Matrix Computation

Yu-Hao Liang

September 2024

Contents

Chapter 1

基本矩陣分析

1.1 符號、基本定義與性質

Denote

$$A \in \mathbb{K}^{m\times n},\ \text{where } \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C} \Longleftrightarrow A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{K}.$$

Then we define

(i) **Product of Matrices** [矩陣乘積] ($\mathbb{K}^{m\times n} \times \mathbb{K}^{n\times p} \to \mathbb{K}^{m\times p}$): Let $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{n\times p}$.

$$AB = C = [c_{ij}] \in \mathbb{K}^{m\times p} \quad \Longrightarrow \quad c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad \forall\, i = 1,\dots,m,\ j = 1,\dots,p$$

(ii) **Transpose** [轉置] ($\mathbb{R}^{m\times n} \to \mathbb{R}^{n\times m}$):

$$C = A^T,\ \text{ where } \ c_{ij} = a_{ji} \in \mathbb{R}$$

(iii) **Conjugate transpose** [共軛轉置] ($\mathbb{C}^{m\times n} \to \mathbb{C}^{n\times m}$):

$$C = A^* \text{ or } C = A^H.\ \text{ where } \ c_{ij} = \overline{a_{ji}} \in \mathbb{C}$$

(iv) **Differentiation** [微分] ($\mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$):

$$C(t) = [c_{ij}(t)] \quad \Longrightarrow \quad \dot{C}(t) = [\dot{c}_{ij}(t)]$$

(v) **Inverse** [反矩陣] ($\mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}$): If $A, B \in \mathbb{K}^{n\times n}$ satisfy $AB = I$, then B is called the **inverse** of A and is denoted by A^{-1}.

- If A^{-1} exists, then A is said to be **nonsingular** [非奇異的] or **invertible** [可逆的]; Otherwise, A is said to be **singular** [奇異的] or **not invertible** [不可逆的].
- It can be proved that A is nonsingular $\Longleftrightarrow \det(A) \neq 0$.

(vi) **Inner product** [內積]: The inner product of $x, y \in \mathbb{K}^n$ is defined by

$$\langle x, y\rangle := x^T y = \sum_{i=1}^{n} x_i y_i = y^T x \in \mathbb{R},$$

$$\langle x, y\rangle := x^* y = \sum_{i=1}^{n} \overline{x_i} y_i = y^* x \in \mathbb{C}.$$

- $\langle x, y\rangle = \|x\|_2 \cdot \|y\|_2 \cos(\theta)$, where θ is the angle between x and y.

(vii) **Some basis operations**:

- Let $A \in \mathbb{K}^{m\times n}$ and $x \in \mathbb{K}^n$. Then
  - Form 1:
  $$y = Ax \quad\Longrightarrow\quad y = [y_i], \text{ where } y_i = \sum_{j=1}^{n} a_{ij}x_j,\ i = 1, \ldots, m.$$
  - Form 2:
  $$y = Ax = \begin{bmatrix} a_1 \mid a_2 \mid \cdots \mid a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ v_n \end{bmatrix} \quad\Longrightarrow\quad y = x_1a_1 + x_2v_2 + \cdots + x_na_n. \tag{1.1.1}$$
  - Form 3:
  $$y = Ax = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} x \quad\Longrightarrow\quad y = \begin{bmatrix} \langle r_1, x\rangle \\ \langle r_2, x\rangle \\ \vdots \\ \langle r_m, x\rangle \end{bmatrix}. \tag{1.1.2}$$
- Let $x \in \mathbb{K}^m$ and $y \in \mathbb{K}^n$. Then
$$xy^* = \begin{bmatrix} x_1\overline{y_1} & \cdots & x_1\overline{y_n} \\ \vdots & \ddots & \vdots \\ x_m\overline{y_1} & \cdots & x_m\overline{y_n} \end{bmatrix} \in \mathbb{K}^{m\times n}. \tag{1.1.3}$$
- Let $A \in \mathbb{K}^{m\times p}$ and $B \in \mathbb{K}^{p\times n}$. Then
  - Form 1:
  $$C = AB \quad\Longrightarrow\quad C = [c_{ij}], \text{ where } c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj},\ i = 1, \ldots, m,\ j = 1, \ldots, n.$$
  - Form 2:
  $$C = AB = A\begin{bmatrix} b_1 \mid b_2 \mid \cdots \mid b_n \end{bmatrix} \quad\Longrightarrow\quad C = \begin{bmatrix} Ab_1 \mid Ab_2 \mid \cdots \mid Ab_n \end{bmatrix}. \tag{1.1.4}$$
  - Form 3:
  $$C = AB = \begin{bmatrix} c_1 \mid c_2 \mid \cdots \mid c_p \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix} \quad\Longrightarrow\quad C = c_1r_1 + c_2r_2 + \cdots + c_pr_p. \tag{1.1.5}$$

1.1.1 擾動下的反矩陣

Theorem 1.1.1 (Sherman-Morrison Formula)**.** Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. Then, for any $u, v \in \mathbb{R}^n$, if $v^TA^{-1}u + 1 \neq 0$, then
$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}. \tag{1.1.6}$$

Proof.

$$
\begin{aligned}
&(A+uv^T)[A^{-1}-A^{-1}u(1+v^TA^{-1}u)^{-1}v^TA^{-1}] \\
&= I + \frac{1}{1+v^TA^{-1}u}[uv^TA^{-1}(1+v^TA^{-1}u) - uv^TA^{-1} - uv^TA^{-1}uv^TA^{-1}] \\
&= I + \frac{1}{1+v^TA^{-1}u}[u(v^TA^{-1}u)v^TA^{-1} - uv^TA^{-1}uv^TA^{-1}] = I.
\end{aligned}
$$

□

Theorem 1.1.2 (Sherman-Morrison-Woodbury Formula)**.** Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. Then, for any $U, V \in \mathbb{R}^{n\times k}$, if $(I + V^TA^{-1}U)$ is invertible, then

$$
(A+UV^T)^{-1} = A^{-1} - A^{-1}U(I+V^TA^{-1}U)^{-1}V^TA^{-1}. \tag{1.1.7}
$$

Example 1.1.1.

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

1.1.2 秩與正交

Let $A \in \mathbb{R}^{m\times n}$. Consider the mapping

$$
\begin{aligned}
T : \mathbb{R}^n &\to \mathbb{R}^m \\
x &\mapsto y = Ax
\end{aligned}
$$

Definition 1.1.1 (Range, Null Space, Rank, Nullity)**.** Let $A \in \mathbb{R}^{m\times n}$. Then we define

(i) The **range space** [值域] of A by

$$
R(A) = \{y \in \mathbb{R}^m \mid y = Ax \text{ for some } x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m. \tag{1.1.8}
$$

- $R(A)$ is a subspace in $\mathbb{R}^m$.

(ii) The **null space/kernel** [零空間/零核] of A by

$$
N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} \subseteq \mathbb{R}^n. \tag{1.1.9}
$$

- $N(A)$ is the subspace in $\mathbb{R}^n$.

(iii) The **rank** [秩] of A by

$$
\operatorname{rank}(A) = \dim[R(A)] = \text{The number of maximal linearly independent columns of } A \tag{1.1.10}
$$

(iv) The **nullity** of A by

$$
\operatorname{nullity}(A) = \dim[N(A)] = \text{The number of maximal linearly independent vectors of } N(A). \tag{1.1.11}
$$

Theorem 1.1.3 (Dimension Theorem)**.** For any $A \in \mathbb{R}^{m\times n}$,

$$
\operatorname{nullity}(A) + \operatorname{rank}(A) = n \text{ (column number)}. \tag{1.1.12}
$$

Theorem 1.1.4 (Equivalence of Nonsingular)**.** Let $A \in \mathbb{R}^{n\times n}$. Then the following are equivalent.

(i) A is nonsingular.

(ii) $Ax = 0$ has only the solution $x = 0$.

(iii) $Ax = b$ has a unique solution, for any $b \in \mathbb{R}^n$.

(iv) $N(A) = \{0\}$.

(v) $R(A) = \mathbb{R}^n$.

(vi) $\operatorname{rank}(A) = n$.

(vii) $\operatorname{nullity}(A) = 0$.

(viii) $\det(A) \neq 0$.

Definition 1.1.2 (Orthogonal, Orthonormal, Orthogonal Complement)**.** Let set $S = \{x_1, \ldots, x_p\} \subseteq \mathbb{R}^n$. Then

(i) S is said to be **orthogonal** [正交] if $x_i^T x_j = 0$, $\forall\, i \neq j$.

(ii) S is said to be **orthonormal** [單範正交] if $x_i^T x_j = 0$, $\forall\, i \neq j$, and $x_i^T x_i = \|x_i\|_2 = 1$, $\forall\, i = 1, \ldots, p$.

(iii) The **orthogonal complement** [正交補] of a S is defined by

$$S^\perp = \{y \in \mathbb{R}^n \mid y^T x = 0, \text{ for } x \in S\} = \text{orthogonal complement of } S. \tag{1.1.13}$$

- $S^\perp$ is a subspace in $\mathbb{R}^n$.

Theorem 1.1.5 (Relation between Range and Null Spaces of A and A^T)**.** For any $A \in \mathbb{R}^{m\times n}$,

(i) $R(A^T) \perp N(A)$ and $R(A) \perp N(A^T)$.

(ii) $\operatorname{rank}(A) = \operatorname{rank}(A^T)$.

Other Definitions:

	$A \in \mathbb{R}^{n\times n}$	$A \in \mathbb{C}^{n\times n}$
對稱性	**Symmetric**: $A^T = A$	**Hermitian**: $A^* = A (A^H = A)$
反對稱性	**Skew-symmetric**: $A^T = -A$	**Skew-Hermitian**: $A^* = -A$
Positive definite [正定]	$x^T Ax > 0$, $x \neq 0$	$x^* Ax > 0$, $x \neq 0$
Non-negative definite [非負定]	$x^T Ax \geq 0$	$x^* Ax \geq 0$
Indefinite [不定]	$(x^T Ax)(y^T Ay) < 0$ for some x, y	$(x^* Ax)(y^* Ay) < 0$ for some x, y
Normal [正規]	$A^* A = AA^*$	
正交性	**Orthogonal** [正交]: $A^T A = AA^T = I$	**Unitary** [么正]: $A^* A = AA^* = I$

Example 1.1.2. If $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ is skew-symmetric, then show that

(i) $a_{ii} = 0$ for all $i = 1, \ldots, n$.

(ii) $x^T Ax = 0$ for all $x \in \mathbb{R}^n$.

Example 1.1.3 (Cayley transformation)**.** If A is skew-symmetric, then $I - A$ is nonsingular and $(I - A)^{-1}(I + A)$ is orthogonal (**Cayley transformation** of A).

Proof. Hint: A matrix A is singular $\Longrightarrow Av = 0$ for some $v \neq 0$. □

1.1.3 特殊矩陣

Definition 1.1.3. Let $A \in \mathbb{R}^{m\times n}$. Then we say the matrix A is

(i) **diagonal** [對角的] if $a_{ij} = 0$ for all $i \neq j$. Sometimes, to save the notations, we denote

$$D = \begin{bmatrix} d_1 & 0 & \cdots & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & d_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & d_n \end{bmatrix} = \operatorname{diag}(d_1, \ldots, d_n)$$

(ii) **(strictly) upper triangular** [(嚴格) 上三角] if $a_{ij} = 0$ for $i > j$ (or $i \geq j$);

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

(iii) **(strictly) lower triangular** [(嚴格) 下三角] if $a_{ij} = 0$ for $i < j$ (or $i \leq j$);

$$\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{bmatrix}$$

(iv) **tridiagonal** [三對角的] if $a_{ij} = 0$ for $|i - j| > 1$.

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix}$$

(v) **upper bi-diagonal** [上雙對角的] if $a_{ij} = 0$ for $i > j$ or $j > i + 1$.

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ 0 & a_{22} & a_{23} & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

(vi) **upper Hessenberg** [上海森伯格] if $a_{ij} = 0$ for $i > j + 1$. (Note: the lower case is the same as above.)

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix}$$

(vii) **sparse** [稀疏矩陣] if at most n^{1+r} elements of $A \in \mathbb{R}^{m\times n}$ is nonzero, where $r < 1$ (usually between 0.2 and 0.5).

- For example, If $n = 1000, r = 0.9$, then $n^{1+r} = 501,187$.

1.1.4 特徵值特徵、向量

Definition 1.1.4 (Eigenvalues and Eigenvectors)**.** Let $A \in \mathbb{C}^{n\times n}$. If

$$Ax = \lambda x, \tag{1.1.14}$$

where $\lambda \in \mathbb{C}$ and $x \neq 0$. Then λ is called an **eigenvalue** [特徵值] of A and x is called its corresponding **eigenvector** [特徵向量]. We also call (λ, x) be an **eigenpair** [特徵對] of A.

Definition 1.1.5 (Spectrum and Spectrum Radius)**.** Let $A \in \mathbb{C}^{n\times n}$. Then we define

(i) $\sigma(A) :=$ Spectrum (譜) of $A =$ The set of eigenvalues of A.

(ii) $\rho(A) :=$ Sprctral Radius (譜半徑) of $A = \max\{|\lambda| : \lambda \in \sigma(A)\}$.

Process 1 (To find (Exact) Eigenvalues and Eigenvectors of a Matrix)**.** Let $A \in \mathbb{C}^{n\times n}$.

(i) Note that $\lambda \in \sigma(A) \Leftrightarrow \det(A - \lambda I) = 0$.

(ii) Define the **characteristic polynomial** [特徵多項式] of A by $p(\lambda) = \det(\lambda I - A)$.

(iii) Factor

$$p(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{m(\lambda_i)} \tag{1.1.15}$$

where $\lambda_i \neq \lambda_j$ (for $i \neq j$) and $\sum\limits_{i=1}^{s} m(\lambda_i) = n$.

(iv) Then the set of **eigenvalues** of A is $\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_s\}$.

(v) Note that $Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$.

(vi) For each eigenvalue $\lambda_i, i = 1, \ldots, s$, the set E that collects all its corresponding **eigenvectors** is

$$E = \{v \neq 0 \mid (A - \lambda_i I)v = 0\} = \text{null}(A - \lambda_i I) - \{0\}. \tag{1.1.16}$$

- E is also called the eigenspace [特徵空間] of λ_i.

Definition 1.1.6 (Algebraic Multiplicity and Geometric Multiplicity)**.** Let $A \in \mathbb{C}^{n\times n}$ and suppose the characteristic polynomial $p(\lambda)$ of A is

$$p(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{m(\lambda_i)}.$$

Then, for each eigenvalue of $\lambda_i, i = 1, \ldots, s$, we say

(i) $m(\lambda_i)$ is the **algebraic multiplicity** [代數重數] of λ_i.

(ii) $n(\lambda_i) := \text{nullity}(A - \lambda_i I)$ is the **geometric multiplicity** [幾何重數] of λ_i.

Theorem 1.1.6 (Relation between Algebraic and Geometric Multiplicities)**.** Let $A \in \mathbb{C}^{n\times n}$. For each eigenvalue λ_i of A, its algebraic multiplicity $m(\lambda_i)$ and geometric multiplicity $n(\lambda_i)$ satisfy

$$1 \leq n(\lambda_i) \leq m(\lambda_i). \tag{1.1.17}$$

Corollary 1.1.7. Let $A \in \mathbb{C}^{n\times n}$. Then A has n **linearly independent** eigenvectors if and only if

$$n(\lambda_i) = m(\lambda_i) \tag{1.1.18}$$

for all eigenvalue λ_i, $i = 1, \ldots, s$, of A.

1.1.5 對角化

Definition 1.1.7 (Diagonalizable)**.** Let $A \in \mathbb{C}^{n\times n}$. We say A is **diagonalizable** [可對角化的] if there is an **invertible matrix** $V \in \mathbb{C}^{n\times n}$ and **diagonal matrix** $D \in \mathbb{C}^{n\times n}$ such that

$$A = VDV^{-1} \ (\Longleftrightarrow V^{-1}AV = D \Longleftrightarrow AV = VD, D: \text{ invertible}). \tag{1.1.19}$$

- Note that, if we let $D = \mathrm{diag}\left(\begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}\right)$ and $V = \begin{bmatrix} v_1 \mid v_2 \mid \cdots \mid v_n \end{bmatrix}$, then

$$\begin{aligned} A = VDV^{-1} &\Leftrightarrow AV = VD,\ V\text{: invertible} \\ &\Leftrightarrow A\begin{bmatrix} v_1 \mid v_2 \mid \cdots \mid v_n \end{bmatrix} = \begin{bmatrix} v_1 \mid v_2 \mid \cdots \mid v_n \end{bmatrix} \cdot \mathrm{diag}(\begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}),\ V\text{: invertible} \\ &\Leftrightarrow \begin{bmatrix} Av_1 \mid Av_2 \mid \cdots \mid Av_n \end{bmatrix} = \begin{bmatrix} d_1v_1 \mid d_2v_2 \mid \cdots \mid d_nv_n \end{bmatrix},\ V\text{: invertible} \\ &\Leftrightarrow Av_1 = dv_1, Av_2 = dv_2 \cdots, Av_n = dv_n,\ \{v_1, v_2, \cdots, v_n\}\text{: independent} \\ &\Leftrightarrow Av_1 = dv_1, Av_2 = dv_2 \cdots, Av_n = dv_n,\ \{v_1, v_2, \cdots, v_n\}\text{: independent} \\ &\Leftrightarrow (d_i, v_i), i = 1, \ldots, n, \text{eigenpair of } A, \text{where}\{v_1, v_2, \cdots, v_n\}\text{: independent.} \end{aligned}$$

Theorem 1.1.8 (Conditions for Diagonalizable)**.** Let $A \in \mathbb{C}^{n\times n}$. The following statements are equivalent:

(i) A is diagonalizable.

(ii) There is an **invertible matrix** $V \in \mathbb{C}^{n\times n}$ and **diagonal matrix** $D \in \mathbb{C}^{n\times n}$ such that

$$A = VDV^{-1} \ (\Longleftrightarrow V^{-1}AV = D \Longleftrightarrow AV = VD, D: \text{ invertible}).$$

(iii) For all eigenvalues λ_i, $i = 1, \ldots, s$, of A, $n(\lambda_i) = m(\lambda_i)$.

Remark 1. We say a matrix $A \in \mathbb{C}^{n\times n}$ is **degenerated** [退化的] or **non-diagonalizable** [不可對角化的] if it is not diagonalizable.

Definition 1.1.8 (Diagonalization)**.** Let $A \in \mathbb{C}^{n\times n}$. Then the process to express $A = VDV^{-1}$ for some invertible V and diagonal matrix D is called the **diagonalization** [對角化] of A.

Process 2 (To find Diagonalization of a Matrix)**.** Let $A \in \mathbb{C}^{n\times n}$.

(i) Find the **characteristic polynomial** [特徵多項式] of A: $p(\lambda) = \det(\lambda I - A)$.

(ii) Find all **eigenvalues** $\lambda_i, i = 1, \ldots, s$, of A by factoring $p(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{m(\lambda_i)}$.

(iii) For each eigenvalue $\lambda_i, i = 1, \ldots, s$, find $n(\lambda_i)$'s **linearly independent eigenvectors** by finding

$$\begin{aligned} \{v \neq 0 \mid (A - \lambda_i I)v = 0\} &= \mathrm{null}(A - \lambda_i I) - \{0\} \\ &= \mathrm{span}(v_1^{(i)}, v_2^{(i)}, \cdots, v_{n(\lambda_i)}^{(i)}) - \{0\}. \end{aligned}$$

(iv) If there is some eigenvalue λ_i of A such that $n(\lambda_i) < m(\lambda_i)$, then A can Not be **diagonalized**. Otherwise, by letting

$$\begin{aligned} V &= \begin{bmatrix} v_1^{(i)}, \cdots, v_{n(\lambda_i)}^{(i)} \mid \ldots \mid v_1^{(s)}, \cdots, v_{n(\lambda_s)}^{(s)} \end{bmatrix} \\ D &= \begin{bmatrix} \lambda_1, \cdots, \lambda_1 \mid \cdots \mid \lambda_s, \cdots, \lambda_s \end{bmatrix}. \end{aligned}$$

Then $A = VDV^{-1}$.

Remark 2. Only diagonalizable matrix can be derived the diagonalization $A = VDV^{-1}$, where $V \in \mathbb{C}^{n\times n}$ is **invertible** and $D \in \mathbb{C}^{n\times n}$ is **diagonal**.

1.1.6 Schur 分解

Theorem 1.1.9 (Schur Lemma; Schur Decomposition)**.** For any $A \in \mathbb{C}^{n\times n}$, there is an **unitary matrix** $U \in \mathbb{C}^{n\times n}$ and **upper triangular matrix** $R \in \mathbb{C}^{n\times n}$ such that

$$A = URU^*. \tag{1.1.20}$$

- Any **unitary matrix** U satisfies $UU^* = U^*U = I$. Equivalently, $U^{-1} = U^*$.
- The diagonal elements of R are just eigenvalues of A.

Theorem 1.1.10 (Unitary/Orthogonal Diagonalizability)**.**

(i) Let $A \in \mathbb{C}^{n\times n}$. Then A is **normal** (i.e., $AA^* = A^*A$) $\Longleftrightarrow$

$$A = UDU^*, \tag{1.1.21}$$

for some **unitary matrix** $U \in \mathbb{C}^{n\times n}$ and **complex** diagonal matrix $D \in \mathbb{C}^{n\times n}$.

- **Normal matrix** has n **orthonormal** eigenvectors $\{v_1, v_2, \cdots, v_n\} \subseteq \mathbb{C}^n$.

(ii) Let $A \in \mathbb{C}^{n\times n}$. Then A is **Hermitian** (i.e., $A^* = A$) $\Longleftrightarrow$

$$A = UDU^*, \tag{1.1.22}$$

for some **unitary matrix** $U \in \mathbb{C}^{n\times n}$ and **real** diagonal matrix $D \in \mathbb{R}^{n\times n}$.

- **Hermitian matrix** has n **orthonormal** eigenvectors $\{v_1, v_2, \cdots, v_n\} \subseteq \mathbb{C}^n$. And all their eigenvalues are real.

(iii) Let $A \in \mathbb{R}^{n\times n}$. Then A is **symmetric** (i.e., $A^T = A$) $\Longleftrightarrow$

$$A = QDQ^T, \tag{1.1.23}$$

for some **orthogonal matrix** $Q \in \mathbb{R}^{n\times n}$ and **real** diagonal matrix $D \in \mathbb{R}^{n\times n}$.

- **Symmetric matrix** has n **orthonormal** eigenvectors $\{v_1, v_2, \cdots, v_n\} \subseteq \mathbb{R}^n$. And all their eigenvalues are real.

1.1.7 Jordan 分解

$A \in \mathbb{C}^{n\times n}$ is diagonalizable $\Longleftrightarrow A = VDV^{-1}$.

Theorem 1.1.11. For any **(degenerated)** $A \in \mathbb{C}^{n\times n}$, there are **invertible** $V \in \mathbb{C}^{n\times n}$ and **Jordan matrix** $J =\in \mathbb{C}^{n\times n}$ such that

$$A = VJV^{-1}, \tag{1.1.24}$$

which is called the **Jordan decomposition** [Jordan 分解] of A.

- Jordan matrix $J = \operatorname{diag}(J_1, J_2, \cdots, J_k)$, where

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \text{ or } \begin{bmatrix} \lambda_i & 0 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 0 \\ & & & \lambda_i \end{bmatrix}. \tag{1.1.25}$$

1.1.8 實對角化/實 Schur 分解/實 Jordan 分解

Theorem 1.1.12 (Real Diagonalization; Real Schur Decomposition; Real Jordan Decomposition)**.**

(i) For any diagonalizable $A \in \mathbb{R}^{n\times n}$, there is an **invertible matrix** $V \in \mathbb{R}^{n\times n}$ and **block diagonalizable matrx** [塊對角] $D_B \in \mathbb{R}^{n\times n}$ such that

$$A = VD_BV^{-1}.$$

(ii) For any $A \in \mathbb{R}^{n\times n}$, there is an **orthogonal matrix** $U \in \mathbb{R}^{n\times n}$ and **quasi-upper triangular** [類上三角] $R_B \in \mathbb{R}^{n\times n}$ such that

$$A = UR_BU^*.$$

- An example of quasi-upper triangular:

$$\begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}.$$

(iii) For any $A \in \mathbb{R}^{n\times n}$, there is an **invertible matrix** $V \in \mathbb{R}^{n\times n}$ and **real (block) Jordan matrix** $J \in \mathbb{R}^{n\times n}$ such that

$$A = VJV^{-1}.$$

- An example of real (block) Jordan matrix:

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