Chapter 2 An Introduction to Matrix Analysis
矩陣的秩

Cheng-Fang Su

(National Yang Ming Chiao Tung University)

Jun 11, 2025

基本觀念

- 秩（rank）是矩陣最重要的數值特徵之一，反映了其列（或行）向量的線性獨立性與所張成的空間維度。理解秩不僅有助於掌握解線性方程組的條件，更能幫助我們辨識一個矩陣在資訊表示上的「有效維度」。
- 在影像處理中，矩陣的秩可以幫助我們回答以下問題：
  (1) 這些影像畫面彼此「有多大差異」?
  (2) 有哪些重複性高、可壓縮的內容？
  (3) 哪些部分是前景中真正的異常變化？
- 這些觀察都可以透過矩陣的秩來量化，而這也正是「背景為低秩、前景為稀疏」這類模型的數學依據。

矩陣秩的定理

- 我們在前面曾定義過，矩陣 $A \in \mathbb{R}^{m\times n}$ 的秩 $\mathrm{rank}(A)$ 是其列向量或行向量中線性獨立向量的最大數目；事實上，我們也可以將矩陣轉換為列階梯形式（row echelon form）後，其非零列的個數就是此矩陣的秩。
- 從先前一些例子，我們觀察到下列的定理成立：

Theorem (行秩等於列秩)

對任意矩陣 $A \in \mathbb{R}^{m\times n}$，其行向量所張成的子空間與列向量所張成的子空間具有相同維度，即：

$$\mathrm{rank}_{\mathrm{row}}(A) = \mathrm{rank}_{\mathrm{col}}(A)\,.$$

- 因此，矩陣秩只需計算行或列其中一者即可。

例題：以列階梯形式判斷矩陣的秩

Example

考慮矩陣：

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 0 \end{bmatrix},$$

請以列階梯形式判斷矩陣的秩。

例題：以列階梯形式判斷矩陣的秩

對矩陣 A 進行初等列運算，第 i 列以 R_i 來表示：

(1) 第一列乘以 -2 倍再加到第二列：

$$\boxed{R_2} \to -2R_1 + R_2 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

(2) 第一列乘以 -1 倍再加到第三列：

$$\boxed{R_3} \to -R_1 + R_3 = \begin{bmatrix} 0 & -1 & -3 \end{bmatrix}.$$

(3) 交換第二列與第三列，得到：

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -3 \\ 0 & 0 & 0 \end{bmatrix}. \tag{1}$$

這個計算過程就叫做「將矩陣轉換為列階梯形式（row echelon form）」，最後的矩陣 (1) 有兩列為非零列，因此：

$$\mathrm{Rank}(A) = 2.$$

這表示矩陣 A 的列或行只含有兩個線性獨立的向量，其所張成的空間為一個 $\mathbb{R}^3$ 中的 2 維平面。

以列階梯形式求矩陣的秩

Example

考慮矩陣：

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 9 \\ 0 & 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 4}.$$

我們先將第一列乘以 -2 加到第二列，得到：

$$\boxed{R_2} \rightarrow R_2 - 2R_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

交換第二列與第三列，得到最終列階梯形式為：

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

有三列為非零列，所以 $\mathrm{Rank}(B) = 3$，這表示矩陣 B 中有三個線性獨立的行（或列）向量，其所張成的是一個三維子空間。 □

為什麼需要低秩背景與稀疏前景模型？

- 在日常的監視錄影、醫學影像或自動駕駛中，我們經常需要從畫面中分離出真正重要的變化。例如教室的背景牆壁、黑板、桌椅都不會改變，但走進畫面的人，是重要的前景資訊，這些畫面可以視為由兩部分構成：

  (1) 背景：穩定、重複、類似的東西（例如靜止的場景）；

  (2) 前景：出現、移動、擾動的東西（例如人物或車輛）。

低秩背景與稀疏前景模型的數學表示法

- 從數學的角度來看，我們可以將一張影像（或一組影像序列）表示為一個矩陣 M，並試著將它分解成兩個部分：

$$M = L + S, \quad (\text{低秩與稀疏結構分解模型})$$

其中：

(1) 矩陣 L 代表「背景部分」，畫面內容穩定、變化少，結構重複性高，因此在數學上稱為低秩（low-rank）；

(2) 矩陣 S 則代表「前景或擾動部分」，例如移動中的人影、閃動的光點，這些只出現在局部位置，因此被視為稀疏（sparse）。

這種分解方法被稱為低秩–稀疏分解（low-rank–sparse decomposition）

低秩背景與稀疏前景模型的圖像直觀

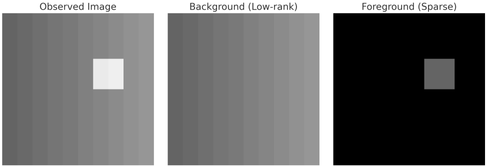


Figure 1: 將畫面分解為低秩背景與稀疏前景

- Observed Image 是原始觀察到的影像；Background (Low-rank) 是穩定的背景結構；Foreground (Sparse) 則為只有局部變化的前景（例如人物或動態物件）。

背景內容重複 $\Rightarrow$ 資訊可以壓縮 $\Rightarrow$ 可由少數方向（基底）張成;
前景變化快速 $\Rightarrow$ 難以預測、無法壓縮 $\Rightarrow$ 形成稀疏擾動.

低秩背景與稀疏前景模型的討論題

- 假設我們有一個簡單的 5×6 灰階影像序列矩陣 M，且每一行 $\vec{v_i}$ 是一張影像，此矩陣由 6 張影像所構成：

$$M = \begin{bmatrix} \vec{v_1} & \vec{v_2} & \cdots & \vec{v_6} \end{bmatrix} = \begin{bmatrix} 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 255 \end{bmatrix}.$$

- 從 M 中可觀察到，每一行是一張 5×1 的影像，而第 6 張影像（即第 6 行）在第 5 個像素有一個異常值 255，這代表這個位置所對應的影像位置，有前景物體正經過這裡。

低秩背景與稀疏前景模型的討論題

接著，我們可將 M 分解為 L 與 S：

(a) 背景矩陣（低秩部分）：

$$L = \begin{bmatrix} 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 & 100 \end{bmatrix} \Rightarrow \operatorname{rank}(L) = 1\,.$$

這個矩陣代表「每一張影像都幾乎相同」，所有行向量都一樣，所以秩為 1；注意到秩為 1 就是一種低秩的情形。

低秩背景與稀疏前景模型的討論題

接著，我們可將 M 分解為 L 與 S：

(b) 稀疏矩陣（前景異常部分）：

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 155 \end{bmatrix}$$

其中只有一個非零值 $255 - 100 = 155$，表示在第 6 張影像的第 5 個像素，有一個前景物體造成的變動，此矩陣其餘位置皆為零，我們將這樣的矩陣稱為稀疏矩陣（sparse）。

低秩背景與稀疏前景模型的討論題

- 在這個例子中，我們可以看到向量 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_6}$ 是線性相依的，其中

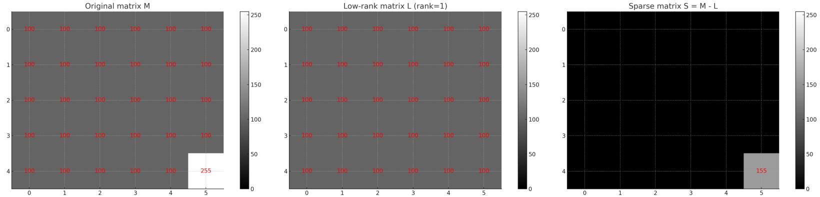
$$\vec{v_1} = \vec{v_2} = \cdots = \vec{v_5},$$

而 $\vec{v_6}$ 提供了一個與 $\vec{v_1}$ 不同的方向資訊，所以這 6 個向量的張成空間為

$$\begin{aligned}\mathrm{span}\{\vec{v}_1, \ldots, \vec{v}_6\} &= \left\{a_1 \vec{v}_1 + \cdots + a_6 \vec{v}_6 \;\middle|\; a_i \in \mathbb{R}, 1 \leqslant i \leqslant 6\right\} \\ &= \left\{a \vec{v}_1 + b \vec{v}_6 \;\middle|\; a, b \in \mathbb{R}\right\},\end{aligned}$$

- 向量 $\vec{v}_1$ 與 $\vec{v}_6$ 線性獨立，所以矩陣 M 的秩為 2（這對矩陣來說也是低秩的情形），而 $\mathrm{span}\{\vec{v}_1, \ldots, \vec{v}_6\}$ 是一個低維子空間（維度為 2）。

低秩背景與稀疏前景模型的討論題



(a) Original matrix M (b) Low-rank L (rank $= 1$) (c) Sparse matrix $S = M - L$

Figure 2: Matrix decomposition $M = L + S$ with rank-1 background: original data M, low-rank background L, and sparse anomaly S.

低秩矩陣的意義

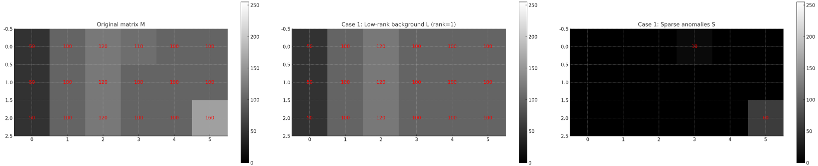
- 從數學上來說，秩代表這個矩陣能被多少組線性獨立的向量「建構」出來。如果某個矩陣的秩為 1，就表示它的所有行（或列）向量都落在一個共同方向上。由於，矩陣不會有秩比 1 還低的情形，所以秩為 1 的矩陣是一種低秩矩陣。
- 從影像來讀解的話，試想一段背景穩定的監視器影片畫面，我們從影片中擷取出的每張影像通常只會有局部的變化，此時，當我們把這多張畫面向量化後組成一個大矩陣時，這個大矩陣的每一列或每一行往往彼此相似，向量間線性相依的程度很高，矩陣的秩就會很低。
- 上述這樣的矩陣就稱為「低秩矩陣（low-rank matrix）」，它表示這些資料內部具有高度冗餘性。
- 在影像處理中，冗餘性常常表現為畫面中大量重複或變化不大的部分，所以，從向量的角度來看，我們可用較少的方向描述矩陣的主要結構。

例題

- 有一個 3×6 灰階影像序列矩陣 M，且每一行 $\vec{v_i}$ 是一張影像，此矩陣由 6 張影像所構成：

$$M = \begin{bmatrix} \vec{v_1} & \vec{v_2} & \cdots & \vec{v_6} \end{bmatrix} = \begin{bmatrix} 50 & 100 & 120 & {\color{red}110} & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & {\color{red}160} \end{bmatrix}.$$

可以看出 $\vec{v_2} = \vec{v_5} = 2\,\vec{v_1}$ 且 $\vec{v_3} = 2.4\,\vec{v_1}$，然後 $\vec{v_1}$、$\vec{v_4}$、$\vec{v_6}$ 三個向量線性獨立，所以矩陣 M 的每一個行向量之間線性相依的程度很高。



原始矩陣 M Case 1: 背景 rank=1 L Case 1: 稀疏矩陣 S

Figure 3: 以秩 1 背景模型進行分解後的三張圖比較：原始影像矩陣 M、低秩背景 L (rank=1)、稀疏矩陣 S。

例題

- 這 6 個向量的張成空間可表示為

$$\begin{aligned}\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_6\} &= \left\{a_1\,\vec{v}_1+\cdots+a_6\,\vec{v}_6 \;\middle|\; a_i\in\mathbb{R},1\leqslant i\leqslant 6\right\}\\ &= \left\{a\,\vec{v}_1+b\,\vec{v}_4+c\,\vec{v}_6 \;\middle|\; a,b,c\in\mathbb{R}\right\}.\end{aligned}$$

- 顯然地，張成空間維度為 3，意味著矩陣 M 的秩為 3，且能進行更進一步的分解：

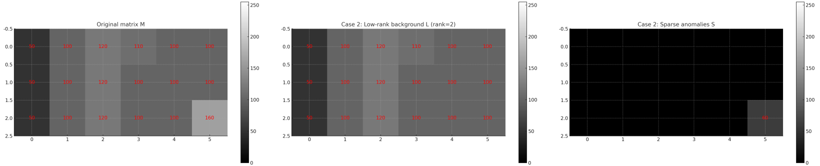
$$M=L+S=\begin{bmatrix}50&100&120&100&100&100\\50&100&120&100&100&100\\50&100&120&100&100&100\end{bmatrix}+\begin{bmatrix}0&0&0&10&0&0\\0&0&0&0&0&0\\0&0&0&0&0&60\end{bmatrix},$$

矩陣 L 顯然秩為 1（也是低秩），它是背景，而矩陣 S 可以讓我們看到影片在什麼時間點的什麼位置有局部變化。

例題

- 這個例子有另一種可以解讀的方式，假設我們取 $\vec{v_4} = \begin{bmatrix} 110 \\ 100 \\ 100 \end{bmatrix}$，值 110 可能是有窗戶的存在，導致影片在某個時間點只在那個位置光影有微小的變化，此時我們的分解可寫為

$$M = L + S = \begin{bmatrix} 50 & 100 & 120 & 110 & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & 100 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 60 \end{bmatrix}.$$



(a) Original matrix M (b) Low-rank L (rank=2) (c) Sparse anomalies S

Figure 4: Matrix decomposition under Case 2 (assumed background of rank 2): $M = L + S$.

例題

$$M = L + S = \begin{bmatrix} 50 & 100 & 120 & \color{red}{110} & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & 100 \\ 50 & 100 & 120 & 100 & 100 & 100 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \color{red}{60} \end{bmatrix}.$$

- 矩陣 L 秩為 2（也仍是低秩），而矩陣 S 還是可以讓我們看到影片在什麼時間點的什麼位置有局部變化。
- 所以，背景 L 怎麼取與我們實際要處理的問題有關，這也就是為什麼有時我們取 L 秩為 1 的情形，有時只單純需要 L 是低秩（不一定要秩為 1）就好，完全看實際的問題來決定。

低秩矩陣在數學上的定義

Definition (低秩矩陣, Low-Rank Matrix)

若矩陣 $A \in \mathbb{R}^{m\times n}$ 的秩滿足 $\mathrm{rank}(A) = r \ll \min(m, n)$，則稱 A 為低秩矩陣。

- 假設我們在一分鐘內用同一個角度拍下 20 張空教室的畫面，每張畫面大小為 100×100 像素。我們將這些畫面向量化，組成矩陣 $M \in \mathbb{R}^{10000\times 20}$。
- 在教室背景完全不變的前提下，那麼這 20 張畫面就會是完全一樣的，矩陣 M 的每一行都是相同的向量，矩陣 M 的秩為 1。
- 即使這一分鐘內教室的光線稍有不同，一般而言秩也不會太高，這意味著我們可以用很少的「方向」來描述整個教室的樣貌。

稀疏矩陣的觀念與定義

Definition (稀疏矩陣, Sparse Matrix)

設矩陣 $A \in \mathbb{R}^{m\times n}$，若其中非零元素的數量遠小於總元素數 mn，則稱 A 為稀疏矩陣。

- 在影像處理或資料科學中，我們常常遇到這樣的情況：「一個矩陣中絕大多數的元素都是 0，只有少數幾個地方是非零數值。」這種矩陣我們稱為稀疏矩陣（sparse matrix），也可以想成，稀疏矩陣就是「幾乎都是空白的資料表格」。
- 為什麼會出現這種矩陣呢？想像我們正在看一段監視器影片，大部分時間都是一張空教室的背景，只有有時候有人經過。如果我們把這些畫面中「有改變」的部分（例如人物輪廓）挑出來、記錄在一個矩陣中，會發現大部分像素值仍然是 0，只有人的位置會有數值。這就形成了一個稀疏的矩陣。

稀疏矩陣的例子

- 比方說，

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 255 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 128 \end{bmatrix},$$

這是某張影像的前景遮罩：其中非零值表示人物出現的位置，其他地方都是背景（用 0 表示）。整體來看幾乎都是 0，只有「重要的點」才是非零。

矩陣稀疏程度的量化

我們可以定義一個數字來描述這個矩陣有多稀疏，這就是所謂的稀疏度（sparsity level）：

$$\text{sparsity} = 1 - \frac{\text{非零元素個數}}{\text{總元素個數}}.$$

以上面矩陣為例，它是 4×4 的矩陣，共 16 個元素，只有 2 個非零值，所以：

$$\text{sparsity} = 1 - \frac{2}{16} = 0.875.$$

這代表有 87.5% 的位置都是 0，是一個高度稀疏的矩陣。

Example

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

此矩陣中僅有兩個非零元素，因此 sparsity 為 $1 - \frac{2}{16} = 0.875$，表示有 87.5% 的元素為零。

結論

- 稀疏矩陣之所以重要，是因為它們可以更有效地處理與儲存資料。
- 由於矩陣中大部分的元素都是 0，只需要記錄少數有意義的位置，就能大幅節省記憶體空間。
- 稀疏矩陣讓我們更容易找出影像中的重點，像是人物、物體等通常只佔據畫面的一小部分，對應到的正是那些非零值。
- 稀疏結構有助於資料的壓縮，因為我們只需要處理真正有資訊的地方。