**7.3.3. Fourier Transform 的運用.** 回到一維的 heat equation，當傳導熱的棒子無限長，我們就不能用 Fourier series 處理。當棒子的一端在 $x=0$，另一端延伸到正無限大，此時可用 Fourier sine（或 cosine）transform；而當棒子的兩端各延伸到負無限大和正無限大時，就可以用 Fourier transform 處理。

首先我們看一端在 $x=0$ 的情況。我們考慮 boundary condition 為 $x=0$ 處溫度一直維持零度的情況，即 $u(0,t)=0$, $\forall t\geq 0$。而我們考慮的 initial condition 為 $u(x,0)=f(x)$, $\forall x\geq 0$。注意此時由 boundary condition，我們有 $f(0)=0$。由 separable variables 的方法 $u(x,t)=F(x)G(t)$，我們同樣得到 $F(x),G(t)$ 需符合 ODE：$F''=kF(x)$ 以及 $G'(t)=c^2kG(t)$。不過和前面不同的是，由於 boundary condition 限制較少，我們無法確定 $k$ 的範圍。雖然 $k\geq 0$ 時會推得 $u(x,t)$ 會隨時間 $t$ 遞增或 $x$ 遞增這類不合物理現象的解，所以我們可以不考慮，但即使 $k<0$，用 Fourier series 處理，必需假設 $f(x)$ 是週期函數才可行。所以在此情況，我們必須轉為用 Fourier sine（或 cosine）integral 處理。既是如此再加上 Fourier transform 對微分的作用相對簡單，我們就直接套用 Fourier sine transform。

首先注意，由於 $u(x,t)$ 是兩變數函數，我們取 Fourier sine transform 是將 $u$ 視為 $x$ 的函數（將 $t$ 視為常數）處理，也就是說 $\mathscr{F}_s(u)$ 會是一個以 $w,t$ 為變數的函數 $\hat{u}_s(w,t)$。由 heat equation $\frac{\partial u}{\partial t}=c^2\frac{\partial^2 u}{\partial x^2}$ 先對等式右邊取 Fourier sine transform 並利用其線性以及對二次微分的性質可得

$$\mathscr{F}_s(c^2\frac{\partial^2 u}{\partial x^2})=c^2(-w^2\mathscr{F}_s(u)+w\sqrt{\frac{2}{\pi}}u(0,t))=-c^2w^2\mathscr{F}_s(u)=-c^2w^2\hat{u}_s. \tag{7.1}$$

注意這裡我們已套用 boundary condition $u(0,t)=0$。接下來我們必須了解 Fourier sine transform 套用在 heat equation 的左邊 $\frac{\partial u}{\partial t}$ 其作用為何。依定義

$$\mathscr{F}_s(\frac{\partial u}{\partial t})=\sqrt{\frac{2}{\pi}}\int_0^\infty\frac{\partial u}{\partial t}(x,t)\sin wx\,dx=\sqrt{\frac{2}{\pi}}\frac{\partial}{\partial t}\int_0^\infty u(x,t)\sin wx\,dx=\frac{\partial}{\partial t}\mathscr{F}_s(u)=\frac{\partial\hat{u}_s}{\partial t}. \tag{7.2}$$

注意這裡類似 separable variables 的概念，我們假設 $u(x,t)$ 會使得上式中的積分和偏微是可交換的。

從 heat equation 等式兩邊取 Fourier sine transform（即式子 (7.1), (7.2)）我們得到 $\hat{u}_s(w,t)$ 所需符合的 PDE：

$$\frac{\partial\hat{u}_s}{\partial t}+c^2w^2\hat{u}_s=0.$$

由於這個偏微僅牽涉到變數 $t$，之前提過，我們可以將之視為變數 $t$ 的一階 ODE 來處理，解得

$$\hat{u}_s(w,t)=C(w)e^{-c^2w^2t}. \tag{7.3}$$

我們可以利用 initial condition：$u(x,0)=f(x)$ 決定 $C(w)$。依定義

$$\hat{u}_s(w,0)=\sqrt{\frac{2}{\pi}}\int_0^\infty u(x,0)\sin wx\,dx=\sqrt{\frac{2}{\pi}}\int_0^\infty f(x)\sin wx\,dx=\hat{f}_s(w).$$

所以將式子 (7.3) 代 $t=0$ 得 $C(w)=\hat{f}_s(w)$。確定 $\hat{u}_s(w,t)$ 後，我們就可以用 inverse Fourier sine transform 解得 $u(x,t)$，亦即

$$u(x,t)=\sqrt{\frac{2}{\pi}}\int_0^\infty \hat{u}_s(w,t)\sin wx\,dw=\sqrt{\frac{2}{\pi}}\int_0^\infty \hat{f}_s(w)e^{-c^2w^2t}\sin wx\,dw.$$

**Question 7.7.** 解 *heat equation*，其中 *boundary condition* 為 $u(0,t)=0,\ \forall t\geq 0$ 以及 *initial condition* 為 $u(x,0)=\begin{cases} k, & \text{if } \ 0<x<1; \\ 0, & \text{if } \ x>1. \end{cases}$ 。可直接套用 *Question 6.15*（課本習題 *11.7.17*）的結果。

最後，我們看另一種情況，即棒子的兩端各延伸到負無限大和正無限大的情況。此時已無 boundary condition，我們考慮 initial condition 為 $u(x,0)=f(x),\ \forall x\in\mathbb{R}$。由於 $f(x)$ 未必是奇或偶函數，我們必須用 Fourier transform 處理。

同樣的，取 Fourier transform 是將 $u$ 視為 $x$ 的函數（將 $t$ 視為常數）處理，也就是說 $\mathscr{F}(u)$ 會是一個以 $w,t$ 為變數的函數 $\hat{u}(w,t)$。由 heat equation $\dfrac{\partial u}{\partial t}=c^2\dfrac{\partial^2 u}{\partial x^2}$，等式兩邊套用 Fourier transform，和前面一樣利用 Fourier transform 線性以及對二次微分的性質，以及假設 $u(x,t)$ 會使得積分和偏微是可交換的情況下，我們得到 $\hat{u}(w,t)$ 所需符合的 PDE：

$$\frac{\partial \hat{u}}{\partial t}+c^2w^2\hat{u}=0.$$

由於這個偏微僅牽涉到變數 $t$，之前提過，我們可以將之視為變數 $t$ 的一階 ODE 來處理，解得

$$\hat{u}(w,t)=C(w)e^{-c^2w^2t}.$$

我們利用 initial condition：$u(x,0)=f(x)$ 得 $C(w)=\hat{f}(w)$，也因此利用 inverse Fourier transform 解得

$$u(x,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \hat{u}(w,t)e^{iwx}\,dw=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \hat{f}(w)e^{-c^2w^2t}e^{iwx}\,dw.$$

我們可以用 convolution 將 $u(x,t)$ 這個積分式再化簡。若可以找到函數 $g(x)$ 使其 Fourier transform $\hat{g}(w)=\dfrac{1}{\sqrt{2\pi}}e^{-c^2w^2t}$，此時 $u(x,t)=\displaystyle\int_{-\infty}^\infty \hat{f}(w)\hat{g}(w)e^{iwx}\,dw$，我們就可以利用式子 (6.18) 得到 $u=f*g(x)$。在課本 Sec. 11.10 的 Table III 中我們知 $\mathscr{F}(e^{-ax^2})=\dfrac{1}{\sqrt{2a}}e^{-w^2/(4a)}$. 所以只要令 $c^2t=\dfrac{1}{4a}$，即 $a=\dfrac{1}{4c^2t}$，可得 $\mathscr{F}(e^{-x^2/(4c^2t)})=c\sqrt{2t}e^{-c^2w^2t}$。因此，若令 $g(x)=\dfrac{e^{-x^2/(4c^2t)}}{2c\sqrt{\pi t}}$ 就可得 $\hat{g}(w)=\dfrac{1}{\sqrt{2\pi}}e^{-c^2w^2t}$。最後利用 convolution 的定義得

$$u(x,t)=f*g(x)=\frac{1}{2c\sqrt{\pi t}}\int_{-\infty}^\infty f(\tau)e^{-(x-\tau)^2/(4c^2t)}\,d\tau.$$

**Example 7.3.5** (課本 Example 12.7.1)**.** 假設兩端無限延伸的棒子，其初始溫度函數為 $f(x)=\begin{cases} k, & \text{if } |x|<1; \\ 0, & \text{if } |x|>1. \end{cases}$ 可得 $u(x,t)=\dfrac{k}{2c\sqrt{\pi t}}\displaystyle\int_{-1}^1 e^{-(x-\tau)^2/(4c^2t)}\,d\tau$. 再利用變數變換 $z=\dfrac{\tau-x}{2c\sqrt{t}}$，我們有 $d\tau=2c\sqrt{t}\,dz$，故得 $u(x,t)=\dfrac{k}{\sqrt{\pi}}\displaystyle\int_{(-1-x)/(2c\sqrt{t})}^{(1-x)/(2c\sqrt{t})} e^{-z^2}\,dz.$ ♯

**Question 7.8.** 做課本習題 *12.7.13.*

Chapter 8

# 線性代數

這一章將用簡單快速的方式介紹線性代數（linear algebra）基本的概念。從大家最熟悉的解聯立方程開始，連結矩陣的運算以及行列式，一直到對角化的課題。內容著重於處理問題的原理及方法，而略去正式的證明。若需要對線性代數有更進一步的了解，可參考我的網站上提供的“線性代數初步”講義。

## 8.1. System of Linear Equations

當我們有多個 $n$ 元一次的方程式要討論它們的共同解時, 就稱為解一次聯立方程組 (system of linear equations). 一般抽象的表示法

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

表示有 $m$ 個 $n$ 元一次方程式所成的方程組 (system of $m$ linear equations in $n$ variables)。這裡 $a_{ij}, b_i$ 皆為實數, 這些實數才是真正影響到聯立方程組的因素, 所以我們也可特別把它們標明出來, 令

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}; \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},
$$

然後將上面的聯立方程組用 $A\mathbf{x} = \mathbf{b}$ 來表示. 矩陣 $A$ 中的每一個 $a_{ij}$ 稱為 $A$ 的一個 *entry*. 因為 $A$ 的每一個 entry 對應到聯立方程組中某個未知數的係數, 通常我們會稱矩陣 $A$ 為此聯立方程式的係數矩陣. 一個矩陣的一個橫排稱為一個 row (列), 而一個豎排稱為一個 column (行). 我們算 row 時是從上而下來數的；而算 column 是由左而右來數的。大家可以看出矩陣 $A$ 的 row 對應的就是此聯立方程組的方程式, 第一個 row 對應到第一個方程式, 依此類推。而 column 對應到的是方程組的未知數, 第一個 column 對應到的是未知數 $x_1$ 的係數, 依此類推。因為這裡是由 $m$ 個方程式而且每個方程式有 $n$ 個未知數所組成的聯立方

程組, 所以 $A$ 共有 $m$ 個 row 以及 $n$ 個 column, 我們稱這樣的矩陣為 $m\times n$ matrix. 注意這裡 $\mathbf{x}$ 表示是一個未知的向量而且我們將向量 $\mathbf{x},\mathbf{b}$ 都寫成 column vector (行向量) 是為了配合將來矩陣乘法的寫法. 目前大家只要記住這也是聯立方程式的一種表示法即可.

例如解聯立方程組

$$\begin{aligned} 3x_1-2x_2+9x_4 &= 4 \\ 2x_1+2x_2-4x_4 &= 6 \end{aligned} \tag{8.1}$$

我們就可以表成

$$\begin{bmatrix} 3 & -2 & 0 & 9 \\ 2 & 2 & 0 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

注意這裡係數矩陣多出 $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 這個 column 因為 $x_3$ 的係數為 0.

**8.1.1. Elementary Row Operations.** 過去學習解一次聯立方程組的方法不外加減消去法或高斯消去法, 它們的原理都是一樣的, 即利用以下三種基本方法:

(1) 變換式子的順序

(2) 將某一式乘上一非零實數

(3) 將某一式乘上一實數後加到另一式上

利用這三種基本方法將方程式的某些變數消去, 最後求出解來. 我們將介紹一個有系統的方法來解聯立方程組, 把這三種基本方法看成是對矩陣的運算.

當我們要解

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

這一個聯立方程組時, 先寫出如下的 augmented matrix (增廣矩陣)

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

例如式子 (8.1) 中的聯立方程組所對應的 augmented matrix 為

$$\left[\begin{array}{cccc|c} 3 & -2 & 0 & 9 & 4 \\ 2 & 2 & 0 & -4 & 6 \end{array}\right]$$

換言之, 若我們要解 $A\mathbf{x}=\mathbf{b}$ 這一個聯立方程組, 就要寫下 $[A\mid\mathbf{b}]$ 這一個 matrix. 反之一個 augmented matrix $[A\mid\mathbf{b}]$ 就對應到一個聯立方程組 $A\mathbf{x}=\mathbf{b}$.

接下來我們將如加減消去法的三種步驟, 利用所謂的 elementary row operation (基本列運算) 處理這個 augmented matrix. 所謂 elementary row operation 即表示對矩陣進行如下三種的列運算:

(1) 將矩陣的某兩個 row 對調

(2) 將矩陣的某一個 row 乘上一非零實數

(3) 將矩陣的某一個 row 乘上一實數後加到另一個 row.

為了方便起見, 我們將上面 (1), (2), (3) 三種 elementary row operation 分別稱為 *type 1*, *type 2* 以及 *type 3* 的 elementary row operation.

**Example 8.1.1.** 考慮矩陣

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

將 $A$ 的第一、第二兩個 row 互換 (即做一個 type 1 的 elementary row operation), 可得

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

而將 $B$ 的第二個 row 乘上 2 (即做一個 type 2 的 elementary row operation), 可得

$$C = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 2 & 4 & 6 & 8 \\ 4 & 0 & 1 & 2 \end{bmatrix}.$$

而將 $C$ 的第三個 row 乘上 $-3$ 加到第一個 row (即做一個 type 3 的 elementary row operation), 可得

$$D = \begin{bmatrix} -10 & 1 & -4 & -3 \\ 2 & 4 & 6 & 8 \\ 4 & 0 & 1 & 2 \end{bmatrix}.$$

$\sharp$

**8.1.2. 解聯立方程組.** 大家應很容易看出一個 augmented matrix 經過上節所提的 elementary row operation 後所得的 augmented matrix 所對應的聯立方程組就是大家熟悉的加減消去法的三種步驟所得的方程組. 利用加減消去法最常遇到的問題就是 (尤其在處理未知數很多的方程組時), 常常做了幾次後, 混亂到不知道那些式子是消過了以及那些式子還可以再進一步消減. 還有就是, 到底要將方程組的式子消到哪種地步時, 才可以解出方程組. 關於第一個問題, 我們可以理解用矩陣的表示法就可以把這消去的過程記錄下來. 而接下來我們要探討的就是第二個問題, 也就是將矩陣化成某種特定的形式就可以解出方程式來.

我們的目的就是要將 augmented matrix $[A \mid \mathbf{b}]$ 中的係數矩陣 $A$ 利用這三種 elementary row operation 化成所謂的 *echelon form*.

我們先解釋一下何謂 echelon form. 首先我們將矩陣每一個 row 從左到右來看第一個不為 0 的項稱為這個 row 的 *leading entry*. 因為係數矩陣中的每一個 entry 對應到聯立方程組中某個 variable (未知數) 的係數, 所以 leading entry 若是 variable $x_i$ 的係數, 我們就說這個 leading entry 發生在 $x_i$ 的位置. 要注意, 這也等同於這個 leading entry 是位於從左到

右算來第 $i$ 個 column. 例如矩陣

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 0 & 0 & 5 & 0 & 2 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}$$

第一個 row 的 leading entry 為 1 且發生在 $x_1$ 的位置, 而第二個 row 和第三個 row 的 leading entry 分別為 5 和 1 且發生的位置皆在 $x_3$。

所謂一個矩陣是 echelon form 表示這個矩陣沒有 leading entry 的 row (即該 row 每一項皆為 0) 必需在最下方, 而有 leading entry 的 row 其 leading entry 所在位置從上到下來看是往右移的. 換言之, 若上一個 row 的 leading entry 所在的位置是 $x_i$, 而下一個 row 的 lading entry 所在的位置是 $x_j$, 則必需 $i < j$. 例如上一個矩陣並非 echelon form, 因為第 3 個 row 和第 2 個 row 的 leading entry 的位置皆為 $x_3$, 並未右移. 另外矩陣

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

都不是 echelon form, 因為前一個矩陣全為 0 的 row 並未置於最下方, 而後一個矩陣第 3 個 row 的 leading entry 在第 2 個 row 的 leading entry 的左方. 至於矩陣

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

就是 echelon form. 當一個矩陣是 echelon form 時, 我們稱每一個 row 的 leading entry 為 *pivot*, 而 pivot 所在的位置我們稱為 *pivot variable*.

我們要強調, 絕不會有 pivot 的個數多於方程組 variables (未知數) 的個數的情形發生. 這是因為當係數矩陣 $A$ 是 echelon form 時, 每一個 column 最多僅能有一個 pivot (因為不能有兩個 leading entry 在同一個 variable), 所以 pivot 的個數不能多於 column 的個數. 而 $A$ 的 column 個數表示的就是此聯立方程組 variables 的個數, 因此 pivot 的個數不會多於 variables 的個數. 另一方面依定義每一個 row 最多僅能有一個 pivot, 所以 pivot 的個數也不會多於該方程組的方程式個數 (即係數矩陣 row 的個數).

很容易就可以發現一個矩陣利用 elementary row operations 化成的 echelon form 的方法有很多種，而所化成的 echelon form 也有可能不同。不過同一個矩陣不管用哪些 elementary row operations 化成的 echelon form 它們的 pivot variables 皆會相同（證明從略)。也因此一個矩陣化為 echelon form 其 pivot 的個數是固定的, 我們特別有以下的定義.

**Definition 8.1.2.** 假設 $A$ 為一矩陣. 若 $A$ 利用 elementary row operations 化為 echelon form 後其 pivot 的個數為 $r$, 我們稱 $r$ 為 $A$ 的 *rank*. 用 $\operatorname{rank}(A) = r$ 來表示.

一個矩陣的 rank 是該矩陣很重要的資訊，以後我們會探討它很多的性質。由於它的重要性，我們特別在這裡先介紹它的定義，讓大家先熟悉一下。

當我們將 augmented matrix $[A \mid \mathbf{b}]$ 利用 elementary row operation 將之化成 $[A' \mid \mathbf{b}']$ 且 $A'$ 為 echelon form 後. $A'$ 有兩種情形. 一種情形為 $A'$ 每一個 row 皆不全為 0; 另一種為 $A'$ 有些 row 全為 0. 我們分別依這兩種情形來討論聯立方程組的解.

(1) $A'$ 每一個 row 皆不全為 0: 此時聯立方程組為 *consistent*, 即一定有解. 我們又可細分成兩種情況.

(a) 第一種情況是每一個變數 (variable) $x_i$ 皆為 pivot variable. 亦即 pivot 的個數等於方程組未知數的個數 (即係數矩陣 $A$ 的 column 個數). 例如

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & 1 \end{array}\right]$$

此時 echelon form 的 pivot variable 分別為 $x_1, x_2, x_3$ 恰就是聯立方程組的未知數 $x_1, x_2, x_3$. 在這種情況之下此聯立方程組會有唯一解, 而且我們可利用從下往上 “代回” 的方式求得解. 例如前面的 augmented matrix 所對應的聯立方程組為

$$\begin{array}{rrrcl} 2x_1 & +x_2 & +x_3 & = & 4 \\ & 3x_2 & +x_3 & = & 2 \\ & & -x_3 & = & 1 \end{array}$$

所以我們從最下面的 $-x_3 = 1$ 可得 $x_3 = -1$. 再將 $x_3 = -1$ 代入其上一式 $3x_2 + x_3 = 2$, 得 $3x_2 - 1 = 2$, 即 $x_2 = 1$. 最後將 $x_3 = -1, x_2 = 1$ 代入 $2x_1 + x_2 + x_3 = 4$, 得 $x_1 = 2$. 故得其解為 $x_1 = 2, x_2 = 1, x_3 = -1$.

(b) 第二種情況是有些 variable $x_i$ 不是 pivot variable. 也就是方程組未知數的個數多於 pivot 的個數. 例如

$$\left[\begin{array}{cccc|c} 2 & 1 & 3 & 1 & 4 \\ 0 & 3 & 3 & 1 & 2 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right]$$

此時 echelon form 的 pivot variable 分別為 $x_1, x_2, x_4$ 少於立方程組的未知數 $x_1, x_2, x_3, x_4$. 在此情形之下此聯立方程組會有無窮多解. 要得到這種方程組所有的解, 首先我們要找到 *free variables.* 所謂 free variable 指的是方程組不是 pivot variable 的 variable. 例如前面這個例子, $x_3$ 就是 free variable. Free variable 意指它可以任意取值, 所以找到 free variables 後你可以給它們任意的參數, 然後再利用如上一情況中由下往上代回的方式找到聯立方程組所有的解. 例如上一個 augmented matrix 所對應的聯立方程組為

$$\begin{array}{rrrrcl} 2x_1 & +x_2 & +3x_3 & +x_4 & = & 4 \\ & 3x_2 & +3x_3 & +x_4 & = & 2 \\ & & & -x_4 & = & 1 \end{array}$$

首先令 free variable $x_3$ 為一參數 $t$ (表示它可以是任意實數 $t \in \mathbb{R}$). 接著我們從最下面的 $-x_4 = 1$ 可得 $x_4 = -1$. 再將 $x_3 = t, x_4 = -1$ 代入其上一式 $3x_2 + 3x_3 + x_4 = 2$, 得 $3x_2 + 3t - 1 = 2$, 即 $x_2 = 1 - t$. 最後將 $x_2 = 1 - t, x_3 = t, x_4 = -1$ 代入 $2x_1 + x_2 + 3x_3 + x_4 = 4$, 得 $x_1 = 2 - t$. 故得其解為 $x_1 = 2 - t, x_2 =$

$1-t, x_3=t, x_4=-1$, 其中 $t$ 為任意實數. 因為 $t$ 可以是任意實數, 由此我們也知此方程組有無窮多解.

(2) $A'$ 有些 row 全為 0: 此時聯立方程組可能無解, 我們分成兩種情況:

(a) $A'$ 有一個 row 全為 0 但 $\mathbf{b}'$ 在該 row 不為 0. 例如

$$[A' \mid \mathbf{b}'] = \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$A'$ 最後一個 row 皆為 0, 但 $\mathbf{b}'$ 在該 row 的位置為 1. 在此情形之下聯立方程組為 *inconsistent*, 即無解. 例如上一個 augmented matrix 其最後一個 row 所對應的方程式為

$$0x_1+0x_2+0x_3=1$$

但不管 $x_1, x_2, x_3$ 代任何的實數都無法滿足 $0x_1+0x_2+0x_3=1$, 所以此方程組無解.

(b) $A'$ 全為 0 的 row, $\mathbf{b}'$ 在該 row 亦為 0. 例如

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{array}\right], \quad \left[\begin{array}{cccc|c} 2 & 1 & 3 & 1 & 4 \\ 0 & 3 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

這兩個 augmented matrices 皆為這種情形. 在此情形之下聯立方程組一定是 consistent. 事實上在此情形我們可以忽略全為 0 的 row, 例如前兩個 augmented matrices 所對應的方程組和

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 3 & 2 \end{array}\right], \quad \left[\begin{array}{cccc|c} 2 & 1 & 3 & 1 & 4 \\ 0 & 3 & 3 & 1 & 2 \end{array}\right]$$

所對應的方程組一樣. 所以我們可依前面 (1) $A'$ 每一個 row 皆不全為 0 的情況找出聯立方程組所有的解.

我們總結, 當 $A$ 是 echelon form 時, 找出 $A\mathbf{x}=\mathbf{b}$ 所有的解的方法. 首先, 當 $A$ 有一個 row 全為 0 但 $\mathbf{b}$ 在該 row 不為 0 時, 我們知該方程組無解, 所以我們僅需討論其他的情況 (即有解的情況). 此時我們先挑出 free variable (即非 pivot variable). 由於 free variable 可以任意取值, 一般來說我們會用一些參數表示之 (注意不同的 free variable 要用不同的參數代號). 接著, 我們由下而上, 從最大編號的 pivot variable 開始, 利用 free variables 的那些參數將它的值寫下來, 再依序寫出所有 pivot variables 的值.

**Example 8.1.3.** Solve the linear system

$$\begin{array}{rrrrcr} x_1 & -2x_2 & +1x_3 & -1x_4 & = & 4 \\ 2x_1 & -3x_2 & +4x_3 & -3x_4 & = & -1 \\ 3x_1 & -5x_2 & +5x_3 & -4x_4 & = & 3 \\ -x_1 & +1x_2 & -3x_3 & +2x_4 & = & 5. \end{array}$$

此聯立方程組的 augmented matrix 為

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 2 & -3 & 4 & -3 & -1 \\ 3 & -5 & 5 & -4 & 3 \\ -1 & 1 & -3 & 2 & 5 \end{array}\right].$$

第二, 三, 四 row 的 leading entry 需被消去. 故將第一 row 分別乘上 $-2,-3,1$ 加到第二, 三, 四 row 得

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & -1 & -2 & 1 & 9 \end{array}\right].$$

接下來第三, 四 row 的 leading entry 需要消去, 所以將第二 row 分別乘上 $-1,1$ 加到第三, 四 row 得

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -1 & 4 \\ 0 & 1 & 2 & -1 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

這是 echelon form. 由於係數矩陣全為 0 的第三, 四 row 全為 0, 知此 linear system 為 consistent.

事實上此 linear system 的 pivot variables 為 $x_1,x_2$, 而 free variables 為 $x_3,x_4$. 我們可以令 $x_4=r$, $x_3=s$, 代入第二 row 表示的 $x_2+2x_3-x_4=-9$, 得 $x_2=-9+r-2s$. 再代入第一 row 表示的 $x_1-2x_2+x_3-x_4=4$, 得 $x_1=-14+3r-5s$. 故知此 linear system 的解為

$$(x_1,x_2,x_3,x_4)=(-14+3r-5s,-9+r-2s,s,r), r,s\in\mathbb{R}.$$

通常我們習慣寫成 column vector 且將 $r,s$ 提出. 故將解寫成

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -14 \\ -9 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, r,s\in\mathbb{R}.$$

♯

## 8.2. 矩陣的運算

在本節中我們將簡單地回顧有關於矩陣的定義。若一矩陣由 $m$ 個 row 和 $n$ 個 column 的數所組成, 我們便稱該矩陣為一個 $m\times n$ matrix. 特別的, 一個 $n\times n$ matrix (即 row 的個數等於 column 的個數), 我們稱之為“方陣” *square matrix*. 在本講義中，我們用 $M_{m\times n}$ 來表示所有實係數的 $m\times n$ 矩陣所成的集合，且大寫的英文字母來表示一個矩陣. 例如考慮

$$A=\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 5 & 8 \\ 2 & 1 & 1 & 0 \end{bmatrix}, \tag{8.2}$$

則 $A\in M_{3\times 4}$。當我們要抽象地描述一般矩陣時, 我們也常用 $A=[a_{ij}]$ 這樣的方法來描述. 這種表示法意指 $A$ 中在第 $i$ 個 row 和 $j$ 個 column 的位置我們用 $a_{ij}$ 來表示, 並稱之為此矩陣

的 $(i,j)$-th entry. 因此當我們說 $A=[a_{ij}]$ 為 $m\times n$ 矩陣, 這表示 $1\leq i\leq m$ 且 $1\leq j\leq n$. 例如對於式子 (8.2) 中的矩陣 $A$, 若 $A=[a_{ij}]$, 則

$$\begin{array}{llll} a_{11}=1, & a_{12}=0, & a_{13}=2, & a_{14}=3, \\ a_{21}=0, & a_{22}=1, & a_{23}=5, & a_{24}=8, \\ a_{31}=2, & a_{32}=1, & a_{33}=1, & a_{34}=0. \end{array}$$

另外為了方便起見, 我們也會將矩陣 $A$ 的每一個 row 和 column 用向量的方法來表示, 這些稱為 $A$ 的 row vectors 和 column vectors. 在本講義我們會將矩陣 $A=[a_{ij}]$ 第 $i$ 個 row 所成的 row vector 用 ${}_i\mathbf{a}$ 來表示, 而第 $j$ 個 column 所成的 column vector 用 $\mathbf{a}_j$ 來表示. 例如對於式子 (8.2) 中的矩陣 $A$, 我們有

$${}_1\mathbf{a}=\begin{bmatrix}1 & 0 & 2 & 3\end{bmatrix},\quad {}_2\mathbf{a}=\begin{bmatrix}0 & 1 & 5 & 8\end{bmatrix},\quad {}_3\mathbf{a}=\begin{bmatrix}2 & 1 & 1 & 0\end{bmatrix}$$

以及

$$\mathbf{a}_1=\begin{bmatrix}1\\0\\2\end{bmatrix},\quad \mathbf{a}_2=\begin{bmatrix}0\\1\\1\end{bmatrix},\quad \mathbf{a}_3=\begin{bmatrix}2\\5\\1\end{bmatrix},\quad \mathbf{a}_4=\begin{bmatrix}3\\8\\0\end{bmatrix}.$$

注意由於我們也想將向量看成是一個矩陣, 這裡的 row vectors 和 column vectors 都用矩陣的形式呈現.

我們也延伸向量加法與係數積的定義來定義矩陣的加法與係數積. 也就是說只有同為 $m\times n$ 階矩陣我們才定義它們之間的加法, 且兩矩陣相加表示將這兩個矩陣在相同位置的數加起來. 而一個實數乘上一個矩陣即為將該矩陣每一個位置上的數乘上該實數. 具體來說我們有以下的定義.

**Definition 8.2.1.** 假設 $A=[a_{ij}]$, $B=[b_{ij}]$ 皆為 $m\times n$ matrix. 定義 $A+B=[c_{ij}]$, 其中對所有的 $1\leq i\leq m$ 以及 $1\leq j\leq n$ 皆有 $c_{ij}=a_{ij}+b_{ij}$. 對任意實數 $r$, 我們定義 $rA=[d_{ij}]$ 其中對所有的 $1\leq i\leq m$ 以及 $1\leq j\leq n$ 皆有 $d_{ij}=ra_{ij}$.

Definition 8.2.1 告訴我們若

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

則

$$A+B=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

且

$$rA=\begin{bmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & \cdots & ra_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \cdots & ra_{mn} \end{bmatrix}$$

接著我們定義矩陣間的乘法. 首先回顧當我們要解聯立方程組

$$\begin{array}{rrcr} x_1 & +2x_2 & = & 1 \\ -x_1 & +1x_2 & = & 2 \\ 2x_1 & -2x_2 & = & 3 \end{array}$$

我們會把它寫成

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

所以若我們定義矩陣 $\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix}$ 和 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的乘法為

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

那麼聯立方程組與矩陣的關係就不只是為了列式方便而已, 聯立方程組和矩陣的運算產生了緊密的關係. 從這個角度出發, 我們有以下定義.

**Definition 8.2.2.** 設 $A = [a_{ij}]$ 為 $m \times n$ matrix 以及 $\mathbf{b} = [b_j]$ 為 $n \times 1$ matrix (即 $\mathbb{R}^n$ 中的 column vector). 若 $\mathbf{a}_i$ 表示 $A$ 的 $i$-th column 則定義

$$A\mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_n\mathbf{a}_n.$$

注意依此定義, 必需 $A \in M_{m\times n}$ 的 column 的個數 $n$ 等於 $\mathbf{b} \in M_{n\times 1}$ 的 row 的個數 $n$, 才能定義 $A\mathbf{b}$ 且此時 $A\mathbf{b}$ 會是 $m \times 1$ matrix (即 $\mathbb{R}^m$ 中的 column vector). 觀察此 column vector, 我們有

$$A\mathbf{b} = b_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + b_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1a_{11} + b_2a_{12} + \cdots + b_na_{1n} \\ b_1a_{21} + b_2a_{22} + \cdots + b_na_{2n} \\ \vdots \\ b_1a_{m1} + b_2a_{m2} + \cdots + b_na_{mn} \end{bmatrix}. \tag{8.3}$$

也就是說 $A\mathbf{b}$ 這一個 $m \times 1$ matrix 的 $i$-th entry 為 ${}_i\mathbf{a}\,\mathbf{b}$ 也就是 $A$ 的 $i$-th row ${}_i\mathbf{a}$ 和 $\mathbf{b}$ 的內積.

現在我們將矩陣乘法推廣到更一般的情況, 當 $A = [a_{ij}]$ 是一個 $m \times n$ matrix, $B = [b_{jk}]$ 是一個 $n \times l$ matrix. 由於對 $B$ 的每一個 column vector $\mathbf{b}_k \in M_{n\times 1}$, $1 \le k \le l$, 我們已定義了 $A\mathbf{b}_k$ 為何, 現在我們定義 $AB$ 為 $m \times l$ matrix, 其中 $AB$ 的 $k$-th column vector 為 $A\mathbf{b}_k$. 我們大致上有以下的圖示.

$$A \begin{bmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_l \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_l \\ | & | & & | \end{bmatrix}$$

由於 $A\mathbf{b}_k$ 為 $m \times 1$ matrix, 依此定義確實 $AB$ 為 $m \times l$ matrix. 現在我們來看正式的定義.

**Definition 8.2.3.** 設 $A=[a_{ij}]$ 為 $m\times n$ matrix 以及 $B=[b_{jk}]$ 為 $n\times l$ matrix, 則定義 $AB=C=[c_{ik}]$ 為 $m\times l$ matrix, 其中對於 $1\le k\le l$, $C$ 的 $k$-th column $\mathbf{c}_k$ 為

$$\mathbf{c}_k=A\mathbf{b}_k=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}\begin{bmatrix} b_{1k}\\ b_{2k}\\ \vdots\\ b_{nk}\end{bmatrix}=b_{1k}\mathbf{a}_1+b_{2k}\mathbf{a}_2+\cdots+b_{nk}\mathbf{a}_n. \tag{8.4}$$

由此定義, 我們知對於 $1\le i\le m$, $1\le k\le l$, $AB$ 的 $(i,k)$-th entry 應為其 $k$-th column (即 $A\mathbf{b}_k$) 從上往下算的第 $i$ 個 entry。此即 $A$ 的 $i$-th row ${}_i\mathbf{a}$ 和 $B$ 的 $k$-th column $\mathbf{b}_k$ 看成向量後取內積. 換言之, 若 $AB=[c_{ik}]$, 則 $AB$ 的 $(i,k)$-th entry $c_{ik}$ 為

$$c_{ik}={}_i\mathbf{a}\,\mathbf{b}_k=a_{i1}b_{1k}+a_{i2}b_{2k}+\cdots+a_{in}b_{nk}=\sum_{j=1}^{n}a_{ij}b_{jk}. \tag{8.5}$$

再次強調一次, 並不是任取兩個矩陣都可以定義乘法, 必須是左邊矩陣的 column 個數和右邊矩陣的 row 個數相同才能相乘.

**Example 8.2.4.** 令

$$A=\begin{bmatrix} -2 & 4\\ 3 & 6\\ 2 & 2\end{bmatrix}, B=\begin{bmatrix} 4 & -1 & 2 & 5\\ 3 & 0 & 1 & 1\end{bmatrix}$$

考慮矩陣乘法 $AB$. 依定義矩陣 $AB$ 的 3-rd column 為

$$A\mathbf{b}_3=\begin{bmatrix} -2 & 4\\ 3 & 6\\ 2 & 2\end{bmatrix}\begin{bmatrix} 2\\ 1\end{bmatrix}=2\mathbf{a}_1+1\mathbf{a}_2=2\begin{bmatrix} -2\\ 3\\ 2\end{bmatrix}+1\begin{bmatrix} 4\\ 6\\ 2\end{bmatrix}=\begin{bmatrix} 0\\ 12\\ 6\end{bmatrix}.$$

所以 $AB$ 的 $(2,3)$ entry 為 12 等於 $A$ 的 2-nd row 和 $B$ 的 3-rd column 看成 $\mathbb{R}^2$ 中的向量所得的內積, 即 $(3,6)\cdot(2,1)=12$. 事實上我們有

$$AB=\begin{bmatrix} -2 & 4\\ 3 & 6\\ 2 & 2\end{bmatrix}\begin{bmatrix} 4 & -1 & 2 & 5\\ 3 & 0 & 1 & 1\end{bmatrix}=\begin{bmatrix} 4 & 2 & 0 & -6\\ 30 & -3 & 12 & 21\\ 14 & -2 & 6 & 12\end{bmatrix}.$$

$\sharp$

接下來我們來看一些矩陣加法和乘法相關的性質：假設 $A,A'\in M_{m\times n}$, $B,B'\in M_{n\times l}$ 以及 $C\in M_{l\times k}$ 且 $r,r'\in\mathbb{R}$. 則

(1) $(rA+r'A')B=rAB+r'A'B$ 以及 $A(rB+r'B')=rAB+r'AB'$

(2) $(AB)C=A(BC)$

我們要強調的是矩陣乘法雖具有許多和實數乘法類似的性質, 但它卻沒有交換律. 事實上有可能 $A$ 乘以 $B$ 有定義, 但 $B$ 卻不能乘以 $A$, 例如 $A\in M_{2\times 3}$, $B\in M_{3\times 4}$ 的情形. 也有可能即使 $A$ 乘以 $B$ 和 $B$ 乘以 $A$ 都有定義, 但由於乘了以後階數不同, 仍會使得 $AB\neq BA$, 例如 $A\in M_{2\times 3}, B\in M_{3\times 2}$ 的情形. 僅有在 $A,B$ 為同階方陣（行、列數相等的矩陣）時, 才有可能使得 $AB$ 和 $BA$ 的階數相同. 但此時仍有可能 $AB\neq BA$, 例如

$$A=\begin{bmatrix} a & b\\ c & d\end{bmatrix}, B=\begin{bmatrix} 1 & 0\\ 0 & -1\end{bmatrix}, AB=\begin{bmatrix} a & -b\\ c & -d\end{bmatrix}, BA=\begin{bmatrix} a & b\\ -c & -d\end{bmatrix},$$

這種情形只有在 $b=c=0$ 時, 才會使得 $AB=BA$. 所以在處理矩陣乘法時要特別小心. 例如當 $A,B$ 為同階方陣時可推得 $(A-B)(A+B)=A^2-AB+BA-B^2$, 但由於可能 $AB\neq BA$, 我們不見得會有 $(A-B)(A+B)=A^2-B^2$.

另一個關於矩陣的重要運算就是取 transpose (即轉置矩陣) 的概念。對於一個 $m\times n$ matrix, 簡單來說其轉置矩陣就是將此矩陣的 row 與 column 的腳色互換, 也就是說將 row vectors 依序換成 column vectors. 我們有以下的定義.

**Definition 8.2.5.** 給定 $A\in M_{m\times n}$. 定義 $A^{\mathrm{t}}\in M_{n\times m}$, 其中對於 $1\leq i\leq m$, $A^{\mathrm{t}}$ 的 $i$-th column 就是將 $A$ 的 $i$-th row 寫成 column vector. 我們稱 $A^{\mathrm{t}}$ 為 $A$ 的 *transpose*.

**Example 8.2.6.** 令

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

依定義 $A^{\mathrm{t}}$ 應為 $3\times 2$ matrix. 其中 $A^{\mathrm{t}}$ 的第一個 column 為 $A$ 的第一個 row $\begin{bmatrix}1&2&3\end{bmatrix}$ 寫成 column vector, 即 $\begin{bmatrix}1\\2\\3\end{bmatrix}$. 同理 $A^{\mathrm{t}}$ 的第二個 column 為 $A$ 的第二個 row $\begin{bmatrix}-1&-2&-3\end{bmatrix}$ 寫成 column vector, 即 $\begin{bmatrix}-1\\-2\\-3\end{bmatrix}$. 故得

$$A^{\mathrm{t}}=\begin{bmatrix}1&-1\\2&-2\\3&-3\end{bmatrix}.$$

注意, $A^{\mathrm{t}}$ 的 1-st, 2-nd 和 3-rd row 也恰為 $A$ 的 1-st, 2-nd 和 3-rd column 寫成 row 而得. $\sharp$

現在我們來看矩陣取 transpose 的基本性質. 假設 $A,B$ 為 $m\times n$ matrix, $C$ 為 $n\times l$ matrix. 我們有以下之性質.

(1) $(A^{\mathrm{t}})^{\mathrm{t}}=A$.

(2) $(A+B)^{\mathrm{t}}=A^{\mathrm{t}}+B^{\mathrm{t}}$.

(3) $(AC)^{\mathrm{t}}=C^{\mathrm{t}}A^{\mathrm{t}}$.

一個方陣 $A$, 若滿足 $A^{\mathrm{t}}=A$, 我們稱 $A$ 為"對稱矩陣" *symmetric matrix*。例如"對角矩陣" *diagonal matrix*(對角線以外皆為 0 的矩陣)就是 symmetric matrix。以後我們會學到 symmetric matrix 的重要性質,所以當一個矩陣不是對稱時,我們希望一些方式利用它得到對稱矩陣。例如當方陣 $A$ 不是對稱時,我們可以考慮 $A+A^{\mathrm{t}}$,此時由轉置的性質 $(A+A^{\mathrm{t}})^{\mathrm{t}}=A^{\mathrm{t}}+(A^{\mathrm{t}})^{\mathrm{t}}=A^{\mathrm{t}}+A$,可知 $A+A^{\mathrm{t}}$ 為 symmetric。甚至當 $A$ 不是方陣時,我們也可考慮矩陣 $A^{\mathrm{t}}A$,此時 $(A^{\mathrm{t}}A)^{\mathrm{t}}=A^{\mathrm{t}}(A^{\mathrm{t}})^{\mathrm{t}}=A^{\mathrm{t}}A$,所以 $A^{\mathrm{t}}A$ 會是對稱矩陣。同理 $AA^{\mathrm{t}}$ 也會是對稱矩陣。不過要注意 $A^{\mathrm{t}}A$ 和 $AA^{\mathrm{t}}$ 一般來說是不一樣的矩陣(甚至當 $A$ 是 $m\times n$ 矩陣時,一個是 $n$ 階方陣,另一個是 $m$ 階方陣)。

## 8.3. Vector Space

這裡我們僅探討 $\mathbb{R}^n$ 這一個向量空間（vector space）及其子空間（subspace）。在此我們將 $\mathbb{R}^n$ 中的向量看成是 $n\times 1$ 的矩陣，所以它們可以相加，也可以乘上 $\mathbb{R}$。另外最重要的是，所有的 $m\times n$ 矩陣，都可以乘上 $\mathbb{R}^n$ 上的向量，得到一個 $\mathbb{R}^m$ 上的向量。

在 $\mathbb{R}^n$ 上的一個子集合 $V$，如果 $V$ 中任兩個向量 $\mathbf{v}_1,\mathbf{v}_2$ 的線性組合，即 $c_1\mathbf{v}_1+c_2\mathbf{v}_2$ 仍在 $V$ 中，則我們稱 $V$ 是一個向量空間。依此定義 $\mathbb{R}^n$ 就是一個向量空間。當 $V$ 是 $\mathbb{R}^n$ 中的一個向量空間，有時為了特別強調它在 $\mathbb{R}^n$ 中，我們便稱之為 $\mathbb{R}^n$ 的子空間（subspace）。例如在 $\mathbb{R}^2$ 中的向量 $\begin{bmatrix} x \\ y \end{bmatrix}$ 中，符合 $x-y=0$ 的向量所成的集合，就是一個向量空間；但符合 $x-y=1$ 的向量所成的集合，就不是一個向量空間。因為 $\mathbf{v}_1=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\mathbf{v}_2=\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ 都在這集合中，但 $\mathbf{v}_1+\mathbf{v}_2=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 就不在其中。由於當 $V$ 時向量空間時，若 $\mathbf{v}\in V$，則因 $\mathbf{v}+(-1)\mathbf{v}$ 為零向量，所以也會在 $V$ 中。因此一個向量空間一定會有零向量。前述符合 $x-y=1$ 的向量所成的集合，因為零向量不符合，因此知道它不是向量空間。不過要注意，一個集合即使有零向量，未必會是向量空間。

在 $\mathbb{R}^n$ 中最容易得到向量空間的方法，是找 $\mathbb{R}^n$ 中的一些向量 $\mathbf{v}_1,\dots,\mathbf{v}_k$，然後考慮所有 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 的線性組合（linear combination）所成的集合, 我們用 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 來表示, 亦即

$$\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=\{\sum_{i=1}^{n} c_i\mathbf{v}_i \mid c_1,\dots,c_n\in\mathbb{R}\}.$$

很容易驗證 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 會是 vector space（即 $\mathbb{R}^n$ 的 subspace）。

給定一個 $m\times n$ matrix $A$，我們可以得到兩個與 $A$ 有關的 vector space，其定義如下：

**Definition 8.3.1.** 假設 $A=\begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}$ 為以 $\mathbb{R}^m$ 中的向量 $\mathbf{a}_1,\dots,\mathbf{a}_n$ 為 column vectors 的 $m\times n$ matrix.

(1) 我們稱 $\mathrm{Span}(\mathbf{a}_1,\dots,\mathbf{a}_n)$ 為 $A$ 的 *column space*, 且用 $\mathrm{Col}(A)$ 來表示 $A$ 的 column space.

(2) 我們稱聯立方程組 $A\mathbf{x}=\mathbf{0}$ 所有解所成的集合為 $A$ 的 *null space* 且用 $N(A)$ 表示 $A$ 的 null space. 即 $N(A)=\{\mathbf{u}\in\mathbb{R}^n \mid A\mathbf{u}=\mathbf{0}\}$.

要注意當 $A\in M_{m\times n}$, 則 $A$ 的 column space $\mathrm{Col}(A)$ 會是 $\mathbb{R}^m$ 的 subspace, 而 $A$ 的 null space $N(A)$ 會是 $\mathbb{R}^n$ 的 subspace (請自行檢驗). 不過要注意，並不是所有聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的所有解所成的集合會是 vector space。當 $\mathbf{b}$ 不是零向量時，很容易檢查零向量不會是解所以不會是 vector space。只有當 $\mathbf{b}=\mathbf{0}$ 時才會是 vector space，此時我們特別稱這樣的聯立方程為 *homogeneous linear system*。

矩陣 $A$ 的 column space 之所以重要，主要是能夠幫我們判斷聯立方程 $A\mathbf{x}=\mathbf{b}$ 是否有解。假設 $\mathbf{c}=\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 為一解, 此即表示 $A\mathbf{c}=\mathbf{b}$. 利用矩陣乘法定義得

$$c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n=\mathbf{b},$$

其中 $\mathbf{a}_1,\ldots,\mathbf{a}_n$ 為 $A$ 的 column vectors. 換句話說, $\mathbf{b}$ 可以寫成 $A$ 的 column vectors 的線性組合，用符號來表示就是 $\mathbf{b}\in\operatorname{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$。反之, 若 $\mathbf{b}\in\operatorname{Span}(\mathbf{a}_1,\ldots,\mathbf{a}_n)$, 表示存在 $c_1,\ldots,c_n\in\mathbb{R}$ 使得 $\mathbf{b}=c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n$. 故得 $x_1=c_1,\ldots,x_n=c_n$ 為 $A\mathbf{x}=\mathbf{b}$ 的一組解. 我們證得了聯立方程組 $A\mathbf{x}=\mathbf{b}$ 有解若且唯若 $\mathbf{b}\in\operatorname{Col}(A)$。

當聯立方程 $A\mathbf{x}=\mathbf{b}$ 有解，下一個課題就是討論其解是否唯一，此時 $A$ 的 null space 就可以幫我們回答此問題。首先觀察，假設 $\mathbf{x}_1,\mathbf{x}_2$ 都是 $A\mathbf{x}=\mathbf{b}$ 的解, 意即 $A\mathbf{x}_1=A\mathbf{x}_2=\mathbf{b}$. 由此得 $A(\mathbf{x}_1-\mathbf{x}_2)=\mathbf{b}-\mathbf{b}=\mathbf{0}$。因此 $\mathbf{x}_1-\mathbf{x}_2$ 會是 $A\mathbf{x}=\mathbf{0}$ 的一組解. 因此若 $N(A)=\{\mathbf{0}\}$ 則 $\mathbf{x}_1=\mathbf{x}_2$，也就是 $A\mathbf{x}=\mathbf{b}$ 的解唯一（這個概念其實在談 linear ODE homogeneous 的解和 nonhomogeneous 解的關係時我們也談過）。因此可知當聯立方程 $A\mathbf{x}=\mathbf{b}$ 有解時，$N(A)=\{\mathbf{0}\}$ 等同於有唯一解。

要了解一個 vector space 最好的方法就是找到它的一組 basis（基底）。要成為一組 basis 首先便是這些向量能夠展成該向量空間，也就是說要了解 $V$ 這個向量空間，首先就必須找到 $\mathbf{v}_1,\ldots,\mathbf{v}_k$ 使得 $\operatorname{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_k)=V$。當然了這些 $\mathbf{v}_i$ 越多越容易展成 $V$。找到最少的向量展成 $V$，就是 basis 的基本概念。要最少就必須這些向量之間沒有線性關係，也就是所謂的線性獨立（linearly independent）。若 $\mathbf{v}_1,\ldots,\mathbf{v}_k$ 能展成 $V$ 且它們是 linearly independent，我們便稱 $\mathbf{v}_1,\ldots,\mathbf{v}_k$ 為 $V$ 的一組 basis。

給定矩陣 $A$ 如何找到 $\operatorname{Col}(A)$ 的一組 basis 呢？由於 $\operatorname{Col}(A)$ 依定義是由它的 column vectors 展成，我們需要的是如何挑到 linearly independent 的 column vectors。所以我們必須先了解如何判定一些向量是否為 linearly independent。當給定 $\mathbb{R}^m$ 上一組向量 $\mathbf{v}_1,\ldots,\mathbf{v}_n$。我們可以考慮由 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 column vector 所組成的 $m\times n$ 矩陣 $A$，然後考慮 $N(A)$。如果 $N(A)=\{\mathbf{0}\}$，表示聯立方程組 $A\mathbf{x}=\mathbf{0}$ 只有 $\mathbf{0}$ 這組解，這就表示 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 這些向量之間沒有線性關係，因此它們是線性獨立。反之若聯立方程 $A\mathbf{x}=\mathbf{0}$ 有非零的解，表示它們之間有關。所以我們可以用 $N(A)$ 是否為 $\{\mathbf{0}\}$ 來決定它們是否為線性獨立。注意 $N(A)=\{\mathbf{0}\}$，表示聯立方程組 $A\mathbf{x}=\mathbf{0}$ 沒有非零的解，也就是將 $A$ 利用基本列運算後，所得的 echelon form 沒有 free variable。這就是判斷是否為 independent 的好方法。

現在回到找 $\operatorname{Col}(A)$ 的 basis 的問題。我們首先將 $A$ 利用 elementary row operation 化為 echelon form $A'$. 假設 $A'$ 的 pivot variables 為 $x_{i_1},\ldots,x_{i_r}$, 若捨去 $A'$ free variable 所在的那些 column，則就沒有 free variables，所以 $\mathbf{a}'_{i_1},\ldots,\mathbf{a}'_{i_r}$ 為 linearly independent 不過我們是要原來 $A$ 的 column vectors 由於 $A\mathbf{x}=\mathbf{0}$ 和 $A'\mathbf{x}=\mathbf{0}$ 有相同的解集合，所以對應到 $A$ 的 column vectors $\mathbf{a}_{i_1},\ldots,\mathbf{a}_{i_r}$ 亦為 linearly independent. 這些向量就是 $\operatorname{Col}(A)$ 的 basis。我們看以下的例子。

**Example 8.3.2.** 考慮

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

將 $A$ 的 2-nd row 分別乘上 $-2$, $-1$, $-1$ 加至 1-st, 3-rd 和 4-th row, 然後再將 1-st, 2-nd rows 交換得

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

接著將 2-nd row 分別乘上 $-1,-2$ 加至 3-rd 和 4-th row 得

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

最後將 3-rd row 乘上 $-1$ 加至 4-th row, 得 echelon form

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

我們很容易看出 $A'$ 的 3 個 pivot 所在的 column vectors $\mathbf{a}'_1,\mathbf{a}'_2,\mathbf{a}'_4$ 為 linearly independent. 最後回到 $A$，得到 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$ 是 $A$ 的 column space 的一組 basis. ♯

接下來談如何找到 $N(A)$ 的一組 basis。回顧我們找 $A\mathbf{x}=\mathbf{0}$ 的解集合的方法為, 利用 elementary row operations 將 $A$ 化為 echelon form $A'$. 此時 $A'\mathbf{x}=\mathbf{0}$ 的解集合就是 $A\mathbf{x}=\mathbf{0}$ 的解集合, 也就是說 $A$ 和 $A'$ 有相同的 null space. 接著我們找出 free variables, 再將每個 free variable 代入不同參數, 從下往上推得出一組解。最後將這些參數提出，寫成一些向量的線性組合。注意這些向量在對應的 free variable 其值為 1 其他 free variables 的值為 0，所以它們是 linearly independent。因此形成 $N(A)$ 的一組 basis。我們看以下的例子。

**Example 8.3.3.** 考慮 Example 8.3.2 的矩陣 $A$ 以及利用基本列運算所得的 echelon form $A'$。知 $x_6,x_5,x_3$ 為 free variables 且分別設其為任意的實數 $r,s,t$, 可解得 $x_4=-2r-s$, $x_2=-4r-2s-t$, $x_1=2r+s$. 也就是說 $A$ 的 null space 中的向量都可以寫成

$$\begin{bmatrix} 2r+s \\ -4r-2s-t \\ t \\ -2r-s \\ s \\ r \end{bmatrix}=r\begin{bmatrix} 2 \\ -4 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}+s\begin{bmatrix} 1 \\ -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}+t\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}=r\mathbf{v}_1+s\mathbf{v}_2+t\mathbf{v}_3.$$

故知 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 $A$ 的 null space 的 spanning vectors, 又 $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ 在 $x_6$ 的位置分別為 $1,0,0$，而在 $x_5$ 的位置分別為 $0,1,0$ 以及在 $x_3$ 的位置分別為 $0,0,1$。所以 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 linearly independent, 因而 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 為 $N(A)$ 的一組 basis。 ♯

找到 vector space 的一組基底後，這組形成基底的向量的個數就稱為此向量空間的維度（dimension），通常我們都用 $\dim(V)$ 來表示 $V$ 的維度。要注意，當 $V_1, V_2$ 為 vector space 且 $V_1 \subseteq V_2$，則 $\dim(V_1)$ 會小於 $\dim(V_2)$ 除非 $V_1 = V_2$。因此若 $V$ 是 $\mathbb{R}^n$ 的 subspace 且 $\dim(V) = n$，便可知 $V = \mathbb{R}^n$。所以維度是一個讓我們辨識一個向量空間的好方法。前面已經知道 $\mathrm{Col}(A)$ 的 basis 是 pivot 所在的 column vector 所形成，而 pivot 的個數我們也提過稱為 $A$ 的 rank（記為 $\mathrm{rank}(A)$），所以 $\dim(\mathrm{Col}(A)) = \mathrm{rank}(A)$。另外我們也知道 $N(A)$ 的基底是由 free variable 所對應的向量所形成，所以 $\dim(N(A))$（一般稱為 $A$ 的 *nullity*）會等於 free variables 的個數。然而 free variables 和 pivot variables 的個數加在一起，就是所有的 variables 的個數，即矩陣 column 的個數。因此我們得到一個很重要有關矩陣的 *dimension theorem* 即，當 $A$ 為 $m \times n$ 矩陣，則 $\dim(\mathrm{Col}(A)) + \dim(N(A)) = n$。此維度定理有時也常用 $\mathrm{rank}(A) + \mathrm{nullity}(A) = n$ 表示。

## 8.4. 方陣

行、列數相等的矩陣稱為方陣（square matrix）是線性代數中很重要的研究對象，主要就是因為它有許多重要的特性。只有方陣，才可談行列式（determinant）；也只有方陣才有可能有反矩陣（inverse matrix）；也只有方陣可以談高次相乘，因此有所謂“對角化”來簡化計算。這一節中我們將分別介紹這些課題。

**8.4.1. 行列式.** 我們先介紹行列式的性質，以及兩種計算行列式的方法，最後談它與解聯立方程的關係。

給定一方陣 $A$，我們用 $\det(A)$ 表示 $A$ 的行列式。當 $A, B$ 為同階方陣時，$A$ 和 $B$ 可以相乘。一個重要的行列式性質就是 $\det(AB) = \det(A)\det(B)$。這個性質提供我們許多簡化計算行列式的方法。例如當 $A$ 為 $n$ 階可逆矩陣，表示存在 $n$ 階方陣 $A^{-1}$ 使得 $AA^{-1} = I_n$，其中 $I_n$ 為 $n$ 階單位矩陣。亦即對角線位置皆為 1，其他位置皆為 0 的矩陣。由於對角矩陣的行列式值就是對角線上的數字相乘，我們有 $\det(I_n) = 1$ 所以知 $A$ 的反矩陣的行列式值與 $A$ 的行列式值互為倒數，即 $\det(A^{-1}) = \dfrac{1}{\det(A)}$，從這裡我們知道若 $A$ 為可逆矩陣則 $\det(A) \neq 0$。事實上也可證明若 $\det(A) \neq 0$，則 $A$ 為可逆矩陣。

利用行列式的性質，我們也可推得一個方陣經由 elementary row operations 後所得矩陣的行列式之間的關係，如下。

(1) 若將 $A$ 做 type 1 的基本列運算所得的矩陣為 $A'$, 則 $\det(A') = -\det(A)$.

(2) 若將 $A$ 做 type 2 的基本列運算，即某個 row 乘上非零實數 $r$ 所得的矩陣為 $A'$, 則 $\det(A') = r\det(A)$.

(3) 若將 $A$ 做 type 3 的基本列運算所得的矩陣為 $A'$, 則 $\det(A') = \det(A)$..

依此我們可以利用 elementary row operation 求 $n \times n$ matrix 的 determinant. 首先我們先用 elementary row operations 將矩陣變為 echelon form. 而且在化為 echelon form 的過程中只用 type 1，即兩 row 交換 (此時 determinant 變號) 以及 type 3 將某個 row 乘上實數 $r$ 加到另一個 row (此時 determinant 不會改變), 這兩種 row operations。由於一個方陣

的 echelon form 一定是一個 *upper triangular matrix* (上三角矩陣), 也就是說矩陣對角線以下的位置皆為 0 (即 $a_{ij}=0$, for $i>j$)，此時其 determinant 就是對角線的乘積。所以我們可由 echelon form 的 determinant 得到原矩陣的 determinant (即做了奇數次變號, 偶數次不變號)。

**Example 8.4.1.** 我們利用 elementary row operation 求 $\begin{bmatrix} 0 & 2 & -1 & 1 \\ 1 & 2 & 0 & 2 \\ 1 & 4 & -2 & 6 \\ 2 & 6 & -1 & 8 \end{bmatrix}$ 的 determinant.

首先將 1-st, 2-nd row 交換得 $\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 2 & -1 & 1 \\ 1 & 4 & -2 & 6 \\ 2 & 6 & -1 & 8 \end{bmatrix}$ (注意此時 determinant 會變號). 接著將 1-st row 分別乘上 $-1,-2$ 加到 3-rd, 4-th row 得 $\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 2 & -1 & 1 \\ 0 & 2 & -2 & 4 \\ 0 & 2 & -1 & 4 \end{bmatrix}$ (注意此時 determinant 不會改變). 最後將 2-nd row 分別乘上 $-1,-1$ 加到 3-rd, 4-th row 得 echelon form $\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ (注意此時 determinant 不會改變)。由於 echelon form 對角線乘積為 $-6$，我們知其 determinant 為 $-6$, 又整個化為 echelon form 的過程中僅用了一次兩 row 交換的 row operation, 故 determinant 僅變號一次, 得知

$$\det\begin{bmatrix} 0 & 2 & -1 & 1 \\ 1 & 2 & 0 & 2 \\ 1 & 4 & -2 & 6 \\ 2 & 6 & -1 & 8 \end{bmatrix} = -\det\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix} = 6.$$

♯

從這個求 determinant 的方法我們發現，當一個 $n$ 階方陣 $A$ 的 rank 小於 $n$ 時，由於 pivot 個數小於 $n$，其化為 echelon form 時一定有一個 row 全為 0，所以對角線上一定有 0，得知 $\det(A)=0$。有此時有 free variable，所以聯立方程組 $A\mathbf{x}=\mathbf{0}$ 一定有非零向量的解 ( 即 $N(A)\neq\{\mathbf{0}\}$ )。所以我們有以下重要性質。

**Theorem 8.4.2.** 假設 $A$ 為 $n$ 階方陣。考慮 $A$ 的行列式 $\det(A)$。當 $\det(A)=0$，與下列皆為等價：

(1) $\operatorname{rank}(A)<n$。

(2) 存在非零向量 $\mathbf{v}$ 使得 $A\mathbf{v}=\mathbf{0}$。

當 $\det(A)\neq 0$，與下列皆為等價：

(1) $\operatorname{rank}(A)=n$。

(2) $A$ 為可逆矩陣。

(3) 對任意 $\mathbf{b}\in\mathbb{R}^n$ 聯立方程組 $A\mathbf{x}=\mathbf{b}$ 一定有解且解唯一。

接下來，我們介紹另一種求行列式值的方法，即所謂“降階”的方式。為了方便起見，我們有以下的定義.

**Definition 8.4.3.** 假設 $A=[a_{ij}]$ 為 $3\times 3$ matrix. 將 $A$ 的 $i$-th row 和 $j$-th column 除去所得的 $2\times 2$ matrix, 稱為 $A$ 的 $(i,j)$ *minor matrix*, 用 $A_{ij}$ 表示. 令 $a'_{ij}=(-1)^{i+j}\det(A_{ij})$, 稱為 $A$ 的 $(i,j)$ cofactor.

依此定義, 我們可將 $\det(A)$ 寫成
$$\det(A)=a_{11}a'_{11}+a_{21}a'_{21}+a_{31}a'_{31}.$$
即
$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}-a_{21}\det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}+a_{31}\det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}.$$
這稱為對 1-st column 展開。也可對 2-nd column 展開 $\det(A)=a_{12}a'_{12}+a_{22}a'_{22}+a_{32}a'_{32}$, 即
$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=-a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}+a_{22}\det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}-a_{32}\det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}.$$
同樣的我們也可對 3-rd column 展開得
$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}-a_{23}\det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix}+a_{33}\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

事實上 determinant 也可對 1-st row 展開求得，即
$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}=a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}-a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}+a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$
因此得知 $\det(A)=\det(A^{t})=a_{11}a'_{11}+a_{12}a'_{12}+a_{13}a'_{13}$, 同理對 2-nd row 和 3-rd row 展開也可求得 determinant。

我們得到了 $3\times 3$ matrix 的 determinant, 也因此由此可定義出 $\mathbb{R}^3$ 中三個向量所張成的平行六面體的“有向體積” *signed volume*。也就是說若將 $\mathbb{R}^3$ 上的三個向量 $\mathbf{u},\mathbf{v},\mathbf{w}$ 寫成 row vectors, 令矩陣 $A$ 為 1-st, 2-nd, 3-rd row 依序為 $\mathbf{u},\mathbf{v},\mathbf{w}$ 的 $3\times 3$ matrix, 則 $\det(A)$ 就是 $\mathbf{u},\mathbf{v},\mathbf{w}$ 三個向量所張成的平行六面體的有向體積。即 $\det(A)$ 的絕對值, 就是這平行六面體的體積. 而 $\det(A)$ 的正負號告訴我們 $\mathbf{u},\mathbf{v},\mathbf{w}$ 這三個向量的方向性. 這裡 $\mathbf{u},\mathbf{v},\mathbf{w}$ 這三個向量正反向我們是用所謂的 *right hand rule*（右手定則）來區分, 意即將右手大拇指指向 $\mathbf{u}$ 的方向, 其餘四個指頭併攏指向 $\mathbf{v}$ 的方向, 若 $\mathbf{w}$ 朝著手掌正面的方向則 $\mathbf{u},\mathbf{v},\mathbf{w}$ 為正向，反之為負向.

給定 $\mathbf{u}=(a_1,a_2,a_3)$, $\mathbf{v}=(b_1,b_2,b_3)\in\mathbb{R}^3$, 我們定義 $\mathbf{u},\mathbf{v}$ 的 *cross product* (外積) $\mathbf{u}\times\mathbf{v}$ 為
$$\mathbf{u}\times\mathbf{v}=\left(\det\begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix},\det\begin{bmatrix} a_3 & a_1 \\ b_3 & b_1 \end{bmatrix},\det\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\right).$$
要注意, 千萬不要將內積和外積弄混了, 兩個向量之內積是一個實數, 而兩個向量之外積仍為向量. 另外 $\mathbf{u}\times\mathbf{v}$ 和 $\mathbf{v}\times\mathbf{u}$ 是不相等的, 除非 $\mathbf{u}\times\mathbf{v}=\mathbf{0}$. 這是因為兩個 row 交換

其 determinant 會變號, 因此依定義 $\mathbf{u}\times\mathbf{v}=-\mathbf{v}\times\mathbf{u}$. 而 $\mathbf{u}\times\mathbf{v}$ 何時會是 $\mathbf{0}$ 呢? 依定義 $\mathbf{u}\times\mathbf{v}=\mathbf{0}$ 若且唯若 $\det\begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix}=\det\begin{bmatrix} a_3 & a_1 \\ b_3 & b_1 \end{bmatrix}=\det\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}=0$, 很容易知道這等同於 $\mathbf{u}=(a_1,a_2,a_3)$, $\mathbf{v}=(b_1,b_2,b_3)$ 為 linearly dependent（即不是線性獨立）.

現考慮 $\mathbf{u}=(a_1,a_2,a_3)$, $\mathbf{v}=(b_1,b_2,b_3)$, $\mathbf{w}=(c_1,c_2,c_3)$ 我們有

$$\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})=c_1\det\begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix}+c_2\det\begin{bmatrix} a_3 & a_1 \\ b_3 & b_1 \end{bmatrix}+c_3\det\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \tag{8.6}$$

由於 $\det\begin{bmatrix} a_3 & a_1 \\ b_3 & b_1 \end{bmatrix}=-\det\begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix}$ 式子 (8.6) 的右式又等同於將 $\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ 對 3-rd row 展開的 determinant, 故得

$$\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})=(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}=\det\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=\det\begin{bmatrix} -\mathbf{u}- \\ -\mathbf{v}- \\ -\mathbf{w}- \end{bmatrix}. \tag{8.7}$$

也就是說 $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}$ 就是 $\mathbf{u},\mathbf{v},\mathbf{w}$ 這三個向量所張成的平行六面體的 signed volume.

特別的, 當 $\mathbf{w}=\mathbf{u}$ 或 $\mathbf{w}=\mathbf{v}$ 時, 由於 $\mathbf{u},\mathbf{v},\mathbf{w}$ 為 row vector 所形成的矩陣有兩個 row 相同, 所以其 rank 小於 3 知其 determinant 為 0（Theorem 8.4.2）。因此由式子 (8.7) 知 $\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v})=\mathbf{v}\cdot(\mathbf{u}\times\mathbf{v})=0$. 也就是說當 $\mathbf{u},\mathbf{v}$ 為 linearly independent 時, $\mathbf{u}\times\mathbf{v}$ 同時會和 $\mathbf{u}$ 以及和 $\mathbf{v}$ 垂直. 而當 $\mathbf{w}=\mathbf{u}\times\mathbf{v}$, 我們有 $(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v})=\|\mathbf{u}\times\mathbf{v}\|^2$. 也就是 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 所張成的平行六面體的 signed volume 為 $\|\mathbf{u}\times\mathbf{v}\|^2$. 考慮 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 所張成的平行六面體以 $\mathbf{u},\mathbf{v}$ 所張的平行四邊形為底, 此時由於 $\mathbf{u}\times\mathbf{v}$ 和 $\mathbf{u}$ 以及和 $\mathbf{v}$ 垂直, 我們得 $\|\mathbf{u}\times\mathbf{v}\|$ 就是此平行六面體的高. 因此由 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 所張成的平行六面體的體積 $\|\mathbf{u}\times\mathbf{v}\|^2$ 為 $\mathbf{u},\mathbf{v}$ 所張的平行四邊形面積乘上高 $\|\mathbf{u}\times\mathbf{v}\|$, 得 $\mathbf{u},\mathbf{v}$ 所張的平行四邊形面積為 $\|\mathbf{u}\times\mathbf{v}\|$. 另外由於 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 所張成的平行六面體的 signed volume 為 $\|\mathbf{u}\times\mathbf{v}\|^2>0$, 我們知 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 利用右手定則為正向. 最後我們將外積的性質歸納如下.

**Theorem 8.4.4.** *給定 $\mathbf{u},\mathbf{v}\in\mathbb{R}^3$. 則 $\mathbf{u}\times\mathbf{w}\neq\mathbf{0}$ 若且唯若 $\mathbf{u},\mathbf{v}$ 為 linearly independent. 此時 $\mathbf{u}\times\mathbf{v}$ 的長度為 $\mathbf{u},\mathbf{v}$ 所張的平行四邊形面積, 且 $\mathbf{u},\mathbf{v}$ 同時與 $\mathbf{u}\times\mathbf{v}$ 垂直, 又 $\mathbf{u},\mathbf{v},\mathbf{u}\times\mathbf{v}$ 利用右手定則為正向.*

*又假設 $\mathbf{w}\in\mathbb{R}$, 則 $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}\neq 0$ 若且唯若 $\mathbf{u},\mathbf{v},\mathbf{w}$ 為 linearly independent. 此時 $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}$ 就是 $\mathbf{u},\mathbf{v},\mathbf{w}$ 這三個向量所張成的平行六面體的 signed volume.*

我們利用降階的方法將求三階行列式降為利用二階行列式來求. 接著我們可利用降階方法將 4 階方陣的行列式，降為 3 階，然後一直下去. 首先我們將 Definition 8.4.3 的定義推廣到一般的情形.

**Definition 8.4.5.** 假設 $A=[a_{ij}]$ 為 $n\times n$ matrix. 將 $A$ 的 $i$-th row 和 $j$-th column 除去所得的 $(n-1)\times(n-1)$ matrix, 稱為 $A$ 的 $(i,j)$ *minor matrix*, 用 $A_{ij}$ 表示. 當 $(n-1)\times(n-1)$ matrix 的 determinant 存在時, 令 $a'_{ij}=(-1)^{i+j}\det(A_{ij})$, 稱為 $A$ 的 $(i,j)$ cofactor.

我們有以下的結果：假設 $A=[a_{ij}]$ 為 $n\times n$ matrix. 令 $a'_{ij}$ 為 $A$ 的 $(i,j)$ cofactor, 則對任意 $k\in\{1,\ldots,n\}$ 皆有 $\det(A)=a_{1k}a'_{1k}+a_{2k}a'_{2k}+\cdots+a_{nk}a'_{nk}$（對 $k$-th column 降階）以及 $\det(A)=a_{k1}a'_{k1}+a_{k2}a'_{k2}+\cdots+a_{kn}a'_{kn}$（對 $k$-th row 降階）。

從行列式可用 column 降階，或用 row 降階的角度來看，我們可得到另一個行列式重要的性質：$\det(A)=\det(A^t)$，也就是取轉置不會改變行列式值。既然 row 和 column 互換不會影響行列式值，我們也可用所謂 column operation 來處理行列式，即兩個 column 互換（行列式差個負號）和某個 column 乘上 $r$ 加到另一個 column（行列式值不變）。

另外要特別提醒的是，一般 2 階、3 階行列式，我們通常會用取對角斜線的方式計算，不過切記，對於 4 階以上這個方法不能用來計算行列式。

我們用 elementary row operations 的方法和降階的方法都可求 determinant。然而在計算 determinant 時，這兩種方法都可混用。一般來說用 row operation 或 column operation 來求 determinant 較快。不過若發現有的 row 或 column 有很多 0 的 entry，則對該 row 或 column 降階，也可幫助我們較快算出 determinant。我們看以下的例子。

**Example 8.4.6.** 我們求 $A=\begin{bmatrix}2&-1&3&5\\0&2&1&2\\0&5&3&3\\4&-2&7&8\end{bmatrix}$ 的 determinant. 首先觀察 $A$ 的 1-st column 僅有兩個 entry 不為 0, 所以利用 1-st row 乘上 $-2$ 加到 4-th row 得 $B=\begin{bmatrix}2&-1&3&5\\0&2&1&2\\0&5&3&3\\0&0&1&-2\end{bmatrix}$ (此時 $\det(A)=\det(B)$). 現因 $B$ 的 1-st column 僅有一個非 0 entry, 我們對 1-st column 降階展開得 $\det(B)=2\det(C)$ 其中 $C$ 為 $B$ 的 $(1,1)$ minor matrix, 即 $C=\begin{bmatrix}2&1&2\\5&3&3\\0&1&-2\end{bmatrix}$. 接著將 $C$ 的 2-nd column 乘上 2 加到 3-rd column 得 $D=\begin{bmatrix}2&1&4\\5&3&9\\0&1&0\end{bmatrix}$ (此時 $\det(C)=\det(D)$). 最後對 $D$ 的 3-rd row 展開得 $\det(D)=(-1)^{3+2}\det\begin{bmatrix}2&4\\5&9\end{bmatrix}=2$. 故知 $\det(A)=\det(B)=2\det(C)=2\det(D)=4$. ♯

最後我們談行列式的應用，如何幫助我們解聯立方程組。首先考慮 $n\times n$ matrix $A=[a_{ij}]$ 對於 $j\in\{1,\ldots,n\}$ 令 $\mathbf{a}_j$ 表示 $A$ 的 $j$-th column. 現對於 $\mathbf{b}=\begin{bmatrix}b_1\\\vdots\\b_n\end{bmatrix}$，若 $x_1=c_1,\ldots,x_n=c_n$ 為聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解，令 $\mathbf{c}=\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$。對於任意 $k\in\{1,\ldots,n\}$, 考慮 $C_k$ 為將 identity matrix $I_n$ 的 $k$-th column 用 $\mathbf{c}$ 取代的 $n\times n$ matrix. 亦即當 $j\neq k$ 時, $C_k$ 的 $j$-th column 為

$\mathbf{e}_j$, 而 $C_k$ 的 $k$-th column 為 $\mathbf{c}$. 現考慮 $AC_k$, 依矩陣乘法的定義, 我們有

$$AC_k = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} | & & c_1 & & | \\ \mathbf{e}_1 & \cdots & \vdots & \cdots & \mathbf{e}_n \\ | & & c_n & & | \end{bmatrix} = \begin{bmatrix} | & & | & & | \\ \mathbf{a}_1 & \cdots & c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n & \cdots & \mathbf{a}_n \\ | & & | & & | \end{bmatrix}. \quad (8.8)$$

也就是說當 $j \neq k$ 時, $AC_k$ 的 $j$-th column 為 $\mathbf{a}_j$, 而 $AC_k$ 的 $k$-th column 為 $c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n$. 然而 $\mathbf{x}=\mathbf{c}$ 為聯立方程組 $A\mathbf{x}=\mathbf{b}$ 的一組解, 亦即

$$\begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1\mathbf{a}_1+\cdots+c_n\mathbf{a}_n = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \mathbf{b}. \quad (8.9)$$

因此若令 $B_k$ 表示將 $A$ 的 $k$-th column 用 $\mathbf{b}$ 取代的 $n \times n$ matrix, 即

$$B_k = \begin{bmatrix} | & & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{b} & \cdots & \mathbf{a}_n \\ | & & | & & | \end{bmatrix},$$

則結合式子 (8.8) (8.9), 我們有 $AC_k = B_k$. 因此由行列式乘法性質，得 $\det(A)\det(C_k) = \det(B_k)$. 然而對 $C_k$ 的 $k$-th row 展開, 我們有 $\det(C_k) = (-1)^{k+k}c_k\det(I_{n-1}) = c_k$. 因此得證以下之定理.

**Theorem 8.4.7** (Cramer's Rule)**.** *假設 $A$ 為 $n \times n$ matrix 且 $\mathbf{b} \in \mathbb{R}^n$ 為 column vector. 對於 $k \in \{1,\ldots,n\}$ 令 $B_k$ 表示將 $A$ 的 k-th column 用 $\mathbf{b}$ 取代的 $n \times n$ matrix. 若 $x_1 = c_1,\ldots,x_n = c_n$ 為聯立方程組 $A\mathbf{x} = \mathbf{b}$ 的一組解, 則 $c_k\det(A) = \det(B_k)$.*

**Example 8.4.8.** 考慮 Example 8.4.6 中的矩陣 $A = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 0 & 2 & 1 & 2 \\ 0 & 5 & 3 & 3 \\ 4 & -2 & 7 & 8 \end{bmatrix}$. 令 $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, 前面已知 $\det(A) = 4 \neq 0$, 我們可用 Cramer's rule 解聯立方程組 $A\mathbf{x} = \mathbf{b}$. 此時將 $\mathbf{b}$ 置換於 $A$ 的 1-st column, 得 $B_1 = \begin{bmatrix} 1 & -1 & 3 & 5 \\ 0 & 2 & 1 & 2 \\ 0 & 5 & 3 & 3 \\ 0 & -2 & 7 & 8 \end{bmatrix}$. 同理得 $B_2 = \begin{bmatrix} 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 3 \\ 4 & 0 & 7 & 8 \end{bmatrix}$, $B_3 = \begin{bmatrix} 2 & -1 & 1 & 5 \\ 0 & 2 & 0 & 2 \\ 0 & 5 & 0 & 3 \\ 4 & -2 & 0 & 8 \end{bmatrix}$, $B_4 = \begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 5 & 3 & 0 \\ 4 & -2 & 7 & 0 \end{bmatrix}$. 利用降階, 我們得 $\det(B_1) = 42$, $\det(B_2) = 12$, $\det(B_3) = -16$, $\det(B_4) = -4$. 故由 Cramer's rule 知 $x_1 = 21/2$, $x_2 = 3$, $x_3 = -4$, $x_4 = -1$ 是聯立方程組 $A\mathbf{x} = \mathbf{b}$ 之唯一解。 $\sharp$

**8.4.2. 反矩陣.** 這裡我們提供兩個求反矩陣的方法。一種是利用 elementary row operation 另一種是利用 Cramer's Rule。

當 $A$ 為 $n$ 階方陣，若要使用 elementary row operation 求反矩陣，首先寫下增廣矩陣 $[A|I_n]$ 然後利用 elementary row operations 將之化為 $[I_n|E]$, 此時會有 $EA = I_n$, 因此 $E$ 就是 $A$ 的反矩陣 $A^{-1}$。整個過程是先將 $A$ 化為 echelon form。如果 pivot 個數少於 $n$，由前 Theorem 8.4.2 知 $A$ 沒有反矩陣。而 pivot 個數為 $n$ 時，再利用 type 2 的方法讓 echelon form 對角線上的 entry 皆為 1，再利用 type 3 的方法由下而上，將對角線上方的 entry 皆變為 0 就得到 $I_n$ 了。我們看以下的例子.

**Example 8.4.9.** 考慮矩陣

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

我們要決定是否 $A$ 是否為 invertible. 若為 invertible, 要找出 $A^{-1}$.

我們考慮增廣矩陣 $[A|I_4]$, 利用 elementary row operation 將 $A$ 的部分轉換成 echelon form. 首先將 1-st row 分別乘上 $-1$, 3 加至 3-rd, 4-th row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 2 & 0 & 0 & 1 & 0 \\ -3 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 3 & -4 & 3 & 0 & 0 & 1 \end{array}\right].$$

接著將 3-rd row 乘上 3/2 加至 4-th row 得

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{3}{2} & 0 & \frac{3}{2} & 1 \end{array}\right].$$

此時增廣矩陣左半部為 echelon form, 其 pivot 的個數為 4, 故知 $A$ 為 invertible. 我們繼續將左半部化為 $I_4$ 便可得到 $A^{-1}$.

先將 4-th row 乘以 2, 然後將所得的增廣矩陣的 4-th row 分別乘上 $-3$, $-4$, 1 加至 3-rd, 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 3 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 4 & 0 & 3 & 2 \\ 0 & -1 & -3 & 0 & -12 & 1 & -12 & -8 \\ 0 & 0 & -2 & 0 & -10 & 0 & -8 & -6 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right].$$

接著將 3-rd row 乘以 $-1/2$, 然後將所得的增廣矩陣的 3-rd row 分別乘上 3, $-1$ 加至 2-nd 和 1-st row, 即

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 4 & 0 & 3 & 2 \\ 0 & -1 & -3 & 0 & -12 & 1 & -12 & -8 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 & 0 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 & 0 & 3 & 2 \end{array}\right].$$

最後將增廣矩陣的 2-nd row 乘上 $-1$. 此時所得增廣矩陣左半部為 $I_4$, 故其右半部為 $A^{-1}$, 即

$$A^{-1} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -3 & -1 & 0 & -1 \\ 5 & 0 & 4 & 3 \\ 3 & 0 & 3 & 2 \end{bmatrix}.$$

$\sharp$

接下來我們利用 Cramer's rule 找反矩陣. 假設 $n$ 階方陣 $A$ 有反矩陣且 $C$ 為 $A$ 的 inverse, 則由 $AC=I_n$, 依矩陣乘法定義我們知 $C$ 的 $j$-th column $\begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$ 需滿足 $A\begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$ 等於 $I_n$ 的 $j$-th column $\mathbf{e}_j$. 也就是說 $C$ 的 $j$-th column 為聯立方程組 $A\mathbf{x}=\mathbf{e}_j$ 的解. 因此 $C$ 的 $(i,j)$-th entry $c_{ij}$ 應為聯立方程組 $A\mathbf{x}=\mathbf{e}_j$ 的解中 $x_i$ 之值. 故由 Cramer's rule 知 $c_{ij}=\det(A(j,i))/\det(A)$, 其中 $A(j,i)$ 表示將 $A$ 的 $i$-th column 用 $\mathbf{e}_j$ 取代的 $n\times n$ matrix. 然而利用對 $A(j,i)$ 的 $i$-th column 展開求 $\det(A(j,i))$, 我們得 $\det(A(j,i))=(-1)^{j+i}\det(A_{ji})=a'_{ji}$. 也就是說 $c_{ij}$ 就是 $A$ 的 $(j,i)$ cofactor (注意 $i,j$ 位置交換) 除以 $\det(A)$. 為了方便起見我們有以下的定義.

**Definition 8.4.10.** 假設 $A=[a_{ij}]$ 為 $n\times n$ matrix, 對於任意 $i,j\in\{1,\ldots,n\}$ 令 $a'_{ij}$ 為 $A$ 的 $(i,j)$ cofactor. 考慮 $n\times n$ matrix $A'$ 其 $(i,j)$-th entry 為 $a'_{ij}$. 我們稱 $A'$ 為 $A$ 的 *cofactor matrix* 而稱 $A'$ 的 transpose $(A')^{\mathrm{t}}$ 為 $A$ 的 *adjoint matrix*, 用 $\operatorname{adj}(A)$ 來表示.

注意 $\operatorname{adj}(A)$ 是將 $A$ 的 cofactor 所成的矩陣 $A'$ 取轉置而得, 千萬不要忘記取轉置. 另外要注意不管一個 $n\times n$ matrix 是否為 invertible, 皆可定義其 adjoint matrix.

我們回到剛才 $A$ 為 invertible 的情況. 假設 $C$ 為其 inverse. 依 $\operatorname{adj}(A)$ 的定義, 我們得到 $C$ 的 $(i,j)$-th entry 就是 $\operatorname{adj}(A)$ 的 $(i,j)$-th entry 除以 $\det(A)$. 因此依矩陣係數積的定義, 我們有 $C=\frac{1}{\det(A)}\operatorname{adj}(A)$. 得證了以下的定理.

**Proposition 8.4.11.** 假設 $A$ 為 $n\times n$ *invertible matrix.* 則

$$A^{-1}=\frac{1}{\det(A)}\operatorname{adj}(A).$$

**Example 8.4.12.** 考慮 Example 8.4.6 中的矩陣 $A=\begin{bmatrix} 2 & -1 & 3 & 5 \\ 0 & 2 & 1 & 2 \\ 0 & 5 & 3 & 3 \\ 4 & -2 & 7 & 8 \end{bmatrix}$. 在 Example 8.4.8 中我們解出 $A\mathbf{x}=\mathbf{e}_1$ 的解, 事實上就是 $A$ 的反矩陣 $A^{-1}$ 的 1-st column. 在 Example 8.4.8 中的 $B_1=\begin{bmatrix} 1 & -1 & 3 & 5 \\ 0 & 2 & 1 & 2 \\ 0 & 5 & 3 & 3 \\ 0 & -2 & 7 & 8 \end{bmatrix}$ 就是將 $A$ 的 1-st column 用 $\mathbf{e}_1$ 取代所得的矩陣 $A(1,1)$. 若我們將 $B_1$ 的 1-st column 展開求 $\det(B_1)$ 得 $\det(B_1)=(-1)^{1+1}\det\begin{bmatrix} 2 & 1 & 2 \\ 5 & 3 & 3 \\ -2 & 7 & 8 \end{bmatrix}=42$ 就是 $A$ 的 $(1,1)$ cofactor. 同樣的 $B_2=\begin{bmatrix} 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 3 \\ 4 & 0 & 7 & 8 \end{bmatrix}$ 就是將 $A$ 的 2-nd column 用 $\mathbf{e}_1$ 取代所得的矩陣 $A(1,2)$. 若我們將 $B_2$ 的 2-nd column 展開求 $\det(B_2)$ 得 $\det(B_2)=(-1)^{1+2}\det\begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 3 \\ 4 & 7 & 8 \end{bmatrix}=12$ 就是 $A$ 的 $(1,2)$ cofactor. 同理得 $B_3$ 是將 $A$ 的 3-rd column 用 $\mathbf{e}_1$ 取代所得的矩陣 $A(1,3)$

且 $\det(B_3)=(-1)^{1+3}\det\begin{bmatrix}0&2&2\\0&5&3\\4&-2&8\end{bmatrix}=-16$ 就是 $A$ 的 $(1,3)$ cofactor. 而 $B_4$ 是將 $A$ 的 4-th column 用 $\mathbf{e}_1$ 取代所得的矩陣 $A(1,4)$ 且 $\det(B_4)=(-1)^{1+4}\det\begin{bmatrix}0&2&1\\0&5&3\\4&-2&7\end{bmatrix}=-4$ 就是 $A$ 的 $(1,4)$ cofactor. 注意這裡求出的 cofactor 其實對應到 $A$ 的 cofactor 所成的矩陣 $A'$ 會是 $A'$ 的 1-st row. 我們求出 $A$ 其他的 cofactor 會有

$$a'_{21}=(-1)^{2+1}\det\begin{bmatrix}-1&3&5\\5&3&3\\-2&7&8\end{bmatrix}=-64,\quad a'_{22}=(-1)^{2+2}\det\begin{bmatrix}2&3&5\\0&3&3\\4&7&8\end{bmatrix}=-18,$$

$$a'_{23}=(-1)^{2+3}\det\begin{bmatrix}2&-1&5\\0&5&3\\4&-2&8\end{bmatrix}=20,\quad a'_{24}=(-1)^{2+4}\det\begin{bmatrix}2&-1&3\\0&5&3\\4&-2&7\end{bmatrix}=10.$$

以及 $a'_{31}=26,a'_{32}=8,a'_{33}=-8,a'_{34}=-4,a'_{41}=-20,a'_{42}=-6,a'_{43}=8,a'_{44}=2$. 因此得

$$A'=\begin{bmatrix}42&12&-16&-4\\-64&-18&20&10\\-26&8&-8&-4\\-20&-6&8&2\end{bmatrix},\quad \operatorname{adj}(A)=\begin{bmatrix}42&-64&-26&-20\\12&-18&8&-6\\-16&20&-8&8\\-4&10&-4&2\end{bmatrix}$$

也因此得

$$A^{-1}=\frac{1}{\det(A)}\operatorname{adj}(A)=\frac{1}{4}\begin{bmatrix}42&-64&-26&-20\\12&-18&8&-6\\-16&20&-8&8\\-4&10&-4&2\end{bmatrix}.$$

$\sharp$

由上面例子可以看出利用 adjoint matrix 求反矩陣非常複雜, 所以在實際求反矩陣的情況還是利用前面 elementary row operation 的方法會比較快.

最後我們提兩個反矩陣的基本性質。第一個是有關於乘法，當 $A,B$ 皆為 $n$ 階方陣且都有反矩陣時，表示 $\det(A),\det(B)$ 皆不為 0，因此由行列式乘法性質 $\det(AB)=\det(A)\det(B)\neq 0$。也就是說 $AB$ 也會有反矩陣。由於 $(AB)(B^{-1}A^{-1})=AI_nA^{-1}=I_n$，我們知 $(AB)^{-1}=B^{-1}A^{-1}$（注意位置相反）。這告訴我們當我們知道 $A$ 和 $B$ 的反矩陣，此時若要求 $AB$ 的反矩陣，不必用求反矩的方法重做，只要將 $B^{-1}$ 右邊乘上 $A^{-1}$ 即可。另一個是關於轉置。當 $A$ 有反矩陣時，由於 $\det(A^{\mathrm{t}})=\det(A)\neq 0$，我們知 $A^{\mathrm{t}}$ 也有反矩陣。事實上由轉置相乘的性質，$A^{\mathrm{t}}\cdot(A^{-1})^{\mathrm{t}}=(A^{-1}\cdot A)^{\mathrm{t}}=I_n^{\mathrm{t}}=I_n$。所以 $(A^{-1})^{\mathrm{t}}$ 就是 $A^{\mathrm{t}}$ 的反矩陣。也就是說直接將 $A$ 的反矩陣取轉置，就是 $A$ 的轉置矩陣的反矩陣。

**8.4.3. 對角化.** 我們先介紹“固有向量”（eigenvector）的概念，再說明何謂“可對角化矩陣”以及如何將之對角化。

給定一個 $n$ 階方陣 $A$，所謂 $A$ 的 eigenvector 表示是一個非零向量 $\mathbf{v}$ 使得 $A\mathbf{v}$ 和 $\mathbf{v}$ 是平行的。排除零向量的原因是，對任意方陣 $A\mathbf{0}=\mathbf{0}$ 必然成立，所以沒有必要去談。然而當 $\mathbf{v}\neq\mathbf{0}$，和 $\mathbf{v}$ 平行的向量都是 $\lambda\mathbf{v}$ 這樣的形式，所以說 $\mathbf{v}$ 是 $A$ 的 eigenvector 等同於存在

$\lambda \in \mathbb{R}$ 使得 $A\mathbf{v} = \lambda\mathbf{v}$，這個 $\lambda$ 就稱為 $A$ 的 eigenvector $\mathbf{v}$ 所對應的 eigenvalue (固有值)。所以一般 eigenvector $\mathbf{v}$ 和其相對應的 eigenvalue $\lambda$ 要成對出現，這樣才能確定 $A\mathbf{v} = \lambda\mathbf{v}$。

**Example 8.4.13.** 考慮

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

我們有

$$A\mathbf{v}_1 = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2\mathbf{v}_1.$$

所以 $\mathbf{v}_1$ 是 $A$ 的一個 eigenvector, 而 $-2$ 是其對應的 eigenvalue. 同樣的

$$A\mathbf{v}_2 = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 20 \end{bmatrix} = 5 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 5\mathbf{v}_2.$$

所以 $\mathbf{v}_2$ 是 $A$ 的一個 eigenvector, 而 5 是其對應的 eigenvalue. 然而

$$A\mathbf{v}_3 = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 22 \end{bmatrix} \not\in \mathrm{Span}(\begin{bmatrix} 4 \\ 3 \end{bmatrix}).$$

所以 $\mathbf{v}_3$ 不是 $A$ 的一個 eigenvector. $\sharp$

另外要注意的是雖然零向量不會是 eigenvector，不過 0 可能會是 eigenvalue。也就是說若 $\mathbf{v} \neq \mathbf{0}$，滿足 $A\mathbf{v} = \mathbf{0} = 0\mathbf{v}$，則 $\mathbf{v}$ 會是 $A$ 的一個 eigenvalue 為 0 的 eigenvector。換言之，若方陣 $A$ 的 null space $N(A)$ 不是零向量空間，則 $N(A)$ 中所有非零向量都是 $A$ 的一個 eigenvalue 為 0 的 eigenvector。也因此我們可以說 0 是 $A$ 的 eigenvalue 等同於 $N(A)$ 中存在非零向量，亦即 $N(A) \neq \{\mathbf{0}\}$。

要如何找到一個方陣的 eigenvector 及其對應的 eigenvalue 呢? 其實一般的找法是先找到 eigenvalue, 然後再找出與其對應的 eigenvector. 首先觀察若 $\lambda \in \mathbb{R}$ 是 $A$ 的 eigenvalue, 表示存在一個非零向量 $\mathbf{v}$ 使得 $A\mathbf{v} = \lambda\mathbf{v}$. 由於 $I_n\mathbf{v} = \mathbf{v}$, 所以看成矩陣的運算 $\lambda\mathbf{v} = (\lambda I_n)\mathbf{v}$. 因此 $A\mathbf{v} = \lambda\mathbf{v}$ 就等同於 $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$. 換言之, $\lambda$ 是 $A$ 的 eigenvalue 等同於 $N(A - \lambda I_n)$ 中有非零向量，亦即 $N(A - \lambda I_n) \neq \{\mathbf{0}\}$ 故由 Theorem 8.4.2 知這也等同於 $\det(A - \lambda I_n) = 0$. 總言之, 要找到 $A$ 的 eigenvalue $\lambda$ 就是要找到 $\lambda$ 滿足 $\det(A - \lambda I_n) = 0$.

要怎樣找到 $\lambda$ 滿足 $\det(A - \lambda I_n) = 0$ 呢? 假設 $A = [a_{ij}]$, 若我們將 $t$ 視為變數, 考慮 $\det(A - tI_n)$. 由於

$$A - tI_n = \begin{bmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{bmatrix}$$

利用數學歸納法, 我們可以證明 $\det(A - tI_n)$ 會是一個以 $t$ 為變數的 $n$ 次實係數多項式. 而若 $t = \lambda$ 為此多項式的一實數根, 則 $\lambda$ 就會滿足 $\det(A - \lambda I_n) = 0$, 也就是說 $\lambda$ 就會是 $A$ 的一個 eigenvalue. 反之, 若 $\lambda$ 是 $A$ 的一個 eigenvalue, 就表示 $t = \lambda$ 會是多項式 $\det(A - tI_n)$ 的一個根. 由此可知多項式 $\det(A - tI_n)$ 可以讓我們完全掌握 $A$ 的 eigenvalue, 我們因而給它一個特別的定義.

**Definition 8.4.14.** 假設 $A$ 為 $n$ 階方陣，考慮以 $t$ 為變數的多項式 $p_A(t) = \det(A - tI_n)$. 我們稱 $p_A(t)$ 為 $A$ 的 *characteristic polynomial* (特徵多項式)。

從上面的討論我們知道 $\lambda$ 為 characteristic polynomial $p_A(t)$ 的一個根若且唯若 $\lambda$ 為 $A$ 的 eigenvalue.

**Example 8.4.15.** 考慮 $B = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$, 此時 $A$ 的 characteristic polynomial 為

$$p_B(t) = \det(B - tI_3) = \det \begin{bmatrix} -1-t & 4 & 2 \\ 1 & 3-t & 1 \\ -1 & 2 & 2-t \end{bmatrix}.$$

對第一個 row 降階求行列式得

$$p_A(t) = (-1-t)\det \begin{bmatrix} 3-t & 1 \\ 2 & 2-t \end{bmatrix} - 4\det \begin{bmatrix} -1 & 1 \\ -1 & 2-t \end{bmatrix} + 2\det \begin{bmatrix} -1 & 3-t \\ -1 & 2 \end{bmatrix}.$$

化簡可得 $p_A(t) = -(t-1)^2(t-2)$. 也因此 $t = 1$ 和 $t = 2$ 為 $A$ 的 characteristic polynomial 的二實根, 也因此得 $A$ 有兩個 eigenvalues $1, 2$.

從多項式根與係數的關係，特徵多項式的一些係數是很好決定的。事實上 $A$ 的特徵多項式為 $t$ 為變數的 $n$ 次多項式. 其 $t^n$ 項係數為 $(-1)^n$，其 $t^{n-1}$ 項係數為 $(-1)^{n-1}\operatorname{tr}(A)$（我們用 $\operatorname{tr}(A)$ 表示 $A$ 對角線之和）而特徵多項式常數項係數為 $\det(A)$。這個方法求二階方陣的特徵多項式特別方便。

**Example 8.4.16.** 考慮 Example 8.4.13 中 $A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ 的 characteristic polynomial $p_A(t)$. 由於 $\operatorname{tr}(A) = 1 + 2 = 3$ 以及 $\det(A) = 2 - 12 = -10$, 利用前面所述可得

$$p_A(t) = (-1)^2 t^2 + (-1)3t + (-10) = t^2 - 3t - 10.$$

事實上利用 characteristic polynomial 的定義直接計算可得

$$p_A(t) = \det \begin{bmatrix} 1-t & 3 \\ 4 & 2-t \end{bmatrix} = (1-t)(2-t) - 12 = t^2 - 3t - 10.$$

分解後可得 $-2, 5$ 為 $A$ 的 eigenvalues.

接下來我們介紹一個和 eigenvalue 有關的定義. 若 $\lambda$ 是 $A$ 的 eigenvalue. 由於 $t = \lambda$ 會是 $A$ 的特徵多項式 $p_A(t) = \det(A - tI_n)$ 的一個根. 由因式定理知 $t - \lambda$ 會整除 $p_A(t)$. 若 $(t-\lambda)^m$ 可整除 $p_A(t)$, 但 $(t-\lambda)^{m+1}$ 不能整除 $p_A(t)$, 則我們稱 eigenvalue $\lambda$ 的 *algebraic multiplicity*（代數重根數）為 $m$。例如 Example 8.4.15 中 $B$ 有兩個 eigenvalue 1 和 2, 其中 eigenvalue 1 的 algebraic multiplicity 為 2, 而 eigenvalue 2 的 algebraic multiplicity 為 1. 而 Example 8.4.16 中 $A$ 的兩個 eigenvalue $-2, 5$ 其 algebraic multiplicity 皆為 1.

我們了解了如何找 eigenvalue 之後, 接下來便是要找出這些 eigenvalue 所對應的 eigenvectors. 假設 $\lambda$ 為 $n$ 階方陣 $A$ 的一個 eigenvalue. 由於 $\det(A - \lambda I_n) = 0$, 我們知聯立方程組 $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ 存在非零的解. 現假設 $\mathbf{v} \in \mathbb{R}^n$ 為非零向量且為 $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ 的一組解. 此即表示 $\mathbf{v}$ 滿足 $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$, 亦即 $A\mathbf{v} = \lambda\mathbf{v}$. 故此時 $\mathbf{v}$ 為 $A$ 的一個以 $\lambda$ 為 eigenvalue

的 eigenvector. 反之, 若 $\mathbf{v}$ 為 $A$ 的一個以 $\lambda$ 為 eigenvalue 的 eigenvector, 則 $\mathbf{x}=\mathbf{v}$ 必為 $(A-\lambda I_n)\mathbf{x}=\mathbf{0}$ 的一組非零解。因此我們只要掌握 $n\times n$ matrix $A-\lambda I_n$ 的 null space 中的非零向量就會是 $A$ 相對於 $\lambda$ 的 eigenvector。因此我們有以下的定義.

**Definition 8.4.17.** 假設 $A$ 為 $n$ 階方陣且 $\lambda\in\mathbb{F}$ 為 $A$ 的一個 eigenvalue. 則 $A-\lambda I_n$ 的 null space 稱為 $A$ 對於 eigenvalue $\lambda$ 的 *eigenspace*. 我們用 $E_A(\lambda)$ 來表示.

由於 eigenspace 是由 eigenvector 組成的向量空間，它能讓我們了解 eigenvectors，因此我們定義 $E_A(\lambda)$ 的 dimension 為 eigenvalue $\lambda$ 的 *geometric multiplicity* (幾何重根數). 要注意 eigenvalue $\lambda$ 的 algebraic multiplicity 無法讓我們知道 $\lambda$ 所對應的 eigenvectors 的多寡, 而是 $\lambda$ 的 geometric multiplicity 可以提供這一個訊息.

**Example 8.4.18.** 考慮 $A=\begin{bmatrix}1&3\\4&2\end{bmatrix}$, $B=\begin{bmatrix}-1&4&2\\-1&3&1\\-1&2&2\end{bmatrix}$. 由前面 Example 8.4.15, 8.4.16 我們已計算出 $A$ 和 $B$ 的 characteristic polynomial 分別為 $p_A(t)=(x+2)(x-5)$, $p_B(t)=-(t-1)^2(t-2)$. 接下來我們分別計算 $A$ 和 $B$ 的 eigenspace.

首先考慮 $A$ 對於 eigenvalue $-2$ 的 eigenspace, 亦即找出 $A-(-2I_2)=\begin{bmatrix}3&3\\4&4\end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix}1&1\\0&0\end{bmatrix}$. 可得 $E_A(1)=\mathrm{Span}(\begin{bmatrix}1\\-1\end{bmatrix})$. 也就是說 $A$ 對於 eigenvalue 為 $-2$ 的 eigenvector 就是那些和 $\begin{bmatrix}1\\-1\end{bmatrix}$ 平行的 nonzero vector. 由於 $\dim(E_A(-2))=1$, 我們也得到 $A$ 對於 eigenvalue $-2$ 的 geometric multiplicity 為 1. 至於 $A$ 對於 eigenvalue 5 的 eigenspace, 亦即找出 $A-5I_2=\begin{bmatrix}-4&3\\4&-3\end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix}-4&3\\0&0\end{bmatrix}$. 因此得 $E_A(5)=\mathrm{Span}(\begin{bmatrix}3\\4\end{bmatrix})$. 也就是說 $A$ 對於 eigenvalue 為 5 的 eigenvector 就是那些和 $\begin{bmatrix}3\\4\end{bmatrix}$ 平行的 nonzero vector, 我們也得到 $A$ 對於 eigenvalue 5 的 geometric multiplicity 為 1. 在 Example 8.4.13 中我們舉出 $A$ 的 eigenvector 的例子其實是這樣得到的.

接著考慮 $B$ 對於 eigenvalue 1 的 eigenspace, 亦即找出 $B-I_3=\begin{bmatrix}-2&4&2\\-1&2&1\\-1&2&1\end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix}1&-2&-1\\0&0&0\\0&0&0\end{bmatrix}$. 可得 $E_B(1)=\mathrm{Span}(\begin{bmatrix}2\\1\\0\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix})$. 也就是說 $B$ 對於 eigenvalue 為 1 的 eigenvector 就是那些由 $\begin{bmatrix}2\\1\\0\end{bmatrix}$

和 $\begin{bmatrix}1\\0\\1\end{bmatrix}$ 的 linear combination 所得的 nonzero vector. 例如 $\mathbf{v}=\begin{bmatrix}4\\1\\2\end{bmatrix}=\begin{bmatrix}2\\1\\0\end{bmatrix}+2\begin{bmatrix}1\\0\\1\end{bmatrix}$ 就滿足

$$B\mathbf{v}=\begin{bmatrix}-1&4&2\\-1&3&1\\-1&2&2\end{bmatrix}\begin{bmatrix}4\\1\\2\end{bmatrix}=\begin{bmatrix}4\\1\\2\end{bmatrix}=\mathbf{v}.$$

由於 $\dim(E_B(1))=2$, 我們也得到 $B$ 對於 eigenvalue 1 的 geometric multiplicity 為 2. 至於 $B$ 對於 eigenvalue 2 的 eigenspace, 亦即找出 $B-2I_3=\begin{bmatrix}-3&4&2\\-1&1&1\\-1&2&0\end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix}1&-2&0\\0&1&-1\\0&0&0\end{bmatrix}$. 因此得 $E_B(2)=\mathrm{Span}(\begin{bmatrix}2\\1\\1\end{bmatrix})$. 也就是說 $B$ 對於 eigenvalue 為 2 的 eigenvector 就是那些和 $\begin{bmatrix}2\\1\\1\end{bmatrix}$ 平行的 nonzero vector, 我們也得到 $B$ 對於 eigenvalue 2 的 geometric multiplicity 為 1. ♯

當 $n$ 階方陣 $A$ 若存在 $n$ 個線性獨立的 eigenvectors, 則稱 $A$ 為"可對角化" *diagonalizable*。我們將探討如何判斷一個方陣是否是 diagonalizable，並說明如何將一個可對角化的矩陣對角化。

要如何知道 $A$ 是否為可對角化呢? 從其定義, 我們知道它必須要有夠多的 eigenvectors。首先要有夠多的 eigenvectors 就表示要有夠多的 eigenvalues, 所以 $A$ 的特徵多項式必需可以在實數中完全分解，也就是 $p_A(t)=(-1)^n(t-\lambda_1)^{m_1}\cdots(t-\lambda_k)^{m_k}$ 其中 $\lambda_1,\ldots,\lambda_k$ 為相異實數。依定義 $m_i$ 就是 $\lambda_i$ 的 algebraic multiplicity 而且因 $p_A(t)$ 的次數為 $n$, 我們有 $m_1+\cdots+m_k=n$。另外每個 eigenvalue, 其 geometric multiplicity 會小於等於其 algebraic multiplicity。所以這裡 $A$ 的 eigenvectors 要夠多, 最好的狀況就是每一個 eigenvalue 其 geometric multiplicity 等於其 algebraic multiplicity，亦即 $\dim(E_A(\lambda_i))=m_i$。此時我們令 $\mathbf{v}_{i,1},\ldots,\mathbf{v}_{i,m_i}$ 為 $E_A(\lambda_i)$ 的一組 basis，將這 $k$ 組 vectors 收集在一起後，由 eigenvector 的特性，這 $m_1+\cdots+m_k=n$ 個 eigenvectors 會線性獨立。所以此時 $A$ 為可對角化。

由上可知，要檢查一個方陣是否可對角化，不只要檢查特徵多項式可否完全分解，還要檢查每個 eigenvalue 的代數重根數與幾何重根數是否相等。不過在 eigenvalue 為特徵多項式的單根（即代數重根數為 1）時，由於幾何重根數必大於 0（因對應的 eigenvector 必存在），又幾何重根數小於等於其代數重根數，所以此時我們不必檢查就可知此 eigenvalue 的幾何重根數一定等於其代數重根數（皆為 1）.

為甚麼稱為可對角化呢？這是因為若 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是線性獨立且皆為 $A$ 的 eigenvectors，又假設它們所對應的 eigenvalues 分別為 $\lambda_1,\ldots,\lambda_n$. 亦即 $A\mathbf{v}_1=\lambda_1\mathbf{v}_1,\ldots,A\mathbf{v}_n=\lambda_n\mathbf{v}_n$. 此時由矩陣乘法的定義我們有

$$A\begin{bmatrix}|&|&&|\\\mathbf{v}_1&\mathbf{v}_2&\cdots&\mathbf{v}_n\\|&|&&|\end{bmatrix}=\begin{bmatrix}|&|&&|\\A\mathbf{v}_1&A\mathbf{v}_2&\cdots&A\mathbf{v}_n\\|&|&&|\end{bmatrix}=\begin{bmatrix}|&|&&|\\\lambda_1\mathbf{v}_1&\lambda_2\mathbf{v}_2&\cdots&\lambda_n\mathbf{v}_n\\|&|&&|\end{bmatrix}.$$

另一方面若考慮 $(i,i)$-th entry 為 $\lambda_i$ 的 $n\times n$ diagonal matrix $D$(即對角線第 $i$ 個位置為 $\lambda_i$ 而對角線外其餘位置皆為 0), 則我們有

$$\begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} D = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_3\mathbf{v}_n \\ | & | & & | \end{bmatrix}.$$

因此若令 $C = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$, 則我們有 $AC = CD$. 現又因 $C$ 的 column 之間為 linearly independent 且有 $n$ 個 column, 我們得 $C$ 的 rank 為 $n$, 因此由 $C$ 為 $n\times n$ matrix 得知 $C$ 為 invertible (參見 Theorem 8.4.2). 因此我們可將 $AC = CD$ 改寫成 $D = C^{-1}AC$。這個表示法就稱為將 $A$ 對角化。

**Example 8.4.19.** 我們考慮矩陣 $A = \begin{bmatrix} 0 & 3 & 1 \\ -1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$. 經計算可得它們有相同的 characteristic polynomial $-(t-1)^2(t-2)$. 也因此 $A,B$ 的 eigenvalue 1 其 algebraic multiplicity 皆為 2, 而 eigenvalue 2 的 algebraic multiplicity 皆為 1. 由於 eigenvalue 2 的 algebraic multiplicity 為 1, 我們知其 geometric multiplicity 亦為 1, 所以我們僅要檢查 eigenvalue 1 的 geometric multiplicity 即可.

矩陣 $A$ 對於 eigenvalue 1 的 eigenspace, 即 $A - I_3 = \begin{bmatrix} -1 & 3 & 1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. 可得 $E_A(1) = \mathrm{Span}(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix})$. 也就是說 $A$ 對於 eigenvalue 為 1 的 eigenvector 就是那些和 $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ 平行的 nonzero vector, 我們也得到 $A$ 對於 eigenvalue 1 的 geometric multiplicity 為 1. 因其 geometric multiplicity 不等於 algebraic multiplicity, 可得 $A$ 不是 diagonalizable matrix. 回顧在 Example 8.4.18 中我們計算過 $B$ 在 eigenvalue 1 和 eigenvalue 2 的 geometric multiplicity 皆等於其 algebraic multiplicity, 所以 $B$ 為 diagonalizable matrix. 我們看如何將 $B$ 對角化.

由於 $B$ 對於 eigenvalue 為 1 和 2 的 eigenspace 分別為 $E_B(1) = \mathrm{Span}(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix})$ 和 $E_B(2) = \mathrm{Span}(\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix})$, 可得 $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ 就是一組由 $B$ 的 eigenvectors 所形成的 $\mathbb{R}^3$ 的

basis. 因此若令 $C=\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ 以及 $D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, 則

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

再由 $C$ 為 invertible, 得 $C^{-1}BC=D$. ♯

將可對角化的矩陣對角化有許多應用，例如若 $A$ 為可對角化，我們可以利用對角化求 $A$ 的高次方。由於存在可逆矩陣 $Q$ 使得 $Q^{-1}AQ$ 為對角矩陣 $D$。換言之, 我們可以將 $A$ 寫成 $A=QDQ^{-1}$. 也因此我們可得

$$A^2=(QDQ^{-1})(QDQ^{-1})=QD^2Q^{-1}.$$

同理對任意 $m\in\mathbb{N}$, 我們有 $A^m=QD^mQ^{-1}$。寫成這樣有什麼好處呢？因為 $D$ 為對角矩陣 $\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, 我們很可以容易算出 $D^m$, 即 $\begin{bmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{bmatrix}$. 因此只要知道 $Q$ 和 $Q^{-1}$, 我們就可以很輕易算出 $A^m$ (即 $QD^mQ^{-1}$), 而不必真正將 $A$ 乘到 $m$ 次方了.

**Example 8.4.20.** 考慮實矩陣 $B=\begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$. 在 Example 8.4.19 我們算出 $Q^{-1}BQ=D$, 其中 $Q=\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ 以及 $D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. 由於 $Q^{-1}=\begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & 0 \\ -1 & 2 & 1 \end{bmatrix}$, 我們得

$$B^5=QD^5Q^{-1}=\begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 32 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & 0 \\ -1 & 2 & 1 \end{bmatrix}=\begin{bmatrix} -61 & 124 & 62 \\ -31 & 63 & 31 \\ -31 & 62 & 32 \end{bmatrix}.$$

**8.4.4. Spectral Theorem.** 我們曾提及在檢查一個矩陣是否為 diagonalizable 時, 對於 algebraic multiplicity 為 1 的 eigenvalue 我們就不必檢查其 geometric multiplicity 了。也因此若 $A$ 的 characteristic polynomial 可以完全分解且其根皆為單根（無重根）, 則 $A$ 一定為 diagonalizable. 另外還有一種矩陣不必檢查就知道一定是 diagonalizable, 就是對稱矩陣。對稱矩陣一定可以對角化, 更重要的是它們都可以“正交對角化”（*orthogonal diagonalizable*）。這就是 *spectral theorem*。這個定理在數學和物理方面都有很重要的應用，不過我們不去探討它的應用，而著重於說明如何將對稱矩陣正交對角化。

所謂一個 $n$ 階方陣可以正交對角化，指的就是不只可以對角化（即找到 $n$ 個線性獨立的 eigenvectors），而且這些 eigenvectors 兩兩互相垂直（*orthogonal eigenvectors*）。所以當我們拿到一個對稱矩陣，由於它一定可以對角化，我們不必擔心它的特徵多項式可否完全分解，也不必擔心幾何重根數是否等於代數重根數，直接利用前述的方法找到所有 eigenvalue 的 eigenspace 的一組基底，形成一組線性獨立的 eigenvectors。接下來的重點便是如何利用這一組 eigenvectors 找到另一組兩兩互相垂直的 eigenvectors。

首先要注意的是由於考慮的是對稱矩陣，不同 eigenvalue 的 eigenvectors 一定互相垂直。所以我們只要專注於同樣的 eigenvalue 的 eigenvectors 如何讓它們兩兩互相垂直。也就是如何讓找出的 eigenspace 的基底兩兩互相垂直。這裡要用到的便是所謂的 Gram-Schmidt process：假設 $\mathbf{v}_1,\ldots,\mathbf{v}_k$ 為 eigenvalue $\lambda$ 的 eigenspace 的一組 basis. 令

$$\mathbf{w}_1=\mathbf{v}_1,\quad \mathbf{w}_2=\mathbf{v}_2-\frac{\mathbf{v}_2\cdot\mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1,\quad \mathbf{w}_3=\mathbf{v}_3-\frac{\mathbf{v}_3\cdot\mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1-\frac{\mathbf{v}_3\cdot\mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2,\quad \ldots$$

這樣一直下去, 即對於 $i=1,\ldots,k-1$ 令

$$\mathbf{w}_{i+1}=\mathbf{v}_{i+1}-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2-\cdots-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_i\rangle}{\|\mathbf{w}_i\|^2}\mathbf{w}_i,$$

則 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 依然會是 eigenvalue $\lambda$ 的 eigenvector 且兩兩互相垂直 (稱為此 eigenspace 的 *orthogonal basis*)。接著令 $\mathbf{u}_i=\frac{1}{\|\mathbf{w}_i\|}\mathbf{w}_i$ 則 $\mathbf{u}_1,\ldots,\mathbf{u}_k$ 不止兩兩互質而且長度為 1 (稱為此 eigenspace 的一組 *orthonormal basis*)。現假設 $\mathbf{u}_1,\ldots,\mathbf{u}_n$ 為 $n$ 階方陣 $A$ 所有的 eigenspace 的 orthonormal basis 所組成的 $n$ 個 eigenvectors。此時若 $\mathbf{u}_i$ 所對應的 eigenvalue 為 $\lambda_i$ 且令 $Q=\begin{bmatrix}| & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & |\end{bmatrix}$ 則可得 $AQ=QD$ 其中 $D$ 為 $(i,i)$-th entry 為 $\lambda_i$ 的對角矩陣，也就是說我們可以將 $A$ 對角化成 $Q^{-1}AQ=D$。由於 $\mathbf{u}_1,\ldots,\mathbf{u}_n$ 多了兩兩互相垂直以及長度為 1 這兩個條件，很容易檢查 $Q^{\mathrm{t}}Q=I_n$, 也因此知 $Q^{\mathrm{t}}=Q^{-1}$。就因為這個特性，我們可以將 $A$ 對角化成 $Q^{\mathrm{t}}AQ=D$，這就是對稱矩陣 $A$ 的正交對角化。

**Example 8.4.21.** (1) 考慮 symmetric matrix $A=\begin{bmatrix}0&1&1\\1&1&0\\1&0&1\end{bmatrix}$. 我們有 $A$ 的 characteristic polynomial 為 $p_A(t)=-(t+1)(t-1)(t-2)$. 所以 $A$ 有三個相異的 eigenvalues, $-1,1,2$. 知道 $A$ 必能對角化，而且因 eigenvalue 皆相異，知它們所對應的 eigenvector 會倆倆互相垂直。事實上我們可求出對應到 $-1,1,2$ 的 eigenvector 分別為

$$\mathbf{v}_1=\begin{bmatrix}-2\\1\\1\end{bmatrix},\quad \mathbf{v}_2=\begin{bmatrix}0\\-1\\1\end{bmatrix},\quad \mathbf{v}_3=\begin{bmatrix}1\\1\\1\end{bmatrix}.$$

很容易檢查它們確實倆倆互相垂直。此時對於 $i=1,2,3$ 令 $\mathbf{u}_i=\frac{1}{\|\mathbf{v}_i\|}\mathbf{v}_i$, 我們得

$$\mathbf{u}_1=\frac{1}{\sqrt{6}}\begin{bmatrix}-2\\1\\1\end{bmatrix},\quad \mathbf{u}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\-1\\1\end{bmatrix},\quad \mathbf{u}_3=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

倆倆互相垂直且長度為 1，故可將 $A$ 正交對角化成

$$\begin{bmatrix}-\frac{2}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\0&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\end{bmatrix}\begin{bmatrix}0&1&1\\1&1&0\\1&0&1\end{bmatrix}\begin{bmatrix}-\frac{2}{\sqrt{6}}&0&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\end{bmatrix}=\begin{bmatrix}-1&0&0\\0&1&0\\0&0&2\end{bmatrix}.$$

(2) 考慮 symmetric matrix $B=\begin{bmatrix}5&-4&-2\\-4&5&-2\\-2&-2&8\end{bmatrix}$. 我們有 $B$ 的 characteristic polynomial 為 $p_B(t)=-t(t-9)^2$. 所以 $B$ eigenvalues 為 $0,9$. 知道 $B$ 必能對角化, 我們

知 $\dim(E_B(0))=1$, $\dim(E_B(9))=2$. 事實上 $\mathbf{v}_1=\begin{bmatrix}2\\2\\1\end{bmatrix}$ 為 $E_B(0)=N(B)$ 的 basis, 而 $\mathbf{v}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}-1\\0\\2\end{bmatrix}$ 為 $E_B(9)$ 的 basis. 而且由 $\mathbf{v}_1$ 的 eigenvalue 與 $\mathbf{v}_2,\mathbf{v}_3$ 的 eigenvalue 不同，我們知 $\mathbf{v}_1\cdot\mathbf{v}_2=\mathbf{v}_1\cdot\mathbf{v}_3=0$，很容易檢查它們確實成立。不過 $\mathbf{v}_2\cdot\mathbf{v}_3=1\neq 0$, 我們必須利用 Gram-Schmidt process 將 $\mathbf{v}_2,\mathbf{v}_3$ 換成 $E_B(9)$ 的一組 orthogonal basis. 令 $\mathbf{w}_2=\mathbf{v}_2$ 且

$$\mathbf{w}_3=\mathbf{v}_3-\mathrm{Proj}_{\mathbf{w}_2}\mathbf{v}_3=\begin{bmatrix}-1\\0\\2\end{bmatrix}-\frac{1}{2}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\frac{1}{2}\begin{bmatrix}-1\\-1\\4\end{bmatrix}.$$

此時令 $\mathbf{u}_1=\frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1$, $\mathbf{u}_2=\frac{1}{\|\mathbf{w}_2\|}\mathbf{w}_2$, $\mathbf{u}_3=\frac{1}{\|\mathbf{w}_3\|}\mathbf{w}_3$ 我們得

$$\mathbf{u}_1=\frac{1}{\sqrt{3}}\begin{bmatrix}2\\2\\1\end{bmatrix},\quad \mathbf{u}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix},\quad \mathbf{u}_3=\frac{1}{3\sqrt{2}}\begin{bmatrix}-1\\-1\\4\end{bmatrix}$$

倆倆互相垂直且長度為 1，故可將 $B$ 正交對角化成

$$\begin{bmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\-\frac{1}{3\sqrt{2}}&-\frac{1}{3\sqrt{2}}&\frac{4}{3\sqrt{2}}\end{bmatrix}\begin{bmatrix}5&-4&-2\\-4&5&-2\\-2&-2&8\end{bmatrix}\begin{bmatrix}\frac{2}{3}&-\frac{1}{\sqrt{2}}&-\frac{1}{3\sqrt{2}}\\\frac{2}{3}&\frac{1}{\sqrt{2}}&-\frac{1}{3\sqrt{2}}\\\frac{1}{3}&0&\frac{4}{3\sqrt{2}}\end{bmatrix}=\begin{bmatrix}0&0&0\\0&9&0\\0&0&9\end{bmatrix}.$$

♯