4.2 Diagonalization of a matrix

這一節我們要看看矩陣的角類似的問題

Theorem 11: Suppose $A \in M_{n\times n}(\mathbb{F})$ has n linearly independent eigenvectors $v_1, v_2, \cdots, v_n$. Let

$$Q = (v_1 \ v_2 \ \cdots \ v_n)$$

Then $Q^{-1}AQ$ is a diagonal matrix D, i.e.,

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

where λ_i: eigenvalue of A corresponding to v_i.

Definition 12: Q: eigenvector matrix.

D: eigenvalue matrix.

記憶方法: $A = QDQ^{-1}$

or $AQ = QD$

利用 eigenvalue, eigenvector 的定義來記.

proof of Theorem 11:

$$AQ = A(v_1 \ v_2 \ \cdots \ v_n) = (Av_1 \ Av_2 \ \cdots \ Av_n)$$
$$= (\lambda_1 v_1 \ \lambda_2 v_2 \ \cdots \ \lambda_n v_n)$$

$$QD = (v_1 \ v_2 \ \cdots \ v_n)\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = (\lambda_1 v_1 \ \lambda_2 v_2 \ \cdots \ \lambda_n v_n)$$

$\Rightarrow AQ = QD$

$\Rightarrow D = Q^{-1}AQ$ or $A = QDQ^{-1}$.

Example 13: (1) $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$.

由 Example 9(1) 可得：A has two eigenvalues $\lambda_1 = 3$, $\lambda_2 = -1$.

and the corresponding eigenvectors $a\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $b\begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $a, b \neq 0$.

Let

$$Q = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

(注意对应的顺序)

then

$$Q^{-1} = \frac{1}{-4}\begin{pmatrix} -2 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/4 \\ 1/2 & -1/4 \end{pmatrix},$$

and

$$Q^{-1}AQ = \begin{pmatrix} 1/2 & 1/4 \\ 1/2 & -1/4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 3/2 & 3/4 \\ -1/2 & 1/4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = D.$$

(2) $A = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$.

By Example 10(2), A has two eigenvalues $i, -i$,

and the corresponding eigenvectors $a\begin{pmatrix} 1 \\ i-1 \end{pmatrix}$, $b\begin{pmatrix} 1 \\ -1-i \end{pmatrix}$, $a, b \neq 0$

then

$$Q = \begin{pmatrix} 1 & 1 \\ -1+i & -1-i \end{pmatrix}, \quad D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

或

$$Q = \begin{pmatrix} 1 & 1 \\ -1-i & -1+i \end{pmatrix}, \quad D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

Remark 14: (1) Q is not unique (since eigenvectors are not unique)

(2) Not all matrices are diagonalizable. e.g.,

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

is not diagonalizable, 原因稍後再講.

Theorem 15: $A \in M_{n\times n}(\mathbb{F})$. $\lambda_1, \lambda_2, \cdots, \lambda_k$: distinct eigenvalues of A.

$v_1, v_2, \cdots, v_k$: eigenvectors of A corresponding to $\lambda_1, \lambda_2, \cdots, \lambda_k$, respectively.

$\Longrightarrow \{v_1, v_2, \cdots, v_k\}$: linearly independent.

不同的 eigenvalues 對應的 eigenvectors 為 linearly independent.

proof: 利用 mathematical induction 証明.

$k=1$: OK.

Suppose it holds for $k-1$

(k−1 個相異的 eigenvalues 對應的 eigenvectors 為 linearly indep.)

For the case of k distinct eigenvalues.

Suppose

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0. \qquad (*)$$

$$\begin{aligned}
\Longrightarrow 0 &= (A-\lambda_k I)(a_1 v_1 + a_2 v_2 + \cdots + a_k v_k) \\
&= a_1(A-\lambda_k I)v_1 + a_2(A-\lambda_k I)v_2 + \cdots + a_k(A-\lambda_k I)v_k \\
&= a_1(Av_1 - \lambda_k I v_1) + a_2(Av_2 - \lambda_k I v_2) + \cdots + a_k(Av_k - \lambda_k I v_k) \\
&= a_1(\lambda_1-\lambda_k)v_1 + a_2(\lambda_2-\lambda_k)v_2 + \cdots + a_{k-1}(\lambda_{k-1}-\lambda_k)v_k
\end{aligned}$$

$\underset{k-1\text{ 的 case}}{\Longrightarrow}$ $a_1(\lambda_1-\lambda_k) = a_2(\lambda_2-\lambda_k) = \cdots = a_{k-1}(\lambda_{k-1}-\lambda_k) = 0$

$\underset{\text{distinct eigenvalues}}{\Longrightarrow}$ $a_1 = a_2 = \cdots = a_{k-1} = 0.$

$\underset{(*)}{\Longrightarrow}$ $a_k = 0$

$\Longrightarrow \{v_1, v_2, \cdots, v_k\}$: linearly independent.

Corollary 16: $A \in M_{n\times n}(\mathbb{F})$ has exactly n distinct eigenvalues

$\Longrightarrow$ A is diagonalizable.

Example 17: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$.

the characteristic polynomial of $A = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$

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

$\Longrightarrow$ A has eigenvalues $0, 2$

$\Longrightarrow$ A is diagonalizable.

Definition 18: A polynomial $f(x)$ splits over $\mathbb{F}$ if $\exists\ c, a_1, a_2, \cdots, a_n \in \mathbb{F}$ (not necessarily distinct) such that

$$f(x) = c(x-a_1)(x-a_2)\cdots(x-a_n).$$

Example 19: $x^2 - 1 = (x+1)(x-1)$: split over $\mathbb{R}$

$(x^2+1)(x-2)$: does not split over $\mathbb{R}$.

Theorem 20: A: diagonalizable

$\Longrightarrow$ the characteristic polynomial of A splits.

proof: A : diagonalizable

$\Longrightarrow$ $\exists$ invertible matrix Q and diagonal matrix D s.t.

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

$\Longrightarrow$ $f(\lambda) = \det(A - \lambda I) = \det(QDQ^{-1} - \lambda I)$

$= \det(QDQ^{-1} - Q(\lambda I)Q^{-1})$

$= \det(Q(D-\lambda I)Q^{-1}) = \det(Q)\det(D-\lambda I)\det(Q^{-1})$

$= \det(Q)\det(D-\lambda I)\cdot \dfrac{1}{\det(Q)}$

$= \det(D - \lambda I)$

$$= \begin{vmatrix} \lambda_1 - t & & & 0 \\ & \lambda_2 - t & & \\ & & \ddots & \\ 0 & & & \lambda_n - t \end{vmatrix}$$

$$= (-1)^n (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n).$$

Definition 21: λ: eigenvalue of a matrix A with characteristic polynomial $f(t)$.
The (algebraic) multiplicity of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of $f(t)$.

Example 22: $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}$

$\Rightarrow$ the characteristic polynomial $f(t) = -(t-3)^2(t-4)$

$\Rightarrow$ $\lambda = 3$: eigenvalue of A with multiplicity 2.

$\lambda = 4$: eigenvalue of A with multiplicity 1.

Definition 23: λ: eigenvalue of $A \in M_{n\times n}(\mathbb{F})$.
Define

$$E_\lambda = \{x \in \mathbb{F}^n : Ax = \lambda x\} = N(A - \lambda I_n)$$

The set E_λ is called the eigenspace (固有空間) of A corresponding to the eigenvalue λ.

Remark 24: $E_\lambda = \{0\} \cup \{$ eigenvectors of A corresponding to $\lambda\}$.
$\dim(E_\lambda)$ = the maximal number of linearly independent eigenvectors of A corresponding to λ.

Example 25: As in Example 9(2).

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -3 \end{pmatrix}$$

$\lambda = 0$: eigenvalue with multiplicity 2

$$E_0 = \left\{ a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} : a \in \mathbb{F} \right\}$$

$\dim E_0 = 1$

$\lambda = -3$: eigenvalue with multiplicity 1

$$E_{-3} = \left\{ a \begin{bmatrix} -5 \\ -12 \\ 9 \end{bmatrix} : a \in \mathbb{F} \right\}$$

$\dim E_{-3} = 1.$

Theorem 26: λ : eigenvalue of A ($\in M_{n\times n}(\mathbb{F})$) with multiplicity m

$\Longrightarrow \quad 1 \le \dim(E_\lambda) \le m.$

Theorem 27: $A \in M_{n\times n}(\mathbb{F})$

A : diagonalizable $\Longleftrightarrow$ the following two conditions hold:

(1) the characteristic polynomial splits.

(2) for each eigenvalue λ of A.

the multiplicity of $\lambda = \dim(E_\lambda) = \text{nullity}(A - \lambda I_n)$

$= n - \text{rank}(A - \lambda I_n).$

(好處: "if and only if," 不用計算所有的 eigenvectors)

Example 28: (1) Let

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \in M_{3\times 3}(\mathbb{R})$$

the characteristic polynomial of $A = -(t-3)^2(t-4)$

multiplicity of $\lambda = 3$: 2.

$$\dim(E_3) = \text{nullity}(A-3I) = 3 - \text{rank}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 3-2 = 1$$

$\neq$ multiplicity of $\lambda = 3 = 2$

$\Rightarrow$ A is not diagonalizable.

(2) Let

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

the characteristic polynomial of $A = \begin{vmatrix} -\lambda & -2 & -3 \\ 1 & 3-\lambda & 3 \\ 0 & 0 & 1-\lambda \end{vmatrix}$

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

$\lambda_1 = 1$: multiplicity $= 2$.

$$\dim(E_1) = 3 - \text{rank}(A-I) = 3 - \text{rank}\begin{bmatrix} -1 & -2 & -3 \\ 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

$= 3 - 1 = 2. =$ multiplicity of λ_1

$\lambda_2 = 2$: multiplicity $= 1$.

$\dim(E_2) = 1$ due to Theorem 26. $=$ multiplicity of λ_2

(multiplicity = 1 的都不用檢驗)

$\Rightarrow$ A is diagonalizable.