2

線性變換與矩陣表現

線性代數理論中一個重要的研究是要徹底了解向量空間之間的線性變換, 這一章即以線性變換為主體給出重要的結果。單元 2.1 首先給出線性變換的定義還有基本的觀察, 並給出線性變換的例子, 當中有一些線性變換可以用歐氏空間中的幾何變換像是伸縮、反射、投影、旋轉了解它。

單元 2.2 則是探討線性變換的性質。給定一個線性變換, 定義域與對應域分別都是向量空間, 假設兩者維度皆有限, 而各自有重要的子空間稱為零核空間與值域, 維度定理告知而這兩個空間的維度和與定義域的維度一致。此外, 我們從映射的一對一與映成的關係觀察線性變換, 則有更好的結果; 也就是說, 一對一與映成這兩個性質在有限維向量空間的線性變換下是等價的。對於一個線性變換, 只要告知基底如何變換, 那麼這個線性變換也完全確定。

對於一個抽象的線性變換, 前面的討論都是從數學形式上的操作了解它, 若要具體呈現一個線性變換, 則是引進矩陣將線性變換轉化成矩陣的運算, 故單元 2.3 是要介紹線性變換的矩陣表現。在這個理論中我們要先確定的是: 給定兩個向量空間, 則所有在這兩個向量空間的線性變換所成的集合將形成一個向量空間, 於是線性變換的加法與純量乘法, 透過線性變換的矩陣表現後, 會對應於矩陣的加法與純量乘法。

單元 2.4 則是在討論線性變換的合成, 而合成變換對應於矩陣表現則是兩個矩陣的乘法。而單元 2.5 是將矩陣單獨拿出來研究, 暫時忘掉矩陣和線性變換的關係, 專注於矩陣運算的規則並探討矩陣的結構。至於矩陣的左乘變換之介紹, 初看之下只是把前面的討論再用一種新的記號表示, 而這麼做的目的是下一章討論矩陣基本列運算時可以把理論說明清楚。

單元 2.6 則是介紹可逆變換, 並給出可逆變換的矩表現是原線性變換之矩陣表現的反矩陣之結果。從兩個向量空間可逆變換的存在性引進向量空間同構的概念, 在單元 2.7 將證明兩個有限維向量空間是同構的等價條件是這兩個向量空間的維度相同。此外, 我們還可以再把線性變換與矩陣變換之間的關係再看得更清楚。

我們從坐標變換的過程得知線性代數的應用。單元 2.8 是以圓錐曲線的分類為例介紹坐標變換的矩陣表現, 關於這部分的理論則是要用矩陣正交對角化的觀點才有辦法完整說明。

單元 2.9 要介紹對偶空間。對偶空間的應用層面很廣, 除了可以用來回答轉置矩陣的意義, 還可以解釋有限維向量空間與其雙對偶空間的同構關係; 此外, 微積分理論中單變數函數的微分 (differential) 與多變數函數的全微分 (total differential) 若要確實將這個概念解釋清楚的話, 也是必須透過對偶空間才有辦法說明清楚。

2.1 線性變換

這一個單元要探討兩個向量空間之間的轉換關係。因為向量空間具有向量加法與純量乘法的結構, 我們感興趣的問題是: 什麼樣的轉換關係可以對於兩個向量空間的運算有一致性或協調性? 這個問題的答案是本單元所要介紹的線性變換。首先給出線性變換的定義。

定義 1 (第 65 頁)**.** 假設 $(V, \mathbb{F})$ 與 $(W, \mathbb{F})$ 是兩個向量空間, 考慮一個映射 (mapping) $T : V \to W$ 滿足以下兩個條件:

(A) 對任意 $\mathbf{u}, \mathbf{v} \in V$ 都有 $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$。

(B) 對任意 $\mathbf{u} \in V$ 以及任何 $c \in \mathbb{F}$ 都有 $T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u})$。

我們說映射 $T : V \to W$ 是一個由 V 映至 W 的 *線性變換* (linear transformation from V to W)。

這裡我們先給出線性變換的幾個觀察:

- 在討論線性變換時, 兩個向量空間所分佈的體 $\mathbb{F}$ 必須相同。這是因為線性變換的第二個條件 $T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u})$ 當中的 c 是同一個數, 而等式左邊的 c 告知它必須屬於分佈於向量空間 V 上的 $\mathbb{F}$, 等式右邊表明這個 c 也必須屬於分佈於向量空間 W 上的 $\mathbb{F}$。

- 由於向量空間 $(V, \mathbb{F})$ 帶有加法運算 $+$ 與純量乘法運算 $\cdot$, 而向量空間 $(W, \mathbb{F})$ 也帶有加法運算 $+$ 與純量乘法運算 $\cdot$, 注意到線性變換的第一個條件 $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, 等式左邊的 $+$ 是向量空間 V 的向量加法, 而等式右邊的 $+$ 是向量空間 W 的向量加法; 至於第二個條件 $T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u})$, 等式左邊的 $\cdot$ 是向量空間 V 的純量乘法, 而等式右邊的 $\cdot$ 是向量空間 W 的純量乘法。

  在代數上, 我們會用以下的交換圖 (commute diagram) 以強調線性變換在結構上的關係:

$$
\begin{tikzcd}
\mathbf{u}, \mathbf{v} \arrow[r, "T"] \arrow[d, "(V,+)"'] \arrow[dr, phantom, "\circlearrowright"] & T(\mathbf{u}), T(\mathbf{v}) \arrow[d, "(W,+)"] \\
\mathbf{u}+\mathbf{v} \arrow[r, "T"] & \begin{array}{c} T(\mathbf{u}+\mathbf{v}) \\ = T(\mathbf{u}) + T(\mathbf{v}) \end{array}
\end{tikzcd}
\qquad
\begin{tikzcd}
c, \mathbf{u} \arrow[r, "T"] \arrow[d, "(V,\cdot)"'] \arrow[dr, phantom, "\circlearrowright"] & c, T(\mathbf{u}) \arrow[d, "(W,\cdot)"] \\
c \cdot \mathbf{u} \arrow[r, "T"] & \begin{array}{c} T(c \cdot \mathbf{u}) \\ = c \cdot T(\mathbf{u}) \end{array}
\end{tikzcd}
$$

圖 2.1: 用交換圖的方式理解線性變換。

  從這個觀點來看, 各位應意識到線性變換的特色是在說明映射 (mapping) T 與運算 (operations) $+$ 或 $\cdot$ 操作雖然先後順序不同, 但最終結果要求一致。

- 若 $T : V \to W$ 是線性變換, 則 $T(\mathbf{0}) = \mathbf{0}$。這是因為取 $\mathbf{u} \in V$ 以及 $0 \in \mathbb{F}$, 則

$$T(0 \cdot \mathbf{u}) = 0 \cdot T(\mathbf{u}) \Rightarrow T(\mathbf{0}) = \mathbf{0}。$$

  同樣地, 這裡也應注意到等式左、右兩邊的 $\mathbf{0}$ 分別代表向量空間 V 與向量空間 W 的零向量。

- $T: V \to W$ 是線性變換若且唯若對任意 $\mathbf{u}, \mathbf{v} \in V$ 以及對任何 $c \in \mathbb{F}$ 都有
$$T(c \cdot \mathbf{u} + \mathbf{v}) = c \cdot T(\mathbf{u}) + T(\mathbf{v})。$$

  證明: ($\Rightarrow$) 已知 $T: V \to W$ 是線性變換, 則對任意 $\mathbf{u}, \mathbf{v} \in V$ 以及任何 $c \in \mathbb{F}$ 都有
$$T(c \cdot \mathbf{u} + \mathbf{v}) = T(c \cdot \mathbf{u}) + T(\mathbf{v}) = c \cdot T(\mathbf{u}) + T(\mathbf{v})。$$

  ($\Leftarrow$) 已知: 對任意 $\mathbf{u}, \mathbf{v} \in V$ 以及任何 $c \in \mathbb{F}$ 都有 $T(c \cdot \mathbf{u} + \mathbf{v}) = c \cdot T(\mathbf{u}) + T(\mathbf{v})$; 特別地, 取 $c = 1 \in \mathbb{F}$, 得到 $T(1 \cdot \mathbf{u} + \mathbf{v}) = 1 \cdot T(\mathbf{u}) + T(\mathbf{v})$, 即 $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$; 此外, 取 $\mathbf{v} = \mathbf{0}$, 得到 $T(c \cdot \mathbf{u} + \mathbf{0}) = c \cdot T(\mathbf{u}) + T(\mathbf{0})$, 因為 $T(\mathbf{0}) = \mathbf{0}$, 所以 $T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u})$。因此 $T: V \to W$ 是一個線性變換。 □

- 我們可以用數學歸納法 (Mathematical Induction) 得到以下結論: $T: V \to W$ 是線性變換若且唯若對任意 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ 以及對任何 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 都有
$$T\left(\sum_{i=1}^{n} c_i \cdot \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i \cdot T(\mathbf{v}_i)。$$

介紹完線性變換的定義還有幾個基本性質後, 以下給出線性變換的例子。

例 **2.** 考慮 $T: \mathbb{R}^2 \to \mathbb{R}^2$ 定義如下:
$$T(a_1, a_2) = (c_{11} \cdot a_1 + c_{12} \cdot a_2, c_{21} \cdot a_1 + c_{22} \cdot a_2),$$
其中 $c_{11}, c_{12}, c_{21}, c_{22} \in \mathbb{R}$。證明: T 是一個線性變換。

證明:

- 對任意 $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$, 都有
$$\begin{aligned}
T(\mathbf{u} + \mathbf{v}) &= T(u_1 + v_1, u_2 + v_2) \\
&= (c_{11} \cdot (u_1 + v_1) + c_{12} \cdot (u_2 + v_2), c_{21} \cdot (u_1 + v_1) + c_{22} \cdot (u_2 + v_2)) \\
&= (c_{11} \cdot u_1 + c_{11} \cdot v_1 + c_{12} \cdot u_2 + c_{12} \cdot v_2, c_{21} \cdot u_1 + c_{21} \cdot v_1 + c_{22} \cdot u_2 + c_{22} \cdot v_2) \\
&= (c_{11} \cdot u_1 + c_{12} \cdot u_2, c_{21} \cdot u_1 + c_{22} \cdot u_2) + (c_{11} \cdot v_1 + c_{12} \cdot v_2, c_{21} \cdot v_1 + c_{22} \cdot v_2) \\
&= T(u_1, u_2) + T(v_1, v_2) = T(\mathbf{u}) + T(\mathbf{v})。
\end{aligned}$$

- 對任意 $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ 以及對任何 $c \in \mathbb{R}$, 都有
$$\begin{aligned}
T(c \cdot \mathbf{u}) &= T(c \cdot u_1, c \cdot u_2) = (c_{11} \cdot (c \cdot u_1) + c_{12} \cdot (c \cdot u_2), c_{21} \cdot (c \cdot u_1) + c_{22} \cdot (c \cdot u_2)) \\
&= (c_{11} \cdot c \cdot u_1 + c_{12} \cdot c \cdot u_2, c_{21} \cdot c \cdot u_1 + c_{22} \cdot c \cdot u_2) \\
&= c \cdot (c_{11} \cdot u_1 + c_{12} \cdot u_2, c_{21} \cdot u_1 + c_{22} \cdot u_2) = c \cdot T(u_1, u_2) = c \cdot T(\mathbf{u})。
\end{aligned}$$

- 由上述討論得知 T 是一個線性變換。

□

各位可以從這個例子中體會到: 若一個映射的每個分量表示都是將定義域的分量做線性組合的話, 那麼這個映射會是一個線性變換。

幾何學上有幾個重要的線性變換, 這裡以 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的轉換為例進行說明, 注意到這些概念可以推廣至一般歐氏空間 $\mathbb{R}^n$ 的情形, 各位在看完以下的討論後可以思考如何將它們類推。

定義 3 (第 66 頁)**.**

(A1) 給定 $\lambda \in \mathbb{R}$, 定義 $D_{x,\lambda} : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$D_{x,\lambda}(a_1, a_2) = (\lambda \cdot a_1, a_2),$$

稱 $D_{x,\lambda}$ 是 對於 x-軸的伸縮變換 (dilation about the x-axis)。

(A2) 給定 $\lambda \in \mathbb{R}$, 定義 $D_{y,\lambda} : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$D_{y,\lambda}(a_1, a_2) = (a_1, \lambda \cdot a_2),$$

稱 $D_{y,\lambda}$ 是 對於 y-軸的伸縮變換 (dilation about the y-axis)。

(A3) 給定 $\lambda \in \mathbb{R}$, 定義 $D_{\lambda} : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$D_{\lambda}(a_1, a_2) = (\lambda \cdot a_1, \lambda \cdot a_2),$$

稱 D_{λ} 是 對於坐標原點的伸縮變換 (dilation about the origin)。

(B1) 定義 $R_x : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$R_x(a_1, a_2) = (a_1, -a_2),$$

稱 R_x 是 對於 x-軸的反射變換 (reflection about the x-axis)。

(B2) 定義 $R_y : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$R_y(a_1, a_2) = (-a_1, a_2),$$

稱 R_y 是 對於 y-軸的反射變換 (reflection about the y-axis)。

(C1) 定義 $P_x : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$P_x(a_1, a_2) = (a_1, 0),$$

稱 P_x 是 對於 x-軸的投影變換 (projection on the x-axis)。

(C2) 定義 $P_y : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$P_y(a_1, a_2) = (0, a_2),$$

稱 P_y 是 對於 y-軸的投影變換 (projection on the y-axis)。

(D) 給定 $\theta \in \mathbb{R}$, 定義 $R_{\theta} : \mathbb{R}^2 \to \mathbb{R}^2$ 為

$$R_{\theta}(a_1, a_2) = (\cos\theta \cdot a_1 - \sin\theta \cdot a_2, \sin\theta \cdot a_1 + \cos\theta \cdot a_2),$$

稱 R_{θ} 是一種 旋轉角爲 θ 的旋轉變換 (rotation counterclockwise by θ)。

我們從 例 **2** 的討論就可以清楚地看出這些幾何變換都是線性變換, 現將這些幾何變換與 例 **3** 中 $c_{11}, c_{12}, c_{21}, c_{22}$ 的值以矩陣的方式寫出:

$$D_{x,\lambda}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} \qquad D_{y,\lambda}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}$$

$$D_{\lambda}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \qquad R_{x}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$R_{y}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad P_{x}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$P_{y}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad R_{\theta}: \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

我們會在之後的單元仔細說明這些線性變換為什麼要用矩陣的方式呈現, 而且線性變換要怎麼樣具體地用矩陣表現, 以及矩陣的每一行所代表的意義。各位此時此刻應該要多花心思在那些線性變換表示法與幾何意義之間的關係, 這裡用圖形的方式說明這些變換的概念。

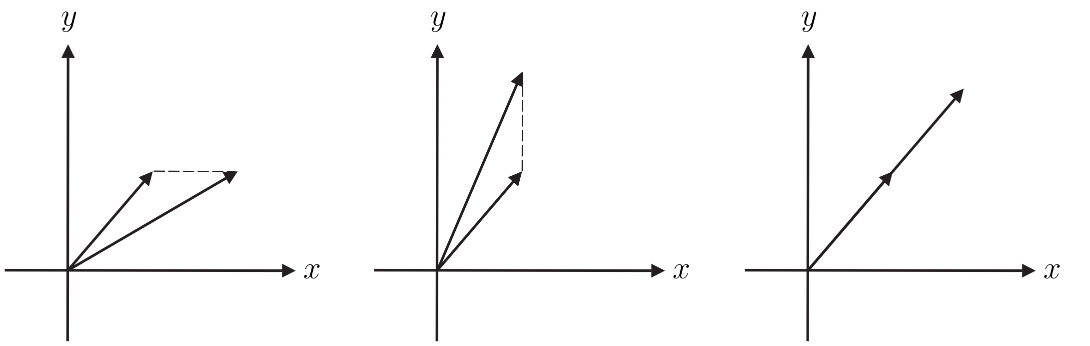

圖 2.2: 左: 對於 x-軸的伸縮變換。中: 對於 y-軸的伸縮變換。右: 對於坐標原點的伸縮變換。

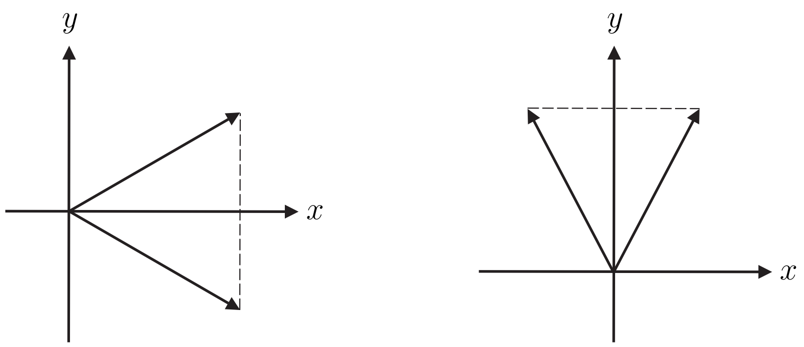

圖 2.3: 左: 對於 x-軸的反射變換。右: 對於 y-軸的反射變換。

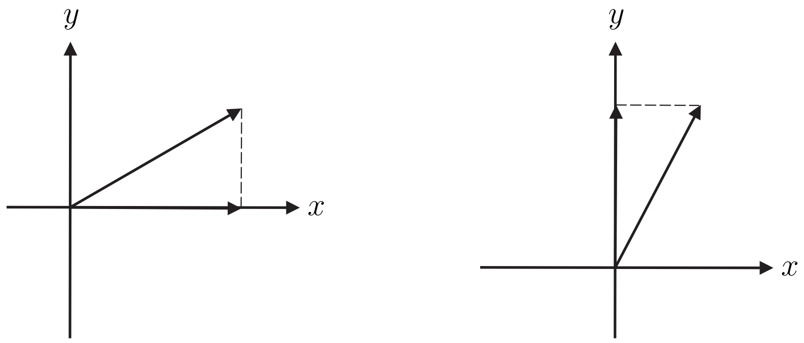

圖 2.4: 左: 對於 x-軸的投影變換。右: 對於 y-軸的投影變換。

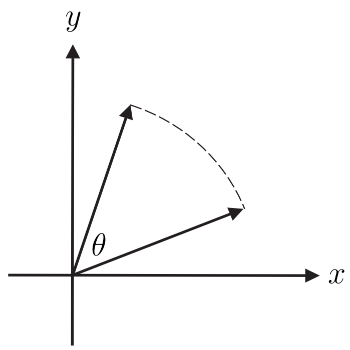

圖 2.5: 旋轉角為 θ 的旋轉變換 (以逆時針為正向)。

以下兩個線性變換在線性代數的理論中經常出現而且常被拿出來探討及使用。

定義 4 (單位變換; 零變換, 第 67 頁)**.**

(A) 一個線性變換 $I_V : V \to V$ 滿足對所有 $\mathbf{v} \in V$ 都有 $I_V(\mathbf{v}) = \mathbf{v}$, 我們說 I_V 是一個 單位變換 (identity transformation)。

(B) 一個線性變換 $T_{\mathbf{0}} : V \to W$ 滿足對所有 $\mathbf{v} \in V$ 都有 $T_{\mathbf{0}}(\mathbf{v}) = \mathbf{0}$, 我們說 T_0 是一個 零變換 (zero transformation)。

單位變換與零變換都是線性變換, 證明如下:

(A1) 對任意 $\mathbf{u}, \mathbf{v} \in V$, 都有 $I_V(\mathbf{u}+\mathbf{v}) = \mathbf{u}+\mathbf{v} = I_V(\mathbf{u}) + I_V(\mathbf{v})$。

(A2) 對任意 $\mathbf{u} \in V$ 與任何 $c \in \mathbb{F}$, 都有 $I_V(c \cdot \mathbf{u}) = c \cdot \mathbf{u} = c \cdot I_V(\mathbf{u})$。

(A3) 由 (A1) 與 (A2) 知: I_V 是一個線性變換。

(B1) 對任意 $\mathbf{u}, \mathbf{v} \in V$, 都有 $T_{\mathbf{0}}(\mathbf{u}+\mathbf{v}) = \mathbf{0}+\mathbf{0} = T_{\mathbf{0}}(\mathbf{u}) + T_{\mathbf{0}}(\mathbf{v})$。

(B2) 對任意 $\mathbf{u} \in V$ 與任何 $c \in \mathbb{F}$, 都有 $T_{\mathbf{0}}(c \cdot \mathbf{u}) = \mathbf{0} = c \cdot \mathbf{0} = c \cdot T_{\mathbf{0}}(\mathbf{u})$。

(B3) 由 (B1) 與 (B2) 知: $T_{\mathbf{0}}$ 是一個線性變換。

2.2 線性變換的性質

關於線性變換理論, 其中一個主題是想要了解向量空間 (定義域中的所有向量) 經過線性變換之後的結果。為此, 我們提出以下兩個概念。

定義 1 (第 67 頁)**.** 假設 $T : V \to W$ 是一個線性變換,

(A) 集合 $N(T) = \{\mathbf{v} \in V | T(\mathbf{v}) = \mathbf{0}\}$ 稱為線性變換 $T : V \to W$ 的 零核空間 (null space) 或稱為 核 (kernel)。

(B) 集合 $R(T) = \{T(\mathbf{v}) \in W | \mathbf{v} \in V\}$ 稱為線性變換 $T : V \to W$ 的 值域 (range) 或稱為 像 (image)。

例 2. 試求以下線性變換的零核空間與值域:

(A) 單位變換 $I_V : V \to V$。

(B) 零變換 $T_{\mathbf{0}} : V \to W$。

解.

(A1) 因為 $I_V(\mathbf{0}) = \mathbf{0}$, 而且對任意 $\mathbf{u} \neq \mathbf{0}$ 都有 $I_V(\mathbf{u}) = \mathbf{u} \neq \mathbf{0}$, 所以 $N(I_V) = \{\mathbf{0}\}$。

(A2) 對任意 $\mathbf{u} \in V$, 取 $\mathbf{u} \in V$ 使得 $I_V(\mathbf{u}) = \mathbf{u}$, 所以 $R(I_V) = V$。

(B1) 因為對任意 $\mathbf{u} \in V$ 都有 $T_{\mathbf{0}}(\mathbf{u}) = \mathbf{0}$, 所以 $N(T_{\mathbf{0}}) = V$。

(B2) 對於 $\mathbf{0} \in W$, 取 $\mathbf{0} \in V$ 使得 $T_{\mathbf{0}}(\mathbf{0}) = \mathbf{0}$, 對任意 $\mathbf{w} \neq \mathbf{0}$, 不存在 $\mathbf{u} \in V$ 使得 $T(\mathbf{u}) = \mathbf{w}$, 所以 $R(T_{\mathbf{0}}) = \{\mathbf{0}\}$。

關於一個線性變換 $T : V \to W$ 的零核空間與值域, 實際上這兩個集合有更好的性質; 也就是說, 以下定理將告知零核空間是向量空間 V 的子空間, 而值域是 W 的子空間。

定理 3 (第 68 頁)**.** 假設 $T : V \to W$ 是一個線性變換, 則 $N(T)$ 是向量空間 V 的子空間, 而 $R(T)$ 是向量空間 W 的子空間。

證明:

(A) 對任何 $\mathbf{u}, \mathbf{v} \in N(T) \subset V$ 以任何 $c \in \mathbb{F}$, 因為

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

而且

$$T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u}) = c \cdot \mathbf{0} = \mathbf{0},$$

因此 $\mathbf{u} + \mathbf{v} \in N(T)$ 而且 $c \cdot \mathbf{u} \in N(T)$, 所以 $N(T)$ 是 V 的子空間。

(B) 對任何 $\mathbf{w}, \mathbf{z} \in R(T) \subset W$, 存在 $\mathbf{u}, \mathbf{v} \in V$ 使得 $T(\mathbf{u}) = \mathbf{w}$ 且 $T(\mathbf{v}) = \mathbf{z}$。考慮 $\mathbf{u} + \mathbf{v} \in V$, 因為

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{w} + \mathbf{z},$$

所以 $\mathbf{w} + \mathbf{z} \in R(T)$。

對任何 $\mathbf{w} \in R(T)$ 以及任何 $c \in \mathbb{F}$, 存在 $\mathbf{u} \in V$ 使得 $T(\mathbf{u}) = \mathbf{w}$。考慮 $c \cdot \mathbf{u} \in V$, 因為

$$T(c \cdot \mathbf{u}) = c \cdot T(\mathbf{u}) = c \cdot \mathbf{w},$$

所以 $c \cdot \mathbf{w} \in R(T)$。由上討論得知: $R(T)$ 是 W 的子空間。

□

各位在看 定理 **3** 的時候應該要覺得很自然, 比方說先思考零核空間, 若有一個方向 $\mathbf{u}$ 透過 T 打到 $\mathbf{0}$, 因為線性的關係, 平行於 $\mathbf{u}$ 的整個方向全部都會打到 $\mathbf{0}$, 所以所有經過 T 打到 $\mathbf{0}$ 的那些向量的線性組合透過 T 也都會對應到 $\mathbf{0}$, 因此零核空間是一個定義域的子空間。至於值域, 若有一個向量 $\mathbf{u}$ 經過 T 對到 $\mathbf{w}$, 線性性質告知平行 $\mathbf{w}$ 的整個方向都會被對應到; 對於值域中的向量之線性組合, 因為值域告知每一項的向量其源頭也都分別有向量透過 T 對應, 所以考慮那些向量以相同的係數進行線性組合, 透過線性變換的意義而得到經由 T 之後的向量也在值域中, 故值域是對應域的子空間。

接下來想要探討的問題是: 對於線性變換 $T : V \to W$, 若我們在向量空間 V 中選取一組基底, 因為基底告知向量空間 V 中的任何一個向量 $\mathbf{v}$ 都可以用基底之線性組合唯一表示, 可見得 $T(\mathbf{v})$ 的結果會和基底的向量逐一透過 T 而得到的向量有關; 換言之, 下面的定理欲說明子空間 $R(T)$ 和 V 中的基底透過 T 而得的向量之關係。

定理 **4** (第 68 頁). 假設 $T : V \to W$ 是線性變換, 而 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 的一組基底, 則

$$R(T) = \operatorname{span}(\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\})。$$

證明: 對任何 $\mathbf{w} \in R(T)$, 存在 $\mathbf{v} \in V$ 使得 $T(\mathbf{v}) = \mathbf{w}$, 因為 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 的一組基底, 所以存在 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得

$$\mathbf{v} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i,$$

因為 $T : V \to W$ 是線性變換, 所以

$$\mathbf{w} = T(\mathbf{v}) = T\left(\sum_{i=1}^{n} c_i \cdot \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i \cdot T(\mathbf{v}_i) \in \operatorname{span}(\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}),$$

所以 $R(T) \subset \operatorname{span}(\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\})$。

對於 $i = 1, 2, \ldots, n$, $T(\mathbf{v}_i) \in R(T)$, 因為 定理 **3** 告知 $R(T)$ 是 W 的一個子空間, 所以 $\operatorname{span}(\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}) \subset R(T)$。

綜合上述討論, 我們得到 $R(T) = \operatorname{span}(\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\})$。 □

接下來感興趣的問題是想要知道一個線性變換的零核空間還有值域的大小, 因為兩者都是子空間, 所以它們都可以定義維度, 這裡先幫它們的維度取名字。

定義 5 (第 69 頁)**.** 假設 $T: V \to W$ 是線性變換, 若 $N(T)$ 與 $R(T)$ 皆為有限維度的子空間, 我們稱 $N(T)$ 的維度為線性變換 T 的 核次數 (nullity), 記為 $\text{nullity}(T)$; 稱 $R(T)$ 的維度為線性變換 T 的 秩 (rank), 記為 $\text{rank}(T)$。

以下要介紹的是線性變換的一個重要定理, 稱為維度定理。它是在形容一些向量空間的維度之間有一個恆等式。

定理 6 (維度定理, Dimension Theorem, 第 70 頁)**.** 假設 V 與 W 是分佈於 $\mathbb{F}$ 上的有限維向量空間, 而 $T: V \to W$ 是一個線性變換, 則

$$\text{nullity}(T) + \text{rank}(T) = \dim(V)。$$

證明: 假設 $\dim(V) = n$ 以及 $\text{nullity}(T) = \dim(N(T)) = k$。令 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ 是 $N(T)$ 的一組基底, 首先將 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ 擴張成 V 的一組基底 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ (詳見單元 1.9 的 推論 **9**)。以下討論將證明集合 $S = \{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\}$ 形成 $R(T)$ 的一組基底。若這件事情成立的話, 那麼就有 $\text{rank}(T) = \dim(R(T)) = n - k$, 於是 $\text{nullity}(T) + \text{rank}(T) = \dim(V)$ 成立。

為證明集合 $S = \{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\}$ 形成 $R(T)$ 的一組基底, 首先證明 S 生成 $R(T)$: 利用定理 **4** 以及對於 $i = 1, 2, \ldots, k$ 有 $T(\mathbf{v}_i) = 0, i = 1, 2, \ldots, k$, 得到

$$\begin{aligned} R(T) &= \text{span}\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_k), T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\} \\ &= \text{span}\{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\} = \text{span}(S)。\end{aligned}$$

再證集合 S 是線性獨立: 假設

$$\sum_{i=k+1}^{n} c_i \cdot T(\mathbf{v}_i) = \mathbf{0}, \quad \text{其中 } c_{k+1}, \ldots, c_n \in \mathbb{F},$$

因為 T 是線性變換, 所以

$$T\left(\sum_{i=k+1}^{n} c_i \cdot \mathbf{v}_i\right) = \mathbf{0},$$

故 $\sum\limits_{i=k+1}^{n} c_i \mathbf{v}_i \in N(T)$。因此存在 $c_1, c_2, \ldots, c_k \in \mathbb{F}$ 使得

$$\sum_{i=k+1}^{n} c_i \cdot \mathbf{v}_i = \sum_{i=1}^{k} c_i \cdot \mathbf{v}_i,$$

現將式子改寫成

$$\sum_{i=1}^{k} c_i \cdot \mathbf{v}_i + \sum_{i=k+1}^{n} (-c_i) \cdot \mathbf{v}_i = \mathbf{0},$$

因為 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ 是向量空間 V 的一組基底, 所以對於 $i = 1, 2, \ldots, n$ 都有 $c_i = 0$。因此 S 是線性獨立。 □

各在看完維度定理, 應該要有一個直觀的見解: 假設向量空間 V 的維度是 n, 對於這 n 個方向, 有某 k 個方向對於線性變換 T 而言是退化的, 也就是 $N(T)$ 的那個子空間全部都會打到零向量, 剩下的 $n-k$ 的方向就是好的, 它們不退化, 於是透過 T 的轉換後, 仍然是一個好的 $n-k$ 維的空間。

關於線性變換的議題, 以下想要從映射的理論探討其關係。這裡我們複習映射是一對一還有映成的概念。假設 $f: X \to Y$ 是一個映射,

(A) 若對任意 $x_1 \neq x_2$ 都有 $f(x_1) \neq f(x_2)$, 我們說映射 f 是 **一對一** (one-to-one)。

(B) 若對任意 $y \in Y$, 存在 $x \in X$ 使得 $f(x) = y$, 則稱映射 f 是 **映成** (onto)。

任給一個映射 f, 若要證明它是一對一或是映成, 都必須按照定義的方式處理。而我們現在想討論的是線性變換 $T: V \to W$, 因為這個映射具有非常好的結構 (線性結構) 在內, 所以關於線性變換 T 若要驗檢它是一對一或是映成, 可以從零核空間或是值域的方式重新認識。

定理 7 (第 71 頁)**.** *假設 $T: V \to W$ 是線性變換, 則 T 是一對一若且唯若 $N(T) = \{\mathbf{0}\}$。*

證明: ($\Rightarrow$) 假設 $T: V \to W$ 是一對一的線性變換, 令 $\mathbf{u} \in N(T)$, 則 $T(\mathbf{u}) = \mathbf{0}$, 因為 $T(\mathbf{0}) = \mathbf{0}$, 所以 $T(\mathbf{u}) = T(\mathbf{0})$, 由 T 是一對一得知 $\mathbf{u} = \mathbf{0}$, 因此 $N(T) = \{\mathbf{0}\}$。
($\Leftarrow$) 假設 $N(T) = \{\mathbf{0}\}$, 考慮 $\mathbf{u}, \mathbf{v} \in V$ 滿足 $T(\mathbf{u}) = T(\mathbf{v})$, 因為

$$T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0},$$

所以 $\mathbf{u}-\mathbf{v} \in N(T) = \{\mathbf{0}\}$, 得到 $\mathbf{u}-\mathbf{v} = \mathbf{0}$, 即 $\mathbf{u} = \mathbf{v}$, 於是 T 是一對一的線性變換。 □

若 V 和 W 都是有限維向量空間, 那麼線性變換 $T: V \to W$ 關於一對一還有映成的討論會有更好的刻畫。

定理 8 (第 71 頁)**.** *假設 V 與 W 是兩個有限維且維度相同的向量空間, 而 $T: V \to W$ 是線性變換, 則以下敘述彼此等價:*

(A) T 是一對一。

(B) T 是映成。

(C) $\operatorname{rank}(T) = \dim(\mathrm{V})$。

證明: (A) $\Leftrightarrow$ (C): 根據維度定理 (Dimension Theorem, 定理 **6**):

$$\operatorname{nullity}(T) + \operatorname{rank}(T) = \dim(V),$$

由 定理 **7** 知道: T 是一對一若且唯若 $N(T) = \{\mathbf{0}\}$, 若且唯若 $\operatorname{nullity}(T) = 0$, 代入維度定理的式子後得到 T 是一對一若且唯若 $\operatorname{rank}(T) = \dim(V)$。
(C) $\Leftrightarrow$ (B): 因為 V 和 W 的維度相同, 所以 $\operatorname{rank}(T) = \dim(V)$ 表示 $\dim(R(T)) = \dim(W)$, 因為 $R(T)$ 是 W 的子空間, 又有 $\dim(R(T)) = \dim(W)$, 所以 $R(T) = W$ (詳見單元 1.9 的 定理 **6**), 這個結果即說明 T 是映成。 □

這裡注意到我們是在有限維向量空間的線性變換進行討論, 若 V 不是有限維向量空間 (無窮維向量空間), 就算 $T: V \to V$ 是線性變換, 那麼 T 是一對一與 T 是映成不見得彼此等價。

以下舉例說明這些定理的應用。

例 **9** (第 72 頁)**.** 考慮線性變換 $T: \mathbb{F}^2 \to \mathbb{F}^2$, 定義如下:

$$T(a_1, a_2) = (a_1 + a_2, a_1 - a_2),$$

首先確定 $N(T)$: 透過解聯立方程式

$$\begin{cases} a_1 + a_2 = 0 \\ a_1 - a_2 = 0 \end{cases}$$

得到 $(a_1, a_2) = (0, 0)$, 所以 $N(T) = \{\mathbf{0}\}$, 故 T 是一對一。由 定理 **8** 得知 T 是映成。

例 **10** (第 72 頁)**.** 考慮線性變換 $T: P_2(\mathbb{R}) \to \mathbb{R}^3$ 定義如下:

$$T(a_0 + a_1 x + a_2 x^2) = (a_0, a_1, a_2),$$

首先確定 $N(T)$: 因為 $(a_1, a_2, a_3) = (0, 0, 0)$ 表示 $a_0 + a_1 x + a_2 x^2 = \mathbf{0}$, 故 $N(T) = \{\mathbf{0}\}$, 所以 T 是一對一。

考慮 $S = \{2 - x + 3x^2, x + x^2, 1 - 2x^2\}$, 則若要證明 S 在 $P_2(\mathbb{R})$ 上是線性獨立, 只要確定

$$T(S) = \{(2, -1, 3), (0, 1, 1), (1, 0, -2)\}$$

在 $\mathbb{R}^3$ 中是線性獨立即可。考慮方程式 $a_1(2, -1, 3) + a_2(0, 1, 1) + a_3(1, 0, -2) = 0$, 寫成聯立方程式的樣子即為

$$\begin{cases} 2a_1 + a_3 = 0 \\ -a_1 + 2a_2 = 0 \\ 3a_1 + a_2 - 2a_3 = 0, \end{cases}$$

將第一式乘以 2 與第三式相加, 得到 $7a_1 + a_2 = 0$, 而這個式子乘以 2 再與第二式相減後得到 $15a_1 = 0$, 所以 $a_1 = 0$, 最後分別代回第二式與第一式則得 $a_2 = 0$ 與 $a_3 = 0$。故 $T(S)$ 在 $\mathbb{R}^3$ 中線性獨立。

至於為什麼 $T(S)$ 是線性獨立可以推得 S 也是線性獨立呢? 這需要用到有限維向量空間的線性變換 T 是一對一與映成的等價性, 討論如下: 已知: $\{\mathbf{w}_i\}_{i=1}^n$ 是線性獨立, 即 $\sum\limits_{i=1}^n a_i \cdot \mathbf{w}_i = \mathbf{0}$ 可得 $a_i = 0, i = 1, 2, \ldots, n$, 因為 T 是一對一且映成, 所以對於每個 $\mathbf{w}_i, i = 1, 2, \ldots, n$, 分別存在唯一 $\mathbf{v}_i, i = 1, 2, \ldots, n$ 使得 $T(\mathbf{v}_i) = \mathbf{w}_i$。

考慮 $\sum\limits_{i=1}^n c_i \cdot \mathbf{v}_i = \mathbf{0}$, 因為

$$\sum_{i=1}^n c_i \cdot \mathbf{w}_i = \sum_{i=1}^n c_i \cdot T(\mathbf{v}_i) = T\left(\sum_{i=1}^n c_i \cdot \mathbf{v}_i\right) = T(\mathbf{0}) = \mathbf{0}$$

得知 $c_i = 0, i = 1, 2, \ldots, n$, 所以 $\{\mathbf{v}_i\}_{i=1}^n$ 也是線性獨立。

線性變換最重要的性質之一就是線性變換可以完全由對於基底的作用而決定。

定理 11 (第 72 頁)**.** 假設 V 與 W 是分佈於體 $\mathbb{F}$ 的向量空間, 並且假設 V 為有限維度, 取 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 的一組基底, 則對 W 中任意一組向量 $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$, 恰存在一個線性變換 $T : V \to W$ 使得 $T(\mathbf{v}_i) = \mathbf{w}_i, i = 1, 2, \ldots, n$。

證明: 對於 $\mathbf{v} \in V$, 存在唯一 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得 $\mathbf{v} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$。考慮映射 $T : V \to W$ 為

$$T(\mathbf{v}) = \sum_{i=1}^{n} c_i \cdot \mathbf{w}_i,$$

以下逐一驗證 T 的性質:

- <u>T 是線性映射</u>: 對任何 $\mathbf{u}, \mathbf{v} \in V$, 記 $\mathbf{u} = \sum_{i=1}^{n} a_i \cdot \mathbf{v}_i$ 以及 $\mathbf{v} = \sum_{i=1}^{n} b_i \cdot \mathbf{v}_i$, 其中 $a_1, a_2, \ldots, a_n \in \mathbb{F}$ 與 $b_1, b_2, \ldots, b_n \in \mathbb{F}$, 則

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T\left(\sum_{i=1}^{n} (a_i + b_i) \cdot \mathbf{v}_i\right) = \sum_{i=1}^{n} (a_i + b_i) \cdot \mathbf{w}_i \\ &= \sum_{i=1}^{n} a_i \cdot \mathbf{w}_i + \sum_{i=1}^{n} b_i \cdot \mathbf{w}_i = T(\mathbf{u}) + T(\mathbf{v})。\end{aligned}$$

  對任何 $\mathbf{u} \in V$ 與任何 $c \in \mathbb{F}$, 記 $\mathbf{u} = \sum_{i=1}^{n} a_i \cdot \mathbf{v}_i$, 其中 $a_1, a_2, \ldots, a_n \in \mathbb{F}$, 則

$$\begin{aligned} T(c \cdot \mathbf{u}) &= T\left(c \cdot \sum_{i=1}^{n} a_i \cdot \mathbf{v}_i\right) = T\left(\sum_{i=1}^{n} (c \cdot a_i) \cdot \mathbf{v}_i\right) = \sum_{i=1}^{n} (c \cdot a_i) \cdot \mathbf{w}_i \\ &= c \cdot \sum_{i=1}^{n} a_i \cdot \mathbf{w}_i = c \cdot T(\mathbf{u})。\end{aligned}$$

  由上討論得知 T 是線性變換。

- 根據定義, 對於 $i = 1, 2, \ldots, n$, $T(\mathbf{v}_i) = \mathbf{w}_i$ 成立。

- <u>映射 T 是唯一的</u>: 假設有另一個線性變換 $U : V \to W$ 滿足對於 $i = 1, 2, \ldots, n$ 都有 $U(\mathbf{v}_i) = \mathbf{w}_i$, 對所有 $\mathbf{v} \in V$, 記 $\mathbf{v} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$, 其中 $c_1, c_2, \ldots, c_n \in \mathbb{F}$, 則

$$U(\mathbf{v}) = \sum_{i=1}^{n} c_i \cdot U(\mathbf{v}_i) = \sum_{i=1}^{n} c_i \cdot \mathbf{w}_i = T(\mathbf{v}),$$

  因此 $U = T$。

$\square$

推論 12 (第 73 頁)**.** 假設 V 與 W 是向量空間, 其中 V 是有限維度的, 假設 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 的一組基底, 若 $T_1, T_2 : V \to W$ 為線性變換, 而且對所有 $i = 1, 2, \ldots, n$ 都有 $T_1(\mathbf{v}_i) = T_2(\mathbf{v}_i)$, 則 $T_1 = T_2$。

證明: 對任何 $\mathbf{u} \in V$, 存在唯一 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得 $\mathbf{u} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$, 因為

$$T_1(\mathbf{u}) = T_1\left(\sum_{i=1}^{n} c_i \cdot \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i \cdot T_1(\mathbf{v}_i) = \sum_{i=1}^{n} c_i \cdot T_2(\mathbf{v}_i) = T_2\left(\sum_{i=1}^{n} c_i \cdot \mathbf{v}_i\right) = T_2(\mathbf{u}),$$

所以 $T_1 = T_2$。 $\square$

上面的推論是在說: 對於定義域是有限維空間的兩個線性變換, 只要驗證基底對應到的向量完全一致, 就可以得知對每個向量透過這兩個線性變換而得的結果相同, 於是這兩個變換就一模一樣。其原因在於: 任何向量都是透過基底的線性組合完全決定, 而線性變換保持每個係數不變所致。

例 **13.** 考慮線性變換 $T : \mathbb{R}^2 \to \mathbb{R}^2$ 如下:

$$T(a_1, a_2) = (2a_2 - a_1, 3a_1),$$

假設 $U : \mathbb{R}^2 \to \mathbb{R}^2$ 是另一個線性變換, 如果 U 滿足 $U(1, 2) = (3, 3) = T(1, 2)$ 以及 $U(1, 1) = (1, 3) = T(1, 1)$, 因為 $\{(1, 2), (1, 1)\}$ 是 $\mathbb{R}^2$ 的基底, 所以 $U = T$。

2.3 線性變換的矩陣表現

這個單元的主要目的是想利用基底還有矩陣的觀點認識線性變換。我們從第 1 章的討論知道: 基底是撐起向量空間的骨架。如果基底視為一個集合看待的話, 集合內的元素並沒有順序可言; 也就是說, 假設 $\{\mathbf{u}_1, \mathbf{u}_2\}$ 是向量空間 $(V^2, \mathbb{F})$ 的一組基底, 則 $\{\mathbf{u}_1, \mathbf{u}_2\}$ 與 $\{\mathbf{u}_2, \mathbf{u}_1\}$ 視為相同的基底。然而, 當我們要討論向量空間之間的線性變換這個主題時, 特別是要用矩陣具體表達一個線性變換, 那麼就會需要將基底給出一個順序。於是我們要先給出有序基底的定義, 將向量的順序也納入考慮。

定義 **1** (第 79 頁)**.** 假設 $(V, \mathbb{F})$ 是有限維向量空間, 我們說 V 的一組 *有序基底* (ordered basis) 是指 V 上賦予一個明確順序的基底。

例 **2** (第 79 頁)**.** 在 $\mathbb{R}^3$ 中, 若 $\beta = \{\mathbf{e}_1 = (1,0,0), \mathbf{e}_2 = (0,1,0), \mathbf{e}_3 = (0,0,1)\}$ 視為一組有序的標準基底時, 它和 $\gamma = \{\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2\}$ 這組有序基底視為不同, 即 $\beta \neq \gamma$。

給定向量 $\mathbf{v}$ 及向量空間 V 上的一組有序基底 β, 則 $\mathbf{v}$ 可用 β 的線性組合唯一表示, 以下我們先學習如何將這件事用矩陣表達。

定義 **3** (第 80 頁)**.** 令 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是有限維向量空間 $(V, \mathbb{F})$ 的一組有序基底, 對於 $\mathbf{v} \in V$, 令 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 是唯一的一組純量滿足

$$\mathbf{v} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i,$$

我們定義 $\mathbf{v}$ *對於有序基底* β *的坐標向量* (coordinate vector of $\mathbf{v}$ relative to β) 為

$$[\mathbf{v}]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix};$$

也就是說, 先將向量 $\mathbf{v}$ 寫成對於有序基底 β 的線性組合之後, 把係數以行向量的方式排列。

例 **4** (第 80 頁)**.** 令 $V = P_2(\mathbb{R})$, 記 $\beta = \{1, x, x^2\}$ 是 V 的一組標準有序基底。考慮 $f(x) = 3 + 2x - 5x^2$, 則

$$[f]_\beta = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix},$$

若取另一組有序基底 $\gamma = \{x^2, x, 1\}$, 則

$$[f]_\gamma = \begin{bmatrix} -5 \\ 2 \\ 3 \end{bmatrix}。$$

現在要介紹線性變換的矩陣表現。假設 V 與 W 是有限維向量空間, 而 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 與 $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ 分別為 V 與 W 的有序基底。假設 $T : V \to W$ 是一個線性變換, 對於 $j = 1, 2, \ldots, n$, 因為 $T(\mathbf{v}_j) \in W$, 而 γ 是 W 的基底, 所以存在唯一一組純量 $A_{ij} \in \mathbb{F}$, 其中 $i = 1, 2, \ldots, m$ 以及 $j = 1, 2, \ldots, n$ 使得

$$T(\mathbf{v}_j) = \sum_{i=1}^{m} A_{ij} \cdot \mathbf{w}_i。$$

定義 5 (第 80 頁)**.** 若 A 是以分量 A_{ij} 定義的 $m \times n$ 矩陣, 則稱 A 是 線性變換 T 對於有序基底 β 及 γ 的矩陣表現 (matrix representation of T in the ordered bases β and γ), 記成 $A = [T]_\beta^\gamma$。若 $V = W$ 且 $\beta = \gamma$, 則線性變換 T 的矩陣表現將簡記為 $[T]_\beta$。

關於線性變換的矩陣表現有以下註記:

- 首先要再次強調的是: 線性變換的矩陣表現和有序基底 β 與 γ 有關, 對於同一個線性變換, 若選取的有序基底不同, 則對應到的矩陣表現會不同。

- 當我們觀察矩陣 A 的第 j 行, 它記錄的是

$$\begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix};$$

  另一方面, 我們知道 $T(\mathbf{v}_j) = \sum\limits_{i=1}^{m} A_{ij} \cdot \mathbf{w}_i$, 所以 A 的第 j 行是 $[T(\mathbf{v}_j)]_\gamma$。若我們將矩陣 A 寫出來, 則為

$$A = [T]_\beta^\gamma = \left[\begin{array}{c|c|c|c} T(\mathbf{v}_1) & T(\mathbf{v}_2) & \cdots & T(\mathbf{v}_n) \end{array}\right]。$$

- 若 $T_1 : V \to W$ 與 $T_2 : V \to W$ 滿足 $[T_1]_\beta^\gamma = [T_2]_\beta^\gamma$, 由單元 2.2 的 推論 **12** 得知: $T_1 = T_2$。

現給出幾個實際例子以再次確定 $[T]_\beta^\gamma$ 的概念。

例 6 (第 81 頁)**.** 考慮線性變換 $T : \mathbb{R}^2 \to \mathbb{R}^3$ 如下:

$$T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2),$$

記 $\beta = \{\bar{\mathbf{e}}_1 = (1, 0), \bar{\mathbf{e}}_2 = (0, 1)\}$ 與 $\gamma = \{\mathbf{e}_1 = (1, 0, 0), \mathbf{e}_2 = (0, 1, 0), \mathbf{e}_3 = (0, 0, 1)\}$ 分別為 $\mathbb{R}^2$ 與 $\mathbb{R}^3$ 的標準有序基底。

因為

$$\begin{aligned} T(1, 0) &= (1, 0, 2) = 1 \cdot \mathbf{e}_1 + 0 \cdot \mathbf{e}_2 + 2 \cdot \mathbf{e}_3 \\ T(0, 1) &= (3, 0, -4) = 3 \cdot \mathbf{e}_1 + 0 \cdot \mathbf{e}_2 - 4 \cdot \mathbf{e}_3, \end{aligned}$$

所以我們註記

$$[T]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}。$$

若我們考慮另一組有序基底 $\gamma' = \{\mathbf{e}_3 = (0,0,1), \mathbf{e}_2 = (0,1,0), \mathbf{e}_1 = (1,0,0)\}$, 這時

$$[T]_\beta^{\gamma'} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{bmatrix}。$$

以下內容想要研究的是線性變換的結構; 也就是說, 以線性變換為研究的對象討論所有線性變換所成的集合, 探討這個集合的結構。以下先給出線性變換的加法與純量乘法之意義。

定義 7 (第 82 頁)**.** 假設 $T: V \to W, T_1 : V \to W$ 與 $T_2 : V \to W$ 是線性變換, 而 $a \in \mathbb{F}$,

- 定義映射 $T_1 + T_2 : V \to W$ 為 $(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v})$, 其中 $\mathbf{v} \in V$。
- 定義映射 $a \cdot T : V \to W$ 為 $(a \cdot T)(\mathbf{v}) = a \cdot T(\mathbf{v})$, 其中 $\mathbf{v} \in V$。

定理 8 (第 82 頁)**.** *假設 V 與 W 是佈於體 $\mathbb{F}$ 的向量空間, 而 T, T_1, T_2 都是由 V 到 W 的線性變換, 則*

(A) $T_1 + T_2 : V \to W$ 是線性變換。

(B) 對所有 $a \in \mathbb{F}$, $a \cdot T : V \to W$ 是線性變換。

(C) 考慮 $\mathcal{L}(V, W) = \{T \,|\, T : V \to W$ 是線性變換$\}$, 則 $\mathcal{L}(V, W)$ 配合上述定義的映射加法 $+$ 與映射的純量乘法 $\cdot$ 下形成一個佈於 $\mathbb{F}$ 的向量空間。

證明:

(A1) 對所有 $\mathbf{u}, \mathbf{v} \in V$, 都有

$$\begin{aligned}(T_1 + T_2)(\mathbf{u} + \mathbf{v}) &= T_1(\mathbf{u} + \mathbf{v}) + T_2(\mathbf{u} + \mathbf{v}) = T_1(\mathbf{u}) + T_1(\mathbf{v}) + T_2(\mathbf{u}) + T_2(\mathbf{v}) \\ &= T_1(\mathbf{u}) + T_2(\mathbf{u}) + T_1(\mathbf{v}) + T_2(\mathbf{v}) = (T_1 + T_2)(\mathbf{u}) + (T_1 + T_2)(\mathbf{v})。\end{aligned}$$

(A2) 對所有 $\mathbf{u} \in V$ 以及對所有 $c \in \mathbb{F}$, 都有

$$\begin{aligned}(T_1 + T_2)(c \cdot \mathbf{u}) &= T_1(c \cdot \mathbf{u}) + T_2(c \cdot \mathbf{u}) = c \cdot T_1(\mathbf{u}) + c \cdot T_2(\mathbf{u}) \\ &= c \cdot (T_1(\mathbf{u}) + T_2(\mathbf{u})) = c \cdot (T_1 + T_2)(\mathbf{u})。\end{aligned}$$

(A3) 由 (A1) 與 (A2) 得知: $T_1 + T_2 : V \to W$ 是一個線性變換。

(B1) 對所有 $\mathbf{u}, \mathbf{v} \in V$, 都有

$$\begin{aligned}(a \cdot T)(\mathbf{u}+\mathbf{v}) &= a \cdot (T(\mathbf{u}+\mathbf{v})) = a \cdot (T(\mathbf{u}) + T(\mathbf{v})) \\ &= a \cdot T(\mathbf{u}) + a \cdot T(\mathbf{v}) = (a \cdot T)(\mathbf{u}) + (a \cdot T)(\mathbf{v})。\end{aligned}$$

(B2) 對所有 $\mathbf{u} \in V$ 以及對所有 $c \in \mathbb{F}$, 都有

$$(a \cdot T)(c \cdot \mathbf{u}) = a \cdot (T(c \cdot \mathbf{u})) = a \cdot c \cdot T(\mathbf{u}) = c \cdot a \cdot T(\mathbf{u}) = c \cdot (a \cdot T)(\mathbf{u})。$$

(B3) 由 (B1) 與 (B2) 得知: $a \cdot T : V \to W$ 是一個線性變換。

(C) 關於集合 $\mathcal{L}(V,W) = \{T \,|\, T : V \to W \text{ 是線性變換}\}$, 由 (A) 和 (B) 的討論知道: 集合 $\mathcal{L}(V,W)$ 關於映射的加法 $+$ 與映射的純量乘法 $\cdot$ 下具有封閉性, 即對任何 $T_1, T_2 \in \mathcal{L}(V,W)$ 都有 $T_1 + T_2 \in \mathcal{L}(V,W)$, 以及對任何 $T \in \mathcal{L}(V,W)$ 與任何 $a \in \mathbb{F}$, 則 $a \cdot T \in \mathcal{L}(V,W)$。

為驗證 $\mathcal{L}(V,W)$ 是一個向量空間, 我們需要逐一檢查向量空間的所有條件:

(V1) 對所有 $T_1, T_2 \in \mathcal{L}(V,W)$ 以及對所有 $\mathbf{u} \in V$, 則

$$(T_1 + T_2)(\mathbf{u}) = T_1(\mathbf{u}) + T_2(\mathbf{u}) = T_2(\mathbf{u}) + T_1(\mathbf{u}) = (T_2 + T_1)(\mathbf{u}),$$

所以 $T_1 + T_2 = T_2 + T_1$。

(V2) 對所有 $T_1, T_2, T_3 \in \mathcal{L}(V,W)$ 以及對所有 $\mathbf{u} \in V$, 則

$$\begin{aligned}((T_1 + T_2) + T_3)(\mathbf{u}) &= (T_1 + T_2)(\mathbf{u}) + T_3(\mathbf{u}) = T_1(\mathbf{u}) + T_2(\mathbf{u}) + T_3(\mathbf{u}) \\ &= T_1(\mathbf{u}) + (T_2 + T_3)(\mathbf{u}) = (T_1 + (T_2 + T_3))(\mathbf{u}),\end{aligned}$$

所以 $(T_1 + T_2) + T_3 = T_1 + (T_2 + T_3)$。

(V3) 考慮 $T_{\mathbf{0}} \in \mathcal{L}(V,W)$ 是零變換, 則對所有 $T \in \mathcal{L}(V,W)$ 以及對所有 $\mathbf{u} \in V$, 都有

$$(T + T_{\mathbf{0}})(\mathbf{u}) = T(\mathbf{u}) + T_{\mathbf{0}}(\mathbf{u}) = T(\mathbf{u}) + \mathbf{0} = T(\mathbf{u}),$$

所以存在 $T_{\mathbf{0}} \in \mathcal{L}(V,W)$ 使得對所有 $T \in \mathcal{L}(V,W)$ 都有 $T + T_{\mathbf{0}} = T$。

(V4) 對任何 $T \in \mathcal{L}(V,W)$, 考慮 $S \in \mathcal{L}(V,W)$, 其中對於 $\mathbf{u} \in V$,

$$S(\mathbf{u}) \overset{\text{定義}}{=\!=} -T(\mathbf{u}),$$

則

$$(T + S)(\mathbf{u}) = T(\mathbf{u}) + S(\mathbf{u}) = T(\mathbf{u}) + (-T(\mathbf{u})) = T(\mathbf{u}) - T(\mathbf{u}) = \mathbf{0} = T_{\mathbf{0}}(\mathbf{u}),$$

所以對任何 $T \in \mathcal{L}(V,W)$, 存在 $S \in \mathcal{L}(V,W)$ 使得 $T + S = T_{\mathbf{0}}$。

(V5) 對於 $1 \in \mathbb{F}$, 對所有 $T \in \mathcal{L}(V, W)$ 以及對所有 $\mathbf{u} \in V$, 都有

$$(1 \cdot T)(\mathbf{u}) = 1 \cdot T(\mathbf{u}) = T(\mathbf{u}),$$

所以 $1 \cdot T = T$。

(V6) 對任何 $a, b \in \mathbb{F}$, 對所有 $T \in \mathcal{L}(V, W)$ 以及對所有 $\mathbf{u} \in V$, 都有

$$\begin{aligned}((a \cdot b) \cdot T)(\mathbf{u}) &= (a \cdot b) \cdot T(\mathbf{u}) = a \cdot b \cdot T(\mathbf{u}) = a \cdot (b \cdot T(\mathbf{u})) = a \cdot (b \cdot T)(\mathbf{u}) \\ &= (a \cdot (b \cdot T))(\mathbf{u}),\end{aligned}$$

所以 $(a \cdot b) \cdot T = a \cdot (b \cdot T)$。

(V7) 對任何 $a \in \mathbb{F}$, 對所有 $T_1, T_2 \in \mathcal{L}(V, W)$ 以及對所有 $\mathbf{u} \in V$, 都有

$$\begin{aligned}(a \cdot (T_1 + T_2))(\mathbf{u}) &= a \cdot ((T_1 + T_2)(\mathbf{u})) = a \cdot (T_1(\mathbf{u}) + T_2(\mathbf{u})) \\ &= a \cdot T_1(\mathbf{u}) + a \cdot T_2(\mathbf{u}) = (a \cdot T_1)(\mathbf{u}) + (a \cdot T_2)(\mathbf{u}) \\ &= (a \cdot T_1 + a \cdot T_2)(\mathbf{u}),\end{aligned}$$

所以 $a \cdot (T_1 + T_2) = a \cdot T_1 + a \cdot T_2$。

(V8) 對任何 $a, b \in \mathbb{F}$, 對所有 $T \in \mathcal{L}(V, W)$ 以及對所有 $\mathbf{u} \in V$, 都有

$$\begin{aligned}((a + b) \cdot T)(\mathbf{u}) &= (a + b) \cdot T(\mathbf{u}) = a \cdot T(\mathbf{u}) + b \cdot T(\mathbf{u}) = (a \cdot T)(\mathbf{u}) + (b \cdot T)(\mathbf{u}) \\ &= (a \cdot T + b \cdot T)(\mathbf{u}),\end{aligned}$$

所以 $(a + b) \cdot T = a \cdot T + b \cdot T$。

綜合上述討論, 我們得知 $\mathcal{L}(V, W)$ 在映射加法 $+$ 與映射的純量乘法 $\cdot$ 下形成一個佈於 $\mathbb{F}$ 的向量空間。

□

註 (第 82 頁). 若 $V = W$, 我們會將 $\mathcal{L}(V, W)$ 簡寫成 $\mathcal{L}(V)$。

不曉得各位有沒有注意到驗證 $\mathcal{L}(V, W)$ 是向量空間的論述, 和第 1 章單元 1.3 的 例 **5** 的過程幾乎是類似的, 只是對應域為 $\mathbb{R}$ 或 $\mathbb{C}$ 的映射我們通常會把它稱為函數, 在探討以映射為主體的代數結構時, 這些概念是一致的。

在這個單元的一開始我們介紹如何透過有序基底將一個線性變換用矩陣的方式表現。現在我們又知道 $\mathcal{L}(V, W)$ 是一個向量空間, 以下就要開始討論關於 $T_1 + T_2$ 還有 $a \cdot T$ 的矩陣表現; 更進一步地, 我們還想了解 $T_1 + T_2$ 的矩陣表現與 T_1 與 T_2 的矩陣表現之關係, 以及 $a \cdot T$ 的矩陣表現與 T 的矩陣表現之關係。

以下定理將給出線性變換之間的加法運算與純量乘法運算和它們的矩陣表現下對於矩陣的加法與純量乘法的運算有一致性。

定理 9 (第 82 頁)**.** 假設 V 與 W 皆為有限維向量空間, 而 β 與 γ 分別為 V 與 W 的一組有序基底, 令 $T, T_1, T_2 : V \to W$ 為線性變換, 任取 $a \in \mathbb{F}$, 則

(A) $[T_1 + T_2]_\beta^\gamma = [T_1]_\beta^\gamma + [T_2]_\beta^\gamma$。

(B) $[a \cdot T]_\beta^\gamma = a \cdot [T]_\beta^\gamma$。

證明: 記有序基底為 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 及 $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, 則分別存在唯一一組純量 $A_{ij} \in \mathbb{F}, B_{ij} \in \mathbb{F}$ 與 $C_{ij} \in \mathbb{F}$, 其中 $i = 1, 2, \ldots, m$ 以及 $j = 1, 2, \ldots, n$ 使得

$$T_1(\mathbf{v}_j) = \sum_{i=1}^m A_{ij} \cdot \mathbf{w}_i, \quad T_2(\mathbf{v}_j) = \sum_{i=1}^m B_{ij} \cdot \mathbf{w}_i \quad \text{及} \quad T(\mathbf{v}_j) = \sum_{i=1}^m C_{ij} \cdot \mathbf{w}_i\text{。}$$

(A) 因為

$$(T_1 + T_2)(\mathbf{v}_j) = T_1(\mathbf{v}_j) + T_2(\mathbf{v}_j) = \sum_{i=1}^m A_{ij} \cdot \mathbf{w}_i + \sum_{i=1}^m B_{ij} \cdot \mathbf{w}_i = \sum_{i=1}^m (A_{ij} + B_{ij}) \cdot \mathbf{w}_i,$$

得知

$$([T_1 + T_2]_\beta^\gamma)_{ij} = A_{ij} + B_{ij} = ([T_1]_\beta^\gamma)_{ij} + ([T_2]_\beta^\gamma)_{ij},$$

這就證明了 $[T_1 + T_2]_\beta^\gamma = [T_1]_\beta^\gamma + [T_2]_\beta^\gamma$。

(B) 因為

$$(a \cdot T)(\mathbf{v}_j) = a \cdot (T(\mathbf{v}_j)) = a \cdot \sum_{i=1}^m C_{ij} \cdot \mathbf{w}_i = \sum_{i=1}^m (a \cdot C_{ij}) \cdot \mathbf{w}_i$$

得知

$$([a \cdot T]_\beta^\gamma)_{ij} = a \cdot C_{ij} = a \cdot ([T]_\beta^\gamma)_{ij}\text{。}$$

這就證明了 $[a \cdot T]_\beta^\gamma = a \cdot [T]_\beta^\gamma$。

□

例 10. 考慮

$$T_1 : \mathbb{R}^2 \to \mathbb{R}^2, \quad T_1(a_1, a_2) = (a_1 + 2 \cdot a_2, 3 \cdot a_1 - a_2, 2 \cdot a_1 + 4 \cdot a_2)$$
$$T_2 : \mathbb{R}^2 \to \mathbb{R}^2, \quad T_2(a_1, a_2) = (4 \cdot a_1 - a_2, a_1 + 2 \cdot a_2, 3 \cdot a_1),$$

取 β 及 γ 是 $\mathbb{R}^2$ 與 $\mathbb{R}^3$ 的標準有序基底, 則

$$[T_1]_\beta^\gamma = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 4 \end{bmatrix} \quad \text{與} \quad [T_2]_\beta^\gamma = \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix},$$

因為 $(T_1+T_2)(a_1,a_2)=(5\cdot a_1+a_2, 4\cdot a_1+a_2, 5\cdot a_1+4\cdot a_2)$, 而

$$[T_1+T_2]_\beta^\gamma=\begin{bmatrix}5 & 1\\ 4 & 1\\ 5 & 4\end{bmatrix},$$

於是兩個線性變換的相加 $[T_1+T_2]_\beta^\gamma=[T_1]_\beta^\gamma+[T_1]_\beta^\gamma$ 與矩陣的加法

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

有一致性。

2.4 線性變換的合成與矩陣乘法

前一個單元是介紹線性變換在給定有序基底的情況下如何以矩陣的方式呈現, 並驗證兩個線性變換的加法與分別對應到的矩陣加法有一致性, 還有一個線性變換乘上純量乘法與相應矩陣進行純量乘法具有一致性。

現在要考慮更複雜的問題, 我們知道一般的映射會去研究它們的合成映射, 而線性變換是一種很特別的映射, 所以線性變換也可以探討其合成映射; 也就是說, 給定 $T : V \to W$ 與 $S : W \to Z$ 是線性變換, 記 $ST : V \to Z$ 是 T 與 S 這兩個線性變換的合成映射, 其中 ST 的映射關係是: 給定 $\mathbf{v} \in V$, 先透過 T 對應到 W 上的一個向量 $T(\mathbf{v})$, 再將向量 $T(\mathbf{v})$ 透過映射 S 對應到 Z 上的一個向量 $S(T(\mathbf{v}))$, 於是, 我們定義 $(ST)(\mathbf{v}) = S(T(\mathbf{v}))$。

這一單元感興趣的議題是關於合成映射的各種性質以及它和各別映射的關係。首先, 我們要驗證的是: $ST : V \to Z$ 是一個線性變換。

定理 1 (第 86 頁)**.** 令 V, W, Z 是向量空間, 而 $T : V \to W$ 與 $S : W \to Z$ 是線性變換, 則合成映射 $ST : V \to Z$ 也是線性變換。

證明:

- 對所有 $\mathbf{u}, \mathbf{v} \in V$, 都有

$$\begin{aligned}(ST)(\mathbf{u}+\mathbf{v}) &= S(T(\mathbf{u}+\mathbf{v})) = S(T(\mathbf{u}) + T(\mathbf{v})) = S(T(\mathbf{u})) + S(T(\mathbf{v})) \\ &= (ST)(\mathbf{u}) + (ST)(\mathbf{v})。\end{aligned}$$

- 對所有 $\mathbf{u} \in V$ 以及對所有 $c \in \mathbb{F}$, 都有

$$(ST)(c \cdot \mathbf{u}) = S(T(c \cdot \mathbf{u})) = S(c \cdot T(\mathbf{u})) = c \cdot S(T(\mathbf{u})) = c \cdot (ST)(\mathbf{u})。$$

- 由上討論得知映射 $ST : V \to Z$ 是一個線性變換。

□

以下繼續探討幾個線性變換的合成映射之性質。

定理 2 (第 87 頁)**.** 令 V, W, Z, X 是向量空間, 而 T, T_1, T_2 是 V 到 W 的線性變換, S, S_1, S_2 是 W 到 Z 的線性變換, U 是 Z 到 X 的線性變換, 則

(A) $S(T_1 + T_2) = ST_1 + ST_2$ 以及 $(S_1 + S_2)T = S_1T + S_2T$。

(B) 對所有 $a \in \mathbb{F}$ 都有 $a \cdot (ST) = (a \cdot S)T = S(a \cdot T)$。

(C) $TI_V = T = I_WT$。

(D) $U(ST) = (US)T$。

證明:

(A1) 對任意 $\mathbf{u} \in V$, 都有

$$\begin{aligned}(S(T_1+T_2))(\mathbf{u}) &= S((T_1+T_2)(\mathbf{u})) = S(T_1(\mathbf{u})+T_2(\mathbf{u})) = S(T_1(\mathbf{u}))+S(T_2(\mathbf{u})) \\ &= (ST_1)(\mathbf{u})+(ST_2)(\mathbf{u}) = (ST_1+ST_2)(\mathbf{u}),\end{aligned}$$

所以 $S(T_1+T_2) = ST_1+ST_2$。

(A2) 對任意 $\mathbf{u} \in V$, 都有

$$\begin{aligned}((S_1+S_2)T)(\mathbf{u}) &= (S_1+S_2)(T(\mathbf{u})) = S_1(T(\mathbf{u}))+S_2(T(\mathbf{u})) = (S_1T)(\mathbf{u})+(S_2T)(\mathbf{u}) \\ &= (S_1T+S_2T)(\mathbf{u}),\end{aligned}$$

所以 $(S_1+S_2)T = S_1T+S_2T$。

(B) 對任意 $\mathbf{u} \in V$, 都有

$$(a\cdot(ST))(\mathbf{v}) = a\cdot(ST)(\mathbf{v}) = a\cdot(S(T(\mathbf{v}))) = a\cdot S(T(\mathbf{v})) = (a\cdot S)(T(\mathbf{v})),$$

這表示 $a\cdot(ST) = (a\cdot S)T$。此外, 因為 S 是線性變換, 所以

$$((a\cdot S)T)(\mathbf{v}) = a\cdot S(T(\mathbf{v})) = S(a\cdot T(\mathbf{v})) = S(a\cdot T)(\mathbf{v}),$$

得到 $(a\cdot S)T = S(a\cdot T)$。

(C) 對任意 $\mathbf{u} \in V$, 都有

$$(TI_V)(\mathbf{v}) = T(I_V(\mathbf{v})) = T(\mathbf{v}) = I_W(T(\mathbf{v})) = (I_WT)(\mathbf{v}),$$

所以 $TI_V = T = I_WT$。

(D) 對任意 $\mathbf{u} \in V$, 都有

$$(U(ST))(\mathbf{u}) = U((ST)(\mathbf{u})) = U(S(T(\mathbf{u}))) = (US)(T(\mathbf{u})) = ((US)T)(\mathbf{u}),$$

所以 $U(ST) = (US)T$。

□

在探討完線性變換的合成映射之基本性質後, 接下來想了解的問題是: 既然線性變換的合成映射 $ST : V \to Z$ 是一個線性變換, 那麼這個合成線性變換該如何用矩陣表達, 進一步想知道 ST 的矩陣表達與 S 和 T 各別的矩陣表達之關係。

關於這個問題的分析如下:

考慮 $T:V\to W$ 與 $S:W\to Z$ 為線性變換, 取 $\alpha=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\},\beta=\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_m\}$ 以及 $\gamma=\{\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_l\}$ 分別為 V,W,Z 的一組有序基底。記 $B=[T]_\alpha^\beta$ 是以有序基底 α 及 β 表示的 T 之矩陣表現, 而 $A=[S]_\beta^\gamma$ 是以有序基底 β 及 γ 表示的 S 之矩陣表現。

現在想要求得以有序基底 α 及 γ 來表示的 ST 之矩陣表現, 並將它記為 $AB=[ST]_\alpha^\gamma$。對於 $j=1,2,\ldots,n$, 現計算

$$\begin{aligned}(ST)(\mathbf{v}_j)=S(T(\mathbf{v}_j))&=S\left(\sum_{k=1}^m B_{kj}\cdot\mathbf{w}_k\right)=\sum_{k=1}^m B_{kj}\cdot S(\mathbf{w}_k)\\&=\sum_{k=1}^m B_{kj}\left(\sum_{i=1}^l A_{ik}\cdot\mathbf{z}_i\right)=\sum_{i=1}^l\left(\sum_{k=1}^m A_{ik}B_{kj}\right)\cdot\mathbf{z}_i\overset{\text{記}}{=}\sum_{i=1}^l C_{ij}\cdot\mathbf{z}_i,\end{aligned}$$

其中 $C_{ij}=\sum\limits_{k=1}^m A_{ik}B_{kj}$。

為了要將合成線性變換的矩陣表現與各別線性變換的矩陣表現有良好的對應關係, 這裡我們定義兩個矩陣相乘的概念。

定義 3 (第 87 頁)**.** 令 A 是 $l\times m$ 矩陣, 而 B 為 $m\times n$ 矩陣, 定義矩陣 A 與 B 的 乘積 (product), 記為 AB, 它是一個 $l\times n$ 的矩陣, 其中

$$(AB)_{ij}=\sum_{k=1}^m A_{ik}B_{kj},\quad \text{其中 } i=1,2,\ldots,l,\ \text{而 } j=1,2,\ldots,n\text{。}$$

這裡注意到: 因為線性變換 $T:V\to W$ 以及 $S:W\to Z$ 必須要用一樣的向量空間 W 才能探討合成線性變換 $ST:V\to Z$, 這件事對應到矩陣的乘法也會給出一些限制; 也就是說, 若矩陣 AB 要有意義, 我們必須約束矩陣 A 和 B 的大小, 在矩陣大小的關係式「$(l\times m)\cdot(m\times n)=(l\times n)$」, 串接的正整數 m 必須相等, 而且是 W 的維度。

以下定理即上述矩陣乘法定義的結果。

定理 4 (第 88 頁)**.** 令 V,W,Z 是有限維向量空間, 而 α,β,γ 分別為 V,W,Z 的有序基底, 假設 $T:V\to W$ 與 $S:W\to Z$ 是線性變換, 則

$$[ST]_\alpha^\gamma=[S]_\beta^\gamma[T]_\alpha^\beta\text{。}$$

證明: 因為

$$([ST]_\alpha^\gamma)_{ij}=C_{ij}=\sum_{k=1}^m A_{ik}B_{kj},$$

而

$$([S]_\beta^\gamma[T]_\alpha^\beta)_{ij}=(AB)_{ij}=\sum_{k=1}^m A_{ik}B_{kj},$$

所以 $[ST]_\alpha^\gamma=[S]_\beta^\gamma[T]_\alpha^\beta$。 □

例 **5.** 以下示範矩陣的乘法運算:

$$\begin{bmatrix} -1 & 2 & 1 \\ 3 & -1 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}=\begin{bmatrix} (-1)\cdot 4+2\cdot 2+1\cdot 3 \\ 3\cdot 4+(-1)\cdot 2+2\cdot 3 \end{bmatrix}=\begin{bmatrix} 3 \\ 16 \end{bmatrix}。$$

註. 這裡注意到, 由於合成函數不具有交換性, 所以矩陣的乘法來說也不具有交換性。舉簡單例子也可以知道此結果:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{以及} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix},$$

即使兩個矩陣 A, B 之乘積 AB 與 BA 皆有定義, 但 $AB = BA$ 這件事未必成立。

例 **6** (第 88 頁)**.** 驗證轉置矩陣的公式: $(AB)^t = B^tA^t$。

證明: 計算

$$(AB)^t_{ij} = (AB)_{ji} = \sum_{k=1}^{m} A_{jk}B_{ki},$$

另一方面,

$$(B^tA^t)_{ij} = \sum_{k=1}^{m}(B^t)_{ik}(A^t)_{kj} = \sum_{k=1}^{m} B_{ki}A_{jk} = \sum_{k=1}^{m} A_{jk}B_{ki},$$

得知 $(AB)^t = B^tA^t$。 □

例 **7** (微積分基本定理在線性代數上的呈現, 第 89 頁)**.** 考慮 $S : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ 為 $S(f)(x) = f'(x)$ 以及 $T : P_2(\mathbb{R}) \to P_3(\mathbb{R})$ 為 $T(f)(x) = \int_0^x f(t)\,\mathrm{d}t$。記 $\alpha = \{1, x, x^2, x^3\}$ 是 $P_3(\mathbb{R})$ 的標準有序基底, 記 $\beta = \{1, x, x^2\}$ 是 $P_2(\mathbb{R})$ 的標準有序基底。

$ST = I_2$ 是 $P_2(\mathbb{R})$ 上的單位變換, 欲說明微積分基本定理, 注意

$$[ST]_\beta = [S]^\beta_\alpha [T]^\alpha_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=[I]_\beta。$$

定理 **8** (第 91 頁)**.** 令 V 與 W 為有限維向量空間, 而 β 與 γ 分別為 V 與 W 的一組有序基底。假設 $T : V \to W$ 是一個線性變換, 則對於 $\mathbf{v} \in V$, 都有

$$[T(\mathbf{v})]_\gamma = [T]^\gamma_\beta [\mathbf{v}]_\beta。$$

證明: 選定 $\mathbf{v} \in V$ 以及 $c \in \mathbb{F}$, 考慮以下兩個線性變換:

$$f : \mathbb{F} \to V, \quad f(c) = c \cdot \mathbf{v}$$
$$g : \mathbb{F} \to W, \quad g(c) = c \cdot T(\mathbf{v}),$$

假設 $\alpha = \{1\}$ 是 $\mathbb{F}$ 的標準有序基底, 注意到 $g = Tf$, 將行向量視為矩陣, 並且利用 定理 **4**, 得到

$$[T(\mathbf{v})]_\gamma = [g(1)]_\gamma = [g]^\gamma_\alpha = [Tf]^\gamma_\alpha = [T]^\gamma_\beta [f]^\beta_\alpha = [T]^\gamma_\beta [f(1)]_\beta = [T]^\gamma_\beta [\mathbf{v}]_\beta。$$

□

2.5 矩陣運算的規則與矩陣的左乘變換

前面的幾個單元都是在關心線性變換及其矩陣表現的關係, 若將矩陣視為 $\mathbb{F}^{m\times n}$ 這個向量空間進行位置重新排列成長方形的樣子, 完全忘掉它和線性變換的關聯, 只關心矩陣所形成的空間配合著運算之結構時, 就值得仔細深究。這一個單元就是以此為出發點綜整出矩陣的運算規則。

首先, 為了日後計算簡便起見, 進引克羅內克符號。

定義 1. 定義 *克羅內克符號* (Kronecker delta) 為

$$\delta_{ij} = \begin{cases} 1 & \text{若 } i = j \\ 0 & \text{若 } i \neq j, \end{cases}$$

而 $n \times n$ 的 *單位矩陣* (identity matrix) I_n 是指 $(I_n)_{ij} = \delta_{ij}$ 的矩陣。

和單元 2.4 的 **定理 2** 對照, 我們可以的證明以下矩陣運算的性質。

定理 2 (第 89 頁)**.** *令 A 為 $m \times n$ 矩陣, B, C 為 $n \times p$ 矩陣, 而 D, E 為 $q \times m$ 矩陣, 則*

(A) $A(B + C) = AB + AC$ 以及 $(D + E)A = DA + EA$。

(B) 對所有 $c \in \mathbb{F}$ 都有 $c(AB) = (cA)B = A(cB)$。

(C) $I_m A = A = AI_n$。

(D) 若 V 是一個 n 維的向量空間, 而 β 是 V 上的一組有序基底, 則 $[I_V]_\beta = I_n$。

證明:

(A1) 計算

$$\begin{aligned}[A(B + C)]_{ij} &= \sum_{k=1}^{n} A_{ik}(B + C)_{kj} = \sum_{k=1}^{n} A_{ik}(B_{kj} + C_{kj}) = \sum_{k=1}^{n} A_{ik}B_{kj} + A_{ik}C_{kj} \\ &= \sum_{k=1}^{n} A_{ik}B_{kj} + \sum_{k=1}^{n} A_{ik}C_{kj} = (AB)_{ij} + (AC)_{ij} = [AB + AC]_{ij},\end{aligned}$$

所以 $A(B + C) = AB + AC$。

(A2) 計算

$$\begin{aligned}[(D + E)A]_{ij} &= \sum_{k=1}^{m} (D + E)_{ik}A_{kj} = \sum_{k=1}^{m} (D_{ik} + E_{ik})A_{kj} = \sum_{k=1}^{m} D_{ik}A_{kj} + E_{ik}A_{kj} \\ &= \sum_{k=1}^{m} D_{ik}A_{kj} + \sum_{k=1}^{m} E_{ik}A_{kj} = (DA)_{ij} + (EA)_{ij} = [DA + EA]_{ij},\end{aligned}$$

所以 $(D + E)A = DA + EA$。

(B1) 計算

$$[c(AB)]_{ij} = c(AB)_{ij} = c \cdot \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{n} cA_{ik}B_{kj} = \sum_{k=1}^{n}(cA_{ik})B_{kj} = [(cA)B]_{ij},$$

所以 $c(AB) = (cA)B$。

(B2) 計算

$$[c(AB)]_{ij} = c(AB)_{ij} = c \cdot \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{n} A_{ik}cB_{kj} = \sum_{k=1}^{n} A_{ik}(cB_{kj}) = [A(cB)]_{ij},$$

所以 $c(AB) = A(cB)$。

(C1) 計算

$$[I_mA]_{ij} = \sum_{k=1}^{m}(I_m)_{ik}A_{kj} = \sum_{k=1}^{m} \delta_{ik}A_{kj} = A_{ij},$$

所以 $I_mA = A$。

(C2) 計算

$$[AI_n]_{ij} = \sum_{k=1}^{n} A_{ik}(I_n)_{kj} = \sum_{k=1}^{n} A_{ik}\delta_{kj} = A_{ij},$$

所以 $AI_n = A$。

(D) 記 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 上的一組有序基底, 因為對所有 $j = 1, 2, \ldots, n$, 都有

$$I_V(\mathbf{v}_j) = \mathbf{v}_j = \sum_{i=1}^{n} \delta_{ij}\mathbf{v}_i,$$

所以

$$([I_V]_\beta)_{ij} = \delta_{ij},$$

於是 $[I_V]_\beta = I_n$。

□

回想在單元 2.3 時, 我們知道對於一個線性變換 $T : V \to W$, 若 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 與 $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ 分別為有限維向量空間 V 與 W 的有序基底時, 則有

$$[T]_\beta^\gamma = \left[\begin{array}{c|c|c|c} T(\mathbf{v}_1) & T(\mathbf{v}_2) & \cdots & T(\mathbf{v}_n) \end{array}\right],$$

由此我們可以看到, 關於一個矩陣, 我們有時候會需要知道特定某一行的資訊, 以下定理即給出如何突顯出矩陣某一行的方法。

定理 3 (第 91 頁)**.** 令 A 是 $m\times n$ 矩陣, 而 B 為 $n\times l$ 矩陣, 對於 $j=1,2,\ldots,l$, 令 $\mathbf{u}_j$ 是矩陣 AB 的第 j 行, $\mathbf{v}_j$ 是矩陣 B 的第 j 行, 而 $\mathbf{e}_j$ 是單位矩陣 I_l 的第 j 行; 也就是說,

$$\begin{aligned} AB &= \left[\begin{array}{c|c|c|c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_l \end{array}\right] \\ B &= \left[\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_l \end{array}\right] \\ I_l &= \left[\begin{array}{c|c|c|c} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_l \end{array}\right], \end{aligned}$$

則有以下結果:

(A) $\mathbf{u}_j = A\mathbf{v}_j$。

(B) $\mathbf{v}_j = B\mathbf{e}_j$。

證明:

(A) 直接計算得到

$$\mathbf{u}_j = \begin{bmatrix} (AB)_{1j} \\ (AB)_{2j} \\ \vdots \\ (AB)_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} A_{1k}B_{kj} \\ \sum_{k=1}^{n} A_{2k}B_{kj} \\ \vdots \\ \sum_{k=1}^{n} A_{mk}B_{kj} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} \begin{bmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{bmatrix} = A\mathbf{v}_j。$$

(B) 直接計算得到

$$\mathbf{v}_j = \begin{bmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{l} B_{1k}\delta_{kj} \\ \sum_{k=1}^{l} B_{2k}\delta_{kj} \\ \vdots \\ \sum_{k=1}^{l} B_{nk}\delta_{kj} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1l} \\ B_{21} & B_{22} & \cdots & B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nl} \end{bmatrix} \begin{bmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{lj} \end{bmatrix} = B\mathbf{e}_j。$$

□

關於矩陣理論, 我們有的時候會把矩陣 $A\in M_{m\times n}(\mathbb{F})$ 想成是一種對於 $\mathbb{F}^n$ 的一個作用, 在這樣的想法下將引出以下的左乘變換。會有這樣的思維為的是在下一章在介紹矩陣的列運算 (矩陣的操作可以對應到解聯立方程式的步驟) 時會有一個很好的觀點把當中的代數結構講清楚。

定義 4 (第 92 頁)**.** 假設 A 是 $m\times n$ 矩陣, 其元素 $A_{ij}\in\mathbb{F}$, 其中 $i=1,2,\ldots,m$ 而 $j=1,2,\ldots,n$。定義 $L_A:\mathbb{F}^n\to\mathbb{F}^m$ 為 $L_A(\mathbf{u})=A\mathbf{u}$; 也就是說, 把 $\mathbf{u}\in\mathbb{F}^n$ 寫成行向量的形式, 然後 $A\mathbf{u}$ 表示 A 與 $\mathbf{u}$ 的乘積, 我們稱 L_A 是一個 *左乘變換* (left-multiplication transformation)。

例 **5.** 假設

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

則 $A \in M_{2\times 4}(\mathbb{R})$, 且 $L_A : \mathbb{R}^4 \to \mathbb{R}^2$。若

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix},$$

則

$$L_A(\mathbf{u}) = \begin{bmatrix} 3 & 2 & 4 & 0 \\ 2 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}。$$

以下先給出左乘變換 L_A 的幾個性質。這裡的諸多性質都只是將之前的結果重新理解成左乘概念下而得的結果。

定理 6 (第 93 頁)**.** 假設 $A \in M_{m\times n}(\mathbb{F})$, 則左乘變換 $L_A : \mathbb{F}^n \to \mathbb{F}^m$ 是一個線性變換。此外, 若 $B \in M_{m\times n}(\mathbb{F})$, 而 β 與 γ 分別為 $\mathbb{F}^n$ 與 $\mathbb{F}^m$ 的標準有序基底, 則有以下結果:

(A) $[L_A]_\beta^\gamma = A$。

(B) $L_A = L_B$ 若且唯若 $A = B$。

(C) $L_{A+B} = L_A + L_B$; 對所有 $c \in \mathbb{F}$ 都有 $L_{cA} = cL_A$。

(D) 若 $T : \mathbb{F}^n \to \mathbb{F}^m$ 是一個線性變換, 則存在唯一矩陣 $C \in M_{m\times n}(\mathbb{F})$ 使得 $T = L_C$。實際上, $C = [T]_\beta^\gamma$。

(E) 若 $E \in M_{n\times l}(\mathbb{F})$, 則 $L_{AE} = L_A L_E$。

(F) 若 $m = n$, 則 $L_{I_n} = I_{\mathbb{F}^n}$。

證明: 首先驗證左乘變換 L_A 是線性變換:

- 對所有 $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$, 則 $L_A(\mathbf{u}+\mathbf{v}) = A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = L_A(\mathbf{u}) + L_A(\mathbf{v})$。
- 對所有 $\mathbf{u} \in \mathbb{F}^n$ 與 $c \in \mathbb{F}$, 則 $L_A(c\mathbf{u}) = A(c\mathbf{u}) = cA\mathbf{u} = cL_A(\mathbf{u})$。
- 由上討論可知左乘變換 $L_A : \mathbb{F}^n \to \mathbb{F}^m$ 是一個線性變換。

以下繼續討論左乘變換的性質:

(A) $[L_A]_\beta^\gamma$ 的第 j 行是 $L_A(\mathbf{e}_j) = A\mathbf{e}_j$, 由 定理 **3** 知道它也是 A 的第 j 行, 所以 $[L_A]_\beta^\gamma = A$。

(B) $(\Rightarrow)$ 若 $L_A = L_B$, 由 (A) 知 $A = [L_A]_\beta^\gamma = [L_B]_\beta^\gamma = B$。

$(\Leftarrow)$ 對所有 $\mathbf{u} \in \mathbb{F}^n$, 都有 $L_A(\mathbf{u}) = A\mathbf{u} = B\mathbf{u} = L_B(\mathbf{u})$, 所以 $L_A = L_B$。

(C) 對所有 $\mathbf{u} \in \mathbb{F}^n$, 都有

$$\begin{aligned}
(L_{A+B}(\mathbf{u}))_j = ((A+B)\mathbf{u})_j &= \sum_{k=1}^{n}(A+B)_{jk}\mathbf{u}_k = \sum_{k=1}^{n}(A_{jk}+B_{jk})\mathbf{u}_k \\
&= \sum_{k=1}^{n}(A_{jk}\mathbf{u}_k + B_{jk}\mathbf{u}_k) = \sum_{k=1}^{n}A_{jk}\mathbf{u}_k + \sum_{k=1}^{n}B_{jk}\mathbf{u}_k \\
&= (A\mathbf{u})_j + (B\mathbf{u})_j = (L_A(\mathbf{u}))_j + (L_B(\mathbf{u}))_j,
\end{aligned}$$

因此 $L_{A+B} = L_A + L_B$。

對所有 $c \in \mathbb{F}$ 以及對所有 $\mathbf{u}$, 則

$$\begin{aligned}
(L_{cA}(\mathbf{u}))_j = ((cA)\mathbf{u})_j &= \sum_{k=1}^{n}(cA)_{jk}\mathbf{u}_k = \sum_{k=1}^{n}cA_{jk}\mathbf{u}_k = c\sum_{k=1}^{n}A_{jk}\mathbf{u}_k \\
&= c(A\mathbf{u})_j = c(L_A(\mathbf{u}))_j,
\end{aligned}$$

所以 $L_{cA} = cL_A$。

(D) 令 $C = [T]_\beta^\gamma$, 因為 $[T(\mathbf{u})]_\gamma = [T]_\beta^\gamma[\mathbf{u}]_\beta$ 以及對所有 $\mathbf{u} \in \mathbb{F}^m$ 都有 $T(\mathbf{u}) = C\mathbf{u} = L_C(\mathbf{u})$, 故 $T = L_C$, 而唯一性由 (B) 可得知。

(E) 對於 $j = 1, 2, \ldots, l$, 由 定理 **3** 知 $(AE)\mathbf{e}_j$ 為 AE 的第 j 行, 而 AE 的第 j 行亦等於 $A(E\mathbf{e}_j)$, 所以 $(AE)\mathbf{e}_j = A(E\mathbf{e}_j)$, 因此

$$L_{AE}(\mathbf{e}_j) = (AE)\mathbf{e}_j = A(E\mathbf{e}_j) = L_A(E\mathbf{e}_j) = L_A(L_E(\mathbf{e}_j)),$$

因此 $L_{AE} = L_A L_E$。

(F) 對所有 $\mathbf{u} \in \mathbb{F}^m$, 則 $L_{I_n}(\mathbf{u}) = I_n(\mathbf{u}) = \mathbf{u}$, 所以 L_{I_n} 是 $\mathbb{F}^n$ 到 $\mathbb{F}^n$ 的單位變換。因此 $L_{I_n} = I_{\mathbb{F}^n}$。

□

到目前為止, 我們已經介紹關於矩陣的加法、純量乘法、兩個矩陣相乘等性質, 關於矩陣的乘法運算, 我們通常會去探討它是否具有結合律。這裡我們可以透過左乘變換關於合成映射的結合律就很容易推論出矩陣乘法的結合律。

定理 7 (矩陣乘法的結合律, 第 93 頁). 假設 A, B, C 為矩陣使得 $A(BC)$ 有意義, 則 $(AB)C$ 有定義, 而且 $A(BC) = (AB)C$。

證明: 首先, 我們要從 $A(BC)$ 有意義這件事推導出 $(AB)C$ 有定義。因為 $A(BC)$ 有意義, 若 A 是 $m \times n$ 矩陣, 那麼 BC 是 $n \times p$ 矩陣。而 BC 是 $n \times p$ 矩陣, 若 B 是 $n \times l$ 矩陣的話, 那麼 C 必須是 $l \times p$ 矩陣。這麼一來, (AB) 是 $m \times l$ 矩陣, 所以 $(AB)C$ 是 $m \times p$ 矩陣, 故有意義。

利用 定理 **6** 的 (E) 與合成映射的結合律, 得知

$$L_{A(BC)} = L_A L_{BC} = L_A(L_B L_C) = (L_A L_B)L_C = L_{AB}L_C = L_{(AB)C},$$

故由 定理 **6** 的 (B) 可知 $A(BC) = (AB)C$。 □

2.6 可逆變換

不論是函數或是映射, 反函數或是逆映射也是一個重要的議題, 若一個映射的逆映射存在, 那麼定義域和值域這兩個系統之間就可以順利地互相轉換, 並且可以說明結構上有一些共通性。這一個單元是要探討可逆變換的議題, 包括逆變換的性質還有可逆變換的矩陣表現之關係。以下先給出線性變換是可逆的概念。

定義 **1** (第 99 頁)**.** 假設 V 與 W 為向量空間, 而 $T : V \to W$ 是線性變換, 若存在一個映射 $S : W \to V$ 使得 $ST = I_V$ 以及 $TS = I_W$, 我們說 T 是 *可逆的* (invertible) 線性變換。此時, 映射 $S : W \to V$ 稱為線性變換 T 的一個 *逆變換* (inverse)。

關於一個線性變換的逆變換, 以下引理證明「若逆變換存在, 則唯一」的性質。

引理 **2.** 若 $T : V \to W$ 是可逆的線性變換, 則線性變換 T 的逆變換唯一。

證明: 假設 $S_1 : W \to V$ 與 $S_2 : W \to V$ 都是 T 的逆變換, 即 S_1, S_2 滿足 $S_1T = I_V, TS_1 = I_W$ 以及 $S_2T = I_V, TS_2 = I_W$。因為對所有 $\mathbf{w} \in W$ 都有

$$\begin{aligned} S_1(\mathbf{w}) &= I_V(S_1(\mathbf{w})) = (S_2T)(S_1(\mathbf{w})) = S_2(T(S_1(\mathbf{w}))) = S_2((TS_1)(\mathbf{w})) \\ &= S_2(I_V(\mathbf{w})) = S_2(\mathbf{w}), \end{aligned}$$

所以 $S_1 = S_2$。 □

若一個線性變換 $T : V \to W$ 是可逆的, 由 *引理* **2** 可知這個逆變換是唯一的, 故將 T 的逆變換記為 T^{-1}。

以下給出一個線性變換是可逆的之等價條件。

引理 **3.** 線性變換 $T : V \to W$ 是可逆的若且唯若 T 是一對一且映成。

證明: ($\Rightarrow$) 因為線性變換 $T : V \to W$ 是可逆的, 存在映射 $S : W \to V$ 使得 $ST = I_V$ 以及 $TS = I_W$。

若有 $\mathbf{v}_1, \mathbf{v}_2 \in V$ 滿足 $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, 則 $S(T(\mathbf{v}_1)) = S(T(\mathbf{v}_2))$, 得到 $(ST)(\mathbf{v}_1) = (ST)(\mathbf{v}_2)$, 因為 $ST = I_V$, 所以 $\mathbf{v}_1 = \mathbf{v}_2$。因此 T 是一對一。

對任何 $\mathbf{w} \in W$, 因為 T 的逆變換 $S : W \to V$ 存在, 考慮 $S(\mathbf{w}) \in V$, 則 $T(S(\mathbf{v})) = (TS)(\mathbf{w}) = I_W(\mathbf{w}) = \mathbf{w}$, 所以 T 映成。

($\Leftarrow$) 若線性變換 $T : V \to W$, 其中 $T(\mathbf{v}) = \mathbf{w}$ 是一對一且映成, 考慮 $S(\mathbf{w}) = \mathbf{v}$ 這個關係, 因為 T 是一對一, 所以 S 是一個映射, 因為 T 是映成, 所以映射 S 的定義域是 W; 也就是說, $S : W \to V$ 是一個映射。

因為對所有 $\mathbf{v} \in V$, 都有 $(ST)(\mathbf{v}) = S(T\mathbf{v}) = S(\mathbf{w}) = \mathbf{v} = I_V(\mathbf{v})$, 所以 $ST = I_V$; 因為對所有 $\mathbf{w} \in W$, 都有 $(TS)(\mathbf{w}) = T(S(\mathbf{w})) = T(\mathbf{v}) = \mathbf{w} = I_W(\mathbf{w})$, 所以 $TS = I_W$。因此 T 是可逆的。 □

回顧逆變換的定義中, 我們只有說 $S : W \to V$ 是一個映射, 並沒有說 S 是線性變換, 而 S 確實是一個線性變換這件事必須加以論述。

定理 4 (第 100 頁)**.** 假設 $T : V \to W$ 是一個可逆的線性變換, 則它的逆變換 $T^{-1} : W \to V$ 也是線性變換。

證明:

- 對任意 $\mathbf{w}_1, \mathbf{w}_2 \in W$, 因為 T 是一對一且映成, 故分別存在唯一 $\mathbf{v}_1, \mathbf{v}_2 \in V$ 使得 $T(\mathbf{v}_1) = \mathbf{w}_1$ 以及 $T(\mathbf{v}_2) = \mathbf{w}_2$, 於是 $\mathbf{v}_1 = T^{-1}(\mathbf{w}_1)$ 且 $\mathbf{v}_2 = T^{-1}(\mathbf{w}_2)$, 得到

$$\begin{aligned} T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) &= T^{-1}(T(\mathbf{v}_1) + T(\mathbf{v}_2)) = T^{-1}(T(\mathbf{v}_1 + \mathbf{v}_2)) = (T^{-1}T)(\mathbf{v}_1 + \mathbf{v}_2) \\ &= (I_V)(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_1 + \mathbf{v}_2 = T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2)。\end{aligned}$$

- 對任意 $\mathbf{w} \in W$ 與 $c \in \mathbb{F}$, 因為 T 是一對一且映成, 故存在唯一 $\mathbf{v}$ 使得 $T(\mathbf{v}) = \mathbf{w}$, 於是 $\mathbf{v} = T^{-1}(\mathbf{w})$, 得到

$$\begin{aligned} T^{-1}(c \cdot \mathbf{w}) &= T^{-1}(c \cdot T(\mathbf{v})) = T^{-1}(T(c \cdot \mathbf{v})) = (T^{-1}T)(c \cdot \mathbf{v}) \\ &= (I_V)(c \cdot \mathbf{v}) = c \cdot \mathbf{v} = c \cdot T^{-1}(\mathbf{w})。\end{aligned}$$

- 由上述討論得知 T^{-1} 是一個線性變換。

□

引理 5 (第 100 頁)**.** 假設 $T : V \to W$ 是一個線性變換, 其中 V 與 W 都是有限維度的向量空間。若 T 是可逆的線性變換, 則 $\dim(V) = \dim(W)$。

證明: 因為 T 是一對一且映成, 所以 $\text{nullity}(T) = 0$, 於是 $\text{rank}(T) = \dim(R(T)) = \dim(W)$, 由維度定理 (Dimension Theorem, 單元 2.2 定理 **6**) 得知 $\dim(V) = \dim(W)$。 □

由線性變換的逆變換之討論, 現在想要了解逆變之矩陣表現, 這裡先定義反矩陣。

定義 6 (第 100 頁)**.** 給定矩陣 $A \in M_{n\times n}(\mathbb{F})$, 若存在矩陣 $B \in M_{n\times n}(\mathbb{F})$ 使得 $AB = BA = I_n$, 則稱 A 是 *可逆的* (invertible) 矩陣, 而 B 稱為矩陣 A 的 *反矩陣* (inverse matrix)。

同樣地, 我們也先證明若矩陣是可逆的, 那麼反矩陣有唯一性。

引理 7. 如果 A 是可逆的矩陣, 則滿足 $AB = BA = I_n$ 的矩陣 B 是唯一的。

證明: 假如 $B, C \in M_{n\times n}(\mathbb{F})$ 滿足 $AB = BA = I_n$ 以及 $AC = CA = I_n$, 則

$$C = CI_n = C(AB) = (CA)B = I_nB = B。$$

□

既然一個可逆矩陣 A 的反矩陣唯一, 之後我們會把 A 的反矩陣用 A^{-1} 表示。

例 **8.** 試求矩陣

$$A = \begin{bmatrix} 3 & 8 \\ 2 & 6 \end{bmatrix}$$

的反矩陣。

解**.** 欲求矩陣 $B \in M_{2\times 2}(\mathbb{R})$ 使得 $AB = BA = I_2$, 假設

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

由 $AB = I_2$, 得到

$$\begin{bmatrix} 3 & 8 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 3b_{11} + 8b_{21} & 3b_{12} + 8b_{22} \\ 2b_{11} + 6b_{21} & 2b_{12} + 6b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

對於 $b_{11}, b_{12}, b_{21}, b_{22}$ 這四個未知數來說, 形成對於 b_{11}, b_{21} 與 b_{12}, b_{22} 的兩個聯立方程組:

$$\begin{cases} 3b_{11} + 8b_{21} = 1 \\ 2b_{11} + 6b_{21} = 0 \end{cases} \quad \text{以及} \quad \begin{cases} 3b_{12} + 8b_{22} = 0 \\ 2b_{12} + 6b_{22} = 1, \end{cases}$$

各別解出 $(b_{11}, b_{21}) = (3, -1)$ 以及 $(b_{12}, b_{22}) = (-4, \frac{3}{2})$, 所以

$$B = \begin{bmatrix} 3 & -4 \\ -1 & \frac{3}{2} \end{bmatrix}。$$

上面的例題只是利用定義以及解聯立方程式的過程求得反矩陣, 關於矩陣的反矩陣如何確實求得是個很重要的議題, 我們在之後的章節提到幾個尋找反矩陣的方法還有理論。

下面的定理將說明逆變換的矩陣表現與原線性變換的矩陣表現之關係。

定理 **9** (第 100 頁)**.** 假設 $T : V \to W$ 是線性變換, 其中 β 與 γ 分別為有限維向量空間 V 與 W 的有序基底。則 T 是可逆的線性變換若且唯若它的矩陣表現 $[T]_\beta^\gamma$ 是可逆的, 並且 $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$。

證明: ($\Rightarrow$) 因為 T 是可逆的線性變換, 由 引理 **5** 得知 $\dim(V) = \dim(W)$, 記 $\dim(V) = n$, 於是 $[T]_\beta^\gamma$ 是一個 $n \times n$ 矩陣。

已知 $T^{-1} : W \to V$ 滿足 $TT^{-1} = I_W$ 以及 $T^{-1}T = I_V$。由合成變換之矩陣表現之關係 (單元 2.4 的 定理 **4**), 得到

因此,

$$I_n = [I_V]_\beta = [T^{-1}T]_\beta = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma$$
$$I_n = [I_W]_\gamma = [TT^{-1}]_\gamma = [T]_\beta^\gamma [T^{-1}]_\gamma^\beta,$$

所以 $[T]_\beta^\gamma$ 是一個可逆矩陣, 由上式亦可得知 $([T]_\beta^\gamma)^{-1} = [T^{-1}]_\gamma^\beta$。

($\Leftarrow$) 若矩陣 $A = [T]_\beta^\gamma$ 是可逆的, 則存在矩陣 $B \in M_{n\times n}(\mathbb{F})$ 使得 $AB = BA = I_n$, 記有序基底 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是 V 的有序基底, $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ 是 W 的有序基底, 由單元 2.2 的 定理 **11** 知: 存在 $S \in \mathcal{L}(W, V)$ 使得

$$S(\mathbf{w}_j) = \sum_{i=1}^{n} B_{ij}\mathbf{v}_i,\ \text{其中}\ j = 1, 2, \ldots, n,$$

於是 $[S]_\gamma^\beta = B$。以下欲證 T 是可逆變換, 並且 $T^{-1} = S$。因為

$$[ST]_\beta = [S]_\gamma^\beta [T]_\beta^\gamma = BA = I_n = [I_V]_\beta,$$

得到 $ST = I_V$。同理,

$$[TS]_\gamma = [T]_\beta^\gamma [S]_\gamma^\beta = AB = I_n = [I_W]_\beta,$$

所以 $TS = I_W$。因此 T 是可逆變換, 並且 $T^{-1} = S$。 $\square$

上面的結果當 $V = W$ 而且 $\beta = \gamma$ 的時候, 就可以得到 V 到 V 自身線性變換的關係。此外, $\mathbb{F}^n$ 到自身的左乘變換也有相應的結果。現將這兩個結果寫成以下推論。

推論 10 (第 100 頁)**.** 設 V 是有限維向量空間, 其中 β 是 V 的一組有序基底, 而 $T : V \to V$ 是線性變換, 則 T 是可逆的線性變換若且唯若 $[T]_\beta$ 是可逆矩陣。此時, $[T^{-1}]_\beta = ([T]_\beta)^{-1}$。

推論 11 (第 100 頁)**.** 假設 $A \in M_{n\times n}(\mathbb{F})$, 則 A 是可逆矩陣若且唯若它的左乘變換 L_A 是可逆的, 並且 $(L_A)^{-1} = L_{A^{-1}}$。

2.7 同構變換

前一個單元主要是介紹可逆的線性變換, 並得到可逆線性變換的幾個重要性質, 包括它對於有序基底的矩陣表現之結果。這個單元我們要從可逆線性變換的存在性引出兩個向量空間在結構上相同的概念, 以數學的術語來說稱為同構。這裡先給出同構的定義。

定義 **1** (第 102 頁)**.** 假設 V 和 W 為向量空間, 若存在一個可逆的線性變換 $T : V \to W$, 我們說 V 同構於 W (V is isomorphic to W), 而這個線性變換 T 稱為 由 V 映成至 W 的同構變換 (isomorphism)。

同構這個概念是一種等價關係 (equivalence relation); 也就是說, 向量空間的同構會滿足以下三個性質:

- 自反律 (reflexive law): 對任何向量空間 V, V 同構於 V。

  因為單位變換 $I_V : V \to V$ 滿足: 對所有 $\mathbf{v} \in V$ 都有 $I_V I_V(\mathbf{v}) = \mathbf{v} = I_V(\mathbf{v})$, 所以 I_V 是可逆的線性變換, 因此 V 同構於 V。

- 對稱律 (symmetric law): 若 V 同構於 W, 則 W 同構於 V。

  若 V 同構於 W, 則存在 $T : V \to W$ 是可逆的線性變換; 換言之, 存在 $S : W \to V$ 使得 $ST = I_V$ 以及 $TS = I_W$, 由單元 2.6 的 定理 **4** 知: S 是一個線性變換。而這個性質若以 $S : W \to V$ 的觀點來看, 它是一個可逆的線性變換, 於是 W 同構於 V。

- 遞移律 (transitive law): 若 V 同構於 W, W 同構於 Z, 則 V 同構於 Z。

  若 V 同構於 W, 則存在 $T_1 : V \to W$ 與 $S_1 : W \to V$ 滿足 $S_1T_1 = I_V$ 以及 $T_1S_1 = I_W$, 若 W 同構於 Z, 則存在 $T_2 : W \to Z$ 與 $S_2 : Z \to W$ 滿足 $S_2T_2 = I_W$ 以及 $T_2S_2 = I_Z$。考慮合成映射 $T_2T_1 : V \to W$, 因為

  $$(S_1S_2)(T_2T_1) = S_1(S_2T_2)T_1 = S_1(I_W)T_1 = S_1(I_WT_1) = S_1T_1 = I_V$$
  $$(T_2T_1)(S_1S_2) = T_2(T_1S_1)S_2 = T_2(I_V)S_2 = T_2(I_VS_2) = T_2S_2 = I_Z,$$

  而且 T_2T_1 與 S_1S_2 都是線性變換 (詳見單元 2.4 的 定理 **1**), 所以 V 同構於 Z。

我們利用兩個向量空間之間存在可逆的線性變換這個概念定義同構, 它們到底在哪些面向上具有相同的結構呢? 一個可以預期的結果是: 它們的大小相同, 這裡所說的大小是以維度來看待。

定理 **2** (第 103 頁)**.** 假設 V 和 W 皆為有限維度且佈於相同的體 $\mathbb{F}$ 的向量空間, 則 V 同構於 W 若且唯若 $\dim(V) = \dim(W)$。

證明: ($\Rightarrow$) 假設 V 同構於 W, 則存在 $T : V \to W$ 是由 V 映成至 W 的同構變換, 由單元 2.6 的 定理 **5** 之前的引理可知 $\dim(V) = \dim(W)$。

($\Leftarrow$) 假設 $\dim(V)=\dim(W)$, 記 $\beta=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ 與 $\gamma=\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_n\}$ 分別為 V 與 W 的一組基底, 由單元 2.2 的 定理 **4**, 存在一個線性變換 $T:V\to W$ 使得對於 $i=1,2,\ldots,n$ 都有 $T(\mathbf{v}_i)=\mathbf{w}_i$。由單元 2.2 的 定理 **11** 可知

$$R(T)=\mathrm{span}(T(\beta))=\mathrm{span}(\gamma)=W,$$

所以 T 是映成的, 再由單元 2.2 的 定理 **8**, 得知 T 是一對一。於是 V 同構於 W。 □

上面這個定理提供了一個非常好的判斷法則, 只要我們確定兩個有限維向量空間的維度一樣, 那麼它們就會同構。由此我們可以繼續給出一些相關的延伸討論。

推論 **3** (第 103 頁)**.** 若 V 是佈於 $\mathbb{F}$ 的向量空間, 則 V 同構於 $\mathbb{F}^n$ 若且唯若 $\dim(V)=n$。

之前我們已使用矩陣來表示線性變換。以下定理將說明所有線性變換所組成的向量空間可以看成是另一種適當的 $m\times n$ 矩陣的向量空間。

定理 **4** (第 103 頁)**.** 假設 V 及 W 分別為佈於 $\mathbb{F}$ 的有限維度之 n 維與 m 維向量空間, 記 β 與 γ 分別為 V 與 W 的有序基底。考慮映射 $\Phi:\mathcal{L}(V,W)\to M_{m\times n}(\mathbb{F})$ 定義為 $\Phi(T)=[T]_\beta^\gamma$, 則 Φ 是一個同構變換。

證明: 由單元 2.3 的 定理 **9** 得知 Φ 是線性的。以下將證明 Φ 是一對一且映成, 那麼 Φ 是一個同構變換。

給定 A 是 $m\times n$ 的矩陣, 令 $\beta=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ 與 $\gamma=\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_m\}$ 分別是 V 與 W 的有序基底, 由單元 2.2 的 定理 **11** 知: 存在唯一線性變換 $T:V\to W$ 使得

$$T(\mathbf{v}_j)=\sum_{i=1}^{m}A_{ij}\mathbf{w}_i,\quad \text{其中 } j=1,2,\ldots,n,$$

所以 $[T]_\beta^\gamma=A$; 換言之, 我們得到 Φ 是映成。由單元 2.2 的 推論 **12** 得知 Φ 是一對一。因此 Φ 是一個同構變換。 □

推論 **5** (第 104 頁)**.** 假設 V 與 W 分別為 n 及 m 維的向量空間, 則 $\mathcal{L}(V,W)$ 是 mn 維向量空間。

證明: 由 定理 **2** 及 定理 **4** 以及 $\dim(M_{m\times n}(\mathbb{F}))=mn$ 得知 $\mathcal{L}(V,W)$ 是 mn 維向量空間。 □

本單元的最後想要具體給出定義在抽象的有限維向量空間上的線性變換與由 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的線性變換兩者之間的關係。這裡我們需要回顧單元 2.3 曾經定義 $\mathbf{v}$ 對於有序基底 β 的坐標向量: 若 $\beta=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ 是 V 上的有序基底, 給定 $\mathbf{v}\in V$, 則存在唯一 $c_1,c_2,\ldots,c_n\in\mathbb{F}$ 使得 $\mathbf{v}=\sum\limits_{i=1}^{n}c_i\cdot\mathbf{v}_i$, 記

$$[\mathbf{v}]_\beta=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}。$$

定義 6 (第 104 頁)**.** 令 β 是佈於 $\mathbb{F}$ 之 n 維向量空間 V 的有序基底, 定義映射 $\phi_\beta : V \to \mathbb{F}^n, \phi_\beta(\mathbf{v}) = [\mathbf{v}]_\beta$, 其中 $\mathbf{v} \in V$, 我們稱 ϕ_β 是 V 相對於 β 的標準表現 (standard representation of V with respect to β)。

以下先證明 ϕ_β 是一個線性變換:

- 對任意 $\mathbf{u}, \mathbf{v} \in V$, 存在唯一 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 與 $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_n \in \mathbb{F}$ 使得 $\mathbf{u} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$ 以及 $\mathbf{v} = \sum_{i=1}^{n} \bar{c}_i \cdot \mathbf{v}_i$, 而 $\mathbf{u} + \mathbf{v} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i + \sum_{i=1}^{n} \bar{c}_i \cdot \mathbf{v}_i = \sum_{i=1}^{n} (c_i + \bar{c}_i) \cdot \mathbf{v}_i$, 於是
$$\phi_\beta(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} c_1 + \bar{c}_1 \\ c_2 + \bar{c}_2 \\ \vdots \\ c_n + \bar{c}_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \\ \vdots \\ \bar{c}_n \end{bmatrix} = \phi_\beta(\mathbf{u}) + \phi_\beta(\mathbf{v})。$$

- 對任意 $\mathbf{u} \in V$, 與 $c \in \mathbb{F}$, 存在唯一 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得 $\mathbf{u} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$, 而 $c \cdot \mathbf{u} = c \cdot \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i = \sum_{i=1}^{n} (c \cdot c_i) \cdot \mathbf{v}_i$, 於是
$$\phi_\beta(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} c \cdot c_1 \\ c \cdot c_2 \\ \vdots \\ c \cdot c_n \end{bmatrix} = c \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c \cdot \phi_\beta(\mathbf{u})。$$

- 由上討論可知: $\phi_\beta : V \to \mathbb{F}^n$ 是一個線性變換。

定理 7 (第 104 頁)**.** 若 V 是有限維向量空間, 任取 V 上的一組有序基底 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, 則 ϕ_β 是一個同構變換。

證明: 欲證 ϕ_β 是一個同構變換, 只要再證明 ϕ_β 是一對一且映成即可。

- 若 $\phi(\mathbf{u}) = \phi(\mathbf{v})$, 即
$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \\ \vdots \\ \bar{c}_n \end{bmatrix},$$
則 $\mathbf{u} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$ 以及 $\mathbf{v} = \sum_{i=1}^{n} \bar{c}_i \cdot \mathbf{v}_i$, 滿足 $c_i = \bar{c}_i, i = 1, 2, \ldots, n$, 於是 $\mathbf{u} = \mathbf{v}$。所以 ϕ_β 是一對一。

- 任給

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

考慮 $\mathbf{u} = \sum_{i=1}^{n} c_i \cdot \mathbf{v}_i$, 則

$$\phi_\beta(\mathbf{u}) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

因此 ϕ_β 是映成的。

$\square$

令 V 及 W 分別為 n 維與 m 維的向量空間, 且令 $T : V \to W$ 為一線性變換。定義 $A = [T]_\beta^\gamma$, 其中 β 及 γ 分別為 V 與 W 的有序基底。我們現在可使用 ϕ_β 及 ϕ_γ 來了解線性變換 T 以及 $L_A : \mathbb{F}^n \to \mathbb{F}^m$ 之間的關係。

考慮以下流程圖:

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "\phi_\beta"'] & W \arrow[d, "\phi_\gamma"] \\
\mathbb{F}^n \arrow[r, "L_A"] & \mathbb{F}^m
\end{tikzcd}$$

圖 2.6: 線性變換與矩陣變換的關係。

注意到這時有兩種線性變換的合成可以將 V 映至 $\mathbb{F}^m$:

- 由 V 映至 $\mathbb{F}^n$ 的 ϕ_β 與線性變換 L_A 得出合成 $L_A\phi_\beta$。
- 由 V 映至 W 的 T 與 ϕ_γ 得出 $\phi_\gamma T$。

於是 T 和 L_A 的關係為: $L_A\phi_\beta = \phi_\gamma T$。

2.8 坐標變換矩陣

在中學數學曾經探討過標準的圓錐曲線; 也就是說, 在平面 $\mathbb{R}^2$ 上引進坐標軸 x-軸與 y-軸之後, 我們知道:

- 型如 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 的圖形是一個橢圓。
- 型如 $y^2=4cx$ 或是 $x^2=4cy$ 的圖形是一個拋物線。
- 型如 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ 或是 $\frac{y^2}{a^2}-\frac{x^2}{b^2}=1$ 的圖形是一個雙曲線。

上述三種曲線稱為 *標準的圓錐曲線* (standard conic section)。由此出發, 我們也曾探索過一類型的推廣, 就是觀察以下變形:

- 型如 $\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1$ 的圖形是中心為 (h,k) 的橢圓。
- 型如 $(y-k)^2=4c(x-h)$ 或是 $(x-h)^2=4c(y-k)$ 的圖形是焦點為 (h,k) 拋物線。
- 型如 $\frac{(x-h)^2}{a^2}-\frac{(y-k)^2}{b^2}=1$ 或是 $\frac{(y-k)^2}{a^2}-\frac{(x-h)^2}{b^2}=1$ 的圖形是中心為 (h,k) 的雙曲線。

若是比較相對應的方程式, 各位會發現到後者的方程式比起前者的方程式來說多出了關於 x 與 y 的一次項, 這件事情其實告訴我們很重要的原則:

> 對於一個二次式, 以代數的觀點, 我們可以透過配方法將一次項消除。從幾何的觀點, 這個二次式的圖形與標準型相差的是平移的效應。

以下我們要討論另一個問題, 比方說在 $\mathbb{R}^2$ 上滿足 $2x^2-4xy+5y^2=1$ 的圖形是什麼? 更一般地, 綜合前述討論, 這類的問題若以數學的語言形容, 則是想要將二次式 (quadric form) 進行分類: 在 $\mathbb{R}^2$ 上滿足 $Ax^2+2Bxy+Cy^2+Dx+Ey+F=0$ 的圖形是否可以分類? 這個問題的答案是正面的, 而且從這個問題中我們可以看到線性代數發揮了重要的功用。然而, 要解決這個問題, 我們需要更多關於線性代數的工具才有辦法完整地說明, 這一個單元並不是要給出完整的理論, 而是想從一個實際例子的討論, 讓大家知道線性變換在平面幾何上的應用。

回到 $2x^2-4xy+5y^2=1$ 這個特別的例子, 若各位日後學完矩陣正交對角化 (orthogonal diagonalization) 這個理論之後, 就會曉得: 我們可以設定以下坐標變換:

$$\begin{cases} x=\frac{2}{\sqrt{5}}x'-\frac{1}{\sqrt{5}}y' \\ y=\frac{1}{\sqrt{5}}x'+\frac{2}{\sqrt{5}}y', \end{cases}$$

將它代入方程式之後, 得到

$$2\left(\frac{2}{\sqrt{5}}x'-\frac{1}{\sqrt{5}}y'\right)^2-4\left(\frac{2}{\sqrt{5}}x'-\frac{1}{\sqrt{5}}y'\right)\left(\frac{1}{\sqrt{5}}x'+\frac{2}{\sqrt{5}}y'\right)+5\left(\frac{1}{\sqrt{5}}x'+\frac{2}{\sqrt{5}}y'\right)^2=1,$$

繼續整理這個方程式後便得到 $(x')^2+6(y')^2=1$, 或是 $\frac{(x')^2}{1^2}+\frac{(y')^2}{(\frac{1}{\sqrt{6}})^2}=1$, 所以這個方程式從變數 x',y' 的觀點來看圖形是一個橢圓。

這一個單元是想透過這個例子仔細觀察變數變換 (x,y) 與 (x',y') 的關係。首先我們將關係式改寫成矩陣的形式:

$$\begin{cases} x = \frac{2}{\sqrt{5}}x' - \frac{1}{\sqrt{5}}y' \\ y = \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y' \end{cases} \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix},$$

其中 θ 滿足 $\cos\theta = \frac{2}{\sqrt{5}}$ 以及 $\sin\theta = \frac{1}{\sqrt{5}}$。此時, 記

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \left[\begin{array}{c|c} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right] = \left[\begin{array}{c|c} I_{\mathbb{R}^2}(\mathbf{e}_1') & I_{\mathbb{R}^2}(\mathbf{e}_2') \end{array}\right]_{\beta=\{\mathbf{e}_1,\mathbf{e}_2\}},$$

現在要解釋矩陣 Q 的意義: 考慮單位變換 $I_{\mathbb{R}^2}: \mathbb{R}^2 \to \mathbb{R}^2$, 其中對應域的有序基底是 $\beta = \{\mathbf{e}_1, \mathbf{e}_2\}$, 至於定義域的有序基底 $\beta = \{\mathbf{e}_1', \mathbf{e}_2'\}$ 分別是 x'-軸與 y'-軸的方向向量, 因為 $I_{\mathbb{R}^2}$ 的單位變換, 所以 $I_{\mathbb{R}^2}(\mathbf{e}_1') = \mathbf{e}_1', I_{\mathbb{R}^2}(\mathbf{e}_2') = \mathbf{e}_2'$, 於是矩陣 Q 的第一行是 $[\mathbf{e}_1']_\beta$, 第二行是 $[\mathbf{e}_2']_\beta$, 如圖 2.7 所示:

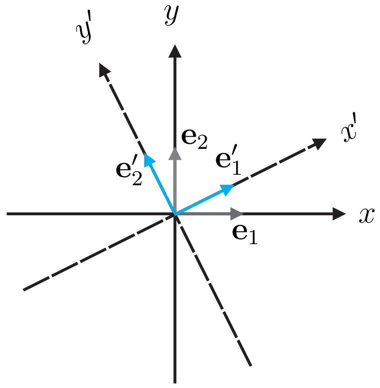

圖 2.7: 坐標變換與矩陣表現的關係。

現在給出另一個關於坐標變換的觀點: 若是想突顯 (x',y') 坐標是將 (x,y) 坐標經過旋轉而得到的結果, 則是將 x-軸與 y-軸設定在對應域, 而我們新找到的 x'-軸與 y-軸設定在定義域, 考慮旋轉變換 R_θ, 於是矩陣 Q 的第一行是 $[R_\theta(\mathbf{e}_1')]_\beta$, 第二行則為 $[R_\theta(\mathbf{e}_2')]_{\beta=\{\mathbf{e}_1,\mathbf{e}_2\}}$, 如圖 2.8 所示。

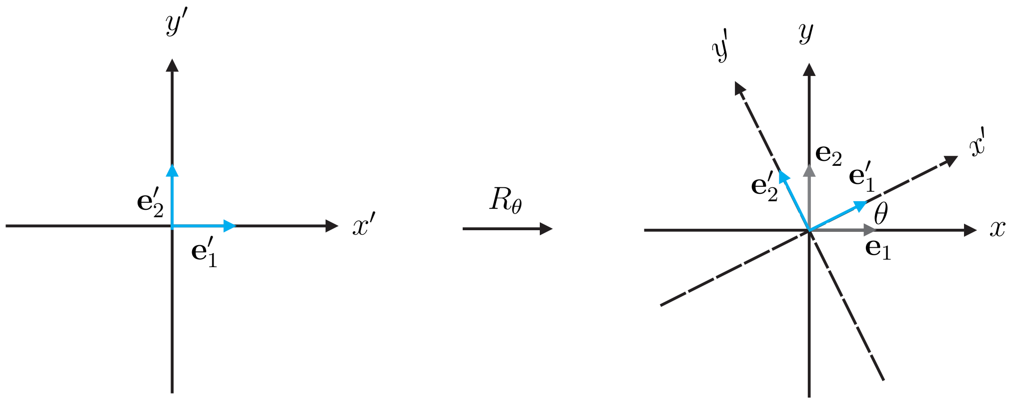

圖 2.8: 坐標變換與矩陣表現的關係。

現給出坐標變換 (以第一種觀點) 的相關結果:

定理 1 (第 111 頁)**.** 令 β 及 β' 是兩組有限維向量空間 V 的有序基底, 記 $Q = [I_V]_{\beta'}^{\beta}$, 則

(A) 矩陣 Q 是可逆的。

(B) 對任何 $\mathbf{v} \in V$, $[\mathbf{v}]_\beta = Q[\mathbf{v}]_{\beta'}$。

證明:

(A) 因為 I_V 是可逆變換, 由單元 2.6 的 **定理 9** 得知: 矩陣 Q 是可逆的。

(B) 對任何 $\mathbf{v} \in V$, 則

$$[\mathbf{v}]_\beta = [I_V(\mathbf{v})]_\beta = [I_V]_{\beta'}^{\beta}[\mathbf{v}]_{\beta'} = Q[\mathbf{v}]_{\beta'}。$$

□

對於一般的線性變換在兩種坐標系統之間的轉換式則有以下結果。

定理 2 (第 112 頁)**.** 假設 $T : V \to V$ 是有限維度之向量空間 V 上的一個線性變換, 令 β 與 β' 為 V 的有序基底。令 Q 表示坐標變換矩陣使得 β'-坐標變換至 β-坐標, 則

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q。$$

證明: 令 I_V 表示 V 上的單位變換, 則 $T = I_V T = T I_V$, 再由單元 2.4 的 **定理 4**, 可得

$$Q[T]_{\beta'} = [I_V]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'} = [I_V T]_{\beta'}^{\beta} = [T I_V]_{\beta'}^{\beta} = [T]_{\beta}^{\beta}[I_V]_{\beta'}^{\beta} = [T]_\beta Q,$$

所以 $[T]_{\beta'} = Q^{-1}[T]_\beta Q$。 □

現以方程式 $2x^2 - 4xy + 5y^2 = 1$ 為例, 聯繫它和 **定理 2** 的關係。這裡先呈現出日後在處理圓錐曲線分類時的流程: 首先, 將方程式改以矩陣乘法的方式呈現

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [1],$$

將中間的矩陣視為一種線性變換; 也就是說,

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [1]$$

而左式的五個矩陣, 將前兩個矩陣相乘、後兩個矩陣相乘, 這個部分即為 (x', y') 與 (x, y) 之間的變數變換, 得到

$$\begin{bmatrix} x' & y' \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = [1],$$

最後再將矩陣乘開並還原得到方程式 $(x')^2 + 6(y')^2 = 1$。

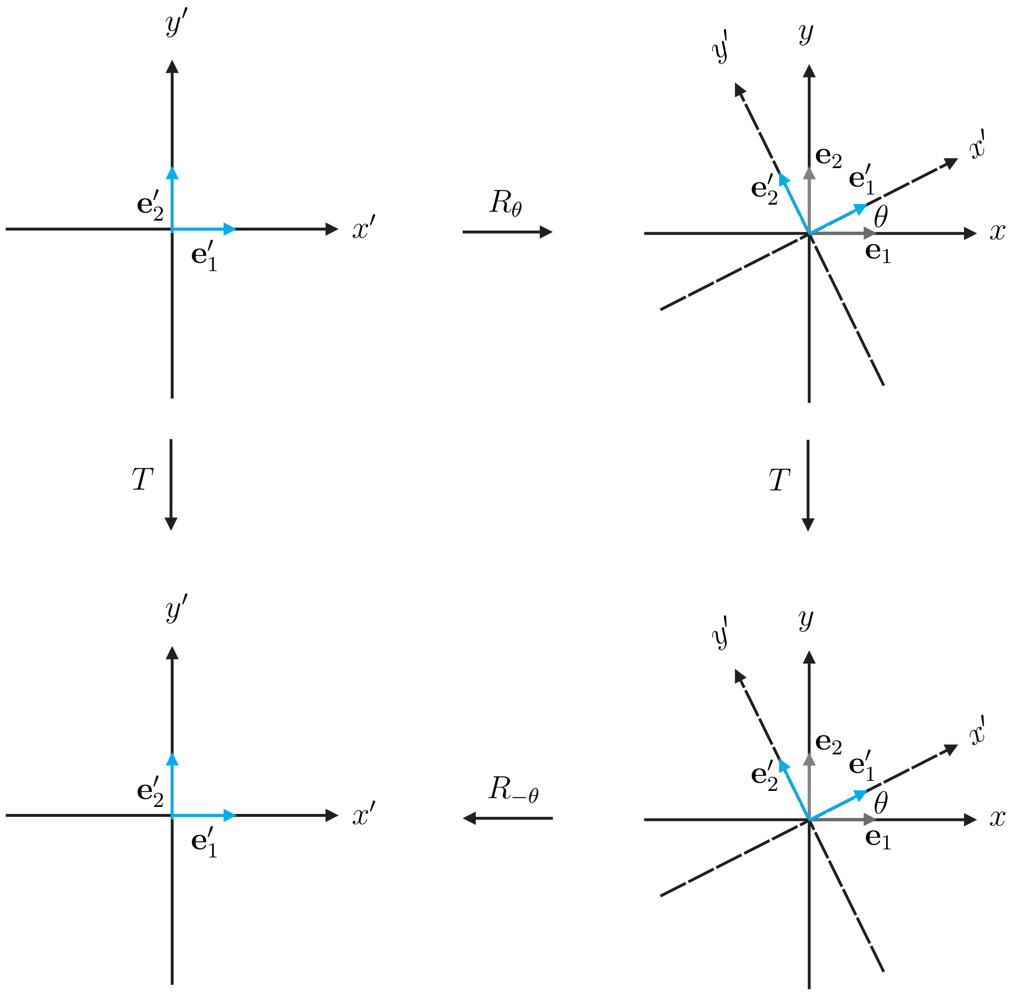

圖 2.9: 線性變換 T 在不同的坐標系統下表現的關係。

在上面的討論中, 原問題是一個比較複雜的方程式 $2x^2 - 4xy + 5y^2 = 1$, 透過變數變換後得到比較單純的方程式 $(x')^2 + 6(y')^2 = 1$。若要和 定理 **2** 所得到的結果相呼應的話, 則觀點是要從最後的式子逐步倒回去理解; 在 定理 **2** 的結論之左式 $[T]_{\beta'}$ 要呈現的是伸縮變換, 其中這個伸縮變換是對於 x' 方向伸長 1 倍、y' 方向伸長 6 倍, 而這個線性變換在 β' 這組基底的表現是簡單又清楚, 但是當我們選不太到好的坐標系的時候, 那麼線性變換的結果就會比較不好掌握, 像是 $[T]_\beta$ 這個矩陣表現就顯得比較複雜, 而 $[T]_{\beta'}$ 與 $[T]_\beta$ 的關係如下:

$$\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

這個矩陣的重現, 可和圖 2.9 對照, 左邊的矩陣是指圖 2.9 左側由上到下的映射 T 之表現, 右邊的三個矩陣相乘, 在圖 2.9 上則是先進行 R_θ 變換 (最右邊的矩陣), 然後再透過線性變換 T (以 β 的方式表現, 它對應於中間的矩陣), 最後再進行 $R_{-\theta}$ 變換 (最左邊的矩陣)。

關於矩陣對角化的理論就留到後面的章節再詳細介紹。

2.9 對偶空間

給定向量空間 $(V,\mathbb{F})$, 本單元想要研究一類特別的空間, 它是將所有 $T: V \to \mathbb{F}$ 為線性變換收集而成的空間, 由單元 2.3 的 **定理 8**, 我們知道它形成一個佈於 $\mathbb{F}$ 的向量空間, 並且以記號 $\mathcal{L}(V,\mathbb{F})$ 表示。現給出關於這個空間的幾個術語。

定義 1 (第 119 頁)**.** 給定佈於 $\mathbb{F}$ 的向量空間 V, 我們將向量空間 $\mathcal{L}(V,\mathbb{F})$ 稱為 V 的 *對偶空間* (dual space), 記為 V^*, 而 $T \in V^*$ 稱為 V *上的線性泛函* (linear functional on V)。

這裡注意到: 如果 V 是有限維向量空間, 由單元 2.7 的 **推論 5**, 得到

$$\dim(V^*) = \dim(\mathcal{L}(V,\mathbb{F})) = \dim(V)\cdot\dim(\mathbb{F}) = \dim(V),$$

根據單元 2.7 的 **定理 2**, 我們知道 V 和 V^* 同構。由此, 我們定義 V^{**} 是 V^* 的對偶空間, 並將 V^{**} 稱為 V 的 *雙對偶空間* (double dual space)。

因為基底是撐起向量空間的骨架, 給定有限維向量空間 V, 假設 $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ 是 V 上的一組有序基底, 我們自然想要問: 對於 V^* 上要如何找到一組有序基底? 以下討論提供一個自然地找法。

定理 2 (第 120 頁)**.** *給定有限維向量空間 V, 假設 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 是一組有序基底, 考慮 $\beta^* = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$, 其中 $\mathbf{f}_i \in V^*, i = 1, 2, \ldots, n$ 滿足*

$$\mathbf{f}_i(\mathbf{v}_j) = \delta_{ij} = \begin{cases} 1 & 若\ i = j \\ 0 & 若\ i \neq j, \end{cases}$$

則 $\beta^* = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 是 V^* 的一組有序基底。

證明: 首先, 對任何 $\mathbf{f} \in V^*$, 目標是要找到 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得

$$\mathbf{f} = \sum_{i=1}^{n} c_i \cdot \mathbf{f}_i。$$

尋找這些係數的方法是將這個線性泛函 $\mathbf{f}$ 分別作用 $\mathbf{v}_j$, 其中 $j = 1, 2, \ldots, n$ 之後, 得到

$$\mathbf{f}(\mathbf{e}_j) = \sum_{i=1}^{n} c_i \cdot \mathbf{f}_i(\mathbf{v}_j) = \sum_{i=1}^{n} c_i \cdot \delta_{ij} = c_j,$$

所以

$$\mathbf{f} = \sum_{i=1}^{n} \mathbf{f}(\mathbf{v}_i) \cdot \mathbf{f}_i,$$

上述討論得到 $V^* \subset \operatorname{span}(\beta^*)$, 又因為 $\operatorname{span}(\beta^*) \subset V^*$, 於是 $\operatorname{span}(\beta^*) = V^*$。又因為 $\dim(V^*) = n$, 所以 β^* 是 V 的一組基底。 $\square$

定義 3 (第 120 頁)**.** 我們說 $\beta^* = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 是 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 的 對偶基底 (dual basis)。

關於對偶空間的理論, 以下給出三個相關的應用: (1) 轉置矩陣對應的線性變換。(2) 向量空間與其雙對偶空間的自然同構。(3) 微積分中的微分 (differential) 之意義。

2.9.1 轉置矩陣

假設 V 與 W 是佈於 $\mathbb{F}$ 的有限維向量空間, 而 β 與 γ 分別為 V 與 W 的有序基底, 在單元 2.7 的 **定理 4** 中我們已經證明線性變換 $T : V \to W$ 與佈於 $\mathbb{F}$ 的 $m \times n$ 矩陣之間存在一對一的關係, 這裡, 我們用記號 $T \leftrightarrow [T]_\beta^\gamma$ 表示這個一對一的關係。

給定一個矩陣表現 $A = [T]_\beta^\gamma$, 現在想要問: 是否存在一個線性變換 S 使得對應於基底的矩陣表現是 A^t? 這個問題的最初步觀察是: 若 $m \neq n$, 則 S 不可能是由 V 至 W 的線性變換。針對這個問題, 若我們從對偶空間的觀點著手時則可順利回答此問題。

定理 4 (第 121 頁)**.** *令 V 及 W 是佈於 $\mathbb{F}$ 的有限維向量空間, 而 β 與 γ 分別為 V 與 W 的有序基底。對任何一個線性變換 $T : V \to W$, 定義映射 $T^t : W^* \to V^*$ 為*

$$T^t(\mathbf{g}) = \mathbf{g}T, \ \text{其中}\ \mathbf{g} \in W^*,$$

則 T^t 是一個線性變換; 此外, $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$。

證明: 給定 $\mathbf{g} \in W^*$, 即 $\mathbf{g} : W \to \mathbb{F}$ 是 W 上的線性泛函, 則 $T^t(\mathbf{g}) = \mathbf{g}T : V \to \mathbb{F}$ 是 V 到 $\mathbb{F}$ 的一個映射, 因此 T^t 是由 W^* 至 V^* 的映射。以下將證明 $T^t : W^* \to V^*$ 是一個線性變換。

- 對任意 $\mathbf{g}_1, \mathbf{g}_2 \in W^*$, 都有

$$T^t(\mathbf{g}_1 + \mathbf{g}_2) = (\mathbf{g}_1 + \mathbf{g}_2)T = \mathbf{g}_1 T + \mathbf{g}_2 T = T^t(\mathbf{g}_1) + T^t(\mathbf{g}_2)。$$

- 對任意 $\mathbf{g} \in W^*$ 以及任何 $c \in \mathbb{F}$, 都有

$$T^t(c \cdot \mathbf{g}) = (c \cdot \mathbf{g})T = c \cdot (\mathbf{g}T) = c \cdot T^t(\mathbf{g})。$$

- 由上述討論, 我們知道 T^t 是一個線性變換。

令 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 與 $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, 記它們的對偶基底為 $\beta^* = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 與 $\gamma^* = \{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\}$。為了方便起見, 令 $A = [T]_\beta^\gamma$, 欲求 $[T^t]_{\gamma^*}^{\beta^*}$ 的第 j 行, 首先將 $T^t(\mathbf{g}_j)$ 表示為 β^* 之線性組合, 由 **定理 2**, 可得

$$T^t(\mathbf{g}_j) = \mathbf{g}_j T = \sum_{k=1}^{n} (\mathbf{g}_j T)(\mathbf{v}_k)\mathbf{f}_k,$$

所以 $[T^t]_{\gamma^*}^{\beta^*}$ 第 i 列第 j 行的項為

$$(\mathbf{g}_j T)(\mathbf{v}_i) = \mathbf{g}_j(T(\mathbf{v}_i)) = \mathbf{g}_j\left(\sum_{k=1}^{n} A_{ki}\mathbf{w}_k\right) = \sum_{k=1}^{n} A_{ki}\mathbf{g}_j(\mathbf{w}_k) = \sum_{k=1}^{n} A_{ki}\delta_{jk} = A_{ji},$$

因此 $[T^t]_{\gamma^*}^{\beta^*} = A^t = ([T]_\beta^\gamma)^t$。 □

在 定理 **4** 中所定義的線性變換 $T^t : W^* \to V^*$ 稱為 $T : V \to W$ 的一個 轉置變換 (transpose)。注意到 T^t 是唯一的線性變換 S 使得 $[S]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$。

2.9.2 向量空間與雙對偶空間的自然同構

這個部分想要證明有限維向量空間 V 和它的雙對偶空間 V^{**} 可以建立一個自然的 (canonical) 同構關係 的同構關係; 也就是說, 存在一個介於 V 與 V^{**} 的同構變換與向量空間的基底選取無關。

任給一個向量 $\mathbf{v} \in V$, 定義 $\hat{\mathbf{v}} : V^* \to \mathbb{F}$ 為 $\hat{\mathbf{v}}(\mathbf{f}) = \mathbf{f}(\mathbf{v})$, 其中 $\mathbf{f} \in V^*$, 首先證明 $\hat{\mathbf{v}}$ 是 V^* 上的線性泛函:

- 對任何 $\mathbf{f}_1, \mathbf{f}_2 \in V^*$, 都有
$$\hat{\mathbf{v}}(\mathbf{f}_1 + \mathbf{f}_2) = (\mathbf{f}_1 + \mathbf{f}_2)(\mathbf{v}) = \mathbf{f}_1(\mathbf{v}) + \mathbf{f}_2(\mathbf{v}) = \hat{\mathbf{v}}(\mathbf{f}_1) + \hat{\mathbf{v}}(\mathbf{f}_2)。$$
- 對任何 $\mathbf{f} \in V^*$ 與任意 $c \in \mathbb{F}$, 都有
$$\hat{\mathbf{v}}(c \cdot \mathbf{f}) = (c \cdot \mathbf{f})(\mathbf{v}) = c \cdot (\mathbf{f}(\mathbf{v})) = c \cdot \hat{\mathbf{v}}(\mathbf{f})。$$
- 由上述討論知道: $\hat{\mathbf{v}} : V^* \to \mathbb{F}$ 是 V^* 上的線性泛函。

而上面的說明告知: $\hat{\mathbf{v}} \in V^{**}$, 而當中也發現 $\mathbf{v} \leftrightarrow \hat{\mathbf{v}}$ 的對應可給出 V 與 V^{**} 之間的自然同構變換; 也就是說, 上述討論並沒有透過基底的方式將 $\mathbf{v}$ 與 $\hat{\mathbf{v}}$ 進行對應。

定理 **5** (第 122 頁)**.** 假設 V 是一個有限維向量空間, 考慮 $\psi : V \to V^{**}$ 定義為 $\psi(\mathbf{v}) = \hat{\mathbf{v}}$, 則 ψ 是一個同構變換。

證明:

(A) 首先證明 ψ 是線性: 對任意 $\mathbf{u}, \mathbf{v} \in V$ 以及對所有 $\mathbf{f} \in V^*$, 則
$$\psi(\mathbf{u} + \mathbf{v})(\mathbf{f}) = \mathbf{f}(\mathbf{u} + \mathbf{v}) = \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v}) = \hat{\mathbf{u}}(\mathbf{f}) + \hat{\mathbf{v}}(\mathbf{f}) = \psi(\mathbf{u})(\mathbf{f}) + \psi(\mathbf{v})(\mathbf{f}),$$
所以 $\psi(\mathbf{u} + \mathbf{v}) = \psi(\mathbf{u}) + \psi(\mathbf{v})$。

對任意 $\mathbf{u} \in V$ 與任何 $c \in \mathbb{F}$, 對所有 $\mathbf{f} \in V^*$, 則
$$\psi(c \cdot \mathbf{u})(\mathbf{f}) = \mathbf{f}(c \cdot \mathbf{u}) = c \cdot \mathbf{f}(\mathbf{u}) = c \cdot \hat{\mathbf{u}}(\mathbf{f}) = c \cdot \psi(\mathbf{u})(\mathbf{f}),$$
所以 $\psi(c \cdot \mathbf{u}) = c \cdot \psi(\mathbf{u})$。

(B) 證明 ψ 是一對一: 假設 $\psi(\mathbf{v})$ 是 V^* 上的零向量, 其中 $\mathbf{v} \in V$, 則 $\hat{\mathbf{v}}(\mathbf{f}) = 0$ 對所有 $\mathbf{f} \in V^*$, 目標是要證 $\mathbf{v} = 0$。假設 $\mathbf{v} \neq \mathbf{0}$, 在 V 上選取一個有序基底 $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 使得 $\mathbf{v}_1 = \mathbf{v}$。令 $\beta^* = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 是 β 的對偶基底, 則 $\mathbf{f}_1(\mathbf{v}) = \mathbf{f}_1(\mathbf{v}_1) = 1 \neq 0$; 換言之, $\psi(\mathbf{v})(\mathbf{f}_1) = \hat{\mathbf{v}}(\mathbf{f}_1) = \mathbf{f}_1(\mathbf{v}) \neq 0$ 得到矛盾。取 $\mathbf{f} = \mathbf{f}_1$ 即為所求。因此 $\mathbf{v} = \mathbf{0}$。

(C) 證明 ψ 是一個同構變換: 由 (B) 與 $\dim(V) = \dim(V^{**})$ 及單元 2.7 的 定理 **2** 得證。

□

回想對偶基底的建構, 先給 V 上的一組有序基底 β 然後得到 V^* 的一組有序基底 β^*, 現在想要問的是: 給定 V^* 上的一組有序基底 β^*, 是否能在 V 上找到一組有序基底 β 使得 β^* 是 β 的對偶基底呢? 由上述的自同構, 我們可以得到以下結果:

定理 **6** (第 123 頁)**.** 假設 V 是一個有窮維向量空間, 其對偶空間為 V^*, 則 V^* 的任何一組有序基底都是 V 上某一組基底的對偶基底。

證明: 令 $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 是 V^* 上的一組有序基底, 對於 V^* 的這個基底, 存在 V^{**} 的對偶基底 $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \ldots, \hat{\mathbf{v}}_n\}$, 即 $\hat{\mathbf{v}}_i(\mathbf{f}_j) = \delta_{ij}$。透過 V 和 V^{**} 是自然同構的關係, 由 $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \ldots, \hat{\mathbf{v}}_n\}$ 可得 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$。此時, 因為 $\hat{\mathbf{v}}_i(\mathbf{f}_j) = \mathbf{f}_j(\mathbf{v}_i) = \delta_{ij}$, 故 $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ 是 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 的對偶基底。 □

雖然這個單元的許多概念可以推廣至 V 是無限維的情況, 但是在 V 與 V^* 與 V^{**} 的同構關係在無限維的向量空間中不見得會成立, 這是泛函分析 (functional analysis) 的理論中討論的重點, 在此並不詳述。

2.9.3 微分與全微分的意義

微積分課程中單變數函數的微分 (differential) 還有多變數函數的全微分 (total differential) 正確理解, 必須從向量空間與對偶基底的觀點進行說明。

先看單變數函數的情況, 給定一個函數 $y(x) : \mathbb{R} \to \mathbb{R}$, 假設函數 $y = y(x)$ 在 $x = x_0$ 是可微分的 (differentiable), 定義函數 $y = y(x)$ *在一點* $x = x_0$ *的微分* (differential of $y = y(x)$ at $x = x_0$)

$$\mathrm{d}y|_{x=x_0} \xlongequal{\text{定義}} y'(x_0)\,\mathrm{d}x : V^1 = \mathbb{R} \to \mathbb{R},$$

它是從向量空間 V^1 映至向量空間 $\mathbb{R}$ 的一個線性泛函 (linear functional)。這個線性泛函的定義域所指的向量空間 V^1 是收集所有在 $x = x_0$ 的求導算子, 它和 $\mathbb{R}$ 同構 (isomorphic), 故寫 $V^1 = \mathbb{R}$, 而它的其中一個基底是 $\{\frac{\mathrm{d}}{\mathrm{d}x}\}$; 至於 $\{\mathrm{d}x : V^1 = \mathbb{R} \to \mathbb{R}\}$ 是 $\{\frac{\mathrm{d}}{\mathrm{d}x}\}$ 的對偶基底, 故有 $\mathrm{d}x(\frac{\mathrm{d}}{\mathrm{d}x}) = 1$, 於是 $\mathrm{d}y|_{x=x_0} = y'(x_0)\,\mathrm{d}x$ 是 $V^1 = \mathbb{R}$ 上的一個線性泛函, 等式右邊是在說明將 $\mathrm{d}y|_{x=x_0}$ 寫成基底 $\{\mathrm{d}x\}$ 的線性組合之結果, 它是向量 $\mathrm{d}x$ 再乘上倍數 $y'(x_0)$ 而得。

到目前為止, 我們只說明了函數 $y = y(x)$ 在一點 $x = x_0$ 微分的意思, 而上面的討論中可發現, 這個線性泛函是隨著點而變的; 也就是說, 函數 $y = y(x)$ 在每一點都對應一個線性泛函, 因為用基底 $\mathrm{d}x$ 寫開來的時候, 前面的係數 $y'(x_0)$ 和位置有關, 所以不同的位置給出不同的線性泛函。現將函數 $y = y(x)$ 在每一點的微分都收集起來, 就定義了可微分函數 $y = y(x)$ *的微分* (differential of $y = y(x)$):

$$\mathrm{d}y = y'(x)\,\mathrm{d}x。$$

總結來說, 函數的微分是在說明函數在每一個點 y 對於 x 的線性增長關係。由線性代數的理論知道: 若要認識一個線性變換, 只要把基底的變換過程說清楚, 那麼任何其它向量的變換都是遵照著線性的方式完全確定。於是 $\mathrm{d}y = y'(x)\,\mathrm{d}x$ 就充分說明了每一點所對應的線性變換之關係。

為了讓各位對微分這個概念有更深刻地認識, 以下再到用圖形的方式解說微分的意義。我們知道: 所有從 $\mathbb{R}$ 到 $\mathbb{R}$ 的線性變換 T, 若在 $\mathbb{R}^2=\mathbb{R}\times\mathbb{R}$ 上呈現其意義, 則 T 會是一條通過原點的直線。如圖 2.10 所示: 給定可微分函數 $y=y(x)$, 對於一點 x, 微分 $\mathrm{d}y=y'(x)\,\mathrm{d}x$ 是一個由 $V^1=\mathbb{R}$ 到 $\mathbb{R}$ 的線性變換, 其中定義域 $V^1=\mathbb{R}$ 是以 x 的位置為零向量橫向延伸而得的向量空間, 而對應域 $\mathbb{R}$ 是以 $y(x)$ 的位置為零向量縱向延伸而得的向量空間。這兩個向量空間的乘積空間 (product space) $\mathbb{R}^2=\mathbb{R}\times\mathbb{R}$ 就如圖形中以 $(x,y(x))$ 為零向量, 灰色軸為坐標而張開的空間。於是微分 $\mathrm{d}y=y'(x)\,\mathrm{d}x$ 在圖形上的呈現是函數 $y=y(x)$ 的圖形在 $(x,y(x))$ 的切線。

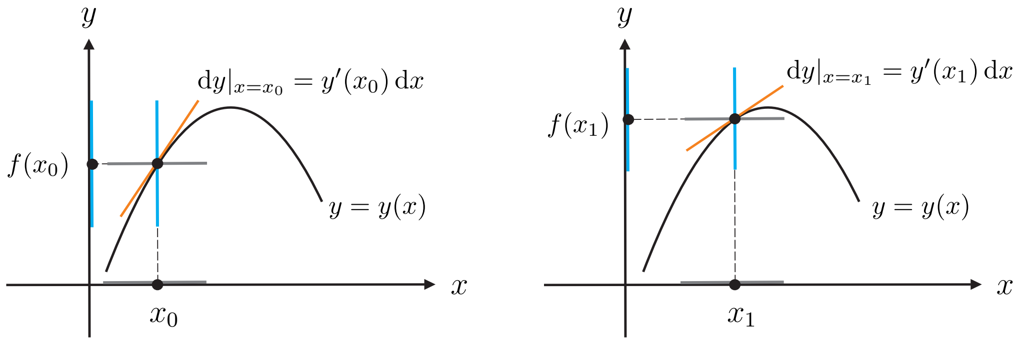

圖 2.10: 微分 $\mathrm{d}y=y'(x)\,\mathrm{d}x:\mathbb{R}\to\mathbb{R}$ 是在每一個點都指定一個線性泛函。

由圖 2.10 也可以看出微分 $\mathrm{d}y$ 在不同點對應到的線性泛函也不同。

看完單變數函數微分後, 現在要討論二變數函數全微分的意義。給定一個二變數函數 $z=z(x,y)$, 假設函數在一點 $p=(x_0,y_0)$ 是可微分的, 這時可定義函數 $z=z(x,y)$ 在一點 $p=(x_0,y_0)$ 的 全微分 (total differential of $z=z(x,y)$ at $p=(x_0,y_0)$) 為

$$\mathrm{d}z|_p \overset{\text{定義}}{=\!=} \frac{\partial z(p)}{\partial x}\,\mathrm{d}x+\frac{\partial z(p)}{\partial y}\,\mathrm{d}y : V^2=\mathbb{R}^2\to\mathbb{R},$$

它是一個定義在 V^2 上的線性泛函, 其中向量空間 V^2 是將所有在 p 點的求導算子收集而成的集合, 這個向量空間的一組基底是 $\{\frac{\partial}{\partial x}|_p,\frac{\partial}{\partial y}|_p\}$, 它的對偶基底是 $\{\mathrm{d}x|_p,\mathrm{d}y|_p\}$; 也就是說, 兩組基底之間滿足以下關係:

$$\begin{cases}\mathrm{d}x\left(\dfrac{\partial}{\partial x}\right)=1\\[2ex] \mathrm{d}x\left(\dfrac{\partial}{\partial y}\right)=0\end{cases}\quad\text{與}\quad\begin{cases}\mathrm{d}y\left(\dfrac{\partial}{\partial x}\right)=0\\[2ex] \mathrm{d}y\left(\dfrac{\partial}{\partial y}\right)=1。\end{cases}$$

而 $\mathrm{d}z|_p=\frac{\partial z(p)}{\partial x}\,\mathrm{d}x+\frac{\partial z(p)}{\partial y}\,\mathrm{d}y$ 要呈現的是函數 $z=z(x,y)$ 在 p 點的線性增長量, 等式右邊就是寫出 $\mathrm{d}z|_p$ 這個線性泛函用基底 $\{\mathrm{d}x,\mathrm{d}y\}$ 進行線性組合的結果。

同樣地, 函數 $z=z(x,y)$ 在每一點的線性增長量都不盡相同, 所以將函數在每一點的全微分都收集起來, 就得到函數 $z=z(x,y)$ 的 全微分 (total differential of $z=z(x,y)$), 用

$$\mathrm{d}z=\frac{\partial z}{\partial x}\,\mathrm{d}x+\frac{\partial z}{\partial y}\,\mathrm{d}y$$

這個記號表示。