Chapter 1 Fundamentals of Linear Algebra 從影像到矩陣表示

Cheng-Fang Su

(National Yang Ming Chiao Tung University)

May 23, 2025

從影像到矩陣表示

- 線性代數（Linear Algebra）是數學中最基本也最具廣泛應用性的課程之一，無論是科學計算、最佳化理論、工程應用、機器學習或量子科技，它都為這些領域提供了極其強大的數學工具。
- 在本門課中，第一章將從一個貼近現實世界的資料形態出發，也就是我們生活中隨處可見的影像，透過一些具體的例子，引導同學們逐步建構出矩陣、矩陣運算、向量結構與子空間等基本概念，為後續的廣泛應用打下良好的數學基礎。

從影像到矩陣表示

- 在現代資訊科學與工程技術中，影像（Iamge）已成為數據資料中極為重要的一種形式。舉凡醫學影像、衛星監測、人工智慧、電腦視覺，甚至社交媒體上的圖片分享，我們日常所接觸到的圖像訊息無不以數位影像的形式儲存與傳遞。
- 然而，要對影像進行處理、分析與理解，就必須先將其轉化為數學物件加以研究，而矩陣（Matrix）正是描述影像最直接也最自然的數學模型。
- 在我們開始處理影像之前，為了進行數學處理，首先需要一種能夠系統地儲存數值的方法；在線性代數中，這樣的表示方式就是「矩陣」。

矩陣的定義與符號

Definition (矩陣, Matrix)

令 $m, n \in \mathbb{N}$。一個 $m \times n$ 的實數矩陣是一個由 m 列 n 行的實數所構成的二維數表，記為

$$A = [a_{ij}] \in \mathbb{R}^{m \times n},$$

其中 $a_{ij} \in \mathbb{R}$ 表示矩陣 A 的第 i 列第 j 行的元素，$1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$。

舉例而言，考慮以下矩陣

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 4 & 3 & 5 \end{bmatrix} \in \mathbb{R}^{2 \times 3},$$

顯然矩陣 A 有 2 列 3 行，屬於 2×3 矩陣。矩陣的元素值分別為

$$a_{11} = 2, \quad a_{12} = -1, \quad a_{13} = 0, \quad a_{21} = 4, \quad a_{22} = 3, \quad a_{23} = 5.$$

數學符號複習

★ 數學符號複習：

(1) $\mathbb{N}$：自然數集、正整數集；

(2) $\mathbb{R}$：實數集；

(3) $\in$：屬於；

(4) $A \in \mathbb{R}^{2\times 3}$ 意即「A 為 2×3 的實數矩陣」；

(5) $a_{ij} \in \mathbb{R}$ 意即「a_{ij} 為一個實數」。

數位影像可表示為矩陣

- 有了矩陣的概念後，我們可以回到影像的數位結構來看：一張影像本質上就是一個由許多像素（Pixel）組成的數值網格，每一格的數值就代表其亮度或顏色資訊。
- 數位影像在本質上就是一個由像素組成的數值陣列，而每個像素值可以用一個實數來表示。因此，我們可以自然地將一張影像表示為矩陣：

(1) 一張灰階影像（Grayscale Image）以矩陣 $I \in \mathbb{R}^{m\times n}$ 表示，其中 I_{ij} 為像素亮度值（如 0–255），代表第 i 列第 j 行的像素。例如我們給定以下矩陣

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 4 & 3 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 3},$$

若將其視為一張灰階影像所對應的像素矩陣，每個元素可對應為一個像素的灰階值。通常灰階影像的像素值範圍為 $[0, 255]$，因此矩陣 A 可視為一張 2×3 像素的灰階影像。

數位影像可表示為矩陣

- 有了矩陣的概念後，我們可以回到影像的數位結構來看：一張影像本質上就是一個由許多像素（Pixel）組成的數值網格，每一格的數值就代表其亮度或顏色資訊。
- 數位影像在本質上就是一個由像素組成的數值陣列，而每個像素值可以用一個實數來表示。因此，我們可以自然地將一張影像表示為矩陣：

(2) 一張彩色影像（Color Image）通常表示為三個矩陣所構成的三維陣列，即

$$I \in \mathbb{R}^{m\times n\times 3}, \quad I = (R, G, B),$$

其中 $R, G, B \in \mathbb{R}^{m\times n}$ 分別代表紅色、綠色與藍色通道的亮度矩陣。這三個通道矩陣沿著第三維（通道維）堆疊，形成一個 $m \times n$ 的空間影像，在幾何上可以想像為一個具有三層的「矩陣長方體」或三維陣列（3D array）。

例題解說

考慮一張 2×3 的灰階影像，其對應矩陣如下：

$$I_{\text{gray}} = \begin{bmatrix} 100 & 150 & 200 \\ 50 & 75 & 125 \end{bmatrix},$$

以及一張 2×3 的彩色影像，其紅、綠、藍三個通道分別如下：

$$R = \begin{bmatrix} 255 & 128 & 64 \\ 0 & 0 & 32 \end{bmatrix}, \quad G = \begin{bmatrix} 100 & 150 & 100 \\ 90 & 100 & 80 \end{bmatrix}, \quad B = \begin{bmatrix} 20 & 30 & 40 \\ 50 & 60 & 70 \end{bmatrix}.$$

試問：

(1) 灰階影像與彩色影像的矩陣形狀（shape）分別為何？

(2) 在彩色影像中，像素位置 $(2,3)$ 的 RGB 值為何？

數位影像表示為矩陣：例題一

- 假設要把灰階影像轉為矩陣，有一張取樣自某張簡單圖片的 4×4 像素影像，每個像素紀錄的亮度值如下：

$$I = \begin{bmatrix} 0 & 64 & 128 & 255 \\ 0 & 64 & 128 & 255 \\ 0 & 64 & 128 & 255 \\ 0 & 64 & 128 & 255 \end{bmatrix}$$

這個矩陣說明了每行都呈現亮度由暗至亮（由左至右遞增），它也正是這張影像的灰階像素矩陣。

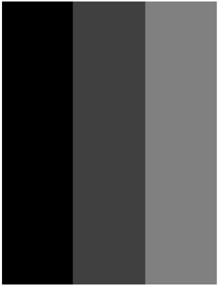
數位影像表示為矩陣：例題一

- 我們也可以試著寫程式呈現它，程式及效果如圖 1 所示。

```python
import numpy as np
import matplotlib.pyplot as plt

# 建立 4x4 灰階矩陣 (左暗右亮)
I = np.array([
    [  0,   64, 128, 255],
    [  0,   64, 128, 255],
    [  0,   64, 128, 255],
    [  0,   64, 128, 255]
], dtype=np.uint8)

# 顯示為灰階影像
plt.imshow(I, cmap='gray', vmin=0, vmax=255)
plt.title("模擬灰階影像 (左暗右亮) ")
plt.axis('off')
plt.show()
```



(a) 模擬灰階影像圖（左暗右亮）　(b) 對應的程式碼片段

Figure 1: 灰階影像與其生成程式碼對照圖

數位影像表示為矩陣：例題二

- 將彩色影像做 RGB 三通道分離。首先給定一張 example.jpg，我們可寫程式分離出三個矩陣 $R, G, B \in \mathbb{R}^{m \times n}$，如圖 2 所示，然後再針對這三個矩陣進行分析。

(a) 原始彩色圖像

(b) RGB 三通道分解圖像

Figure 2: 原始影像與通道分解結果比較

數位影像表示為矩陣：例題二

- 讀取 example.jpg 中的灰階影像並轉為矩陣，如圖 3 所示。

```python
import numpy as np
import matplotlib.pyplot as plt
from PIL import Image

# 讀取灰階影像並轉為矩陣
img = Image.open('example.jpg').convert('L')  # L 模式表示灰階
A = np.asarray(img)

# 顯示影像與對應矩陣維度
plt.imshow(A, cmap='gray')
plt.title(f"Image as Matrix: shape = {A.shape}")
plt.axis('off')
plt.show()
```

(a) 將影像轉換為灰階矩陣的 Python 程式碼

(b) 影像作為矩陣的視覺呈現與維度說明

Figure 3: 從影像轉換為矩陣表示的程式與輸出結果

數位影像表示為矩陣：例題二

- 注意到圖 3 (b) 上方有一排文字：

  Image as Matrix: shape = (1226, 1321).

  它意謂著我們已把灰階影像轉成了一個二維數值矩陣 $A \in \mathbb{R}^{1226\times1321}$，也就是說，每個像素會被記成一個「灰階數值」，範圍通常在 0-255。

- 影像中的每一個像素，對應矩陣中的一個元素，最終我們可以得到一個大小為 1226 列、1321 行的二階矩陣。以數學的方式來表達，這張影像現在對應到以下的矩陣：

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,1321} \\ a_{21} & a_{22} & \cdots & a_{2,1321} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1226,1} & a_{1226,2} & \cdots & a_{1226,1321} \end{bmatrix},$$

  其中每個 $a_{ij} \in [0, 255]$，代表第 i 列第 j 行像素的亮度。

數位影像表示為矩陣：例題三

Example (影像轉矩陣)

給定以下 4×4 的灰階影像片段，其像素亮度值如下：

$$I = \begin{bmatrix} 34 & 67 & 120 & 255 \\ 12 & 45 & 80 & 200 \\ 23 & 60 & 100 & 180 \\ 0 & 30 & 90 & 170 \end{bmatrix}.$$

請說明此矩陣如何代表影像？若將此影像縮放為 0.5 倍亮度，請寫出新的矩陣表示。

此矩陣為 $\mathbb{R}^{4\times4}$ 中的一個元素，每個元素表示對應像素的灰階亮度值。縮放亮度即進行純量乘法：$\tilde{I} = 0.5 \times I$，得到：

$$\tilde{I} = 0.5 \begin{bmatrix} 34 & 67 & 120 & 255 \\ 12 & 45 & 80 & 200 \\ 23 & 60 & 100 & 180 \\ 0 & 30 & 90 & 170 \end{bmatrix} = \begin{bmatrix} 17 & 33.5 & 60 & 127.5 \\ 6 & 22.5 & 40 & 100 \\ 11.5 & 30 & 50 & 90 \\ 0 & 15 & 45 & 85 \end{bmatrix}.$$

觀念整理

- 總結本節所提到的概念：
  - (a) 一張 300×400 的影像即為 $\mathbb{R}^{300\times400}$ 中的一個元素。
  - (b) 每一像素值可視為矩陣元素，亮度的變化可由數值大小詮釋。
  - (c) 後續所有線性代數運算（如加法、乘法、分解）都可直接作用於影像矩陣上。
- 「影像可以表示成矩陣的樣子」，而「操作矩陣就等於操作影像」。