Eigenvalues & Eigenvectors

Eigenvalues and Eigenvectors

- Eigenvalue problem (one of the most important problems in the linear algebra): If A is an $n \times n$ matrix, do there exist nonzero vectors $\mathbf{x}$ in R^n such that $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$?

  (The term eigenvalue is from the German word *Eigenwert*, meaning "proper value")

- Eigenvalue

  A: an $n \times n$ matrix

  λ: a scalar (could be **zero**)

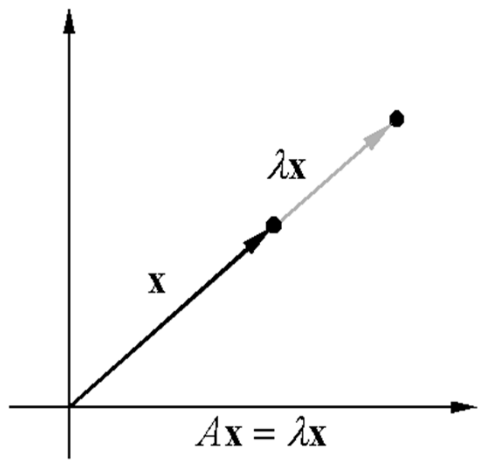
  $\mathbf{x}$: a **nonzero** vector in R^n

Eigenvalue

$$A\mathbf{x} = \lambda\mathbf{x}$$

Eigenvector

Geometric Interpretation



Eigenvalues and Eigenvectors

- Example 1: Verifying eigenvalues and eigenvectors.

$$A=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix} \quad \mathbf{x}_1=\begin{bmatrix}1\\ 0\end{bmatrix} \quad \mathbf{x}_2=\begin{bmatrix}0\\ 1\end{bmatrix}$$

Eigenvalue

$$A\mathbf{x}_1=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}2\\ 0\end{bmatrix}=2\begin{bmatrix}1\\ 0\end{bmatrix}=2\mathbf{x}_1$$

Eigenvector

Eigenvalue

$$A\mathbf{x}_2=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}0\\ -1\end{bmatrix}=-1\begin{bmatrix}0\\ 1\end{bmatrix}=(-1)\mathbf{x}_2$$

Eigenvector

- In fact, for each eigenvalue, it has infinitely many eigenvectors. For $\lambda = 2$, $[3\ 0]^T$ or $[5\ 0]^T$ are both corresponding eigenvectors. Moreover, $([3\ 0] + [5\ 0])^T$ is still an eigenvector.
- The proof is in Theorem. 7.1.

Eigenspace

Theorem 7.1: (The *eigenspace* corresponding to λ of matrix A)

- If A is an $n \times n$ matrix with an eigenvalue λ, then the set of all eigenvectors of λ **together with the zero vector** is a subspace of R^n. This subspace is called the *eigenspace* of λ.

Proof:

$\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors corresponding to λ

$$\text{(i.e., } A\mathbf{x}_1 = \lambda\mathbf{x}_1,\ \ A\mathbf{x}_2 = \lambda\mathbf{x}_2)$$

$$(1)\ A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2 = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$$

(i.e., $\mathbf{x}_1 + \mathbf{x}_2$ is also an eigenvector corresponding to λ)

$$(2)\ A(c\mathbf{x}_1) = c(A\mathbf{x}_1) = c(\lambda\mathbf{x}_1) = \lambda(c\mathbf{x}_1)$$

(i.e., $c\mathbf{x}_1$ is also an eigenvector corresponding to λ)

Since this set is closed under vector addition and scalar multiplication, this set is a subspace of R^n according to Theorem 4.5.

Eigenspace

Eamplex 3: Examples of eigenspaces on the *xy*-plane

- For the matrix *A* as follows, the corresponding eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 1$:

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

Solution

For the eigenvalue $\lambda_1 = -1$, corresponding vectors are any vectors on the x-axis

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

Thus, the eigenspace corresponding to $\lambda = -1$ is the x-axis, which is a subspace of R^2.

For the eigenvalue $\lambda_2 = 1$, corresponding vectors are any vectors on the y-axis

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

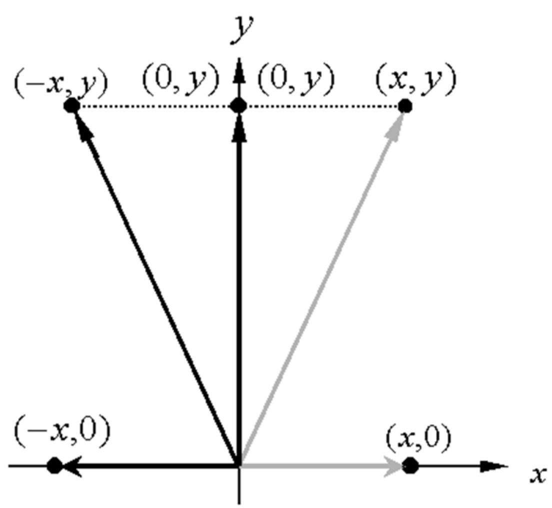
Thus, the eigenspace corresponding to $\lambda = 1$ is the *y*-axis, which is a subspace of R^2.

Eigenspace

- Geometrically speaking, multiplying a vector (x, y) in R^2 by the matrix A in the example, corresponds to a reflection to the y-axis.

$$A\mathbf{v} = A\begin{bmatrix} x \\ y \end{bmatrix} = A\left(\begin{bmatrix} x \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ y \end{bmatrix}\right) = A\begin{bmatrix} x \\ 0 \end{bmatrix} + A\begin{bmatrix} 0 \\ y \end{bmatrix}$$

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



Finding eigenvalues and eigenvectors

Theorem 7.2: (Finding eigenvalues and eigenvectors of a matrix $A \in M_{n\times n}$)

- Let A be an $n\times n$ matrix.

  (1) An eigenvalue of A is a scalar λ such that $\det(\lambda I - A) = 0$.

  (2) The eigenvectors of A corresponding to λ are the nonzero solutions of $(\lambda I - A)\mathbf{x} = \Theta$.

Finding eigenvalues and eigenvectors

- Note: some definitions of the eigenvalue problem:
  - homogeneous system

$$A\mathbf{x} = \lambda\mathbf{x} \;\Rightarrow\; A\mathbf{x} = \lambda I\mathbf{x} \;\Rightarrow\; (\lambda I - A)\mathbf{x} = \Theta$$

$(\lambda I - A)\mathbf{x} = \Theta$ has nonzero solutions for $\mathbf{x}$ iff $\det(\lambda I - A) = 0$

  - Characteristic equation of A:

$$\det(\lambda I - A) = 0$$

  - Characteristic polynomial of $A \in M_{n\times n}$:

$$\det(\lambda I - A) = \left|(\lambda I - A)\right| = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

Finding eigenvalues and eigenvectors

- **Example 4:** Finding eigenvalues and eigenvectors

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Solution: Characteristic equation:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix}$$

$$= \lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2) = 0$$

$$\Rightarrow \lambda = -1, -2$$

Eigenvalue: $\lambda_1 = -1,\ \lambda_2 = -2$

Finding eigenvalues and eigenvectors

(1) $\lambda_1 = -1$ $\quad \Rightarrow (\lambda_1 I - A)\mathbf{x} = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & -4 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad t \neq 0$$

(2) $\lambda_2 = -2$ $\quad \Rightarrow (\lambda_2 I - A)\mathbf{x} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3s \\ s \end{bmatrix} = s\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad s \neq 0$$

Finding eigenvalues and eigenvectors

Example 5:

- Find the eigenvalues and corresponding eigenvectors for the matrix A. What is the dimension of the eigenspace of each eigenvalue?

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

Solution: Characteristic equation:

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

Eigenvalue: $\lambda = 2$

Finding eigenvalues and eigenvectors

The eigenspace of $\lambda = 2$:

$$(\lambda I - A)\mathbf{x} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \neq 0$$

$\left\{ s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \middle| \, s, t \in R \right\}$:the eigenspace of A corresponding to $\lambda = 2$

Thus, the dimension of its eigenspace is 2.

Finding eigenvalues and eigenvectors

- Notes:

  (1) If an eigenvalue λ_1 occurs as a multiple root (k times) for the characteristic polynominal, then λ_1 has multiplicity k.

  (2) The multiplicity of an eigenvalue is greater than or equal to the dimension of its eigenspace.

  - In Example. 5, k is 3 and the dimension of its eigenspace is 2.

Finding eigenvalues and eigenvectors

Example 6:

- Find the eigenvalues of the matrix *A* and find a basis for each of the corresponding eigenspaces

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & -10 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

Solution: Characteristic equation:

$$|\lambda I - A| = \begin{vmatrix} \lambda-1 & 0 & 0 & 0 \\ 0 & \lambda-1 & -5 & 10 \\ -1 & 0 & \lambda-2 & 0 \\ -1 & 0 & 0 & \lambda-3 \end{vmatrix}$$

$$= (\lambda-1)^2(\lambda-2)(\lambda-3) = 0$$

Eigenvalues: $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$

- According to the note on the previous slide, the dimension of the eigenspace of $\lambda_1 = 1$ is at most to be 2.
- For $\lambda_2 = 2$ and $\lambda_3 = 3$, the demensions of their eigenspaces are at most to be 1.

Finding eigenvalues and eigenvectors

$$(1)\ \lambda_1 = 1 \quad \Rightarrow (\lambda_1 I - A)\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 10 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\overset{\text{G.-J.E.}}{\Longrightarrow} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2t \\ s \\ 2t \\ t \end{bmatrix} = s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \quad s, t \neq 0$$

$$\Rightarrow \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}$$ is a basis for the eigenspace corresponding to $\lambda_1 = 1$.

The dimension of the eigenspace of $\lambda_1 = 1$ is 2.

Finding eigenvalues and eigenvectors

(2) $\lambda_2 = 2$ $\qquad \Rightarrow (\lambda_2 I - A)\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -5 & 10 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$$\overset{\text{G.-J.E.}}{\Rightarrow} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 5t \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} 0 \\ 5 \\ 1 \\ 0 \end{bmatrix}, \quad t \neq 0$$

$\Rightarrow \left\{ \begin{bmatrix} 0 \\ 5 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis for the eigenspace corresponding to $\lambda_2 = 2$

The dimension of the eigenspace of $\lambda_2 = 2$ is 1.

Finding eigenvalues and eigenvectors

(3) $\lambda_3 = 3 \quad \Rightarrow (\lambda_3 I - A)\mathbf{x} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & -5 & 10 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$$\overset{\text{G.-J.E.}}{\Rightarrow} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ -5t \\ 0 \\ t \end{bmatrix} = t\begin{bmatrix} 0 \\ -5 \\ 0 \\ 1 \end{bmatrix}, \quad t \neq 0$$

$\Rightarrow \left\{ \begin{bmatrix} 0 \\ -5 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigenspace corresponding to $\lambda_3 = 3$.

The dimension of the eigenspace of $\lambda_3 = 3$ is 1.

Eigenvalues for triangular matrices

Theorem 7.3:

- If A is an $n \times n$ triangular matrix, then its eigenvalues are the entries on its main diagonal.

Eigenvalues for triangular matrices

Example 7:

- Finding eigenvalues for triangular and diagonal matrices

$$\text{(a) } A=\begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 5 & 3 & -3 \end{bmatrix} \qquad \text{(b) } A=\begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

Solution:

$$\text{(a) } |\lambda I - A| = \begin{vmatrix} \lambda-2 & 0 & 0 \\ 1 & \lambda-1 & 0 \\ -5 & -3 & \lambda+3 \end{vmatrix} = (\lambda-2)(\lambda-1)(\lambda+3)=0$$

$$\Rightarrow \lambda_1 = 2,\ \lambda_2 = 1,\ \lambda_3 = -3$$

$$\text{(b) } \lambda_1 = -1,\ \lambda_2 = 2,\ \lambda_3 = 0,\ \lambda_4 = -4,\ \lambda_5 = 3$$

- According to Theorem 3.2, the determinant of a triangular matrix is the product of the entries on the main diagonal .

Eigenvalues and eigenvectors of L.T.

- A number λ is called an eigenvalue of a linear transformation $T:V\rightarrow V$ if there is a nonzero vector $\mathbf{x}$ such that $T(\mathbf{x})=\lambda\mathbf{x}$. The vector $\mathbf{x}$ is called an eigenvector of T corresponding to λ, and the set of all eigenvectors of λ (together with the zero vector) is called the eigenspace of λ.

Eigenvalues and eigenvectors of L.T.

- The definition of linear transformation functions was introduced in Chapter 6.
- The typical example of a linear transformation function is that each component of the resulting vector is the linear combination of the components in the input vector **x**.
  - An example for a linear transformation $T: R^3 \rightarrow R^3$

$$T(x_1, x_2, x_3) = (x_1 + 3x_2, 3x_1 + x_2, -2x_3)$$

Eigenvalues and eigenvectors of L.T.

Example 8: Find eigenvalues and eigenvectors for standard matrices

- Find the eigenvalues and corresponding eigenvectors for

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

- A is the standard matrix for $T(x_1, x_2, x_3) = (x_1 + 3x_2, 3x_1 + x_2, -2x_3)$

Solution

$$|\lambda I - A| = \begin{bmatrix}\lambda-1 & -3 & 0\\ -3 & \lambda-1 & 0\\ 0 & 0 & \lambda+2\end{bmatrix} = (\lambda+2)^2(\lambda-4)=0$$

$\Rightarrow$ eigenvalues $\lambda_1 = 4,\ \lambda_2 = -2$

For $\lambda_1 = 4$, the corresponding eigenvector is $(1,\ 1,\ 0)$.

For $\lambda_2 = -2$, the corresponding eigenvectors are $(1,\ -1,\ 0)$ and $(0,\ 0,\ 1)$.

Eigenvalues and eigenvectors of L.T.

- Notes: The relationship among eigenvalues, eigenvectors, and diagonalization

Let $T: R^3 \rightarrow R^3$ be the linear transformation whose corresponding standard matrix is A in Example 8, and let B' be a basis of R^3 made up of three linearly independent eigenvectors of A found in Example 8, i.e., $B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1,\ 1,\ 0), (1,\ -1,\ 0), (0,\ 0,\ 1)\}$

Then A', the matrix of T relative to the basis B', defined as

$$[[T(\mathbf{v}_1)]_{B'}\ [T(\mathbf{v}_2)]_{B'}\ [T(\mathbf{v}_3)]_{B'}]$$

is diagonal, and the main diagonal entries are corresponding eigenvalues.

$$B' = \{\underbrace{(1,\ 1,\ 0)}_{\text{for } \lambda_1 = 4}, \underbrace{(1,\ -1,\ 0), (0,\ 0,\ 1)}_{\text{for } \lambda_2 = -2}\}$$

Eigenvectors of A

$$A' = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Eigenvalues of A

Diagonalization

- **Diagonalization problem:** For a square matrix A, does there exist an invertible matrix P such that $P^{-1}AP$ is diagonal?
- **Diagonalizable matrix:**
  - Definition 1: A square matrix A is called diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix (i.e., P diagonalizes A)
  - Definition 2: A square matrix A is called diagonalizable if A is similar to a diagonal matrix.

- In Sec. 6.4, two square matrices A and B are **similar** if there exists an invertible matrix P such that $B = P^{-1}AP$.

- Notes: In this section, it is shown that the eigenvalue problem is related closely to the diagonalization problem.

Diagonalization

Theorem. 7.4: Similar matrices have the same eigenvalues

- If A and B are similar $n\times n$ matrices, then they have the same eigenvalues.

Proof:

A and B are similar $\Rightarrow B = P^{-1}AP$

Consider the characteristic equation of B:

For any diagonal matrix in the form of $D = \lambda I$, $P^{-1}DP = D$

$$
\begin{aligned}
|\lambda I - B| &= |\lambda I - P^{-1}AP| = |P^{-1}\lambda IP - P^{-1}AP| = |P^{-1}(\lambda I - A)P| \\
&= |P^{-1}||\lambda I - A||P| = |P^{-1}||P||\lambda I - A| = |P^{-1}P||\lambda I - A| \\
&= |\lambda I - A|
\end{aligned}
$$

Since A and B have the same characteristic equation, they are with the same eigenvalues.

Diagonalization

- **Example 1:** Eigenvalue problems and diagonalization programs

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

Solution:

Characteristic equation:

$$|\lambda I - A| = \begin{vmatrix}\lambda-1 & -3 & 0\\ -3 & \lambda-1 & 0\\ 0 & 0 & \lambda+2\end{vmatrix} = (\lambda-4)(\lambda+2)^2 = 0$$

$$\text{The eigenvalues}: \lambda_1 = 4,\ \lambda_2 = -2,\ \lambda_3 = -2$$

$$(1)\ \lambda = 4 \Rightarrow \text{ the eigenvector } \quad \mathbf{p}_1 = \begin{bmatrix}1\\ 1\\ 0\end{bmatrix}$$

Diagonalization

(2) $\lambda = -2 \Rightarrow$ the eigenvector $\quad \mathbf{p}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{p}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- Note: If $P = [\mathbf{p}_2 \quad \mathbf{p}_1 \quad \mathbf{p}_3]$

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

- The above example can verify Theorem 7.4 since the eigenvalues for both A and $P^{-1}AP$ are the same to be 4, –2, and –2
- The reason why the matrix P is constructed with the eigenvectors of A is demonstrated in Theorem 7.5 (see the next slide).

Diagonalization

Theorem 7.5: Condition for diagonalization

- An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.
  - If there are n linearly independent eigenvectors, it does not imply that there are n distinct eigenvalues. In an extreme case, it is possible to have only one eigenvalue with the multiplicity n, and there are n linearly independent eigenvectors for this eigenvalue.
  - However, if there are n distinct eigenvalues, then there are n linearly independent eivenvectors (see Thm. 7.6), and thus A must be diagonalizable.

Proof: $(\Rightarrow)$

Since A is diagonalizable, there exists an invertible P s.t. $D = P^{-1}AP$ is diagonal.

Let $P = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_n]$ and $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then

$$PD = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$= [\lambda_1\mathbf{p}_1 \ \lambda_2\mathbf{p}_2 \ \cdots \ \lambda_n\mathbf{p}_n]$$

Diagonalization

$AP = PD$ (since $D = P^{-1}AP$)

$[A\mathbf{p}_1\ A\mathbf{p}_2\ \cdots\ A\mathbf{p}_n] = [\lambda_1\mathbf{p}_1\ \lambda_2\mathbf{p}_2\ \cdots\ \lambda_n\mathbf{p}_n]$

$\Rightarrow A\mathbf{p}_i = \lambda_i\mathbf{p}_i,\quad i = 1,\ 2, \ldots,\ n$

(The above equations imply the column vectors $\mathbf{p}_i$ of P are eigenvectors of A, and the diagonal entries λ_i in D are eigenvalues of A)

Because A is diagonalizable $\Rightarrow P$ is invertible

$\Rightarrow$ Columns in P, i.e., $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n$, are linearly independent

Thus, A has n linearly independent eigenvectors.

- $(\Leftarrow)$

Since A has n linearly independent eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \cdots \mathbf{p}_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots \lambda_n$, then

$\Rightarrow A\mathbf{p}_i = \lambda_i\mathbf{p}_i,\quad i = 1,\ 2, \ldots,\ n$

Let $P = [\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n]$

Diagonalization

$$
\begin{aligned}
AP = A[\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n] &= [A\mathbf{p}_1\ A\mathbf{p}_2\ \cdots\ A\mathbf{p}_n] \\
&= [\lambda_1\mathbf{p}_1\ \lambda_2\mathbf{p}_2\ \cdots\ \lambda_n\mathbf{p}_n] \\
&= [\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = PD
\end{aligned}
$$

Since $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n$ are linearly independent

$\Rightarrow P$ is invertible

$\Rightarrow AP = PD \Rightarrow P^{-1}AP = D$

$\Rightarrow A$ is diagonalizable

(according to the definition of the diagonalizable matrix)

Note that $\mathbf{p}_i$'s are linearly independent eigenvectors and the diagonal entries λ_i in the resulting diagonalized D are eigenvalues of A.

Diagonalization

Example 4: A matrix that is not diagonalizable

- Show that the following matrix is not diagonalizable

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

Solution: Characteristic equation:

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

The eigenvalue $\lambda_1 = 1$, and then solve $(\lambda_1 I - A)\mathbf{x} = \Theta$ for eigenvectors

$$\lambda_1 I - A = I - A = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \Rightarrow \text{eigenvector } \mathbf{p}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Since A does not have two linearly independent eigenvectors, A is not diagonalizable.

Diagonalization

- Steps for diagonalizing an $n \times n$ square matrix:

Step 1: Find n linearly independent eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \cdots \mathbf{p}_n$ for A with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

Step 2: Let $P = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_n]$

Step 3:

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

where $A\mathbf{p}_i = \lambda_i \mathbf{p}_i, \quad i = 1, \ 2, \ldots, \ n$

Diagonalization

- Example 5: Diagonalizing a matrix

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

Find a matrix P such that $P^{-1}AP$ is diagonal.

Solution: Characteristic equation:

$$\left|\lambda I-A\right|=\begin{vmatrix}\lambda-1 & 1 & 1\\ -1 & \lambda-3 & -1\\ 3 & -1 & \lambda+1\end{vmatrix}=(\lambda-2)(\lambda+2)(\lambda-3)=0$$

The eigenvalues: $\lambda_1=2,\ \lambda_2=-2,\ \lambda_3=3$

Diagonalization

$$\lambda_1 = 2 \quad \Rightarrow \lambda_1 I - A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t \\ 0 \\ t \end{bmatrix} \quad \Rightarrow \quad \text{eigenvector } \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_2 = -2 \quad \Rightarrow \lambda_2 I - A = \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}t \\ -\frac{1}{4}t \\ t \end{bmatrix} \quad \Rightarrow \quad \text{eigenvector } \mathbf{p}_2 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

Diagonalization

$$\lambda_3 = 3 \quad \Rightarrow \lambda_3 I - A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t \\ t \\ t \end{bmatrix} \quad \Rightarrow \quad \text{eigenvector } \mathbf{p}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3] = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & 4 & 1 \end{bmatrix} \text{ and it follows that}$$

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

Diagonalization

- Note: a quick way to calculate A^k based on the diagonalization technique

$$(1)\ D=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix} \Rightarrow D^k=\begin{bmatrix}\lambda_1^k & 0 & \cdots & 0\\ 0 & \lambda_2^k & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n^k\end{bmatrix}$$

$$(2)\ D=P^{-1}AP \Rightarrow D^k=\underbrace{\underbrace{P^{-1}AP}\ \underbrace{P^{-1}AP}\cdots\underbrace{P^{-1}AP}}_{\text{repeat } k \text{ times}}=P^{-1}A^kP$$

$$A^k=PD^kP^{-1},\ \text{where } D^k=\begin{bmatrix}\lambda_1^k & 0 & \cdots & 0\\ 0 & \lambda_2^k & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n^k\end{bmatrix}$$

Diagonalization

Theorem 7.6: Sufficient conditions for diagonalization

- If an $n \times n$ matrix A has n distinct eigenvalues, then the corresponding eigenvectors are linearly independent and thus A is diagonalizable according to Theorem 7.5.

Proof:

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct eigenvalues and corresponding eigenvectors be $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. In addition, consider that the first m eigenvectors are linearly independent, but the first $m+1$ eigenvectors are linearly dependent, i.e.,

$$\mathbf{x}_{m+1} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m, \qquad (1)$$

where c_i's are not all zero. Multiplying both sides of Eq. (1) by A yields

$$A\mathbf{x}_{m+1} = Ac_1\mathbf{x}_1 + Ac_2\mathbf{x}_2 + \cdots + Ac_m\mathbf{x}_m$$

$$\lambda_{m+1}\mathbf{x}_{m+1} = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_m\lambda_m\mathbf{x}_m \qquad (2)$$

Diagonalization

On the other hand, multiplying both sides of Eq. (1) by λ_{m+1} yields

$$\lambda_{m+1}\mathbf{x}_{m+1} = c_1\lambda_{m+1}\mathbf{x}_1 + c_2\lambda_{m+1}\mathbf{x}_2 + \cdots + c_m\lambda_{m+1}\mathbf{x}_m \qquad (3)$$

Now, subtracting Eq. (2) from Eq. (3) produces

$$c_1(\lambda_{m+1} - \lambda_1)\mathbf{x}_1 + c_2(\lambda_{m+1} - \lambda_2)\mathbf{x}_2 + \cdots + c_m(\lambda_{m+1} - \lambda_m)\mathbf{x}_m = 0$$

Since the first m eigenvectors are linearly independent, we can infer that all coefficients of this equation should be zero, i.e.,

$$c_1(\lambda_{m+1} - \lambda_1) = c_2(\lambda_{m+1} - \lambda_2) = \cdots = c_m(\lambda_{m+1} - \lambda_m) = 0$$

Because all the eigenvalues are distinct, it follows all c_i's equal to 0, which contradicts our assumption that $\mathbf{x}_{m+1}$ can be expressed as a linear combination of the first m eigenvectors. So, the set of n eigenvectors is linearly independent given n distinct eigenvalues, and according to Thm. 7.5, we can conclude that A is diagonalizable.

Diagonalization

❖ Example 7: Determining whether a matrix is diagonalizable

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

Solution: Because A is a triangular matrix, its eigenvalues are

$$\lambda_1 = 1,\ \lambda_2 = 0,\ \lambda_3 = -3.$$

According to Theorem 7.6, because these three values are distinct, A is diagonalizable.

Diagonalization

Example 8: Finding a diagonalized matrix for a linear transformation

- Let $T: R^3 \rightarrow R^3$ be the linear transformation given by

$$T(x_1, x_2, x_3) = (x_1 - x_2 - x_3,\ x_1 + 3x_2 + x_3,\ -3x_1 + x_2 - x_3)$$

Find a basis B' for R^3 such that the matrix for T relative to B' is diagonal.

Solution:

The standard matrix for T is gven by

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

From Example 5 you know that $\lambda_1 = 2,\ \lambda_2 = -2,\ \lambda_3 = 3$ and thus A is diagonalizable. So, these three linearly independent eigenvectors found in Example 5 can be used to form the basis B'. That is

Diagonalization

$$B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(-1,\ 0,\ 1), (1,\ -1,\ 4), (-1,\ 1,\ 1)\}$$

The matrix for T relative to this basis is

$$A' = \left[[T(\mathbf{v}_1)]_{B'}\ \ [T(\mathbf{v}_2)]_{B'}\ \ [T(\mathbf{v}_3)]_{B'} \right]$$
$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

❖ Note that it is not necessary to calculate A' through the above equation, we already know that A' is a diagonal matrix and its main diagonal entries are corresponding eigenvalues of A.

Symmetric Matrices and Orthogonal Diagonalization

- Symmetric matrix:

  A square matrix A is symmetric if it is equal to its transpose:

$$A = A^T$$

- Example 1: Symmetric matrices and nonsymetric matrices

$$A = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 3 & 0 \\ -2 & 0 & 5 \end{bmatrix}$$ (symmetric)

$$B = \begin{bmatrix} 4 & 3 \\ 3 & 1 \end{bmatrix}$$ (symmetric)

$$C = \begin{bmatrix} 3 & 2 & 1 \\ 1 & -4 & 0 \\ 1 & 0 & 5 \end{bmatrix}$$ (nonsymmetric)

Theorem 7.7: Eigenvalues of symmetric matrices

- If A is an $n \times n$ symmetric matrix, then the following properties are true.

  (1) A is diagonalizable (symmetric matrices are guaranteed to have *n* linearly independent eigenvectors and thus be diagonalizable).

  (2) All eigenvalues of A are real numbers.

  (3) If λ is an eigenvalue of A with the multiplicity to be *k*, then λ has *k* linearly independent eigenvectors. That is, the eigenspace of λ has dimension *k*.

- The above theorem is called the **Real Spectral Theorem**, and the set of eigenvalues of A is called the **spectrum** of A.

❖ Example 2:

Prove that a 2 × 2 symmetric matrix is diagonalizable.

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

proof: Characteristic equation:

$$|\lambda I - A| = \begin{vmatrix} \lambda - a & -c \\ -c & \lambda - b \end{vmatrix} = \lambda^2 - (a+b)\lambda + ab - c^2 = 0$$

As a function in λ, this quadratic polynomial function has a nonnegative discriminant as follows:

$$\begin{aligned}(a+b)^2 - 4(1)(ab - c^2) &= a^2 + 2ab + b^2 - 4ab + 4c^2 \\ &= a^2 - 2ab + b^2 + 4c^2 \\ &= \underline{(a-b)^2 + 4c^2} \geq 0\end{aligned}$$

(1) $(a-b)^2+4c^2=0$

$$\Rightarrow a=b,\ c=0$$

$$A=\begin{bmatrix} a & c \\ c & b \end{bmatrix}=\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\ \text{itself is a diagonal matrix.}$$

(2) $(a-b)^2+4c^2>0$

The characteristic polynomial of A has two distinct real roots, which implies that A has two distinct real eigenvalues. According to Thm. 7.6, A is diagonalizable.

Orthogonal matrix

- Orthogonal matrix: A square matrix P is called orthogonal if it is invertible and

$$P^{-1} = P^T \text{ (or } PP^T = P^T P = I\text{)}$$

Orthogonal matrix

- Theorem 7.8: Properties of orthogonal matrices
- An $n \times n$ matrix P is orthogonal if and only if its column vectors form an orthonormal set.

Proof: Suppose the column vectors of P form an orthonormal set, i.e.,

$$P = [\mathbf{p}_1 \ \ \mathbf{p}_2 \ \ \cdots \ \ \mathbf{p}_n], \text{ where } \mathbf{p}_i \cdot \mathbf{p}_j = 0 \text{ for } i \neq j \text{ and } \mathbf{p}_i \cdot \mathbf{p}_i = 1.$$

$$P^T P = \begin{bmatrix} \mathbf{p}_1{}^T\mathbf{p}_1 & \mathbf{p}_1{}^T\mathbf{p}_2 & \cdots & \mathbf{p}_1{}^T\mathbf{p}_n \\ \mathbf{p}_2{}^T\mathbf{p}_1 & \mathbf{p}_2{}^T\mathbf{p}_2 & \cdots & \mathbf{p}_2{}^T\mathbf{p}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p}_n{}^T\mathbf{p}_1 & \mathbf{p}_n{}^T\mathbf{p}_2 & \cdots & \mathbf{p}_n{}^T\mathbf{p}_n \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 \cdot \mathbf{p}_1 & \mathbf{p}_1 \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_1 \cdot \mathbf{p}_n \\ \mathbf{p}_2 \cdot \mathbf{p}_1 & \mathbf{p}_2 \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_2 \cdot \mathbf{p}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p}_n \cdot \mathbf{p}_1 & \mathbf{p}_n \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_n \cdot \mathbf{p}_n \end{bmatrix} = I_n$$

It implies that $P^{-1} = P^T$ and thus P is orthogonal.

Orthogonal matrix

- Ex 5: Show that P is an orthogonal matrix.

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{-2}{3\sqrt{5}} & \frac{-4}{3\sqrt{5}} & \frac{5}{3\sqrt{5}} \end{bmatrix}$$

Solution: If P is a orthogonal matrix, then $P^{-1} = P^T \Rightarrow PP^T = I$

$$PP^T = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{-2}{3\sqrt{5}} & \frac{-4}{3\sqrt{5}} & \frac{5}{3\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{-2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{-4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

Orthogonal matrix

Moreover, let $\mathbf{p}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{-2}{\sqrt{5}} \\ \frac{-2}{3\sqrt{5}} \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{\sqrt{5}} \\ \frac{-4}{3\sqrt{5}} \end{bmatrix}$, and $\mathbf{p}_3 = \begin{bmatrix} \frac{2}{3} \\ 0 \\ \frac{5}{3\sqrt{5}} \end{bmatrix}$,

we can produce $\mathbf{p}_1 \cdot \mathbf{p}_2 = \mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}_2 \cdot \mathbf{p}_3 = 0$ and $\mathbf{p}_1 \cdot \mathbf{p}_1 = \mathbf{p}_2 \cdot \mathbf{p}_2 = \mathbf{p}_3 \cdot \mathbf{p}_3 = 1$.

So, $\{\mathbf{p}_1,\ \mathbf{p}_2,\ \mathbf{p}_3\}$ is an orthonormal set. (Theorem 7.8 can be verified by this example.)

Orthogonal matrix

Theorem 7.9: Properties of symmetric matrices

- Let A be an $n \times n$ symmetric matrix. If λ_1 and λ_2 are distinct eigenvalues of A, then their corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal. (Theorem 7.6 only states that eigenvectors corresponding to distinct eigenvalues are linearly independent)

Proof: $\lambda_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \lambda_1 \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle A\mathbf{x}_1, \mathbf{x}_2 \rangle = (A\mathbf{x}_1)^T \mathbf{x}_2 = (\mathbf{x}_1^T A^T)\mathbf{x}_2$

$$\overset{\text{because } A \text{ is symmetric}}{=} (\mathbf{x}_1^T A)\mathbf{x}_2 = \mathbf{x}_1^T (A\mathbf{x}_2) = \mathbf{x}_1^T (\lambda_2 \mathbf{x}_2) = \langle \mathbf{x}_1, \lambda_2 \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$$

The above equation implies $(\lambda_1 - \lambda_2)\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$, and because $\lambda_1 \neq \lambda_2$, it follows that $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$. So, $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

- For distinct eigenvalues of a symmetric matrix, their corresponding eigenvectors are orthogonal and thus linearly independent to each other.
- Note that there may be multiple $\mathbf{x}_1$ and $\mathbf{x}_2$ corresponding to λ_1 and λ_2.

Orthogonal matrix

- **Orthogonal diagonalization:** A matrix A is orthogonally diagonalizable if there exists an orthogonal matrix P such that $P^{-1}AP = D$ is diagonal.

Theorem 7.10: Fundamental theorem of symmetric matrices

- Let A be an $n \times n$ matrix. Then A is orthogonally diagonalizable and has real eigenvalues if and only if A is symmetric.

Proof:

$(\Rightarrow)$

A is orthogonally diagonalizable

$\Rightarrow D = P^{-1}AP$ is diagonal, and P is an orthogonal matrix s.t. $P^{-1} = P^T$

$\Rightarrow A = PDP^{-1} = PDP^T \Rightarrow A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$

$(\Leftarrow)$

See the next two slides

Orthogonal matrix

- **Orthogonal diagonalization of a symmetric matrix:**

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

  (1) Find all eigenvalues of A and determine the multiplicity of each.

  - According to Theorem 7.9, eigenvectors corresponding to distinct eigenvalues are orthogonal.

  (2) For each eigenvalue of multiplicity 1, choose the unit eigenvector.

  (3) For each eigenvalue of the multiplicity to be $k \geq 2$, find a set of k linearly independent eigenvectors. If this set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is not orthonormal, apply the Gram-Schmidt orthonormalization process.

Orthogonal matrix

It is known that G.-S. process is a kind of linear transformation, i.e., the produced vectors can be expressed as $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$,

i. Since $A\mathbf{v}_1 = \lambda\mathbf{v}_1, A\mathbf{v}_2 = \lambda\mathbf{v}_2, \ldots, A\mathbf{v}_k = \lambda\mathbf{v}_k$,

$\Rightarrow A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = \lambda(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k)$

$\Rightarrow$ The produced vectors through the G.-S. process are still eigenvectors for λ

ii. Since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are orthogonal to eigenvectors corresponding to other different eigenvalues (according to Theorem 7.9), $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ is also orthogonal to eigenvectors corresponding to other different eigenvalues.

(4) The composite of steps (2) and (3) produces an orthonormal set of n eigenvectors. Use these orthonormal and thus linearly independent eigenvectors as column vectors to form the matrix P.

i. According to Thm. 7.8, the matrix P is orthogonal

ii. Following the diagonalization process, $D = P^{-1}AP$ is diagonal

Therefore, the matrix A is orthogonally diagonalizable

Orthogonal matrix

- Ex 7: Determining whether a matrix is orthogonally diagonalizable

	Symmetric matrix	Orthogonally diagonalizable
$A_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	○	○
$A_2 = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 1 & 8 \\ -1 & 8 & 0 \end{bmatrix}$	×	×
$A_3 = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix}$	×	×
$A_4 = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$	○	○

Orthogonal matrix

❖ Example 9: Orthogonal diagonalization

Find an orthogonal matrix P that diagonalizes A.

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

Solution:

(1) $\left|\lambda I - A\right| = (\lambda-3)^2(\lambda+6)=0$

$\lambda_1=-6,\ \lambda_2=3$ (has a multiplicity of 2)

(2) $\lambda_1=-6,\ \mathbf{v}_1=(1,\ -2,\ 2)\ \Rightarrow\ \mathbf{u}_1=\dfrac{\mathbf{v}_1}{\left\|\mathbf{v}_1\right\|}=(\frac{1}{3},\ \frac{-2}{3},\ \frac{2}{3})$

(3) $\lambda_2=3,\ \mathbf{v}_2=(2,\ 1,\ 0),\ \mathbf{v}_3=(-2,\ 4,\ 5)$

orthogonal

Verify Theorem 7.9 that
$\mathbf{v}_1\cdot\mathbf{v}_2=\mathbf{v}_1\cdot\mathbf{v}_3=0$

Orthogonal matrix

- If $\mathbf{v}_2$ and $\mathbf{v}_3$ are not orthogonal, the Gram-Schmidt Process should be performed. Here we simply normalize $\mathbf{v}_2$ and $\mathbf{v}_3$ to find the corresponding unit vectors

$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\tfrac{2}{\sqrt{5}},\ \tfrac{1}{\sqrt{5}},\ 0\right), \quad \mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(\tfrac{-2}{3\sqrt{5}},\ \tfrac{4}{3\sqrt{5}},\ \tfrac{5}{3\sqrt{5}}\right)$$

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} \\ \frac{-2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix} \Rightarrow P^{-1}AP = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3$$

- Note that there are some calculation error in the solution of Example 9 in the text book

Applications of Eigenvalues and Eigenvectors

- The rotation for quadratic equation:

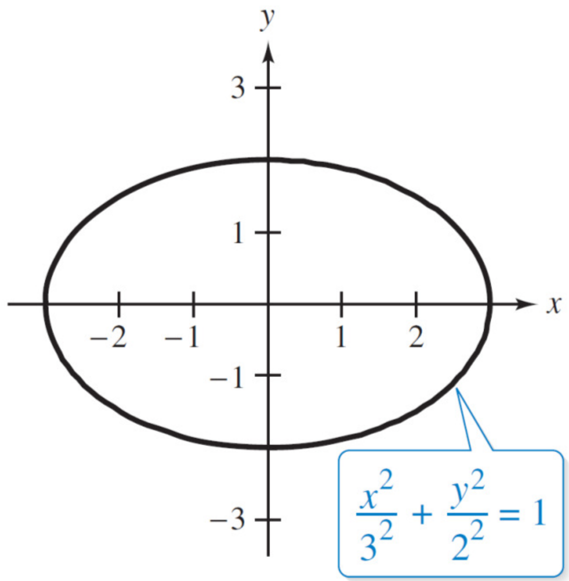
$$ax^2+bxy+cy^2+dx+ey+f=0$$

- Example 5: Identify the graphs of the following quadratic equations

(a) $4x^2+9y^2-36=0$ (b) $13x^2-10xy+13y^2-72=0$

solution

(a) In standard form, we can obtain $\frac{x^2}{3^2}+\frac{y^2}{2^2}=1.$

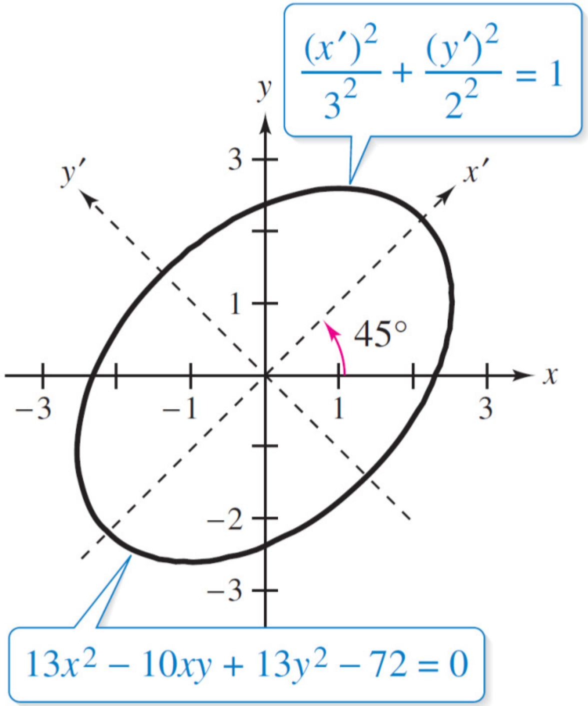


- Since there is no xy-term, it is easy to derive the standard form and it is apparent that this equation represents an ellipse.

Applications of Eigenvalues and Eigenvectors

(b) $13x^2 - 10xy + 13y^2 - 72 = 0$

❖ Since there is a xy-term, it is difficult to identify the graph of this equation. In fact, it is also an ellipse, which is oblique on the xy-plane.



❖ There is a easy way to identify the graph of quadratic equation. The basic idea is to rotate the x- and y-axes to x'- and y'-axes such that there is no more $x'y'$-term in the new quadratic equation.

❖ In the above example, if we rotate the x- and y-axes by 45 degree counterclockwise, the new quadratic equation $\frac{(x')^2}{3^2} + \frac{(y')^2}{2^2} = 1$ can be derived, which represents an ellipse apparently.

❖ In Section 4.8, the rotation of conics is achieved by changing basis, but here the diagonalization technique based on eigenvalues and eignvectors is applied to solving the rotation problem.

Applications of Eigenvalues and Eigenvectors

- **Quadratic form:**

$$ax^2 + bxy + cy^2$$

is the quadratic form associated with the quadratic equation

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

- **Matrix of the quadratic form:**

$$A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

- Note that A is a symmetric matrix.

If we define $X = \begin{bmatrix} x \\ y \end{bmatrix}$, then $X^T AX = ax^2 + bxy + cy^2$. In fact, the quadratic equation can be expressed in terms of X as follows:

$$X^T AX + [d \quad e]\, X + f = ax^2 + bxy + cy^{2+} + dx + ey + f = 0$$

Applications of Eigenvalues and Eigenvectors

Principal Axes Theorem

- For a conic whose equation is $ax^2 + bxy + cy^2 + dx + ey + f = 0$, the rotation to eliminate the xy-term is achieved by $X = PX'$, where P is an orthogonal matrix that diagonalizes A (matrix of the quadratic form). That is,

$$P^{-1}AP = P^T AP = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

where λ_1 and λ_2 are eigenvalues of A. The equation for the rotated conic is given by

$$\lambda_1 (x')^2 + \lambda_2 (y')^2 + \begin{bmatrix} d & e \end{bmatrix} PX' + f = 0.$$

Applications of Eigenvalues and Eigenvectors

Proof:

According to Theorem 7.10, since A is symmetric, we can conclude that there exists an orthogonal matrix P such that $P^{-1}AP = P^TAP = D$ is diagonal.

Replacing X with PX', the quadratic form becomes

$$\begin{aligned} X^T AX &= (PX')^T A(PX') = (X')^T P^T APX' \\ &= (X')^T DX' = \lambda_1 (x')^2 + \lambda_2 (y')^2. \end{aligned}$$

- It is obvious that the new quadratic form in terms of X' has no $x'y'$-term, and the coefficients for $(x')^2$ and $(y')^2$ are the two eigenvalues of the matrix A.
- $X = PX' \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} x' \\ y' \end{bmatrix} = x'\mathbf{v}_1 + y'\mathbf{v}_2 \Rightarrow$ Since $\begin{bmatrix} x \\ y \end{bmatrix}$ and $\begin{bmatrix} x' \\ y' \end{bmatrix}$ are the orignal and new coodinates, the roles of $\mathbf{v}_1$ and $\mathbf{v}_2$ (the eigenvectors of A) are like the basis vectors (or the axis vectors) in the new coordinate system.

Example 6: Rotation of a conic

- Perform a rotation of axes to eliminate the xy-term in the following quadratic equation

$$13x^2 - 10xy + 13y^2 - 72 = 0$$

Solution:

The matrix of the quadratic form associated with this equation is

$$A = \begin{bmatrix} 13 & -5 \\ -5 & 13 \end{bmatrix}$$

The eigenvalues are $\lambda_1 = 8$ and $\lambda_2 = 18$, and the corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Applications of Eigenvalues and Eigenvectors

After normalizing each eigenvector, we can obtain the orthogonal matrix *P* as follows.

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{bmatrix}$$

- According to the results on p. 268 in Ch4, $X=PX'$ is equivalent to rotate the xy-coordinates by 45 degree to form the new $x'y'$-coordinates.

Then by replacing *X* with *PX'*, the equation of the rotated conic is

$$8(x') + 18(y')^2 - 72 = 0,$$

which can be written in the standard form

$$\frac{(x')^2}{3^2} + \frac{(y')^2}{2^2} = 1.$$

- The above equation represents an ellipse on the $x'y'$-plane.

Applications of Eigenvalues and Eigenvectors

- In three-dimensional version:

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

is the quadratic form associated with the equation of quadric surface: $ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0.$

- Matrix of the quadratic form:

$$A = \begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix}$$

- Note that A is a symmetric matrix.

If we define $X = [x\ y\ z]^T$, then

$$X^TAX = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

and the quadratic surface equation can be expressed as

$$X^TAX + \begin{bmatrix} g & h & i \end{bmatrix} X + j = 0$$

Unitary and Hermitian Matrices

Conjugate Transpose of a Complex Matrix

- The **conjugate transpose** of a complex matrix A denoted by A^*, is given by:

$$A^* = \bar{A}^T$$

where the entries of $\bar{A}$ are the complex conjugates of the corresponding entries of A.

- Example: Determine A^* for the matrix

$$A = \begin{bmatrix} 3 + 7i & 0 \\ 2i & 4 - i \end{bmatrix}$$

Solution:

$$\bar{A} = \begin{bmatrix} \overline{3 + 7i} & \overline{0} \\ \overline{2i} & \overline{4 - i} \end{bmatrix} = \begin{bmatrix} 3 - 7i & 0 \\ -2i & 4 + i \end{bmatrix}$$

$$A^* = \bar{A}^T = \begin{bmatrix} 3 - 7i & -2i \\ 0 & 4 + i \end{bmatrix}$$

Unitary and Hermitian Matrices

- Theorem 8.8: Properties of the Conjugate Transpose
- If A and B are complex matrices and k is a complex number, then the following properties are true.
  1. $(A^*)^* = A$
  2. $(A^* + B^*)^* = A + B$
  3. $(kA^*)^* = \bar{k}A$
  4. $(AB)^* = B^* A^*$

Unitary and Hermitian Matrices

- Unitary Matrix: A complex matrix is **unitary** when

$$A^{-1} = A^*$$

- Example 2: Show that the matrix A is unitary.

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

Solution:

$$\begin{aligned} AA^* &= \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1-i & 1+i \\ 1+i & 1-i \end{bmatrix} \\ &= \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Unitary and Hermitian Matrices

- **Theorem 8.9: Unitary Matrices**
- An $n \times n$ complex matrix A is unitary if and only if its row (or column) vectors form an orthonormal set in C^n.

Unitary and Hermitian Matrices

- Hermitian Matrices: A square matrix A is **Hermitian** when

$$A = A^*$$

- Example 3: A is Hermitian,

$$A = \begin{bmatrix} a_1 & b_1 + b_2 i \\ b_1 - b_2 i & d_1 \end{bmatrix}$$

Unitary and Hermitian Matrices

Theorem 8.10: The Eigenvalues of a Hermitian Matrix

- If A is a Hermitian matrix, then its eigenvalues are real numbers.

Proof:

Let λ be an eigenvalue of A and let

$$\mathbf{v} = \begin{bmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ \vdots \\ a_n + b_n i \end{bmatrix}$$

be its corresponding eigenvector. If both sides of the equation $A\mathbf{v} = \lambda\mathbf{v}$ are multiplied by the row vector $\mathbf{v}^*$, then

$$\mathbf{v}^*A\mathbf{v} = \mathbf{v}^*(\lambda\mathbf{v}) = \lambda(\mathbf{v}^*\mathbf{v}) = \lambda(a_1^2 + b_1^2 + a_2^2 + b_2^2 + \cdots + a_n^2 + b_n^2).$$

Furthermore, because

$$(\mathbf{v}^*A\mathbf{v})^* = \mathbf{v}^*A^*(\mathbf{v}^*)^* = \mathbf{v}^*A\mathbf{v}$$

it follows that $\mathbf{v}^*A\mathbf{v}$ is a Hermitian 1×1 matrix. This implies that $\mathbf{v}^*A\mathbf{v}$ is a real number, so λ is real.

Unitary and Hermitian Matrices

- **Example 4:** Find the eigenvalues of the matrix A.

$$A = \begin{bmatrix} 3 & 2-i & -3i \\ 2+i & 0 & 1-i \\ 3i & 1+i & 0 \end{bmatrix}$$

Solution:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 3 & -2+i & 3i \\ -2-i & \lambda & -1+i \\ -3i & -1-i & \lambda \end{vmatrix}$$

$$= (\lambda - 3)(\lambda^2 - 2) - (-2+i)[(-2-i)\lambda - (3i+3)] + 3i[(1+3i) + 3\lambda i]$$

$$= (\lambda^3 - 3\lambda^2 - 2\lambda + 6) - (5\lambda + 9 + 3i) + (3i - 9 - 9\lambda)$$

$$= \lambda^3 - 3\lambda^2 - 16\lambda - 12$$

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

$\Rightarrow \lambda_1 = -1,\ \lambda_1 = -2,\ \lambda_1 = 6$

Unitary and Hermitian Matrices

- **Theorem 8.11: Hermitian Matrices and Diagonalization**
- If A is an $n \times n$ Hermitian matrix, then
  1. eigenvectors corresponding to distinct eigenvalues are orthogonal.
  2. A is unitarily diagonalizable.

proof:

To prove part 1, let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two eigenvectors corresponding to the distinct (and real) eigenvalues λ_1 and λ_2. Because $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, you have the equations shown below for the matrix product $(A\mathbf{v}_1)^*\mathbf{v}_2$.

$$(A\mathbf{v}_1)^*\mathbf{v}_2 = \mathbf{v}_1^*A^*\mathbf{v}_2 = \mathbf{v}_1^*A\mathbf{v}_2 = \mathbf{v}_1^*\lambda_2\mathbf{v}_2 = \lambda_2\mathbf{v}_1^*\mathbf{v}_2$$
$$(A\mathbf{v}_1)^*\mathbf{v}_2 = (\lambda_1\mathbf{v}_1)^*\mathbf{v}_2 = \mathbf{v}_1^*\lambda_1\mathbf{v}_2 = \lambda_1\mathbf{v}_1^*\mathbf{v}_2$$

So,

$$\lambda_2\mathbf{v}_1^*\mathbf{v}_2 - \lambda_1\mathbf{v}_1^*\mathbf{v}_2 = 0$$
$$(\lambda_2 - \lambda_1)\mathbf{v}_1^*\mathbf{v}_2 = 0$$
$$\mathbf{v}_1^*\mathbf{v}_2 = 0 \qquad \text{because } \lambda_1 \neq \lambda_2$$

and this shows that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. Part 2 of Theorem 8.11 is often called the **Spectral Theorem,** and its proof is left to you.

Unitary and Hermitian Matrices

- Example 5: Find a unitary matrix P such that P^*AP is a diagonal matrix where

$$A = \begin{bmatrix} 3 & 2-i & -3i \\ 2+i & 0 & 1-i \\ 3i & 1+i & 0 \end{bmatrix}$$

Solution: Eigenvalues of A are calculated in example 3. The normalized eigenvectors of A are:

$$\|\mathbf{v}_1\| = \|(-1, 1+2i, 1)\| = \sqrt{1+5+1} = \sqrt{7}$$
$$\|\mathbf{v}_2\| = \|(1-21i, 6-9i, 13)\| = \sqrt{442+117+169} = \sqrt{728}$$
$$\|\mathbf{v}_3\| = \|(1+3i, -2-i, 5)\| = \sqrt{10+5+25} = \sqrt{40}$$

$$P = \begin{bmatrix} -\frac{1}{\sqrt{7}} & \frac{1-21i}{\sqrt{728}} & \frac{1+3i}{\sqrt{40}} \\ \frac{1+2i}{\sqrt{7}} & \frac{6-9i}{\sqrt{728}} & \frac{-2-i}{\sqrt{40}} \\ \frac{1}{\sqrt{7}} & \frac{13}{\sqrt{728}} & \frac{5}{\sqrt{40}} \end{bmatrix} \Rightarrow P^*AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Symmetric and Hermitian Matrices

A is a symmetric matrix (real)	A is a Hermitian matrix (complex)
Eigenvalues of are real.	Eigenvalues of are real.
Eigenvectors corresponding to distinct eigenvalues are orthogonal.	Eigenvectors corresponding to distinct eigenvalues are orthogonal.
There exists an orthogonal matrix such that $P^T A P$ is diagonal.	There exists an unitary matrix such that $P^* A P$ is diagonal.