $(A - 2I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, then x_1 and x_2 can be any number except both $x_1 = x_2 = 0$.

Let $x_1 = s$ and $x_2 = t$.

$\therefore$ The corresponding eigenvectors of A for $\lambda = 2$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{where } s^2 + t^2 \neq 0 \text{ (i.e. } s, t \neq 0 \text{ at the same time).}$$

(Note: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are two **linearly independent** eigenvectors corresponding to $\lambda = 2$.)



3 Eigenvalue Results

Definition

Two vectors $\vec{x}, \vec{y}$ are called **orthogonal** if $\vec{x}^T \vec{y} = 0$ or $\vec{x} \cdot \vec{y} = 0$.

For example, $\vec{x} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ and $\vec{y} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ are orthogonal since

$$\vec{x} \cdot \vec{y} = (2)(-1) + (-1)(-1) + (1)(1) = 0.$$

A set of vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is **orthogonal** if and only if each pair of vectors in this set is orthogonal.

For example, the set of vectors $\left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ is orthogonal because $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = 0$,

$$\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0 \text{ and } \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0.$$

Definition

A set of vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is **linearly independent** if the only linear combination of them that equals the zero vector is the trivial one, i.e. all the coefficients are 0.

That is, $c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n = \vec{0}$ if and only if $c_1 = c_2 = \dots = c_n = 0$ is the **only (trivial) solution**.

Vectors that are not linearly independent are **linearly dependent**.

Note that **linear independence** and **orthogonality** are two different concepts.

➤ If $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are orthogonal, then they must be linearly independent.

However, the converse may not be true.

➤ If $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly independent, then they are not necessarily to be orthogonal.

Orthogonal $\Rightarrow$ Linearly independent $\nLeftarrow$
--

Theorem 1

Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof of Theorem 1 (for your reference):

See course notes (p.6 of Chapter 1).



Illustration of Theorem 1

Take the matrix $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ from Example 2. Recall from Example 2 that the eigenvalues of A are $\lambda = 0, 1, 3$, and one of their corresponding eigenvectors are as follows:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ for } \lambda = 0, \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \text{ for } \lambda = 1, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \text{ for } \lambda = 3.$$

Note that the eigenvalues $\lambda = 0, 1, 3$ of A are all **distinct**.

To check if $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly independent or not, we solve the system of equations

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + c_3 \vec{v}_3 = \vec{0},$$

$$\text{i.e. } \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & 0 & -2 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right) \xrightarrow[R_3 - R_1]{R_2 - R_1} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 2 & 0 & 0 \end{array} \right) \xrightarrow{R_3 - 2R_2} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 6 & 0 \end{array} \right).$$

Solving backward gives $c_1 = c_2 = c_3 = 0$, which is the only (trivial) solution.

$\therefore$ The eigenvectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ corresponding to **distinct** eigenvalues $\lambda = 0, 1, 3$ are **linearly independent**. □

As seen in Example 5 before, it is also possible to have a repeated eigenvalue which does lead to more than one **linearly independent** eigenvector. Here is another example:

Example 6 (for your reference)

Let J_n , $n = 2, 3, \dots$ be a square matrix with all its entries being 1, that is,

$$J_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

- (a) Find all eigenvalues of J_n .
- (b) Find n linearly independent eigenvectors.

Solution

First observe that $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ \vdots \\ n \end{pmatrix} = n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$

so $\lambda = n$ is an eigenvalue of J_n and one of its eigenvectors is $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$

In addition, we observe that
$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

so $\lambda = 0$ is an eigenvalue of J_n and one of its eigenvectors is $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}.$

For $\lambda = 0$, we solve $(J_n - 0I)\vec{x} = \vec{0}$, i.e.

$$\left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \end{array} \right) \xrightarrow[\begin{smallmatrix} R_2-R_1 \\ R_3-R_1 \\ \vdots \\ R_n-R_1 \end{smallmatrix}]{\quad} \left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{array} \right).$$

We have $x_1 + x_2 + x_3 + \cdots + x_n = 0$.

Let $x_1 = a_1$, $x_2 = a_2$, ..., $x_{n-1} = a_{n-1}$, where $a_1, a_2, \dots, a_{n-1}$ are arbitrary real numbers.

Then $x_n = -a_1 - a_2 - \cdots - a_{n-1}$.

$\therefore$ The eigenvectors of J_n corresponding to the eigenvalue $\lambda = 0$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + a_{n-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}.$$

The eigenvectors $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}$ are linearly independent, because if we solve the

$$\text{system } c_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \cdots + c_{n-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ -c_1 - c_2 - \cdots - c_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \text{we}$$

find that $c_1 = c_2 = \cdots = c_{n-1} = 0$ is the only solution.

We have found an eigenvalue $\lambda = n$ with corresponding eigenvector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$; and an eigenvalue

$\lambda = 0$ with $n - 1$ linearly independent eigenvectors $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}$.

Our remaining work is to show that these n vectors are linearly independent.

For this, we solve the system $c_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix} + \dots + c_{n-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix} + c_n \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$:

$$\left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 1 & \cdots & 0 & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 & 0 \end{array} \right) \xrightarrow{R_n + (R_1 + R_2 + \cdots + R_{n-1})} \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 1 & \cdots & 0 & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & n & 0 \end{array} \right)$$

Solving backward, we find that $c_1 = c_2 = \dots = c_{n-1} = c_n = 0$ is the only solution.

Hence, $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}$ are n linearly independent eigenvectors of J_n .



Definition

The square matrix A is **symmetric** if and only if $A^T = A$.

E.g. $\begin{pmatrix} 3 & 4 \\ 4 & -2 \end{pmatrix}$ and $\begin{pmatrix} 8 & -2 & 3 \\ -2 & 1 & 0 \\ 3 & 0 & 4 \end{pmatrix}$ are symmetric.

Theorem 2

If A is a **real symmetric** matrix, then

- (a) the eigenvalues are all **real** (but not necessarily distinct),
- (b) eigenvectors corresponding to **distinct** eigenvalues are **orthogonal**.

Proof of Theorem 2 (for your reference):

See course notes (p.10 of Chapter 1).



Illustration of Theorem 2

Again, take $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ from Example 2.

Since $A^T = A$, the matrix A is **symmetric**. Furthermore, all entries in A are **real** numbers.

Thus, A is a **real symmetric** matrix.

Recall from Example 2 that the eigenvalues of A are $\lambda = 0, 1, 3$, and the corresponding eigenvectors are as follows:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ for } \lambda = 0, \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \text{ for } \lambda = 1, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \text{ for } \lambda = 3.$$

Note that the eigenvalues $\lambda = 0, 1, 3$ of A are all **real**. $\leftarrow$ (a) ✓

Then for the **distinct** eigenvalues $\lambda = 0, 1, 3$, their corresponding eigenvectors

$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ and $\vec{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ are **orthogonal**, as we can easily check that

$\vec{v}_1 \cdot \vec{v}_2 = 0$, $\vec{v}_1 \cdot \vec{v}_3 = 0$ and $\vec{v}_2 \cdot \vec{v}_3 = 0$. $\leftarrow$ (b) ✓

□

Theorem 3

If A is an $n \times n$ **real symmetric** matrix, then it is possible to find **n linearly independent eigenvectors** (even though the eigenvalues may not be distinct).

Proof of Theorem 3: Omitted

Illustration of Theorem 3

Consider the matrix $J_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Note that J_3 is **real symmetric**.

$$\begin{aligned} |J_3 - \lambda I| &= \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} - 1 \begin{vmatrix} 1 & 1 \\ 1 & 1-\lambda \end{vmatrix} + 1 \begin{vmatrix} 1 & 1-\lambda \\ 1 & 1 \end{vmatrix} \\ &= (1-\lambda)[(1-\lambda)^2 - 1] - [(1-\lambda) - 1] + 1[1 - (1-\lambda)] \\ &= (1-\lambda)^3 - 3(1-\lambda) + 2 \\ &= 1 - 3\lambda + 3\lambda^2 - \lambda^3 - 3(1-\lambda) + 2 \\ &= -\lambda^3 + 3\lambda^2 \\ &= \lambda^2(3 - \lambda) \end{aligned}$$

Setting $|J_3 - \lambda I| = 0 \Rightarrow \lambda = 3, 0, 0$.

Thus, J_3 has eigenvalue 3 of multiplicity 1 and eigenvalue 0 of multiplicity 2. All these **eigenvalues** are **real**.

For $\lambda = 3$, we solve $(J_3 - 3I)\vec{x} = \vec{0}$:

$$\begin{aligned} \left(\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array} \right) &\xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 1 & -2 & 1 & 0 \\ -2 & 1 & 1 & 0 \end{array} \right) &\xrightarrow{\substack{R_2 - R_1 \\ R_3 + 2R_1}} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{array} \right) \\ &\xrightarrow{R_3 + R_2} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) &\xrightarrow{R_2 / (-3)} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

$$\therefore \begin{cases} x_1 + x_2 - 2x_3 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

Let $x_3 = k$, where $k \neq 0$.

Then $x_2 = x_3 = k$ and $x_1 = -x_2 + 2x_3 = -k + 2k = k$.

$\therefore$ The eigenvectors of J_3 corresponding to $\lambda = 3$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ where } k \neq 0.$$

$\therefore$ For $\lambda = 3$, one of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

For $\lambda = 0$, we solve $(J_3 - 0I)\vec{x} = \vec{0}$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right) \xrightarrow[R_3 - R_1]{R_2 - R_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\therefore x_1 + x_2 + x_3 = 0$$

Let $x_1 = s$ and $x_2 = t$, where $s^2 + t^2 \neq 0$.

Then $x_3 = -s - t$.

$\therefore$ The eigenvectors of J_3 corresponding to the eigenvalue $\lambda = 0$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \text{where } s^2 + t^2 \neq 0.$$

$\therefore$ For $\lambda = 0$, its linearly independent eigenvectors are $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

These results can also be obtained from Example 6 (with $n = 3$).

Next, we check if $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ are linearly independent or not.

We solve the system $c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{array} \right) \xrightarrow[R_3 - R_1]{R_2 - R_1} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & -1 & 0 \end{array} \right) \xrightarrow{R_3 - 2R_2} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -3 & 0 \end{array} \right)$$

Solving backward, we get $c_1 = c_2 = c_3 = 0$ (the **trivial solution**) as its only solution, therefore the **three eigenvectors** are **linearly independent**. □

3.1 Diagonalization

Definition

An $n \times n$ matrix A is said to be **diagonalizable** if there is a non-singular (or invertible) $n \times n$ matrix P such that

$$\boxed{D = P^{-1}AP},$$

where D is a diagonal matrix.

Observe that

$$\begin{aligned}\det(D - \lambda I) &= \det(P^{-1}AP - \lambda P^{-1}IP) \\ &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \det(A - \lambda I) \det(P) \\ &= \frac{1}{\det(P)} \det(A - \lambda I) \det(P) = \det(A - \lambda I)\end{aligned}$$

Therefore, A and $D = P^{-1}AP$ have the same characteristic polynomial and thus they have the same eigenvalues. We also notice that the **eigenvalues of D** are just the **main diagonal entries of D** .

Theorem 4

If an $n \times n$ matrix A has **n linearly independent eigenvectors**, then it is **diagonalizable**.

Before we prove Theorem 4, let's recall the following results on matrix multiplication.

Remark on matrix multiplication:

Let A be a $p \times m$ matrix and B be an $m \times n$ matrix. Recall that the product AB can be evaluated in two different ways:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} = (\overrightarrow{a_1} | \overrightarrow{a_2} | \cdots | \overrightarrow{a_m}), \text{ where } \overrightarrow{a_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{pmatrix}, \dots, \overrightarrow{a_m} = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{pm} \end{pmatrix}.$$

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = (\overrightarrow{b_1} | \overrightarrow{b_2} | \cdots | \overrightarrow{b_n}), \text{ where } \overrightarrow{b_1} = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}, \dots, \overrightarrow{b_n} = \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{mn} \end{pmatrix}.$$

We observe that

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = (A\overrightarrow{b_1} \mid A\overrightarrow{b_2} \mid \cdots \mid A\overrightarrow{b_n})$$

and also

$$AB = (b_{11}\overrightarrow{a_1} + \cdots + b_{m1}\overrightarrow{a_m} \mid b_{12}\overrightarrow{a_1} + \cdots + b_{m2}\overrightarrow{a_m} \mid \cdots \mid b_{1n}\overrightarrow{a_1} + \cdots + b_{mn}\overrightarrow{a_m}).$$

Proof of Theorem 4: (for your reference)

Suppose A is an $n \times n$ matrix, and $\overrightarrow{x_1}, \dots, \overrightarrow{x_n}$ are n **linearly independent** eigenvectors of A with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ such that

$$A\overrightarrow{x_1} = \lambda_1\overrightarrow{x_1}, \dots, A\overrightarrow{x_n} = \lambda_n\overrightarrow{x_n}.$$

Construct an $n \times n$ matrix $P = (\overrightarrow{x_1} \mid \overrightarrow{x_2} \mid \cdots \mid \overrightarrow{x_n})$ with $\overrightarrow{x_1}, \dots, \overrightarrow{x_n}$ as its columns, where $\overrightarrow{x_1}, \dots, \overrightarrow{x_n}$ are the **eigenvectors** of A .

Then

$$\begin{aligned}
AP &= (A\vec{x}_1 \mid A\vec{x}_2 \mid \cdots \mid A\vec{x}_n) = (\lambda_1\vec{x}_1 \mid \lambda_2\vec{x}_2 \mid \cdots \mid \lambda_n\vec{x}_n) \\
&= (\lambda_1\vec{x}_1 + 0\vec{x}_2 + \cdots + 0\vec{x}_n \mid 0\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + 0\vec{x}_n \mid \cdots \mid 0\vec{x}_1 + \cdots + 0\vec{x}_{n-1} + \lambda_n\vec{x}_n) \\
&= (\vec{x}_1 \mid \vec{x}_2 \mid \cdots \mid \vec{x}_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = PD
\end{aligned}$$

where $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n).$

Since $\vec{x}_1, \dots, \vec{x}_n$ are linearly independent, then $\det P \neq 0$ and P^{-1} exist.

Thus, we have $P^{-1}AP = P^{-1}PD = D$, i.e. A is diagonalizable. □

Summary of Theorems 1 to 4

Theorem 1: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 2: If A is a real symmetric matrix, then

- (a) the eigenvalues are all real (but not necessarily distinct),
- (b) eigenvectors corresponding to distinct eigenvalues are orthogonal.

Theorem 3: If A is an $n \times n$ real symmetric matrix, then it is possible to find n linearly independent eigenvectors (even though the eigenvalues may not be distinct).

Theorem 4: If an $n \times n$ matrix A has n linearly independent eigenvectors, then it is diagonalizable.

Useful results: From the above theorems, we notice that

- (i) All $n \times n$ **real symmetric matrices** are **diagonalizable**;
- (ii) If an $n \times n$ matrix has **n distinct eigenvalues**, then it is **diagonalizable**.

Reason for result (i):

If A is an $n \times n$ **real symmetric** matrix, by Theorem 3, it is possible to find **n linearly independent eigenvectors** (even though the eigenvalues may not be distinct). Then by Theorem 4, A is **diagonalizable**.

Reason for result (ii):

If an $n \times n$ matrix A has **n distinct eigenvalues**, by Theorem 1, there must be **n corresponding eigenvectors** which are **linearly independent**. Then by Theorem 4, A is **diagonalizable**.

Example 7

Consider the matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$.

- (a) Is A diagonalizable?
- (b) Express A into the form PDP^{-1} , where D is a diagonal matrix.

Solution

$$\begin{aligned} \text{(a)} \quad |A - \lambda I| &= \begin{vmatrix} 1 - \lambda & 4 \\ 2 & 3 - \lambda \end{vmatrix} = (1 - \lambda)(3 - \lambda) - 8 \\ &= \lambda^2 - 4\lambda - 5 \\ &= (\lambda - 5)(\lambda + 1). \end{aligned}$$

Set $|A - \lambda I| = 0 \Rightarrow \lambda = 5$ or -1 .

$\therefore$ The eigenvalues of A are

$$\lambda_1 = 5 \text{ and } \lambda_2 = -1.$$

Since the **eigenvalues** are all **distinct**, then A is **diagonalizable**.

(b) For $\lambda_1 = 5$, solve $(A - 5I)\vec{x} = \vec{0}$.

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} -4 & 4 & 0 \\ 2 & -2 & 0 \end{array}\right) \xrightarrow{R_2 + \frac{1}{2}R_1} \left(\begin{array}{cc|c} -4 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right) \xrightarrow{\left(-\frac{1}{4}\right) \times R_1} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

This corresponds to the equation $x_1 - x_2 = 0$.

Let $x_2 = s$, then $x_1 = x_2 = s$.

$\therefore$ The eigenvectors of A corresponding to $\lambda_1 = 5$ are

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (s \neq 0).$$

$\therefore$ One of the corresponding eigenvectors of A for $\lambda_1 = 5$ is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

For $\lambda_2 = -1$, solve $(A + I)\vec{x} = \vec{0}$.

Using Gaussian Elimination,

$$\left(\begin{array}{cc|c} 2 & 4 & 0 \\ 2 & 4 & 0 \end{array}\right) \xrightarrow{R_2 - R_1} \left(\begin{array}{cc|c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right) \xrightarrow{\frac{1}{2} \times R_1} \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

This corresponds to the equation $x_1 + 2x_2 = 0$.

Let $x_2 = k$, then $x_1 = -2x_2 = -2k$.

$\therefore$ The eigenvectors of A corresponding to $\lambda_2 = -1$ are

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad (k \neq 0).$$

$\therefore$ One of the corresponding eigenvectors of A for $\lambda_2 = -1$ is $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

Let $P = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$.

As the eigenvectors of A are linearly independent, P^{-1} exists.

To find P^{-1} , we may use the formula or the Gauss-Jordan elimination method.

By the formula,

$$P^{-1} = \frac{1}{\det(P)} \text{adj}(P) = \frac{1}{1 \times 1 - (-2) \times 1} \begin{pmatrix} 1 & -1 \\ -(-2) & 1 \end{pmatrix}^T = \frac{1}{3} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

OR By **Gauss-Jordan elimination method**,

$$(P|I_2) = \left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \xrightarrow{R_2 - R_1} \left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array} \right) \xrightarrow{\frac{1}{3} \times R_2} \left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} \end{array} \right)$$

$$\xrightarrow{R_1 + 2R_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} \end{array} \right)$$

$$\therefore P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Note that $P^{-1}AP = D$, as A is diagonalizable.

$$\text{Then } A = PDP^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Check:

$$PDP^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = A. \quad \square$$

Remark:

If $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 invertible matrix, then

$$P^{-1} = \frac{1}{\det(P)} \operatorname{adj}(P) = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Example 8

Is $A = \begin{pmatrix} 5 & 3 \\ -3 & -1 \end{pmatrix}$ diagonalizable?

Solution

$$|A - \lambda I| = \begin{vmatrix} 5 - \lambda & 3 \\ -3 & -1 - \lambda \end{vmatrix} = (5 - \lambda)(-1 - \lambda) - (-9) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$$

Set $|A - \lambda I| = 0 \Rightarrow \lambda = 2, 2$ (i.e. $\lambda = 2$ is the only eigenvalue.)

For $\lambda = 2$, solve $(A - 2I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 3 & 3 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Using Gaussian Elimination, $\left(\begin{array}{cc|c} 3 & 3 & 0 \\ -3 & -3 & 0 \end{array} \right) \xrightarrow{R_2+R_1} \left(\begin{array}{cc|c} 3 & 3 & 0 \\ 0 & 0 & 0 \end{array} \right) \xrightarrow{\frac{1}{3} \times R_1} \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right)$.

This corresponds to the equation $x_1 + x_2 = 0$.

Let $x_2 = k$, then $x_1 = -x_2 = -k$.

$\therefore$ The eigenvectors of A corresponding to the eigenvalue $\lambda = 2$ are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad k \neq 0.$$

As there is **only one** linearly independent eigenvector, matrix A is **not diagonalizable**. □

Example 9

Diagonalize $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$, if possible, by finding the matrices P , P^{-1} and D such that $P^{-1}AP = D$.

Solution

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 1 - \lambda & 3 & 3 \\ -3 & -5 - \lambda & -3 \\ 3 & 3 & 1 - \lambda \end{vmatrix} \\ &= (1 - \lambda) \begin{vmatrix} -5 - \lambda & -3 \\ 3 & 1 - \lambda \end{vmatrix} - 3 \begin{vmatrix} -3 & -3 \\ 3 & 1 - \lambda \end{vmatrix} + 3 \begin{vmatrix} -3 & -5 - \lambda \\ 3 & 3 \end{vmatrix} \\ &= (1 - \lambda)[(-5 - \lambda)(1 - \lambda) + 9] - 3[-3(1 - \lambda) + 9] + 3[-9 - 3(-5 - \lambda)] \\ &= (1 - \lambda)(\lambda^2 + 4\lambda + 4) - 3(3\lambda + 6) + 3(3\lambda + 6) \\ &= (1 - \lambda)(\lambda + 2)^2 \end{aligned}$$

Setting $|A - \lambda I| = 0$ gives $\lambda = 1$ (of **multiplicity 1**) and $\lambda = -2$ (of **multiplicity 2**) as the eigenvalues.

For $\lambda = 1$, solve $(A - I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

By Gaussian Elimination,

$$\begin{pmatrix} 0 & 3 & 3 & | & 0 \\ -3 & -6 & -3 & | & 0 \\ 3 & 3 & 0 & | & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 3 & 3 & 0 & | & 0 \\ -3 & -6 & -3 & | & 0 \\ 0 & 3 & 3 & | & 0 \end{pmatrix} \xrightarrow{R_2 + R_1} \begin{pmatrix} 3 & 3 & 0 & | & 0 \\ 0 & -3 & -3 & | & 0 \\ 0 & 3 & 3 & | & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 + R_2} \begin{pmatrix} 3 & 3 & 0 & | & 0 \\ 0 & -3 & -3 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \xrightarrow{\substack{R_1/3 \\ R_2/(-3)}} \begin{pmatrix} 1 & 1 & 0 & | & 0 \\ 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

This corresponds to the system

$$\begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{cases}$$

Let $x_3 = s$, then $x_2 = -x_3 = -s$, $x_1 = -x_2 = s$.

$\therefore \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = s \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $s \neq 0$ are the eigenvectors of A corresponding to $\lambda = 1$.

$\therefore$ One of the corresponding eigenvectors for $\lambda = 1$ is $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

For $\lambda = -2$, solve $(A + 2I)\vec{x} = \vec{0}$, i.e. $\begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

By Gaussian Elimination,

$$\left(\begin{array}{ccc|c} 3 & 3 & 3 & 0 \\ -3 & -3 & -3 & 0 \\ 3 & 3 & 3 & 0 \end{array} \right) \xrightarrow[R_3 - R_1]{R_2 + R_1} \left(\begin{array}{ccc|c} 3 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \xrightarrow{R_1/3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

which corresponds to the equation $x_1 + x_2 + x_3 = 0$.

Let $x_2 = s$, $x_3 = t$, then $x_1 = -x_2 - x_3 = -s - t$.

$\therefore$ The eigenvectors corresponding to $\lambda = -2$ are

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -s - t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ where } s^2 + t^2 \neq 0.$$

$\therefore \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ are two **linearly independent** eigenvectors corresponding to $\lambda = -2$.

Since A is a 3×3 matrix and it has **three linearly independent eigenvectors**, then A is **diagonalizable**.

Take $P = \begin{pmatrix} \mathbf{1} & \mathbf{-1} & \mathbf{-1} \\ \mathbf{-1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \end{pmatrix}$, $D = \begin{pmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{-2} & 0 \\ 0 & 0 & \mathbf{-2} \end{pmatrix}$.

By Gauss-Jordan elimination method,

$$\begin{aligned}
 (P|I_3) &= \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow[R_3 - R_1]{R_2 + R_1} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array} \right) \\
 &\xrightarrow{R_2 \leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 \end{array} \right) \xrightarrow[R_2 + 2R_3]{R_1 - R_3} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 \end{array} \right) \\
 &\xrightarrow[(-1) \times R_3]{R_1 + R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{array} \right) = (I_3|P^{-1})
 \end{aligned}$$

$$\therefore P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Then we have $P^{-1}AP = D$, that is

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

□

Diagonalization is useful when it is required to raise a matrix A to a high power, that is, A^m .

If $P^{-1}AP = D$, then $A = PDP^{-1}$.

Then

$$\begin{aligned} A^m &= (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) \\ &= PD \underbrace{P^{-1}P}_{=I_n} D \underbrace{P^{-1}P}_{=I_n} DP^{-1} \cdots PD \underbrace{P^{-1}P}_{=I_n} D \underbrace{P^{-1}P}_{=I_n} DP^{-1} \\ &= PD^m P^{-1} \end{aligned}$$

where

$$\begin{aligned} D &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \\ \Rightarrow D^m &= \begin{pmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^m \end{pmatrix} = \text{diag}(\lambda_1^m, \lambda_2^m, \dots, \lambda_n^m). \end{aligned}$$



An important application of diagonalization occurs in the solution of systems of differential equations.

Example 10 (for your reference)

Consider a system of n first order linear differential equations for the dependent variables $x_1, x_2, \dots, x_n$ as follows.

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases} \dots \dots (E),$$

where the a_{ij} 's are given constants. The system (E) can be written as the following matrix

form $\frac{d\vec{u}}{dt} = A\vec{u} \dots (F)$, where $\vec{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.

Suppose A is diagonalizable, then there exists non-singular matrix P such that $D = P^{-1}AP$, where D is a diagonal matrix.

Let $\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = P^{-1} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, that is, $\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = P^{-1} \vec{u}$. Then

$$\frac{d}{dt} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \frac{d}{dt} P^{-1} \vec{u} = P^{-1} \frac{d\vec{u}}{dt} = P^{-1} A \vec{u} = D P^{-1} \vec{u} = D \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

Observe that D is diagonal so that each row of the system is of the form $\frac{d}{dt} y_i(t) = d_i y_i(t)$, which is easy to solve.

The solution of $\frac{d}{dt} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = D \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$ is easy to obtain. Once $\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$ is found, then

$\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ can be obtained by using $\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = P \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$. □

(Note: Topics on differential equations are not required in this course.)

Example 11

- (a) Use Gauss-Jordan elimination method to compute the inverse of $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.
- (b) Let A be a 3×3 matrix which has eigenvalues $-1, 1, 0$ with corresponding eigenvectors $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, respectively. Using part (a), find A^{10} .

Solution

$$\begin{aligned}
 \text{(a)} \quad & \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{R_2 - R_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\
 & \xrightarrow{R_1 - R_3} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{R_1 - R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\
 \therefore & \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
 \end{aligned}$$

(b) We have $A \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$

Then

$$\begin{aligned}
 A &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \\
 \Rightarrow A^{10} &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}^{10} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} (-1)^{10} & 0 & 0 \\ 0 & 1^{10} & 0 \\ 0 & 0 & 0^{10} \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

□

Example 12

(a) Let B be a 3×3 matrix and $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of B . Write down the characteristic equation of B and show that $\lambda_1 \lambda_2 \lambda_3 = \det B$.

(b) Let $A_{3 \times 3}$ be a real symmetric matrix. Suppose

$$\det A = 1, \quad A \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}, \quad A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}.$$

(i) Use part (a), or otherwise, find the eigenvalues of A .

(ii) Find a matrix P having its columns to be the unit eigenvectors of A and show that $P^T P = I$.

(iii) Show that A is diagonalizable and find the matrix A .

Solution

(a) Let $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$

The characteristic equation is given by $\det(B - \lambda I) = 0$.

$$\begin{aligned}\det(B - \lambda I) &= \begin{vmatrix} b_{11} - \lambda & b_{12} & b_{13} \\ b_{21} & b_{22} - \lambda & b_{23} \\ b_{31} & b_{32} & b_{33} - \lambda \end{vmatrix} \quad (\text{Coefficient of } \lambda^3 \text{ is } -1) \\ &= -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \quad (\because \lambda_1, \lambda_2, \lambda_3 \text{ are the eigenvalues of } B) \\ &= -[\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3) \lambda^2 + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3) \lambda - \lambda_1 \lambda_2 \lambda_3]\end{aligned}$$

$\therefore$ The characteristic equation is

$$-[\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3) \lambda^2 + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3) \lambda - \lambda_1 \lambda_2 \lambda_3] = 0$$

or written as

$$\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3) \lambda^2 + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3) \lambda - \lambda_1 \lambda_2 \lambda_3 = 0.$$

Putting $\lambda = 0$ into

$$\det(B - \lambda I) = -[\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3) \lambda^2 + (\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3) \lambda - \lambda_1 \lambda_2 \lambda_3],$$

we get

$$\det(B) = \lambda_1 \lambda_2 \lambda_3.$$

(b)(i) Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of A .

$$\text{Given that } A \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} \text{ and } A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}.$$

Since $A \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, then $\lambda_1 = -1$ is an eigenvalue of A and one of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

Also, $A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} = (-2) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, then $\lambda_2 = -2$ is an eigenvalue of A and one of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

It is also given that $\det A = 1$. Then using by part (a),

$$\det A = \lambda_1 \lambda_2 \lambda_3 \Rightarrow 1 = (-1)(-2)\lambda_3 \Rightarrow \lambda_3 = \frac{1}{2}.$$

Hence, the eigenvalues of A are $\lambda_1 = -1$, $\lambda_2 = -2$, $\lambda_3 = \frac{1}{2}$.

(**Remark**: Observe that the real symmetric matrix A has distinct simple (of multiplicity 1) eigenvalues, therefore, A has the **orthonormal (orthogonal unit) eigenvectors** which are linearly independent.)

(b)(ii) For $\lambda_1 = -1$, we have $A \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

One of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

Then the corresponding **unit eigenvector** (with magnitude 1) is

$$\frac{1}{\sqrt{1^2 + 2^2 + 1^2}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}.$$

For $\lambda_2 = -2$, we have $A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (-2) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

One of its corresponding eigenvectors is $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

Then the corresponding **unit eigenvector** is

$$\frac{1}{\sqrt{1^2 + 0^2 + (-1)^2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

For $\lambda_3 = \frac{1}{2}$, we have $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, where $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is a corresponding eigenvector.

Since A is real symmetric, then by Theorem 2(b), eigenvectors corresponding to **distinct** eigenvalues are **orthogonal**. Thus, we have

$$\begin{cases} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 0 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0 \end{cases} \Leftrightarrow \begin{cases} x + 2y + z = 0 \\ x - z = 0 \end{cases} \Leftrightarrow \begin{cases} x = z \\ y = \frac{1}{2}(-x - z) = -z \end{cases}$$

Let $z = 1$, then $x = 1$ and $y = -1$.

Then $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ is an eigenvector of A for $\lambda_3 = \frac{1}{2}$.

The corresponding **unit eigenvector** is

$$\frac{1}{\sqrt{1^2 + (-1)^2 + 1^2}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}.$$

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

$$\text{Then } P^T P = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, $P^T P = I$.

(**Remark:** Observe that the matrix P , which is formed by the orthonormal eigenvectors of A , has its inverse that is just its transpose. That is, $P^{-1} = P^T$.)

(b)(iii) Since A is a real symmetric matrix, it is diagonalizable.

That is, there exists an invertible matrix P and diagonal matrix D such that

$$P^{-1}AP = D$$

or written as $AP = PD$,

$$\text{i.e. } A \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

$$P^{-1}AP = D$$

$$\Rightarrow A = PDP^{-1} \underset{\substack{= \\ \text{from (ii)}}}{=} PDP^T$$

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

$$\begin{aligned}
&= \begin{pmatrix} -\frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{2}} & \frac{1}{2\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & -\frac{1}{2\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{2}} & \frac{1}{2\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{6} - 1 + \frac{1}{6} & -\frac{2}{6} + 0 - \frac{1}{6} & -\frac{1}{6} + 1 + \frac{1}{6} \\ -\frac{2}{6} + 0 - \frac{1}{6} & -\frac{4}{6} + 0 + \frac{1}{6} & -\frac{2}{6} + 0 - \frac{1}{6} \\ -\frac{1}{6} + 1 + \frac{1}{6} & -\frac{2}{6} + 0 - \frac{1}{6} & -\frac{1}{6} - 1 + \frac{1}{6} \end{pmatrix} \\
&= \begin{pmatrix} -1 & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -1 \end{pmatrix}
\end{aligned}$$

$$\therefore A = \begin{pmatrix} -1 & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -1 \end{pmatrix}.$$

4 Quadratic Forms

4.1 Introduction

A **quadratic form** is an expression of the form

$$q = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

where a_{ij} 's are constants, $x_1, \dots, x_n$ are the variables.

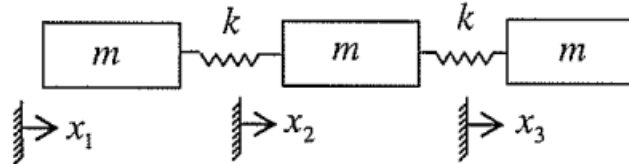
Note that each term in a quadratic form is of **degree two**.

Examples:

- $q = 3x_1^2 + 4x_1x_2 - x_2^2$ is a quadratic form in two variables x_1 and x_2 .
- $q = x_1^2 - x_1x_2 + 3x_1x_3 + 4x_3^2$ is a quadratic form in three variables x_1, x_2 and x_3 .
- $q = 2x^2 + 3y^2$ is a quadratic form in the variables x and y .
- $q = 2x^2 - x + 3y^2$ is not a quadratic form (because the term “ $-x$ ” has degree 1).
- $q = x^2 + 3xy + 1$ is not a quadratic form (because the term “ $+1$ ” has degree 0).

Example 13 (for your reference)

The 3-mass, 2-spring system



has kinetic energy

$$T = m \left(\frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} \right)$$

and potential energy

$$V = m \left[\frac{k(x_2 - x_1)^2}{2} + \frac{k(x_3 - x_2)^2}{2} \right] = mk \left(\frac{x_1^2}{2} + x_2^2 + \frac{x_3^2}{2} - x_1x_2 - x_2x_3 \right).$$

Observe that T is a quadratic form in x_1 , x_2 and x_3 .

V is a quadratic form in x_1 , x_2 and x_3 .

□

Every quadratic form is associated with a symmetric matrix, i.e.

$$q = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad \cdots (*)$$

can be written as

$$q = \underline{x}^T A \underline{x}$$

where $\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $A = [a_{ij}]$ and $A = A^T$ (i.e. symmetric).

The diagonal entries of A are the coefficients a_{ii} appearing in $(*)$, while the off-diagonal entries are half the corresponding coefficients.

Example 14

Express

$$q = 3x_1^2 + 4x_1x_2 - x_2^2$$

in the form $q = \underline{x}^T A \underline{x}$, where A is a symmetric matrix.

Solution

$$\begin{aligned} q &= 3x_1^2 + 4x_1x_2 - x_2^2 \\ &= [x_1 \quad x_2] \underbrace{\begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix}}_{\text{symmetric}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

Remark:

It is also possible to write q in terms of non-symmetric matrices, e.g.

$$q = [x_1 \quad x_2] \underbrace{\begin{bmatrix} 3 & 4 \\ 0 & -1 \end{bmatrix}}_{\text{non-symmetric}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{or} \quad q = [x_1 \quad x_2] \underbrace{\begin{bmatrix} 3 & 1 \\ 3 & -1 \end{bmatrix}}_{\text{non-symmetric}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{etc.} \quad \square$$

Example 15

Express

$$q = x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3$$

in the form $q = \underline{x}^T A \underline{x}$, where A is a symmetric matrix.

Solution

$$q = x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3$$

$$= [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Note: The associated symmetric matrix of q is $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 7 \end{bmatrix}$.

□

If in the general case $a_{ij} \neq a_{ji}$, we define $c_{ij} = \frac{a_{ij} + a_{ji}}{2}$, then $c_{ij} = c_{ji}$ and the coefficient of $(x_i x_j + x_j x_i)$ is $c_{ij} + c_{ji} = a_{ij} + a_{ji}$ as required. Then

$$q = \underline{x}^T C \underline{x} \quad \text{with} \quad C = C^T.$$

For example,

$$\begin{aligned} q &= x_1^2 + 7x_1x_2 - x_2x_1 + 2x_2^2 \\ &= x_1^2 + 6x_1x_2 + 2x_2^2 \\ &= [x_1 \quad x_2] \underbrace{\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}}_{\text{symmetric}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

4.2 Definiteness

It is of interest in many applications to know if the sign of q remains the same whatever the values of the elements of $\underline{x}$.

- $q = \underline{x}^T A \underline{x}$ is **positive definite** if $q > 0$ for all $\underline{x} \neq \underline{0}$.
- $q = \underline{x}^T A \underline{x}$ is **negative definite** if $q < 0$ for all $\underline{x} \neq \underline{0}$.

- $q = \underline{x}^T A \underline{x}$ is **positive semidefinite** if $q \geq 0$ for all $\underline{x} \neq \underline{0}$.
- $q = \underline{x}^T A \underline{x}$ is **negative semidefinite** if $q \leq 0$ for all $\underline{x} \neq \underline{0}$.

- $q = \underline{x}^T A \underline{x}$ is **indefinite** if it can take both positive and negative values.

The associated real symmetric matrix A is also said to be positive (or negative) definite if $q > 0$ (or $q < 0$) respectively.

Note that if q is positive definite, then it is also positive semidefinite. Similarly, if q is negative definite, then it is also negative semidefinite. However, if q is positive semidefinite, it may or may not be positive definite, since q may take on the value 0.

Simple examples:

1. $q = x_1^2 + 3x_2^2$ is positive definite (i.e. $q > 0$ whenever $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \underline{0}$).
2. $q = -4x_1^2 - x_2^2$ is negative definite.
3. $q = 3x_1^2$ is positive semidefinite (i.e. $q = 3x_1^2 = 3x_1^2 + 0x_2^2 \geq 0$ whenever $\underline{x} \neq \underline{0}$).
4. $q = -x_2^2$ is negative semidefinite.
5. $q = 3x_1^2 - x_2^2$ is indefinite.

How about $q = 3x_1^2 + 4x_1x_2 - x_2^2$ or quadratic forms with more than two variables?

We may determine whether q has constant sign or not by “completing the square”.

Example 16

Determine whether the quadratic form

$$q = x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3$$

and its associated symmetric matrix are positive/negative definite, positive/negative semidefinite or indefinite.

Solution

$$\begin{aligned} q &= x_1^2 + 2x_2^2 + 7x_3^2 - 2x_1x_2 + 4x_1x_3 - 2x_2x_3 \\ &= (x_1 - x_2 + 2x_3)^2 - x_2^2 - 4x_3^2 + 4x_2x_3 + 2x_2^2 + 7x_3^2 - 2x_2x_3 \\ &= (x_1 - x_2 + 2x_3)^2 + x_2^2 + 3x_3^2 + 2x_2x_3 \\ &= (x_1 - x_2 + 2x_3)^2 + (x_2 + x_3)^2 + 2x_3^2 \end{aligned}$$

Thus $q > 0$ for any choice of x_1, x_2, x_3 .

$\therefore$ Both q and its associated symmetric matrix $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 7 \end{bmatrix}$ (from Example 15)

are **positive definite**. □

Example 17

Determine whether the quadratic form

$$q = 3x_1^2 + 4x_1x_2 - x_2^2$$

and its associated symmetric matrix are positive/negative definite, positive/negative semidefinite or indefinite.

Solution

$$\begin{aligned} q &= 3x_1^2 + 4x_1x_2 - x_2^2 \\ &= 3 \left(x_1^2 + \frac{4x_1x_2}{3} - \frac{x_2^2}{3} \right) \\ &= 3 \left[\left(x_1 + \frac{2}{3}x_2 \right)^2 - \frac{4x_2^2}{9} - \frac{x_2^2}{3} \right] \\ &= 3 \left(x_1 + \frac{2}{3}x_2 \right)^2 - \frac{7x_2^2}{3} \end{aligned}$$

Thus q and its associated symmetric matrix $A = \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix}$ are **indefinite**. □

We shall now reduce q to a sum of squares using matrix and eigenvalue ideas.

Recall that for a real symmetric matrix, it is always possible to find n linearly independent eigenvectors. If the eigenvalues are distinct, these eigenvectors will also be orthogonal. If they are not distinct, it is in fact still possible to obtain n orthogonal eigenvectors by using the Gram-Schmidt process.

If these eigenvectors $\underline{x}_i$ are normalized such that $\underline{x}_i^T \underline{x}_i = 1$ (orthonormal eigenvectors) then the matrix P whose columns are the orthonormalized eigenvectors will have the property that

$$P^T P = I$$

since the ij -th element of $P^T P$ is

$$\underline{x}_i^T \underline{x}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Hence, $P^T = P^{-1}$.

Any matrix with this property is said to be orthogonal.

Thus, it is possible to diagonalize a symmetric matrix using an orthogonal matrix P , i.e.

$$P^T A P = D.$$

Given a quadratic form

$$q = \underline{x}^T A \underline{x}.$$

We may make the change of variable

$$\underline{x} = P \underline{y}$$

where P is the orthogonal matrix whose columns are the orthonormal eigenvectors of A to give

$$q = (P \underline{y})^T A (P \underline{y}) = \underline{y}^T P^T A P \underline{y} = \underline{y}^T D \underline{y}.$$

Hence,

$$q = \sum_{i=1}^n \lambda_i y_i^2$$

where λ_i 's are the diagonal entries of D and they are the eigenvalues A .

This shows that q may be written as a sum of squares with the coefficients of the squared terms being the eigenvalues of A .

This gives the following result:

Theorem

A quadratic form and its associated symmetric matrix are positive definite iff all the eigenvalues of the associated symmetric matrix are positive.

Example 18

Consider the quadratic form

$$q = 2x_1^2 - 2x_1x_2 + 2x_2^2 = [x_1 \ x_2] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underline{x}^T A \underline{x}.$$

After some calculations (left as exercise), we can find that the eigenvalues and their corresponding eigenvectors of A are

$$\lambda_1 = 1, \quad \underline{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \text{Its normalized eigenvector is } \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

$$\lambda_2 = 3, \quad \underline{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \text{Its normalized eigenvector is } \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Note that $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ are orthogonal, since $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 0.$

Take $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Then $P^T = P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Then $P^T A P = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = D$.

If we make the change of variable $\underline{x} = P\underline{y}$ or $\underline{y} = P^{-1}\underline{x} = P^T\underline{x}$, i.e.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} (y_1 - y_2)/\sqrt{2} \\ (y_1 + y_2)/\sqrt{2} \end{bmatrix},$$

we have

$$q = y_1^2 + 3y_2^2 \quad (= \underline{y}^T D \underline{y})$$

showing that q and A are **positive definite**.

It may be verified directly that

$$\begin{aligned} q &= 2x_1^2 - 2x_1x_2 + 2x_2^2 \\ &= 2\left(\frac{y_1 - y_2}{\sqrt{2}}\right)^2 - 2\left(\frac{y_1 - y_2}{\sqrt{2}}\right)\left(\frac{y_1 + y_2}{\sqrt{2}}\right) + 2\left(\frac{y_1 + y_2}{\sqrt{2}}\right)^2 \end{aligned}$$

$$\begin{aligned} &= (y_1^2 - 2y_1y_2 + y_2^2) - (y_1^2 - y_2^2) + (y_1^2 + 2y_1y_2 + y_2^2) \\ &= y_1^2 + 3y_2^2 \end{aligned}$$

Or

$$q = 2x_1^2 - 2x_1x_2 + 2x_2^2 = \frac{(x_1 + x_2)^2}{2} + \frac{3(x_2 - x_1)^2}{2}.$$

Note that “completing the square” does not necessarily give the same factorization. In this case we might have

$$\begin{aligned} q &= 2x_1^2 - 2x_1x_2 + 2x_2^2 \\ &= 2(x_1^2 - x_1x_2 + x_2^2) \\ &= 2 \left[\left(x_1 - \frac{x_2}{2} \right)^2 - \frac{x_2^2}{4} + x_2^2 \right] \\ &= 2 \left(x_1 - \frac{x_2}{2} \right)^2 + \frac{3x_2^2}{2}. \end{aligned}$$

Example 19

For any constant $c > 0$, the points (x_1, x_2) satisfy the equation

$$2x_1^2 - 2x_1x_2 + 2x_2^2 = c.$$

Form a conic in the plane.

Solution

With the change of variable given in the Example 18, we see that the equation becomes

$$y_1^2 + 3y_2^2 = c$$

which we recognize as the equation of an ellipse.

The transformation $\underline{x} = P\underline{y}$, where $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, represents a rotation of the axes

through 45° in anti-clockwise direction. (See the next page for the explanation.)

Explanation:

Recall from MA1200 (rotation of axes in Chapter 1) that when the x - and y -axes are rotated through an angle θ in the anti-clockwise direction to form the new x' - and y' -axes, we have the following relations

$$\begin{cases} x = x' \cos \theta - y' \sin \theta \\ y = x' \sin \theta + y' \cos \theta \end{cases}$$

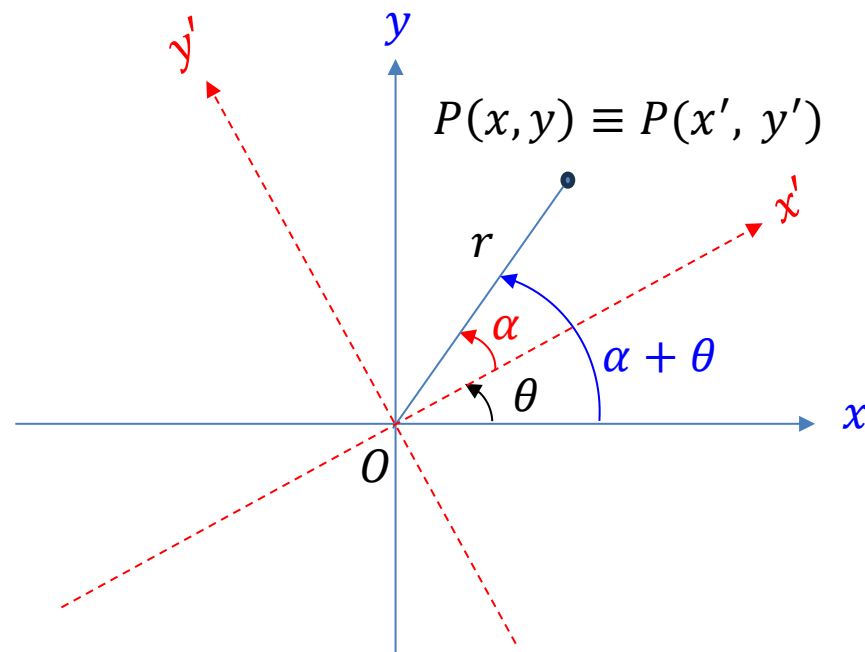
or in matrix form,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

If we put

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix},$$

we get $\begin{cases} \cos \theta = \frac{1}{\sqrt{2}} \\ \sin \theta = \frac{1}{\sqrt{2}} \end{cases} \Rightarrow \theta = 45^\circ. \quad \square$



$$x = r \cos(\alpha + \theta) = r \cos \alpha \cos \theta - r \sin \alpha \sin \theta$$

$$= x' \cos \theta - y' \sin \theta$$

$$y = r \sin(\alpha + \theta) = r \sin \alpha \cos \theta + r \cos \alpha \sin \theta$$

$$= y' \cos \theta + x' \sin \theta$$

Definition

The **k -th order leading principal minor** of the $n \times n$ symmetric matrix $A = [a_{ij}]$ is the determinant of the matrix obtained by deleting the last $(n - k)$ rows and columns of A (where $k = 1, \dots, n$):

$$A_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}.$$

Here are two equivalent definitions that can be used in practice for checking the positive definiteness of a real symmetric matrix A :

- (a) A is positive definite iff all leading principal minors are positive.

That is, let

$$\begin{aligned} A_1 &= a_{11}, \\ A_2 &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \\ A_3 &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \\ &\vdots \\ A_n &= \det(A) \end{aligned}$$

be the leading principal minors, then A is positive definite iff $A_1, A_2, \dots, A_n$ are all positive.

- (b) A is positive definite iff when A has been reduced to echelon form (only by replacement row reduction), all the pivots (diagonal elements) are positive.

Example 20

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

Compute all leading principal minors of A :

$$A_1 = 2 > 0,$$

$$A_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 6 - 1 = 5 > 0,$$

$$A_3 = \begin{vmatrix} 2 & 1 & -2 \\ 1 & 3 & 0 \\ -2 & 0 & 4 \end{vmatrix} = -1 \begin{vmatrix} 1 & -2 \\ 0 & 4 \end{vmatrix} + 3 \begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix} = -(4) + 3(4) = 8 > 0.$$

Since all leading principal minors are positive, it follows that A is **positive definite**.

OR

$$A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 3 & 0 \\ -2 & 0 & 4 \end{bmatrix} \xrightarrow[R_3 + R_1]{R_2 - \frac{1}{2}R_1} \begin{bmatrix} 2 & 1 & -2 \\ 0 & 5/2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{R_3 - \frac{2}{5}R_2} \begin{bmatrix} 2 & 1 & -2 \\ 0 & 5/2 & 1 \\ 0 & 0 & 8/5 \end{bmatrix}$$

Since all the pivots of the echelon form of A are positive, then A is **positive definite**.

Theorem

Let A be an $n \times n$ symmetric matrix and let A_k for $k = 1, \dots, n$ be its leading principal minors. Then

- A is positive definite if and only if $A_k > 0$ for $k = 1, \dots, n$.
- A is negative definite if and only if $(-1)^k A_k > 0$ for $k = 1, \dots, n$.

(That is, A is negative definite if and only if the leading principal minors alternate in sign, starting with negative for A_1 .)

Example 21

$$\text{Let } A = \begin{bmatrix} -3 & 2 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

The leading principal minors of A are

$$A_1 = -3 (< 0),$$

$$A_2 = \begin{vmatrix} -3 & 2 \\ 2 & -3 \end{vmatrix} = 5 (> 0)$$

$$\text{and } A_3 = |A| = \begin{vmatrix} -3 & 2 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & -5 \end{vmatrix} = (-5) \begin{vmatrix} -3 & 2 \\ 2 & -3 \end{vmatrix} = (-5)(5) = -25 (< 0).$$

Thus, A is negative definite.

Remark:

The eigenvalues of A are $\lambda_1 = -5$, $\lambda_2 = \frac{-5+\sqrt{17}}{2}$ and $\lambda_3 = \frac{-5-\sqrt{17}}{2}$.

Since all the eigenvalues of A are negative, it shows that A is negative definite.

Definition

The **k th order principal minors** of an $n \times n$ symmetric matrix A are the determinants of the $k \times k$ matrices obtained by deleting $n - k$ rows and the corresponding $n - k$ columns of A (where $k = 1, \dots, n$).

Note that the k th order leading principal minor of a matrix is one of its k th order principal minors.

E.g. Let $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

- The 2 **first-order principal minors** of A are a and c . (The first-order leading principal minor of A is a).
- The **second-order principal minor** of A is $|A| = ac - b^2$. (The second-order leading principal minor of A is $|A| = ac - b^2$.)

E.g. Let $A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$.

- The 3 **first-order principal minors** of A are a , d and f . (The first-order leading principal minor of A is a).
- The 3 **second-order principal minors** of A are $\begin{vmatrix} a & b \\ b & d \end{vmatrix}$, $\begin{vmatrix} a & c \\ c & f \end{vmatrix}$ and $\begin{vmatrix} d & e \\ e & f \end{vmatrix}$. (The second-order leading principal minor of A is $\begin{vmatrix} a & b \\ b & d \end{vmatrix}$).
- The **third-order principal minor** of A is $|A|$. (The third-order leading principal minor of A is $|A|$).

The following results give the criteria of positive/negative semidefinite:

Proposition

Let A be an $n \times n$ symmetric matrix. Then

- A is **positive semidefinite** iff all its principal minors ≥ 0 .
- A is **negative semidefinite** iff its k th order principal minors $\begin{cases} \leq 0 & \text{for } k \text{ odd} \\ \geq 0 & \text{for } k \text{ even} \end{cases}$.

If A is neither positive/negative definite nor positive/negative semidefinite, then A is indefinite. For example, if the k th order leading principal minor is negative when k is even, then A must be indefinite.

Example 22

Classify the following matrices as positive or negative definite, positive or negative semidefinite, or indefinite.

$$(a) \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(b) \quad A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution

$$(a) \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The 3 first-order principal minors of A are $1 (> 0)$, $1 (> 0)$ and $1 (> 0)$.

The 3 second-order principal minors of A are $\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = \mathbf{0} (\geq \mathbf{0})$, $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 (\geq 0)$

and $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 (\geq 0)$.

The third-order principal minor of A is

$$|A| = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = \mathbf{0 (\geq 0)}$$

Since all principal minors of A are non-negative (i.e. ≥ 0), A is positive semidefinite.

(b) $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Since the second-order leading principal minor of A is $\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3 < 0$, A cannot be positive definite, negative definite, positive semidefinite or negative semidefinite. Then A must be indefinite.