矩陣的定義

定義 1 (矩陣). **矩陣是代數學中一種特殊的表達形式，用以同時表達多位元的表格：**

$$\mathbf{A} = (a_{ij})_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

若元素$a_{ij} \in S$則稱爲S上的$m \times n$矩陣，記$\mathbf{A} \in \mathbb{M}^{m \times n}(S) = \mathbb{M}_S^{m \times n}$。

示例. *以下爲矩陣的例子：*

1. $1 =: [1] \in \mathbb{M}^{1 \times 1}(\mathbb{R})$。

2. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{M}^{2 \times 2}(\mathbb{R})$。

3. $\begin{bmatrix} 1 + 2i & 3 - 4i & i \\ -i & 1 & 2 - i \end{bmatrix} \in \mathbb{M}^{2 \times 3}(\mathbb{C})$。

定義 2 (矩陣的加法與乘法定義). *給定$\mathbf{A}, \mathbf{B} \in \mathbb{M}^{m \times n}$，$\mathbf{C} \in \mathbb{M}^{n \times t}$，則*

1. 加法： $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$。

2. 數乘： $k\mathbf{A} = (ka_{ij})$。

3. 矩陣乘法： $\mathbf{A} \times \mathbf{B} = (\sum_{k=1}^{n} a_{ik} b_{kj})$。

示例. *設$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 2 & 0 \\ 2 & 1 \\ 3 & 3 \end{pmatrix}$，則*

1. $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 1+2 & 1+1 & 0+0 \\ 1+1 & 0+2 & 1+1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \end{pmatrix}$。

2. $4\mathbf{A} = \begin{pmatrix} 4 & 4 & 0 \\ 4 & 0 & 4 \end{pmatrix}$。

3. $\mathbf{A} \times \mathbf{C} = \begin{pmatrix} 1 \times 2 + 1 \times 2 + 0 \times 3 & 1 \times 0 + 1 \times 1 + 0 \times 3 \\ 1 \times 2 + 0 \times 2 + 1 \times 3 & 1 \times 0 + 0 \times 1 + 1 \times 3 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 5 & 3 \end{pmatrix}$。

定義 3 (特殊矩陣). 下列為一些矩陣寫法的共識：

- 零矩陣：$\mathbf{0}_{m \times n} = \mathbf{O}_{m \times n} = (0)_{m \times n}$。

- 一矩陣：$\mathbf{1}_{m \times n} = (1)_{m \times n}$。

- 單位矩陣：$\mathbf{I}_{m \times n} = (\delta_{ij})_{m \times n}$，其中 $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$。

- 向量：

  - 行向量：$\mathbf{x}^T = \vec{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{M}^{1 \times n}$。

  - 列向量：$\mathbf{y} = \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = (y_1, y_2, \ldots, y_m) \in \mathbb{M}^{m \times 1}$

命題 (加法與數乘定則). 設 $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{M}^{m \times n}$，則

1. 加法結合律：$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$。

2. 加法交換律：$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$。

3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$。

4. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$。

5. $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + \mathbf{B}$。

6. $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$。

7. $(ab)\mathbf{A} = a(b\mathbf{A})$

命題 (矩陣乘法定則). 設 $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ 為矩陣，則

1. 乘法結合律：$(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C})$。

2. 分配律I：$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C}$。

3. 分配律II：$\mathbf{D}(\mathbf{A} + \mathbf{B}) = \mathbf{D}\mathbf{A} + \mathbf{D}\mathbf{B}$。

4. $\alpha\mathbf{A}\mathbf{B} = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$。

5. $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$。

定義 4 (矩陣的轉置). 設 $\mathbf{A} = (a_{ij})_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$，則定義 **A 的轉置**

為

$$\mathbf{A}^T = (a_{ji})_{n \times m} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

示例. 設 $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$，則 $\mathbf{A}^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$。

定義 5 (矩陣的跡). 設 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$，則定義 **A 的跡**為

$$\mathrm{tr}(\mathbf{A}) = \sum_{k}^{\min(m,n)} a_{kk} = a_{11} + a_{22} + \cdots + a_{ll}$$

其中 $l = \min(m, n)$。

命題. $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^T)$。

　　證明. 留做習題。 □

定義 6 (綫性組合). 設 $c_1, c_2, \ldots, c_\ell$ 為常數，$\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell \in \mathbb{M}_{m \times n}$，則稱

$$\sum_{k=1}^{\ell} c_k \mathbf{A}_k = c_1 \mathbf{A}_1 + c_2 \mathbf{A}_2 + \cdots + c_\ell \mathbf{A}_\ell$$

為**綫性組合**。

示例. 設 $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$，則可寫

$$\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 3e_1 + 2e_2 + e_1$$

示例. 設 $E_n \in \mathbb{M}^{3 \times 3}$, 對於 $1 \le n \le 9$, $E_n = (e_{ij})_{3 \times 3}$ 而且

$$e_{ij} = \begin{cases} 1 & n = 3(i-1) + j \\ 0 & otherwise \end{cases}$$

, 則可寫

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \sum_{k=1}^{9} k E_k$$

小記. 以上例子便是矩陣與綫性方程組的關係。

命題. 設 $c_1, c_2, \ldots, c_\ell$ 為常數, $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell \in \mathbb{M}_{m \times n}$, 則

- $\operatorname{tr}\left(\sum_{k=1}^{\ell} c_k \mathbf{A}_k\right) = \sum_{k=1}^{\ell} c_k \operatorname{tr}(\mathbf{A}_k)$;

- $\left(\sum_{k=1}^{\ell} c_k \mathbf{A}_k\right)^T = \sum_{k=1}^{\ell} c_k \mathbf{A}_k^T$;

- $\operatorname{tr}\left(\sum_{k=1}^{\ell} c_k \mathbf{A}_k^T\right) = \sum_{k=1}^{\ell} c_k \operatorname{tr}(\mathbf{A}_k)$。

  證明. 留作習題。 □

綫性方程組

綫性方程組主要用以表達兩條或多條同時成立的綫性方程，最具代表性的可數初中的聯立方程：

$$\begin{cases} ax + by & = k_1 \\ cx + dy & = k_2 \end{cases}$$

當然，綫性方程組的含義不僅是二元一次方程，更可以拓展到多元一次(綫性)方程：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

示例. 以下是綫性方程組的例子：

1. $$\begin{cases} x + y & = 0 \\ 2x - 3y & = 0 \end{cases}$$

2. $$\begin{cases} x + y + z & = 0 \\ 2x + 3y + 4z & = 0 \end{cases}$$

3. $$\begin{cases} x - y + z & = 0 \\ 2x + 4z & = 0 \\ -2x - 4y & = 0 \end{cases}$$

解綫性方程組

欲求綫性方程組的解，我們可使**代入法**或**消元法**，其中消元法比代入法的效率更高，因此普遍數學家都會使用消元法解方程。同時此辦法也衍生出矩陣的各項命題。

示例. 求解綫性方程組 $\begin{cases} 2x + 3y & = 8 \\ 6x - 2y & = 9 \end{cases}$

1. 運用代入法求解：從$2x + 3y = 8$可得$x = \dfrac{8 - 3y}{2}$，代入$6x - 2y = 9$得

$$6(\frac{8 - 3y}{2}) - 2y = 9$$
$$3(8 - 3y) - 2y = 9$$
$$24 - 9y - 2y = 9$$
$$11y = 15$$
$$y = \frac{15}{11}$$

再代$y = \dfrac{15}{11}$入$x = \dfrac{8 - 3y}{2}$得

$$x = \frac{8 - 3(\frac{15}{11})}{2}$$
$$= \frac{43}{22}$$

因此綫性方程組的解為$(\dfrac{43}{22}, \dfrac{15}{11})$。

2. 運用消元法求解：從$2x + 3y = 8$三倍後可得$6x + 9y = 24$，則上式減去下式可得

$$
\begin{array}{lr}
(1) \times 3: & 6x + 9y = 24 \\
-)(2): & 6x - 2y = 9 \\
\hline
(1) \times 3 - (2): & 11y = 15 \\
& y = \dfrac{15}{11} \\
& 6x - 2(\dfrac{15}{11}) = 9 \\
& x = \dfrac{43}{22}
\end{array}
$$

示例. 求解綫性方程組 $\begin{cases} 2x + 3y + 4z & = 1 \\ 3x - y + 3z & = 0 \\ x + y + z & = 1 \end{cases}$

1. 運用代入法求解：先從$3x - y + 3z = 0$得出$y = 3x + 3z$，代入其餘兩式可得

$$\begin{cases} 2x + 3(3x + 3z) + 4z & = 1 \\ x + (3x + 3z) + z & = 1 \end{cases} \implies \begin{cases} 11x + 13z & = 1 \\ 4x + 4z & = 1 \end{cases}$$

再從$4x + 4z = 1$得$z = \dfrac{1 - 4x}{4}$，代入$11x + 13z = 1$得

$$11x + 13(\frac{1-4x}{4}) = 1$$
$$44x + 13 - 52x = 4$$
$$-8x = -9$$
$$x = \frac{9}{8}$$

由此可得：

$$z = \frac{1 - 4(\frac{9}{8})}{4}$$
$$= -\frac{7}{8}$$
$$y = 3(\frac{11}{4}) + 3(-\frac{5}{2})$$
$$= \frac{3}{4}$$

2. 運用消元法求解：

$(1):$	$2x + 3y + 4z = 1$
$(2):$	$3x - y + 3z = 0$
$(3):$	$x + y + z = 1$
$(1) - (3) \times 2:$	$y + 2z = -1$
$(2) - (3) \times 3:$	$-4y = -3$
	$y = \dfrac{3}{4}$
	$z = -\dfrac{7}{8}$
	$x = \dfrac{9}{8}$

我們亦容許線性方程組**沒有解**或**有無限解**。若要理解沒有解或有無限多個解，可參考直線的特性：

定義 7 (平行綫). 設L_1和L_2為兩條直綫，並分別以m_1和m_2代表其斜率。若$L_1 // L_2$，則$m_1 = m_2$。

定理 (平行綫的交點數). 設L_1和L_2為一對（歐幾里得幾何）平行綫，則L_1和L_2的交點數只能為0或∞。

證明. 設m爲L_1及L_2的斜率，及c_1和c_2分別爲L_1及L_2的縱軸截距，則

$$L_1 : y = mx + c_1, \; L_2 : y = mx + c_2$$

若$c_1 = c_2$，則$\forall x$, 若 $(x, y_1) \in L_1, (x, y_2) \in L_2$,

$$y_1 = mx + c_1 = mx + c_2 = y_2$$

由於對任意x均成立，此爲無限多交點。

若$c_1 \neq c_2 \implies c_1 - c_2 \neq 0$，則$\forall x$, 若 $(x, y_1) \in L_1, (x, y_2) \in L_2$,

$$y_2 - y_1 = (mx + c_1) - (mx + c_2) = c_1 - c_2 \neq 0 \implies y_1 \neq y_2$$

因此沒有交點。　　　　　　　　　　　　　　　　　　　　　　　　$\square$

衍理. 任何一對歐幾里得二維直綫可有0,1或∞交點。

因此，綫性方程組可有0,1或無限多解。同時，綫性方程組若有兩個解，則會有無限多解。

定理. 設$A \in \mathbb{M}^{m \times n}, \mathbf{b} \in \mathbb{M}^{m \times 1}$使得$(S) : [\mathbf{A}|\mathbf{b}]$爲一綫性方程組，並設$\mathbf{x}, \mathbf{y} \in \mathbb{M}^{m \times 1}$爲相異解，則

$$I := \{t\mathbf{x} + (1-t)\mathbf{y} : t \in \mathbb{R}\}$$

爲方程組的解集。

證明. 對於任意$1 \leq k \leq m$,

$$a_{k1}[tx_1 + (1-t)y_1] + \cdots + a_{kn}[tx_n + (1-t)y_n]$$
$$= t[a_{k1}x_1 + \cdots + a_{kn}x_n] + (1-t)[a_{k1}y_1 + \cdots + a_{kn}y_n]$$
$$= tb_k + (1-t)b_k$$
$$= b_k$$

$$\square$$

我們稱擁有無限多解的方程組的解集爲**通解**。

示例. 以下爲綫性方程組無限多解的例子：

1. 考慮 $\begin{cases} 3x + y = 2 \\ -3x - y = -2 \end{cases}$ ，由於任何符合$3x + y = 2$的點均符合$-3x - y = -2$，因此此方程組有無限多解。解集爲

$$\{(t, 2 - 3t) : t \in \mathbb{R}\}$$

2. 考慮 $\begin{cases} 3x + y + 2z = 0 \\ x + y + 7z = 0 \\ 2x - 5z = 0 \end{cases}$ ，由於上式減中式等於下式，因此此方程組無法解下

去，有無限多解。解集為

$$\{(5t, -19t, 2t) : t \in \mathbb{R}\}$$

以下為綫性方程組沒有解的例子：

1. 考慮 $\begin{cases} 3x + y = 2 \\ -3x - y = 0 \end{cases}$ ，由於任何符合$3x + y = 2$的點均不符合$-3x - y = 0$，

因此此方程組沒有解。解集為

$$\emptyset$$

綫性方程組的可解性I

現記綫性方程組的通常式為

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

定理 (綫性方程組的可解性I). 對於m行n列綫性方程組S，若$m \leq n$，則S有解。

證明. 當m=1時，$a_{11}x_1 + a_{12}x_2 + \cdots + x_{1n}x_n = b_1$，則存在綫性函數$f$使得$x_1 = f(x_2, x_3, \ldots, x_n)$令$a_{11}x_1 + a_{12}x_2 + \cdots + x_{1n}x_n = b_1$成立。

當$1 < m < n$時，已知有函數$f_1, f_2, f_3, \ldots, f_{m-1}$使得

$$x_1 = f_1(x_{m+1}, \ldots, x_n)$$
$$x_2 = f_2(x_{m+1}, \ldots, x_n)$$
$$\vdots = \vdots$$
$$x_m = f_m(x_{m+1}, \ldots, x_n)$$

令 $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases}$ 成立。

則 $m+1 \le n$ 時，存在綫性函數 f_{m+1} 使得

$$x_{m+1} = f_{m+1}(x_1, x_2, x_3, \ldots, x_m, x_{m+2}, \ldots, x_n)$$
$$= f_{m+1}(f_1(x_{m+1}, \ldots, x_n), \ldots, f_m(x_{m+1}, \ldots, x_n), x_{m+2}, \ldots, x_n)$$

則存在綫性函數 F 和 g_{m+1} 使得

$$x_{m+1} = F(x_{m+1}, x_{m+2}, \ldots, x_n)$$
$$= g_{m+1}(x_{m+2}, \ldots, x_n)$$

因此可見存在綫性函數 $g_1, g_2, \ldots, g_m$(若 $m+2 > n$ 視所有 g_k 為常函數)

$$x_1 = f_1(g_{m+1}(x_{m+2}, \ldots, x_n), x_{m+2}, \ldots, x_n)$$
$$= g_1(x_{m+2}, \ldots, x_n)$$
$$x_2 = f_2(g_{m+1}(x_{m+2}, \ldots, x_n), x_{m+2}, \ldots, x_n)$$
$$= g_2(x_{m+2}, \ldots, x_n)$$
$$\vdots = \vdots$$
$$\vdots = \vdots$$
$$x_m = f_m(g_{m+1}(x_{m+2}, \ldots, x_n), x_{m+2}, \ldots, x_n)$$
$$= g_m(x_{m+2}, \ldots, x_n)$$

此證畢方程解的存在性。 □

目前只有存在性能證明，因爲：

1. 未能確定是否有多於一個解；

2. 當 $m > n$ 時，未知是否有解。

下一節會解決上述問題，判斷解的唯一性。

增廣矩陣

爲了使方程組的觀察更方便及易於處理，在此階段引入**增廣矩陣**的概念。

考慮以下綫性方程組：

$$(S): \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

考慮 $\mathbf{A} := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{M}^{m \times n}$ 為**係數矩陣**及 $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ 為**水平向量**，則

可將 (S) 作以下簡化表達：

$$(S): [\mathbf{A}|\mathbf{b}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \vline & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & \vline & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vline & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & \vline & b_m \end{bmatrix}$$

以上 $[\mathbf{A}|\mathbf{b}]$ 稱爲**增廣矩陣**。增廣矩陣的概念在於簡化方程組的表達，因此方程組的運作均可在擴增矩陣中實現。如此引入以下概念：

定義 8 (列階梯形矩陣). 設 $A \in \mathbb{M}^{m \times n}$，若 A 符合以下條件，則稱 A 為**列階梯形矩陣**：

- 若某列有非零元素，則必在所有全零列之上；

- 某列最左邊的非零元素稱爲**首項係數**。某列的首項係數必定比上一列的首項係數更靠右。

示例. 列階梯形矩陣的例子：

$$\begin{bmatrix} 1 & a_{12} & \cdots & a_{1(n-1)} & a_{1n} & \vline & b_1 \\ 0 & 2 & \cdots & a_{2(n-1)} & a_{2n} & \vline & b_2 \\ 0 & 0 & \cdots & 0 & a_{3n} & \vline & b_3 \\ 0 & 0 & \cdots & 0 & 0 & \vline & 0 \end{bmatrix}$$

定義 9 (還原列梯矩陣). 設 $A \in \mathbb{M}^{m \times n}$，若 A 符合以下條件，則稱 A 為**還原列梯矩陣**：

- **A**為列階梯形矩陣，并且；

- **A**的所有首項係數均爲1；

- **A**中所有首項係數的對應行僅有一個非零數字。

示例. 還原列梯矩陣的例子：

$$\left[\begin{array}{ccccc|c} 1 & 0 & \cdots & a_{1(n-1)} & 0 & b_1 \\ 0 & 1 & \cdots & a_{2(n-1)} & 0 & b_2 \\ 0 & 0 & \cdots & 0 & 1 & b_3 \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{array}\right]$$

已知綫性方程組能通過消元法求解，由此可見還原列梯矩陣正是消元法所形成的結果。我們給出以下定義：

定義 10 (初等行(列)變換). 下列三種對矩陣的變換稱爲矩陣的**初等行（列）變換**：

1. 把第i行（列）和第j行（列）互換位置，記爲$r_i \leftrightarrow r_j$（$c_i \leftrightarrow c_j$）；

2. 用非零數α乘以第i行（列），記爲αr_i（αc_i）；

3. 把第i行（列）的α倍加到第j行（列），記爲$\alpha r_i + r_j$（$\alpha c_i + c_j$）。

定義 11. 設$\mathbf{A}, \mathbf{A}' \in \mathbb{M}^{m \times n}$。

我們稱$\mathbf{A}$和$\mathbf{A}'$是**等價**的若$\mathbf{A}$可通過列運算達成$\mathbf{A}'$，記$\mathbf{A} \sim \mathbf{A}'$；若$\mathbf{A}'$為還原列階矩陣而且$\mathbf{A} \sim \mathbf{A}'$，則稱$\mathbf{A}'$為$\mathbf{A}$的還原式，記$\mathbf{A}' = \mathrm{RREF}(\mathbf{A})$。

示例. 設$\mathbf{A} = \left[\begin{array}{ccccc|c} 6 & 9 & 3 & 3 & 9 & 3 \\ 1 & 5 & 4 & 0 & 1 & 9 \\ 0 & 7 & 3 & 2 & 6 & 6 \\ 8 & 1 & 1 & 2 & 3 & 4 \\ 0 & 2 & 7 & 3 & 9 & 2 \\ 7 & 5 & 6 & 0 & 0 & 2 \end{array}\right]$，求$\mathrm{RREF}(\mathbf{A})$。

先將第一行的首項係數求出，由於第二行的首項係數為1，因此我們可以通過簡單的行變換造出適當的等價矩陣。

$$\begin{bmatrix} 6 & 9 & 3 & 3 & 9 & | & 3 \\ 1 & 5 & 4 & 0 & 1 & | & 9 \\ 0 & 7 & 3 & 2 & 6 & | & 6 \\ 8 & 1 & 1 & 2 & 3 & | & 4 \\ 0 & 2 & 7 & 3 & 9 & | & 2 \\ 7 & 5 & 6 & 0 & 0 & | & 2 \end{bmatrix} \quad \begin{matrix} \frac{1}{3}r_1 \\ r_1 \leftrightarrow r_2 \\ \sim \end{matrix} \quad \begin{bmatrix} 1 & 5 & 4 & 0 & 1 & | & 9 \\ 2 & 3 & 1 & 1 & 3 & | & 1 \\ 0 & 7 & 3 & 2 & 6 & | & 6 \\ 8 & 1 & 1 & 2 & 3 & | & 4 \\ 0 & 2 & 7 & 3 & 9 & | & 2 \\ 7 & 5 & 6 & 0 & 0 & | & 2 \end{bmatrix} \quad \begin{matrix} -2r_1 + r_2 \\ -8r_1 + r_4 \\ -7r_1 + r_6 \\ \sim \end{matrix} \quad \begin{bmatrix} 1 & 5 & 4 & 0 & 1 & | & 9 \\ 0 & -7 & -7 & 1 & 1 & | & -17 \\ 0 & 7 & 3 & 2 & 6 & | & 6 \\ 0 & -39 & -31 & 2 & -5 & | & -68 \\ 0 & 2 & 7 & 3 & 9 & | & 2 \\ 0 & -30 & -22 & 0 & -7 & | & -61 \end{bmatrix}$$

隨後將第二及其餘行的首項係數求出：

$$\begin{matrix} r_3 + r_2 \\ 20r_5 + r_4 \\ 15r_5 + r_6 \\ \sim \end{matrix} \begin{bmatrix} 1 & 5 & 4 & 0 & 1 & | & 9 \\ 0 & 0 & -4 & 3 & 7 & | & -11 \\ 0 & 7 & 3 & 2 & 6 & | & 6 \\ 0 & 1 & 109 & 62 & 175 & | & -28 \\ 0 & 2 & 7 & 3 & 9 & | & 2 \\ 0 & 0 & 83 & 45 & 128 & | & -31 \end{bmatrix} \begin{matrix} -5r_4 + r_1 \\ -7r_4 + r_3 \\ -2r_4 + r_5 \\ \sim \end{matrix} \begin{bmatrix} 1 & 0 & -541 & -310 & -874 & | & 149 \\ 0 & 0 & -4 & 3 & 7 & | & -11 \\ 0 & 0 & -760 & -432 & -1219 & | & 174 \\ 0 & 1 & 109 & 62 & 175 & | & -28 \\ 0 & 0 & -211 & -121 & -341 & | & 58 \\ 0 & 0 & 83 & 45 & 128 & | & -31 \end{bmatrix}$$

$$\begin{matrix} -r_2 \\ r_2 \leftrightarrow r_4 \\ \sim \end{matrix} \begin{bmatrix} 1 & 0 & -541 & -310 & -874 & | & 149 \\ 0 & 1 & 109 & 62 & 175 & | & -28 \\ 0 & 0 & -760 & -432 & -1219 & | & 174 \\ 0 & 0 & 4 & -3 & -7 & | & 11 \\ 0 & 0 & -211 & -121 & -341 & | & 58 \\ 0 & 0 & 83 & 45 & 128 & | & -31 \end{bmatrix} \begin{matrix} 135r_4 + r_1 \\ -27r_4 + r_2 \\ 190r_4 + r_2 \\ 53r_4 + r_5 \\ -20r_4 + r_6 \\ \sim \end{matrix} \begin{bmatrix} 1 & 0 & -1 & -715 & -1819 & | & 1634 \\ 0 & 1 & 1 & 143 & 364 & | & -325 \\ 0 & 0 & 0 & -1002 & -2549 & | & 2264 \\ 0 & 0 & 4 & -3 & -7 & | & 11 \\ 0 & 0 & 1 & -280 & -712 & | & 641 \\ 0 & 0 & 3 & 105 & 268 & | & -251 \end{bmatrix}$$

$$\begin{matrix} r_5 + r_1 \\ -r_5 + r_2 \\ -4r_5 + r_4 \\ -3r_5 + r_6 \\ \sim \cdots \sim \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & | & -\frac{148}{551} \\ 0 & 1 & 0 & 0 & 0 & | & \frac{717}{551} \\ 0 & 0 & 1 & 0 & 0 & | & \frac{895}{551} \\ 0 & 0 & 0 & 1 & 0 & | & \frac{3975}{551} \\ 0 & 0 & 0 & 0 & 1 & | & -\frac{2058}{551} \\ 0 & 0 & 0 & 0 & 0 & | & -\frac{6817}{551} \end{bmatrix}$$

詳細步驟參考 *RREF calculator: https://www.emathhelp.net/en/calculators/linear-algebra/reduced-row-echelon-form-rref-calculator/*

小記. 上例可見最後一行的左方為零行，但右方為非零數字，得到$0 = 1$的結論。此情形我們稱方程組無解。

重新考慮綫性方程組的樣式，因應矩陣乘法的特性，我們可以將$S = [\mathbf{A}|\mathbf{b}]$寫作

$$\mathbf{Ax} = \mathbf{b}$$

若存在$\mathbf{x_0} \in \mathbb{M}^{n \times 1}$使得$\mathbf{Ax_0} = \mathbf{b}$成立，則稱$\mathbf{x_0}$為$\mathbf{Ax} = \mathbf{b}$的一個解。

定理. 綫性方程組$S = [\mathbf{A}|\mathbf{b}]$有解當且僅當$\text{RREF}(S) = [\mathbf{A'}|\mathbf{b'}]$有解；並且$S$與$\text{RREF}(S)$同解。

證明. 由於S與$\text{RREF}(S)$等價，故存在若干行變換及對應的變換矩陣$\mathbf{P}$使得

$$\mathbf{Ax_0} = \mathbf{b} \iff \mathbf{PAx_0} = \mathbf{Pb} \iff \mathbf{A'x_0} = \mathbf{b'}$$

$\square$

故本節以下列定理總結：

定理 (綫性方程組的可解性II). 對於m行n列綫性方程組S，若$\text{RREF}(S)$中係數矩陣的非零行數$\leq n$同時大於或等於水平向量的非零行數，則S有解。

行列式

行列式為矩陣運算中一種重要的工具，其概念起源於解二元一次方程組：

$$\begin{cases} ax + by = \alpha \\ cx + dy = \beta \end{cases} \implies \begin{cases} x = \frac{\alpha d - \beta b}{ad - bc} \\ y = \frac{\alpha c - \beta a}{bc - ad} \end{cases}$$

從上可見$ad-bc$的值會完全影響方程的可解性，由此可定義係數矩陣$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}^{2\times 2}$時，$\mathbf{A}$的行列式$\det \mathbf{A} = |\mathbf{A}| = ad - bc$。

命題. 設$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in \mathbb{M}^{3\times 3}$，則$|A| = aei + bfg + cdh - ceg - bdi - afh$.

定義 12 (分塊矩陣). 任何矩陣$\mathbf{A} \in \mathbb{M}^{m\times n}$，都有$\mathbf{A}_1 \in \mathbb{M}^{s\times t}, \mathbf{A}_2 \in \mathbb{M}^{s\times (n-t)}, \mathbf{A}_3 \in \mathbb{M}^{(m-s)\times t}, \mathbf{A}_4 \in \mathbb{M}^{(m-s)\times(n-t)}$使得

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix}$$

定理. $\begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{a} & \mathbf{A} \end{vmatrix} = |\mathbf{A}|$.

證明. 基於行變換可得出$\begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{vmatrix}$。對任意

$$\left[\begin{array}{cc|c} \mathbf{I} & \mathbf{0} & \mathbf{b}_1 \\ \mathbf{0} & \mathbf{A} & \mathbf{b}_2 \end{array}\right]$$

由此可見，解的存在性基於$|\mathbf{A}|$變化。 □

定理. $|\mathbf{A}^T| = |\mathbf{A}|$.

證明. 證明太難，暫且不提。 □

定理. $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} + c_{i1} & \cdots & a_{in} + c_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ c_{i1} & \cdots & c_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix}$.

定理.

$$\begin{vmatrix} a_{11} & \cdots & a_{1l}+c_{1l} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{ml}+c_{ml} & \cdots & a_{mn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1l} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{ml} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & c_{1l} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & c_{ml} & \cdots & a_{mn} \end{vmatrix}.$$

定理.

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ \alpha a_{i1} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix}.$$

定理.

$$\begin{vmatrix} a_{11} & \cdots & \alpha a_{1l} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & \alpha a_{ml} & \cdots & a_{mn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1l} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{ml} & \cdots & a_{mn} \end{vmatrix}.$$

行列式亦可作相應的行列變換，因此有以下定理：

定理. 若行列式有兩行（或列）元素對應相等，則行列式等於 *0*。

衍理. 若行列式有兩行（或列）元素對應成比例，則行列式等於 *0*。

衍理. 行列式經過行列變換時，其值不變。

定理 (Laplace定理). $\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{a} & \mathbf{B} \end{vmatrix} = |\mathbf{A}||\mathbf{B}|.$

Cramer法則

設

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{1n}x_n &= b_2 \\ & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{cases}$$

則稱由係數構成的行列式

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

為方程組的係數行列式。

定理 (Cramer法則). 設綫性方程組的係數行列式$D \neq 0$，則方程組有唯一一組解

$$x_1 = \frac{D_1}{D},\, x_2 = \frac{D_2}{D},\, \cdots,\, x_n = \frac{D_n}{D},$$

其中D_k是把D的第k列元素依次換成常數項$b_1, b_2, \ldots, b_n$而得到的行列式

$$D_k = \begin{vmatrix} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & b_2 & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

其中$1 \leq k \leq n$。

證明. 證明分爲兩部分。

1. 先證明$x_1 = \frac{D_1}{D},\, x_2 = \frac{D_2}{D},\, \cdots,\, x_n = \frac{D_n}{D}$為方程解。設$1 \leq i \leq n$，并有$A_{ij}$為$a_{ij}$的餘子式，

$$\begin{aligned} L.H.S. &= \sum_{k=1}^{n} a_{ik} x_k = \sum_{k=1}^{n} a_{ik} \frac{D_k}{D} = \frac{1}{D} \sum_{k=1}^{n} a_{ik} D_k \\ &= \frac{1}{D} \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{n} b_j A_{jk} = \frac{1}{D} \sum_{j=1}^{n} b_j \sum_{k=1}^{n} a_{ik} A_{jk} \\ &= \frac{1}{D} \sum_{j=1}^{n} b_j \delta_{ij} D \\ &= b_i = R.H.S. \end{aligned}$$

2. 再證明此乃唯一解。設$x_k = c_k$, $k = 1, 2, \ldots$，則

$$D_k = \begin{vmatrix} a_{11} & \cdots & a_{1,k-1} & \sum_{k=1}^{n} a_{1k}c_k & a_{1,k+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & \sum_{k=1}^{n} a_{2k}c_k & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & \sum_{k=1}^{n} a_{nk}c_k & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1,k-1} & a_{1k}c_k & a_{1,k+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & a_{2k}c_k & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & a_{nk}c_k & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix} = c_k D$$

因此$c_k = \dfrac{D_k}{D}$。

$\square$

逆矩陣

設 $\mathbf{A} \in \mathbb{M}^{m \times n}, \mathbf{x} \in \mathbb{M}^{n \times 1}, \mathbf{b} \in \mathbb{M}^{m \times 1}$，設

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

若 $\mathrm{RREF}(\mathbf{A}) = \mathbf{I}$，則存在 $\mathbf{B} \in \mathbb{M}^{m \times n}$ 令

$$[\mathbf{A}|\mathbf{I}] \sim [\mathbf{I}|\mathbf{B}]$$

對符合條件的矩陣 $\mathbf{B}$，我們稱其爲**A的逆矩陣**，記 $\mathbf{B} = \mathbf{A}^{-1}$。

定理 (可逆矩陣). 若 $\mathbf{A} \in \mathbb{M}^{m \times n}$ 符合上述情況，則稱**A為可逆矩陣**。

衍理. 若 $\mathbf{A} \in \mathbb{M}^{m \times n}$ 可逆，則必有方陣 $\mathbf{B}$ 使得

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

示例. 設 $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，則

$$[\mathbf{A}|\mathbf{I}] = \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ ac & ad & 0 & a \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & ad-bc & -c & a \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} a & 0 & 1+\frac{bc}{ad-bc} & \frac{-ab}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

$$\sim \left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

因此，$\mathbf{A}^{-1} = \dfrac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

定義 13 (伴隨矩陣). 設 $\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$，則 $\mathbf{A}$ 的**伴隨矩陣** $\mathrm{adj}\mathbf{A}$ 被

定義爲

$$\mathrm{adj}\mathbf{A}_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{mn} \end{vmatrix}$$

示例 (2×2 矩陣的伴隨矩陣). 設 $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，則

$$\mathrm{adj}\mathbf{A} = \begin{bmatrix} \det[d] & -\det[c] \\ -\det[b] & \det[a] \end{bmatrix} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

由此可見，

$$\mathbf{A}^{-1} = \frac{1}{\det\{\mathbf{A}\}} (\mathrm{adj}\mathbf{A})^T$$

巧合的是，不僅二維方陣如此。

示例 (3×3 矩陣的伴隨矩陣). 設 $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$，則

$$\mathrm{adj}\mathbf{A} = \begin{bmatrix} \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ -\det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \\ \det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{bmatrix}$$

由此可見，

$$\mathbf{A}^{-1} = \frac{1}{\det\{\mathbf{A}\}} (\mathrm{adj}\mathbf{A})^T$$

因此，對於高維方陣而言，其中一個有效求逆矩陣的方法是伴隨矩陣法：

定理 (伴隨矩陣法). 設 $\mathbf{A} \in \mathbb{M}^{n \times n}$，則

$$\mathbf{A}^{-1} = \frac{1}{\det\{\mathbf{A}\}}(\mathrm{adj}\mathbf{A})^T$$

根據本節提及的逆矩陣的概念，我們可以以下方法解綫性方程組：已知綫性方程組 $S : \mathbf{A}\mathbf{x} = \mathbf{b}$ 有唯一解，則

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

矩陣的幾何含義

若考慮向量空間$\mathbb{R}^n$的元素$\vec{x} = (x_1, x_2, \ldots, x_n)$，則對於坐標變換，有以下看法：

Introduction

Mathematics starts from logic and observation of the nature, and splits into many fields of study. The main streams can be thought of as pure Mathematics and applied mathematics. In this article, we will examine a useful object that usually appears in pure geometry and applied learning over physics and computer science - the Transformation Matrix. We will go through its meaning and the derivation. More exploration is recommended if you find it useful.

Complex numbers

If we currently identify the complex plane $\mathbb{C}$ with the Euclidean plane $\mathbb{R}^2$, such that there is a relation $z \sim (x, y)$ by a mapping (function) $z = x + yi$, where i is the number for $i^2 = -1$. So we can see every complex number z with its distance to 0 being equal to $|z| := \sqrt{x^2 + y^2}$ and its argument (angle raising from x-axis) begin equal to $\arg z := \tan^{-1}(\frac{y}{x})$. Let's define $r = |z|$ and $\theta = \arg z$. Then by Taylor expansion centered at $z = 0$, we have

$$
\begin{aligned}
\sin\theta &= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \\
\cos\theta &= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} \\
e^{i\theta} &= \sum_{n=0}^{\infty} \frac{(i\theta)^n}{(n)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \\
&= \cos\theta + i\sin\theta
\end{aligned}
$$

Since for each $z \in \mathbb{C}$, we can write $z = x + yi = r\cos\theta + ir\sin\theta = re^{i\theta}$, we call this the Euler's formula for complex number, infamous for its consequential equation $e^{i\pi} + 1 = 0$.

The matrix for transformation

Recall that if we have a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we can conclude the following equation is generating new coordinate by old coordinate (x, y).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

in which if we put the equations into a system, we will obtain

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases}$$

We should establish that every transformation matrix can be represented by a suitable matrix, first. The converse may not be true, but that's not important to what we should do.

Translation

By a translation we mean moving (x, y) by a constant pair (a, b), and results in $(x + a, y + b)$ as a new coordinate. To do this, we extend our system to

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \\ 1 \end{bmatrix}$$

so that for each coordinate we directly associate with constant, and let the 1 remains itself by zeroing out x and y. The order triple presents the same coordinate by adjoining a 1 at the end of it. We shall call it the translation factor for the third entry, as we could always replace it by any other number for doing translation.

Dilation

A dilation is either enlarging the coordinate or contracting the coordinate. The meaning of enlargement and contraction should be referring to the distance between (x, y) and $(0, 0)$. To do this, we may simply multiply the matrix by a constant $k > 0$.

However, we may want to enlarge different coordinates by different values. We will make this happen if we put things in diagonal, so that the multiplication acts on

respective coordinates only once. It is interesting to see

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha x \\ \beta y \\ 1 \end{bmatrix}$$

Rotation

To rotate the coordinate, we may refer back to the Euler's fomrula and think of changing the value of $\arg z$. One should see that if we rotate the complex number by φ (we could omit the length as it should have no change), we will get the new number

$$z' = e^{i(\theta + \varphi)} = \cos(\theta + \varphi) + i\sin(\theta + \varphi)$$

In this sense, we consider the real part and imaginary part respectively, which corresponds to x and y coordinate. Then

$$\begin{aligned} x' &= \cos(\theta + \varphi) \\ &= \cos\theta\cos\varphi - \sin\theta\sin\varphi \\ &= (x, y) \cdot (\cos\varphi, -\sin\varphi) \\ y' &= \sin(\theta + \varphi) \\ &= \sin\theta\cos\varphi + \sin\theta\cos\varphi \\ &= (x, y) \cdot (\sin\varphi, \cos\varphi) \end{aligned}$$

Then we may write

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\varphi - y\sin\varphi \\ x\sin\varphi + y\cos\varphi \end{bmatrix}$$

To maintain the matrix as 3×3, we may consider

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x\cos\varphi - y\sin\varphi \\ x\sin\varphi + y\cos\varphi \\ 1 \end{bmatrix}$$

We may also identify the real part of the rotational matrix as

$$\begin{bmatrix} \cos\varphi & 0 \\ 0 & \cos\varphi \end{bmatrix}$$

and the imaginary part of the rotational matrix as

$$\begin{bmatrix} 0 & -\sin\varphi \\ \sin\varphi & 0 \end{bmatrix}$$

You may check that

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

Caution on the order of matrix operation

Usually, we would like to combine the use of translation, dilation and rotation as one stuff. However, their positioning means a lot to the resultant transformation. Let us define $\mathcal{T}$ the group of translation, $\mathcal{R}$ the group of rotation, and $\mathcal{D}$ the group of dilation. For $T_{a,b} \in \mathcal{T}$ we mean a translation by a units in x-direction and b units in y-direction; for $R_\theta \in \mathcal{R}$ we mean a rotation by θ; for $D_{\alpha,\beta} \in \mathcal{D}$ we mean a dilation of scale α in x-direction and scale β in y-direction.

We shall see and check that

$$T_{a,b} \circ R_\theta \neq R_\theta \circ T_{a,b}, D_{\alpha,\beta} \circ R_\theta \neq R_\theta \circ D_{\alpha,\beta}, T_{a,b} \circ D_{\alpha,\beta} \neq D_{\alpha,\beta} \circ T_{a,b}$$

In formal reading, we say the transformations are not commutative with one another. This means that whenever we are trying to apply the matrices to our calculation, we must follow the priority of transformation strictly.

Restriction among general geometry

As we see each of the translation, dilation and rotation can be represented using matrix operation, so it is fruitful to use matrices in order to perform transformation in Euclidean geometry. However, for general purpose, since matrix is indeed linear operators, it could not perform any nonlinear operations like inversion and Möbius transformation.

It is enough to use matrix operations to describe high-school geometry, so this is a good place to stop. For who are struggling with M2 trigonometry, it is a good perspective to learn more about the Euler's formula for complex number. You will benefit a lot from using it.

博弈論簡介

博弈論簡介