Chapter 6

Linear Transformations

在本章中我們介紹在 vector spaces 之間重要的函數, 所謂的 linear transformation. 我們會介紹 linear transformation 相關的基本性質. 然後引進其矩陣表示法, 將 linear transformation 與矩陣相連結.

6.1. Basic Properties

在數學中, 函數是我們常常利用來了解所要探討的結構的重要工具. 在線性代數中, 我們要探討的結構就是 vector space, 而 linear transformation 就是幫助我們探討及理解 vector spaces 相互之間的關係的重要函數與工具.

6.1.1. Function. 給定 vector spaces V, W. 若有一個從 V 的向量對應到 V 的向量的對應關係, 即對任意 $\mathbf{v}\in V$, 此對應會將 $\mathbf{v}$ 對應到 W 中一個向量 $\mathbf{w}$. 若對每一個 $\mathbf{v}\in V$, 所對應到的 $\mathbf{w}$ 我們用 $T(\mathbf{v})$ 來表示, 我們就用 $T:V\to W$, $T(\mathbf{v})=\mathbf{w}\in W$, $\forall\mathbf{v}\in V$ 來表示這一個對應關係, 而稱 T 是一個從 V 到 W 的 *function* (函數). 這裡 V 稱為 T 的 *domain* (定義域), 而 W 就稱為 T 的 *codomain* (對應域). 注意依函數的定義, 若 $T:V\to W$ 是一個函數, 則對任意 $\mathbf{v}\in V$, $T(\mathbf{v})$ 一定要是 W 中的一個確定的向量. 也就是說 $T(\mathbf{v})\in W$, 而且不能一下子令 $T(\mathbf{v})=\mathbf{w}$, 一下子又改變成 $T(\mathbf{v})=\mathbf{w}'$, 但 $\mathbf{w}\neq\mathbf{w}'$, 造成不一致的情形發生. 所以當我們在建構一個新的函數時, 一定要確認這一點. 也就是說, 我們必須說明定出來的函數是 *well-defined*.

另外要注意, 函數的定義中並沒有要求定義域中每一個元素都要對應到不同的元素去. 也就是說若 $T:V\to W$ 是一個函數, 是容許在 V 中有兩個向量 $\mathbf{v}\neq\mathbf{v}'$, 但 $T(\mathbf{v})=T(\mathbf{v}')$. 若我們多要求定義域中每一個元素都要對應到不同的元素, 即 V 中任意兩相異向量 $\mathbf{v}\neq\mathbf{v}'$, 所對應的 $T(\mathbf{v})$ 和 $T(\mathbf{v}')$ 要相異 (即 $T(\mathbf{v})\neq T(\mathbf{v}')$), 則我們給這樣的函數一個特殊的名稱, 稱這樣的函數為 *one-to-one* (一對一), 有時也稱為 *injective*. 另外, 函數的定義中也沒有要求對應域中每一個元素都要被對應到. 也就是說若 $T:V\to W$ 是一個函數, 是容許在 W 中有向量 $\mathbf{w}$, 沒有任何 V 的向量會對應到 $\mathbf{w}$ (即不存在 $\mathbf{v}\in V$ 使得 $T(\mathbf{v})=\mathbf{w}$). 若我們多要求對應

域中每一個元素都要被對應到, 即 W 中任意向量 $\mathbf{w}$, 皆可找到 $\mathbf{v}\in V$ 使得 $T(\mathbf{v})=\mathbf{w}$, 則我們給這樣的函數一個特殊的名稱, 稱這樣的函數為 *onto* (映成), 有時也稱為 *surjective*.

當一個函數是 one-to-one 且 onto (此時一般稱為 *bijective*), 那就更特別了. 這時候一定可以找到一個從原來函數的對應域送到原來函數的定義域的反向函數 (我們稱為原函數的 *inverse* (反函數)), 使其合成後會是將每一個元素自己映射到自己的函數 (即所謂的 identity function). 所以此時我們也稱這樣的函數為 *invertible* (可逆函數).

6.1.2. Linear Transformation. 要了解 vector spaces, 若僅是考慮一般的函數, 並無法利用向量之間的運算, 幫助我們了解這些 vector spaces 中向量的結構. 我們需要的函數是能保持向量運算的, 所以有以下之定義.

Definition 6.1.1. 假設 V,W 皆為 vector space over $\mathbb{F}$ 且 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為一個函數, 若 T 滿足對任意 $\mathbf{v}_1,\dots,\mathbf{v}_k\in V$ 以及 $c_1,\dots,c_k\in\mathbb{F}$ 皆有

$$T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k).$$

則稱 T 為一個 *linear transformation*. 有時我們簡稱 T 為 *linear*.

要注意這裡 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_1$ 是 V 中向量的線性組合, 而 $c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$ 是 W 中向量的線性組合, 要區分清楚. 尤其在 $V\neq W$ 時要特別注意. 特別是當 $\mathbf{0}\in V$ 是 V 的 zero vector 時, 依 linear transformation 的定義, 我們有 $T(\mathbf{0})=T(\mathbf{0}+\mathbf{0})=T(\mathbf{0})+T(\mathbf{0})$. 此時兩邊加上 $T(\mathbf{0})$ 的加法反元素, 得 $T(\mathbf{0})$ 應為 W 中的 zero vector. 也就是說一個 linear transformation $T:V\to W$, 會將 V 中的零向量映射到 W 中的零向量. 雖然當 $V\neq W$ 時, V 和 W 的零向量有可能是不同的, 不過一般我們都用 $\mathbf{0}$ 來表示, 而不區分它. 所以我們仍用 $T(\mathbf{0})=\mathbf{0}$ 來表示 linear transformation 會將 V 中的零向量映射到 W 中的零向量. 這個性質雖然簡單, 但相當有用, 我們特別將此性質敘述如下.

Lemma 6.1.2. 假設 $T:V\to W$ 為一個 *linear transformation*. 則 T 會將 V 中的零向量映射到 W 中的零向量, 亦即 $T(\mathbf{0})=\mathbf{0}$.

再次提醒, 這裡兩個 $\mathbf{0}$ 哪一個是 V 的零向量, 哪一個是 W 的零向量, 一定要區分清楚.

依定義要檢查 $T:V\to W$ 是否為 linear transformation, 我們必須考慮 V 中任意有限多個向量的線性組合代入 T 中是否符合 linear transformation 的要求, 感覺起來很麻煩. 事實上, 如同檢查 subspace 的方法 (參見 Corollary 3.3.3), 下一個定理告訴我們, 只要檢查任兩個向量的線性組合即可.

Proposition 6.1.3. 假設 V,W 皆為 *vector space over* $\mathbb{F}$ 且 $T:V\to W$ 為一個函數, 則 T 為 *linear transformation* 若且唯若對任意 $\mathbf{u},\mathbf{v}\in V$, $r\in\mathbb{F}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$.

Proof. $(\Rightarrow)$: 依 T 為 linear 的定義, 對任意 $\mathbf{u},\mathbf{v}\in V$, $r\in\mathbb{F}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$.

$(\Leftarrow)$: 我們要利用對任意 $\mathbf{u},\mathbf{v}\in V$, $r\in\mathbb{F}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$ 這個性質來證明對任意 $\mathbf{v}_1,\dots,\mathbf{v}_k\in V$ 以及 $c_1,\dots,c_k\in\mathbb{F}$ 皆有 $T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$.

我們對向量的個數 k 作數學歸納法. 首先考慮只有一個向量的情形 (即 $k=1$), 我們要證明若 $\mathbf{v}_1\in\mathbb{R}^n$, $c_1\in\mathbb{R}$ 則 $T(c_1\mathbf{v}_1)=c_1T(\mathbf{v}_1)$. 此時考慮 $\mathbf{u}=\mathbf{0}$, $\mathbf{v}=\mathbf{v}_1$, $r=c_1$, 依 Lemma 6.1.2, 我們有 $T(c_1\mathbf{v}_1)=T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})=\mathbf{0}+rT(\mathbf{v})=c_1T(\mathbf{v}_1)$. 得證 $k=1$ 的情形成立. 現假設有 k 個向量時成立, 亦即對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k\in V$ 以及 $c_1,\ldots,c_k\in\mathbb{F}$ 皆有 $T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$. 我們要證明對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k,\mathbf{v}_{k+1}\in V$ 以及 $c_1,\ldots,c_k,c_{k+1}\in\mathbb{F}$ 皆有 $T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k+c_{k+1}\mathbf{v}_{k+1})=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)+c_{k+1}T(\mathbf{v}_{k+1})$. 然而對此時令 $\mathbf{u}=c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k$, $\mathbf{v}=\mathbf{v}_{k+1}$ 以及 $r=c_{k+1}$. 依歸納假設我們有 $T(\mathbf{u})=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$, 故

$$\begin{aligned}&T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k+c_{k+1}\mathbf{v}_{k+1})=\\&\qquad T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)+c_{k+1}T(\mathbf{v}_{k+1}).\end{aligned}$$

故由數學歸納法知 T 為 linear transformation. □

Example 6.1.4. (1) 考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=\begin{bmatrix}x_1+x_2\\x_1-x_3\end{bmatrix}$. 我們驗證 T 是一個 linear transformation. 任取 $\mathbf{u}=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix},\mathbf{v}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}\in\mathbb{R}^3$, 以及 $r\in\mathbb{R}$. 我們有 $\mathbf{u}+r\mathbf{v}=\begin{bmatrix}a_1+rb_1\\a_2+rb_2\\a_3+rb_3\end{bmatrix}$. 故依 T 的定義, 我們有

$$T(\mathbf{u}+r\mathbf{v})=T(\begin{bmatrix}a_1+rb_1\\a_2+rb_2\\a_3+rb_3\end{bmatrix})=\begin{bmatrix}(a_1+rb_1)+(a_2+rb_2)\\(a_1+rb_1)-(a_3+rb_3)\end{bmatrix}=\begin{bmatrix}a_1+a_2+rb_1+rb_2\\a_1-a_3+rb_1-rb_3\end{bmatrix}.$$

另一方面我們有 $T(\mathbf{u})=T(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix})=\begin{bmatrix}a_1+a_2\\a_1-a_3\end{bmatrix}$, $T(\mathbf{v})=T(\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix})=\begin{bmatrix}b_1+b_2\\b_1-b_3\end{bmatrix}$, 故

$$T(\mathbf{u})+rT(\mathbf{v})=\begin{bmatrix}a_1+a_2\\a_1-a_3\end{bmatrix}+r\begin{bmatrix}b_1+b_2\\b_1-b_3\end{bmatrix}=\begin{bmatrix}a_1+a_2+rb_1+rb_2\\a_1-a_3+rb_1-rb_3\end{bmatrix}.$$

得證 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$, 故 T 為 linear transformation.

(2) 考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=\begin{bmatrix}x_1+x_2+1\\x_1-x_3\end{bmatrix}$. 我們說明 T 不是 linear transformation. 依 T 的定義, 我們有 $T(\mathbf{0})=\begin{bmatrix}1\\0\end{bmatrix}\neq\mathbf{0}$, 故由 Lemma 6.1.2 知, T 不是 linear transformation.

(3) 考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=\begin{bmatrix}x_1^2+x_2\\x_1-x_3\end{bmatrix}$. 我們說明 T 不是 linear transformation. 雖然此時 $T(\mathbf{0})=\mathbf{0}$, 但 $T(\begin{bmatrix}1\\0\\0\end{bmatrix})=\begin{bmatrix}1\\1\end{bmatrix}$, 而 $T(\begin{bmatrix}2\\0\\0\end{bmatrix})=\begin{bmatrix}4\\2\end{bmatrix}$ 得

$$T(\begin{bmatrix}2\\0\\0\end{bmatrix})=T(2\begin{bmatrix}1\\0\\0\end{bmatrix})\neq 2T(\begin{bmatrix}1\\0\\0\end{bmatrix}).$$

此與 linear transformation 的條件不符, 故 T 不是 linear transformation.

接下來我們來看一個最常見的 linear transformation, 事實上以後我們會知道所有 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation 都是這樣的形式.

Lemma 6.1.5. *假設 $\mathbb{F}$ 為 field 且 $A \in M_{m\times n}(\mathbb{F})$. 考慮 $T : \mathbb{F}^n \to \mathbb{F}^m$ 定義為: $T(\mathbf{v}) = A\mathbf{v}$, $\forall \mathbf{v} \in V$. 則 T 為一個 linear transformation.*

Proof. 首先我們先檢查 T 是 well-defined, 也就是說 T 確實是一個從 $\mathbb{F}^n$ 映到 $\mathbb{F}^m$ 的函數. 任取 $\mathbf{v} \in \mathbb{F}^n$, 依定義 $T(\mathbf{v}) = A\mathbf{v}$. 然而 A 為 $m \times n$ matrix, 依矩陣乘法定義 $A\mathbf{v}$ 是一個 $m \times 1$ matrix (注意這裡向量都視為 column vector, 所以 $\mathbf{v} \in \mathbb{F}^n$ 為 $n \times 1$ matrix), 故 $A\mathbf{v} \in \mathbb{F}^m$. 得 T 確實是一個從 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 function.

現要證明 T 為 linear, 亦即對任意 $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$ 以及 $r \in \mathbb{F}$, 我們要證明 $T(\mathbf{u} + r\mathbf{v}) = T(\mathbf{u}) + rT(\mathbf{v})$. 不過依 T 的定義 $T(\mathbf{u}) = A\mathbf{u}$, $T(\mathbf{v}) = A\mathbf{v}$, 而 $T(\mathbf{u} + r\mathbf{v}) = A(\mathbf{u} + r\mathbf{v})$. 故依矩陣乘法加法的分配律 (Proposition 2.1.9 和 Proposition 2.1.10) 我們得

$$T(\mathbf{u} + r\mathbf{v}) = A(\mathbf{u} + r\mathbf{v}) = A\mathbf{u} + A(r\mathbf{v}) = A\mathbf{u} + rA\mathbf{v} = T(\mathbf{u}) + rT(\mathbf{v}).$$

□

從 Lemma 6.1.5 我們知道可以造出許多的 linear transformations. 事實上, 我們可以利用現有的 linear transformations 造出更多的 linear transformations. 首先若 T_1, T_2 皆為 V 到 W 的 linear transformation, 我們可以利用 T_1, T_2 造出新的 linear transformation, $T_1 + T_2$. 前面已說過, 要造出新的函數需先說明定義域和對應域是甚麼. 這裡我們定義 $T_1 + T_2$ 仍為 V 到 W 的函數. 對於任意 $\mathbf{v} \in V$, 我們定義 $(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v})$. 依此定義, $T_1 + T_2$ 確實將 V 的向量映射到 W 中. 這是因為依假設, 對任意 $\mathbf{v} \in V$, 我們有 $T_1(\mathbf{v}) \in W$ 以及 $T_2(\mathbf{v}) \in W$, 所以自然有 $(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v}) \in W$. 所以 $T_1 + T_2 : V \to W$ 確實是 well-defined function. 接下來我們要說明若 $T_1 : V \to W$, $T_2 : V \to W$ 皆為 linear transformation, 則 $T_1 + T_2 : V \to W$ 亦為 linear transformation. 也就是說對於任意 $\mathbf{u}, \mathbf{v} \in V$ 以及 $r \in \mathbb{F}$, 我們要證明 $(T_1 + T_2)(\mathbf{u} + r\mathbf{v}) = (T_1 + T_2)(\mathbf{u}) + r(T_1 + T_2)(\mathbf{v})$. 首先依定義我們有

$$(T_1 + T_2)(\mathbf{u} + r\mathbf{v}) = T_1(\mathbf{u} + r\mathbf{v}) + T_2(\mathbf{u} + r\mathbf{v}).$$

接著利用 T_1, T_2 為 linear 我們得

$$T_1(\mathbf{u} + r\mathbf{v}) + T_2(\mathbf{u} + r\mathbf{v}) = T_1(\mathbf{u}) + rT_1(\mathbf{v}) + T_2(\mathbf{u}) + rT_2(\mathbf{v}).$$

另外, 依定義

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

故由向量運算性質, 得證 $(T_1 + T_2)(\mathbf{u} + r\mathbf{v}) = (T_1 + T_2)(\mathbf{u}) + r(T_1 + T_2)(\mathbf{v})$, 亦即 $T_1 + T_2$ 為 V 到 W 的 linear transformation.

Question 6.1. *若 A_1, A_2 皆為 $m \times n$ matrix. 考慮 $T_1 : \mathbb{F}^n \to \mathbb{F}^m$, $T_2 : \mathbb{F}^n \to \mathbb{F}^m$, 分別定義為 $T_1(\mathbf{v}) = A_1\mathbf{v}$, $T_2(\mathbf{v}) = A_2\mathbf{v}$, $\forall \mathbf{v} \in \mathbb{F}^n$. 則 $T_1 + T_2$ 是怎樣的函數?*

給定一個 linear transformation $T:V\to W$, 以及 $c\in\mathbb{F}$, 我們也可定義函數 $cT:V\to W$, 其定義為 $(cT)(\mathbf{v})=c(T(\mathbf{v})), \forall\mathbf{v}\in V$ (也就是說它把每一個 V 的向量 $\mathbf{v}$ 對應到 c 倍的 $T(\mathbf{v})$). 很容易看出 $rT:V\to W$ 確實是一個 function. 事實上, 它也是 linear transformation. 這是因為對於任意 $\mathbf{u},\mathbf{v}\in V$ 以及 $r\in\mathbb{F}$, 我們有

$$(cT)(\mathbf{u}+r\mathbf{v})=c(T(\mathbf{u}+r\mathbf{v}))=cT(\mathbf{u})+rcT(\mathbf{v}).$$

而 $(cT)(\mathbf{u})+r(cT)(\mathbf{v})=c(T(\mathbf{u}))+rc(T(\mathbf{v}))$, 故知 $(cT)(\mathbf{u}+r\mathbf{v})=(cT)(\mathbf{u})+r(cT)(\mathbf{v})$, 得證 $cT:V\to W$ 為 linear transformation.

Question 6.2. 設 A 為 $m\times n$ *matrix.* 考慮 $T:\mathbb{F}^n\to\mathbb{F}^m$ 定義為 $T(\mathbf{v})=A\mathbf{v}, \forall\mathbf{v}\in\mathbb{F}^n$. 則對於 $c\in\mathbb{F}$, cT 是怎樣的函數?

設 $T_1,\cdots,T_k$ 為 V 到 W 的 linear transformations. $c_1,\ldots,c_k\in\mathbb{F}$, 則由前知 $c_1T_1,\ldots,c_kT_k$ 皆為 V 到 W 的 linear transformations. 所以 $c_1T_2+c_2T_2$ 為 linear transformation. 再利用數學歸納法, 得 $c_1T_1+\cdots+c_kT_k$ 為 linear transformation. 因此我們有下面之結果.

Proposition 6.1.6. 設 $T_1,\cdots,T_k$ 為 V 到 W 的 *linear transformations,* $c_1,\ldots,c_k\in\mathbb{F}$ 則 $c_1T_1+\cdots+c_kT_k$ 為 V 到 W 的 *linear transformation.*

另一個產生 linear transformation 的方法就是利用 "合成函數". 若 U,V,W 為 vector spaces 且 $T:U\to V$ 和 $T':V\to W$ 為函數, 由於對任意 $\mathbf{u}\in U$, 依定義 $T(\mathbf{u})\in V$, 也就是說 $T(\mathbf{u})$ 會落在 T' 的定義域中. 所以我們可以將 $T(\mathbf{u})$ 代入 T' 中, 亦即得 $T'(T(\mathbf{u}))\in W$. 這樣的方法幫我們定義出一個從 U 到 W 的函數, 稱之為 T,T' 的 *composite function* (合成函數), 我們用 $T'\circ T$ 來表示. 也就是說 $T'\circ T:U\to W$ 的定義為 $T'\circ T(\mathbf{u})=T'(T(\mathbf{u})), \forall\mathbf{u}\in U$. 我們有下面之結果.

Proposition 6.1.7. 假設 $T:U\to V$ 和 $T':V\to W$ 為 *linear transformation,* 則 $T'\circ T:U\to W$ 亦為 *linear transformation.*

Proof. 已知 $T'\circ T$ 為 function, 我們僅要證明 $T'\circ T$ 為 linear, 亦即對於任意 $\mathbf{u},\mathbf{v}\in U$ 以及 $r\in\mathbb{R}$, 我們有 $(T'\circ T)(\mathbf{u}+r\mathbf{v})=(T'\circ T)(\mathbf{u})+r(T'\circ T)(\mathbf{v})$. 依定義 $(T'\circ T)(\mathbf{u}+r\mathbf{v})=T'(T(\mathbf{u}+r\mathbf{v}))$. 然而因為 T, T' 為 linear, 故有

$$T'(T(\mathbf{u}+r\mathbf{v}))=T'(T(\mathbf{u})+rT(\mathbf{v}))=T'(T(\mathbf{u}))+rT'(T(\mathbf{v})).$$

再由 $T'(T(\mathbf{u}))=(T'\circ T)(\mathbf{u})$ 以及 $T'(T(\mathbf{v}))=(T'\circ T)(\mathbf{v})$ 得證 $T'\circ T$ 為 linear transformation. □

Question 6.3. 設 A 為 $m\times n$ *matrix,* B 為 $k\times m$ *matrix.* 考慮 $T:\mathbb{F}^n\to\mathbb{F}^m$, $T':\mathbb{F}^m\to\mathbb{F}^k$, 分別定義為 $T(\mathbf{v})=A\mathbf{v}, \forall\mathbf{v}\in\mathbb{F}^n$ 且 $T'(\mathbf{w})=B\mathbf{w}, \forall\mathbf{w}\in\mathbb{F}^m$. 則 $T'\circ T$ 是怎樣的函數?

當 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 V 的一組 basis 時, 對任意 $\mathbf{v}\in V$, 皆存在唯一的一組 $c_1,\ldots,c_n\in\mathbb{F}$, 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 現若 $T:V\to W$ 是 linear transformation, 則由 linear transformation 定義知, 此時 $T(\mathbf{v})=c_1T(\mathbf{v}_1)+\cdots+c_nT(\mathbf{v}_n)$. 也就是說, 只要我們知道 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)$ 是 W 中的哪些向量, 則對於任意 $\mathbf{v}\in V$, 我們都可以知道 $T(\mathbf{v})$ 為何. 因此我們有以下的定理.

Theorem 6.1.8. 假設 V,W 為 *vector spaces over* $\mathbb{F}$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$, 為 V 的一組 *basis.* 令 $\mathbf{w}_1,\ldots,\mathbf{w}_n\in W$, 則存在唯一的 *linear transformation* $T:V\to W$, 滿足 $T(\mathbf{v}_1)=\mathbf{w}_1,\ldots,T(\mathbf{v}_i)=\mathbf{w}_i,\ldots,T(\mathbf{v}_n)=\mathbf{w}_n$.

Proof. 首先證明存在性. 定義 $T:V\to W$ 為 $T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$, $\forall c_1,\ldots,c_n\in F$. 我們需說明這是 well-defined function. 也就是說對任意 $\mathbf{v}\in V$, $T(\mathbf{v})$ 皆有定義且 $T(\mathbf{v})\in W$. 然而因 $\mathbf{v}_1,\ldots,\mathbf{v}_n$, 為 V 的一組 basis, 故存在唯一的一組 $c_1,\ldots,c_n\in\mathbb{F}$, 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 故此時得 $T(\mathbf{v})=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n\in W$. 接著我們要說明 T 為 linear transformation, 也就是說對任意 $\mathbf{u},\mathbf{v}\in V$ 以及 $r\in\mathbb{F}$, 我們要證明 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$. 由於 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis, 存在 $c_1,\ldots,c_n$ 以及 $d_1,\ldots,d_n\in\mathbb{F}$ 使得 $\mathbf{u}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 且 $\mathbf{v}=d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n$. 故此時 $\mathbf{u}+r\mathbf{v}=(c_1+rd_1)\mathbf{v}_1+\cdots+(c_n+rd_n)\mathbf{v}_n$. 依 T 的定義得

$$T(\mathbf{u}+r\mathbf{v})=(c_1+rd_1)T(\mathbf{v}_1)+\cdots+(c_n+rd_n)T(\mathbf{v}_n)=(c_1+rd_1)\mathbf{w}_1+\cdots+(c_n+rd_n)\mathbf{w}_n.$$

另一方面

$$\begin{aligned}T(\mathbf{u})+rT(\mathbf{v})&=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)+rT(d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n)=\\&(c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n)+r(d_1\mathbf{w}_1+\cdots+d_n\mathbf{w}_n)=(c_1+rd_1)\mathbf{w}_1+\cdots+(c_n+rd_n)\mathbf{w}_n.\end{aligned}$$

得證 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$.

接著證明唯一性, 我們用反證法. 也就是說若 $T':V\to W$ 是另一個 linear transformation 滿足 $T'(\mathbf{v}_1)=\mathbf{w}_1,\ldots,T'(\mathbf{v}_n)=\mathbf{w}_n$, 且 $T'\neq T$, 則會造成矛盾. 依定義, $T'\neq T$ 表示存在 $\mathbf{v}\in V$ 使得 $T'(\mathbf{v})\neq T(\mathbf{v})$. 此時因存在 $c_1,\ldots,c_n$, 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 故依 T,T' 皆為 linear 的假設, 我們有

$$T'(\mathbf{v})=T'(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=c_1T'(\mathbf{v}_1)+\cdots+c_nT'(\mathbf{v}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n=T(\mathbf{v}).$$

此與 $T'(\mathbf{v})\neq T(\mathbf{v})$ 相矛盾, 證得唯一性. □

要注意 Theorem 6.1.8 中的 $\mathbf{w}_1,\ldots,\mathbf{w}_n\in W$ 是可以任意選取的, 不需要是一組 basis 或是 linear independent. 這個定理, 再次讓我們確定 basis 的重要性. 它告訴我們給定 V 的一組 basis 後, 我們可以將這組 basis 裡的向量對應到 W 中任意的向量, 就會得到一個 V 到 W 的 linear transformation. 更重要的是, 一般來講兩個函數要說明它們是相等的, 我們必須檢查定義域裡的每個元素是否被這兩個函數對應到對應域裡相同的元素. 這個過程是很複雜的, 因為一般來說定義域裡的元素有無窮多個, 我們無法一個一個檢查. 但是 linear transformation 就有這個好處, Theorem 6.1.8 告訴我們僅要檢查兩個 linear transformations 在一組 basis 裡中的那些有限多個向量是一致的, 那麼這兩個 linear transformation 事實上就會是相同的函數.

6.2. Range and Null Space

Linear transformation 由於有保持 linear combination 的特點, 所以它會保持定義域與對應域中的 subspaces. 在這一節中我們便是要探討一個 linear transformation 所得到

的兩個重要的 subspaces, “null space” 和 “range”, 並利用這兩個 subspace 來探討 linear transformation 本身的特點.

首先對於一般的函數 $T: V \to W$, 任取定義域 V 中的子集合 S, 我們很自然的會考慮 $T(S) = \{T(\mathbf{v}) \in W \mid \mathbf{v} \in S\}$ 這一個集合, 它就是將所有 S 中的元素利用 T 映射到 W 後的元素收集起來所得的集合. 同樣的, 對於對應域 W 中的子集合 S', 我們也會考慮 $T^{-1}(S') = \{\mathbf{v} \in V \mid T(\mathbf{v}) \in S'\}$ 這樣的集合, 它就是收集所有在定義域中會映射到 S' 的元素所成的集合. 很容易知道 $T(S)$ 會是對應域 W 中的子集合, 而 $T^{-1}(S')$ 會是定義域 V 中的子集合. 注意當 $\mathbf{w} \in T(S)$ 時, 依定義這表示 $\mathbf{w}$ 是 S 中某個元素經 T 映射所得, 亦即存在 $\mathbf{v} \in V$ 使得 $\mathbf{w} = T(\mathbf{v})$. 所以 $T(S)$ 也可表示成 $T(S) = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}), \text{ for some } \mathbf{v} \in S\}$, 有時為了強調 $T(S)$ 為 W 的子集合, 我們也會用這種表示法.

由於 linear transformation 的特色, 當 $T: V \to W$ 為 linear transformation 時, 我們專注於考慮 V' 為 V 的 subspace 時的情況. 也就是說我們要了解

$$T(V') = \{T(\mathbf{v}) \in W \mid \mathbf{v} \in V'\} = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}), \text{ for some } \mathbf{v} \in V'\}$$

的特性. 同樣的若 W' 為 W 的 subspace, 我們也要了解

$$T^{-1}(W') = \{\mathbf{v} \in V \mid T(\mathbf{v}) \in W'\}$$

的特性. 事實上, 我們有以下之結果.

Proposition 6.2.1. 假設 V, W 皆為 *vector space over* $\mathbb{F}$ 且 $T: V \to W$ 為 *linear transformation.* 若 V' 為 V 的 *subspaces,* 則 $T(V')$ 是 W 的 *subspace.* 另外, 若 W' 為 W 的 *subspaces,* 則 $T^{-1}(W')$ 是 V 的 *subspace.*

Proof. 依定義我們知 $T(V')$ 會是 W 的子集合, 而 $T^{-1}(W')$ 會是 V 的子集合. 故現僅需證明它們為 subspaces, 即利用 Corollary 3.3.3 我們要證明, 若 $\mathbf{w}, \mathbf{w}' \in T(V')$ 且 $r \in \mathbb{F}$, 則 $\mathbf{w} + r\mathbf{w}' \in T(V')$ 以及若 $\mathbf{v}, \mathbf{v}' \in T^{-1}(W')$ 且 $r \in \mathbb{F}$, 則 $\mathbf{v} + r\mathbf{v}' \in T^{-1}(W')$.

首先再強調一次, 當我們說 W 中的一個向量 $\mathbf{w}$ 在 $T(V')$ 時, 表示存在 $\mathbf{v} \in V'$ 使得 $\mathbf{w} = T(\mathbf{v})$. 因此若 $\mathbf{w}, \mathbf{w}' \in T(V')$, 則存在 $\mathbf{v}, \mathbf{v}' \in V'$ 使得 $T(\mathbf{v}) = \mathbf{w}, T(\mathbf{v}') = \mathbf{w}'$. 此時對於 $r \in \mathbb{F}$, 我們有 $\mathbf{w} + r\mathbf{w}' = T(\mathbf{v}) + rT(\mathbf{v}')$. 再利用 T 為 linear, 得 $\mathbf{w} + r\mathbf{w}' = T(\mathbf{v} + r\mathbf{v}')$. 然而依假設 V' 為 V 的 subspace, 故由 $\mathbf{v}, \mathbf{v}' \in V'$ 知 $\mathbf{v} + r\mathbf{v}' \in V'$, 得證 $\mathbf{w} + r\mathbf{w}' = T(\mathbf{v} + r\mathbf{v}') \in T(V')$.

另一方面, 若 $\mathbf{v}, \mathbf{v}' \in T^{-1}(W')$, 表示 $T(\mathbf{v}), T(\mathbf{v}') \in W'$. 此時對於 $r \in \mathbb{F}$, 由於 T 為 linear, 我們有 $T(\mathbf{v} + r\mathbf{v}') = T(\mathbf{v}) + rT(\mathbf{v}')$. 然而依假設 W' 為 W 的 subspace, 故由 $T(\mathbf{v}), T(\mathbf{v}') \in W'$ 知 $T(\mathbf{v}) + rT(\mathbf{v}') \in W'$. 因此由 $T(\mathbf{v} + r\mathbf{v}') = T(\mathbf{v}) + rT(\mathbf{v}') \in W'$, 得證 $\mathbf{v} + r\mathbf{v}' \in T^{-1}(W')$. □

特別的, 在 $V' = V$ 和 $W' = \{\mathbf{0}\}$ 這兩個特殊情況時, 即

$$T(V) = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\} \quad \text{and} \quad T^{-1}(\{\mathbf{0}\}) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

這兩個 subspaces, 對我們了解 T 這個 linear transformation 非常有幫助. 我們先給這兩個 subspace 特殊的名稱.

Definition 6.2.2. 假設 V, W 皆為 vector space over $\mathbb{F}$ 且 $T: V \to W$ 為 linear transformation.

(1) 我們稱 W 的 subspace $T(V)$ 為 T 的 *range* (也稱為 *image*). 通常我們用 $\mathsf{R}(T)$ (或 $\mathrm{im}(T)$) 來表示 T 的 range.

(2) 我們稱 V 的 subspace $T^{-1}(\{\mathbf{0}\})$ 為 T 的 *null space* (也稱為 *kernel*), 通常我們用 $\mathsf{N}(T)$ (或 $\ker(T)$) 來表示 T 的 null space.

首先我們來看 linear transformation 的 range. 假設 $T: V \to W$ 為 linear transformation. 由 Proposition 6.2.1 我們知 T 的 range $\mathsf{R}(T)$ 是 W 的 subspace, 故知 $\dim(\mathsf{R}(T)) \leq \dim(W)$. 若 $\dim(\mathsf{R}(T)) = \dim(W)$, 則依 Proposition 3.6.10 知此時 $\mathsf{R}(T) = W$. 也就是說對於任意的 $\mathbf{w} \in W$, 由於 $\mathbf{w} \in \mathsf{R}(T)$ 依定義存在 $\mathbf{v} \in V$ 使得 $\mathbf{w} = T(\mathbf{v})$. 也就是說此時 T 為 onto. 另一方面, 若 T 為 onto, 則依定義對於任意 $\mathbf{w} \in W$, 存在 $\mathbf{v} \in V$ 使得 $T(\mathbf{v}) = \mathbf{w}$ 故得 $\mathbf{w} \in \mathsf{R}(T)$, 得證 $W \subseteq \mathsf{R}(T)$. 再利用已知 $\mathsf{R}(T) \subseteq W$, 得證 $\mathsf{R}(T) = W$. 我們有以下的性質.

Proposition 6.2.3. 假設 $T: V \to W$ 為 *linear transformation.* 則 T 為 *onto* 若且唯若 $\dim(\mathsf{R}(T)) = \dim(W)$.

一般來說, 我們要判斷一個函數是否為 onto 便是要確認其 range 是否就是 codomain (對應域). 對於一般的函數要確認是否為 onto 有時並不容易. 不過 Proposition 6.2.3 告訴我們對於 linear transformation, 可以直接由它的 range 的 dimension 來判斷是否為 onto. 至於如何知道一個 linear transformation 的 range 呢? 我們有以下的性質.

Proposition 6.2.4. 假設 V, W 皆為 *vector space over* $\mathbb{F}$ 且 $T: V \to W$ 為 *linear transformation* 且 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ 為 V 的一組 *spanning vectors.* 則

$$\mathsf{R}(T) = \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)).$$

Proof. 設 $\mathbf{w} \in \mathsf{R}(T)$, 表示存在 $\mathbf{v} \in V$ 使得 $\mathbf{w} = T(\mathbf{v})$. 又因 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一組 spanning vectors, 故存在 $c_1, \cdots, c_n \in \mathbb{F}$, 使得 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. 因此利用 T 為 linear 得

$$\mathbf{w} = T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n) \in \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)),$$

得證 $\mathsf{R}(T) \subseteq \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n))$.

另一方面, 設 $\mathbf{w} \in \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n))$, 表示存在 $c_1, \cdots, c_n \in \mathbb{F}$, 使得 $\mathbf{w} = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$. 因此利用 T 為 linear 得

$$\mathbf{w} = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) \in \mathsf{R}(T),$$

得證 $\mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)) \subseteq \mathsf{R}(T)$. 因此證明了 $\mathsf{R}(T) = \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n))$. □

Example 6.2.5. (1) 考慮 $T: \mathbb{R}^3 \to \mathbb{R}^2$ 定義為 $T(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_3 \end{bmatrix}$. 在 Example 6.1.4 中我們已知 T 是一個 linear transformation. 考慮定義域 $\mathbb{R}^3$ 的 standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, 我們得 $T(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $T(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $T(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. 由於 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ 為 $\mathbb{R}^2$ 的一組

spanning vectors, 由 Proposition 6.2.4 我們有 $\mathsf{R}(T)=\mathrm{Span}(\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\-1\end{bmatrix})=\mathbb{R}^2$. 故得 T 為 onto.

(2) 考慮 $T:\mathbb{R}^2\to\mathbb{R}^3$ 定義為 $T(\begin{bmatrix}x_1\\x_2\end{bmatrix})=\begin{bmatrix}x_1\\x_1+x_2\\x_1-x_2\end{bmatrix}$. 很容易驗證 T 是一個 linear transformation. 考慮定義域 $\mathbb{R}^2$ 的 standard basis $\{\mathbf{e}_1,\mathbf{e}_2\}$, 我們得 $T(\begin{bmatrix}1\\0\end{bmatrix})=\begin{bmatrix}1\\1\\1\end{bmatrix}$, $T(\begin{bmatrix}0\\1\end{bmatrix})=\begin{bmatrix}0\\1\\-1\end{bmatrix}$. 由 Proposition 6.2.4 我們有 $\mathsf{R}(T)=\mathrm{Span}(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix})$. 由於 $\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix}$ 為 linearly independent, 故得 $\dim(\mathsf{R}(T))=2$. 由 Proposition 6.2.3 知 T 不是 onto.

Question 6.4. 假設 $T:V\to W$ 為 *linear transformation.* 若 $\dim(W)>\dim(V)$, 則 T 有可能是 *onto* 嗎?

Question 6.5. 假設 $\mathbb{F}$ 是 *field* 且 $A\in M_{m\times n}(\mathbb{F})$, 考慮 *linear transformation* $T_A:\mathbb{F}^n\to\mathbb{F}^m$ 定義為 $T_A(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$. 考慮 $\mathbb{F}^n$ 的 *standard basis* $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$. 試說明 $\mathsf{R}(T_A)$ *(*即 T_A 的 *rang)* 和 $\mathrm{Col}(A)$ *(*即 A 的 *column space)* 之間的關係.

要注意是有可能一個 linear transformation T 的 range 為 $\{\mathbf{0}\}$. 此表示 T 將所有定義域的向量皆映射到 $\mathbf{0}$, 亦即 $T(\mathbf{v})=\mathbf{0}$, $\forall\mathbf{v}\in V$. 這樣的 linear transformation, 我們依慣例, 仍稱之為 *zero mapping.*

Question 6.6. 假設 $T:V\to W$ 為 *linear transformation* 且 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 V 的一組 *basis.* 試證明 T 為 *zero mapping* 若且唯若 $T(\mathbf{v}_1)=\cdots=T(\mathbf{v}_n)=\mathbf{0}$.

接下來我們看 null space 與 linear transformation 的關係. 假設 $T:V\to W$ 為 linear transformation. 若 T 為 one-to-one, 由於已知 $T(\mathbf{0})=\mathbf{0}$, 故知不可能有非零的向量 $\mathbf{v}$ 使得 $T(\mathbf{v})=\mathbf{0}$. 因此可得 $\mathsf{N}(T)=\{\mathbf{0}\}$. 其實反過來 $\mathsf{N}(T)=\{\mathbf{0}\}$, 也會使得 T 為 one-to-one, 我們有以下的結果.

Proposition 6.2.6. 假設 $T:V\to W$ 為 *linear transformation.* 則 T 為 *one-to-one* 若且唯若 $\dim(\mathsf{N}(T))=0$, 亦即 $\mathsf{N}(T)=\{\mathbf{0}\}$.

Proof. 我們已知當 T 為 one-to-one 時, 不會有非零向量映射到 $\mathbf{0}$, 故知 $\mathsf{N}(T)=\{\mathbf{0}\}$, 得 $\dim(\mathsf{N}(T))=0$. 反之, 當 $\dim(\mathsf{N}(T))=0$, 即 $\mathsf{N}(T)=\{\mathbf{0}\}$, 此時若 T 不是 one-to-one, 表示存在 $\mathbf{v},\mathbf{v}'\in V$ 滿足 $\mathbf{v}\neq\mathbf{v}'$ 但 $T(\mathbf{v})=T(\mathbf{v}')$. 由於 T 為 linear, 得 $T(\mathbf{v}-\mathbf{v}')=T(\mathbf{v})-T(\mathbf{v}')=\mathbf{0}$, 亦即 $\mathbf{v}-\mathbf{v}'\in T^{-1}(\{\mathbf{0}\})=\mathsf{N}(T)=\{\mathbf{0}\}$. 得到 $\mathbf{v}=\mathbf{v}'$ 之矛盾, 故證得 T 為 one-to-one. □

Example 6.2.7. 我們探討 Example 6.2.5 中的 linear transformation 是否為 one-to-one.

(1) 考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix})=\begin{bmatrix} x_1+x_2\\ x_1-x_3\end{bmatrix}$. 若 $\mathbf{v}=\begin{bmatrix} a\\ b\\ c\end{bmatrix}\in\mathsf{N}(T)$, 表示 $T(\mathbf{v})=T(\begin{bmatrix} a\\ b\\ c\end{bmatrix})=\begin{bmatrix} a+b\\ a-c\end{bmatrix}=\begin{bmatrix} 0\\ 0\end{bmatrix}$, 故得 $a+b=0$ 以及 $a-c=0$, 因此可得 $\begin{bmatrix} 1\\ -1\\ 1\end{bmatrix}\in\mathsf{N}(T)$, 知 $\mathsf{N}(T)\neq\{\mathbf{0}\}$ (事實上 $\mathsf{N}(T)=\mathrm{Span}(\begin{bmatrix} 1\\ -1\\ 1\end{bmatrix})$). 所以知 T 不是 one-to-one.

(2) 考慮 $T:\mathbb{R}^2\to\mathbb{R}^3$ 定義為 $T(\begin{bmatrix} x_1\\ x_2\end{bmatrix})=\begin{bmatrix} x_1\\ x_1+x_2\\ x_1-x_2\end{bmatrix}$. 若 $\mathbf{v}=\begin{bmatrix} a\\ b\end{bmatrix}\in\mathsf{N}(T)$, 表示 $T(\mathbf{v})=T(\begin{bmatrix} a\\ b\end{bmatrix})=\begin{bmatrix} a\\ a+b\\ a-b\end{bmatrix}=\begin{bmatrix} 0\\ 0\\ 0\end{bmatrix}$. 故得 $a=0$, $a+b=0$ 以及 $a-b=0$, 即 $a=b=0$. 因此可得 $\mathsf{N}(T)=\{\mathbf{0}\}$, 所以由 Proposition 6.2.6 知 T 是 one-to-one.

要注意 Proposition 6.2.6 需 T 為 linear transformation 才適用. 例如 $f(x)=x^2$ 的情形, 雖然 $f^{-1}(0)=\{0\}$ (因為只有當 $x=0$ 才會使得 $x^2=0$) 但 $f(x)$ 不是一對一 (例如 $f(1)=f(-1)=1$). 事實上我們知道 $f(x)$ 不是 linear. 所以當 f 不是 linear transformation 時, 不能由 $f^{-1}(\{0\})$ 來判斷是否為 one-to-one.

Question 6.7. *Proposition 6.2.6* 中是哪一個部分需用到 T 為 *linear* 的假設? 是由 T 為 *one-to-one* 推得 $\mathsf{N}(T)=\{\mathbf{0}\}$ 還是由 $\mathsf{N}(T)=\{\mathbf{0}\}$ 推得 T 為 *one-to-one*?

Question 6.8. 假設 $T:V\to W$ 為 *linear transformation*. 在 *Question 6.6* 中我們知道 T 為 *zero mapping* 和 T 的 *range* 的等價關係. 你知道 T 為 *zero mapping* 和 T 的 *null space* 的等價關係嗎?

Question 6.9. 假設 $\mathbb{F}$ 是 *field* 且 $A\in M_{m\times n}(\mathbb{F})$, 考慮 *linear transformation* $T_A:\mathbb{F}^n\to\mathbb{F}^m$ 定義為 $T_A(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$. 試說明 $\mathsf{N}(T_A)$ (即 T_A 的 *null space*) 和 $N(A)$ (即 A 的 *null space*) 之間的關係.

對於一般的函數, 要判斷其是否為 onto 或是 one-to-one 並不容易, 但當 T 為 linear transformation 時, Proposition 6.2.3 和 Proposition 6.2.6 提供我們一個簡便的方法判斷 T 是否為 onto 或是 one-to-one. 也就是僅要確認 $\mathsf{R}(T)$ 和 $\mathsf{N}(T)$ 的維度即可. 既然 $\mathsf{R}(T)$ 和 $\mathsf{N}(T)$ 的維度這麼重要, 我們有以下的定義.

Definition 6.2.8. 假設 V,W 皆為 vector space over $\mathbb{F}$ 且 $T:V\to W$ 為 linear transformation.

(1) 我們稱 T 的 range $\mathsf{R}(T)$ 的維度為 T 的 *rank*, 記做 $\mathrm{rank}(T)$.

(2) 我們稱 T 的 null space $\mathsf{N}(T)$ 的維度為 T 的 *nullity*, 記做 $\mathrm{nullity}(T)$.

Question 6.10. 假設 $\mathbb{F}$ 是 *field* 且 $A\in M_{m\times n}(\mathbb{F})$, 考慮 *linear transformation* $T_A:\mathbb{F}^n\to\mathbb{F}^m$ 定義為 $T_A(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$. 試說明 $\mathrm{rank}(T_A)$ 和 $\mathrm{rank}(A)$ 以及 $\mathrm{nullity}(T_A)$ 和 $\mathrm{nullity}(A)$ 之間的關係.

在 Question 6.10 中我們看到 matrix 的 rank 和 nullity 和 linear transformation 的 rank 和 nullity 關係密切. 對於 matrix 的 rank 和 nullity 我們有所謂的 Dimension Theorem (參見 Theorem 3.7.14), 對於 linear transformation 我們也有以下的定理.

Theorem 6.2.9 (Dimension Theorem)**.** 假設 V,W 皆為 *vector space over* $\mathbb{F}$*,* 其中 V 為 *finite dimensional.* 若 $T:V\to W$ 為 *linear transformation* 則

$$\mathrm{rank}(T)+\mathrm{nullity}(T)=\dim(V).$$

Proof. 假設 $\mathbf{u}_1,\dots,\mathbf{u}_n\in\mathsf{N}(T)$, 為 T 的 null space 的一組 basis. 因為 $\mathbf{u}_1,\dots,\mathbf{u}_n$ 在 V 中且為 linearly independent, 故由 Proposition 3.6.5 知存在 $\mathbf{v}_1,\dots,\mathbf{v}_m\in V$ 使得 $\{\mathbf{u}_1,\dots,\mathbf{u}_n,\mathbf{v}_1,\dots,\mathbf{v}_m\}$ 為 V 的一組 basis. 注意此時 $\mathrm{nullity}(T)=\dim(\mathsf{N}(T))=n$ 且 $\dim(V)=m+n$. 我們要證明 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)\}$ 會是 $\mathsf{R}(T)$ 的一組 basis.

首先證明 $\mathrm{Span}(T(\mathbf{v}_1),\dots,T(\mathbf{v}_m))=\mathsf{R}(T)$. 由 Proposition 6.2.4, 我們知

$$\mathsf{R}(T)=\mathrm{Span}(T(\mathbf{u}_1),\dots,T(\mathbf{u}_n),T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)).$$

然而 $\mathbf{u}_1,\dots,\mathbf{u}_n\in\mathsf{N}(T)$, 亦即 $T(\mathbf{u}_1),\dots,T(\mathbf{u}_n)$ 皆為 W 中的零向量, 故得

$$\mathsf{R}(T)=\mathrm{Span}(T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)).$$

接下來我們證明 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)\}$ 為 linearly independent. 假設 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)\}$ 不是 linearly independent, 亦即存在 $c_1,\dots,c_m\in\mathbb{F}$ 不全為 0 使得 $c_1T(\mathbf{v}_1)+\cdots+c_mT(\mathbf{v}_m)=\mathbf{0}$. 此時由 T 為 linear transformation 知 $T(c_1\mathbf{v}_1+\cdot+c_m\mathbf{v}_m)=\mathbf{0}$, 因此得 $c_1\mathbf{v}_1+\cdot+c_m\mathbf{v}_m\in\mathsf{N}(T)$. 然而 $\{\mathbf{u}_1,\dots,\mathbf{u}_n\}$ 為 $\mathsf{N}(T)$ 的 basis, 故存在 $d_1,\dots,d_n\in\mathbb{F}$ 使得 $c_1\mathbf{v}_1+\cdot+c_m\mathbf{v}_m=d_1\mathbf{u}_1+\cdot+d_n\mathbf{u}_n$, 故得

$$d_1\mathbf{u}_1+\cdots+d_n\mathbf{u}_n-c_1\mathbf{v}_1-\cdots-c_m\mathbf{v}_m=\mathbf{0}.$$

然而 $\{\mathbf{u}_1,\dots,\mathbf{u}_n,\mathbf{v}_1,\dots,\mathbf{v}_m\}$ 為 linearly independent, 故得 $d_1=\cdots=d_m=c_1=\cdots=c_m=0$. 此與 $c_1,\dots,c_m$ 不全為 0 相矛盾, 故得證 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)\}$ 為 linearly independent.

既然 $\{T(\mathbf{v}_1),\dots,T(\mathbf{v}_m)\}$ 是 $\mathsf{R}(T)$ 的一組 basis, 我們有 $\mathrm{rank}(T)=\dim(\mathsf{R}(T))=m$, 故得證 $\mathrm{rank}(T)+\mathrm{nullity}(T)=\dim(V)$. □

前面提過對於一般的函數, 要探討是否為 onto 或是 one-to-one 並不容易. 而對於 linear transformation, 我們可以藉由求其 range 及 null space 這兩個 subspaces 來了解這些問題. 而 Dimension Theorem 告訴我們, 只要了解 range 及 null space 這兩個 subspaces 中其中一個, 就可以了解另一個了.

Question 6.11. 假設 V,W 皆為 *vector space over* $\mathbb{F}$ 且 $\dim(V)=\dim(W)$*.* 若 $T:V\to W$ 為 *linear transformation,* 證明 T 是 *one-to-one* 若且唯若 T 為 *onto.*

6.3. Matrix Representation

給定一個 matrix $A\in M_{m\times n}(\mathbb{F})$, 前面我們已知可以定義出一個 linear transformation $T:\mathbb{F}^n\to\mathbb{F}^m$, 其定義為 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$. 在這一節中, 我們要說明所有的 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformations 都可以寫成這樣的形式, 並將此概念推廣到一般 linear spaces 之間的 linear transformation. 也就是說, 我們將 linear transformation 和 matrix 相連結並推得一些重要的性質.

6.3.1. $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformations. 前面 Theorem 6.1.8 告訴我們給定一個 linear transformation, 只要知道此 linear transformation 將一組 $\mathbb{F}^n$ 的 basis 對應到哪些向量, 就可以唯一確定這一個 linear transformation. 在 $\mathbb{F}^n$ 中, 我們有一個最簡單的 basis, 即 standard basis $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$. 若 $T:\mathbb{F}^n\to\mathbb{F}^m$ 是一個 linear transformation, 由前面所述, 我們僅要知道 $T(\mathbf{e}_1),\dots,T(\mathbf{e}_n)$ 是哪些 $\mathbb{F}^m$ 的 vectors, 就可以知道任意 $\mathbf{v}\in\mathbb{F}^n$, $T(\mathbf{v})$ 為何了.

事實上對於每一個 $\mathbf{v}\in\mathbb{F}^n$, 都可以找到 $c_1,\dots,c_n\in\mathbb{F}$ 使得 $\mathbf{v}=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$, 也就是說此時 $\mathbf{v}$ 的坐標表示法為 $\mathbf{v}=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}$. 因此由 T 為 linear, 得 $T(\mathbf{v})=T(c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n)=c_1T(\mathbf{e}_1)+\cdots+c_nT(\mathbf{e}_n)$. 現若考慮 $m\times n$ matrix A, 其中 A 的 i-th column 為 $T(\mathbf{e}_i)$, 則

$$A\mathbf{v}=\begin{bmatrix}| & | & & |\\T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n)\\| & | & & |\end{bmatrix}\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}=c_1T(\mathbf{e}_1)+\cdots+c_nT(\mathbf{e}_n)=T(\mathbf{v}).$$

也就是說對任意 $\mathbf{v}\in\mathbb{F}^n$, 皆有 $T(\mathbf{v})=A\mathbf{v}$. 因此 T 就等同於將 $\mathbf{v}$ 左邊乘上 A 這一個矩陣這樣的 linear transformation. 我們有以下這一個重要的定理.

Theorem 6.3.1. 給定一個 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 *function* T. 則 T 為 *linear transformation* 若且唯若存在一個 $m\times n$ *matrix* A 使得 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$. 此 $m\times n$ *matrix* A 是唯一的, 事實上對任意 $i=1,\dots,n$, A 的 *i-th column* 為 $T(\mathbf{e}_i)$, 其中 $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ 為 $\mathbb{F}^n$ 的 *standard basis.*

Proof. 由 Lemma 6.1.5 我們知道, 若 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$, 則 T 為 linear transformation. 反之, 若 $T:\mathbb{F}^n\to\mathbb{F}^m$ 為 linear transformation, 如前面所討論的, 我們可以考慮 A 為 i-th column 為 $T(\mathbf{e}_i)$ 的 $m\times n$ matrix, 則由矩陣乘法性質知 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{F}^n$.

現若 B 為 $m\times n$ matrix 亦滿足 $T(\mathbf{v})=B\mathbf{v}$, 依矩陣乘法定義知對任意 $i=1,\dots,n$, $B\mathbf{e}_i$ 為 B 的 i-th column. 但由假設 $B\mathbf{e}_i=T(\mathbf{e}_i)$, 亦即 B 的 i-th column 為 $T(\mathbf{e}_i)$. 因此 B 的所有 column 皆與前述 A 的 column 相一致, 證得唯一性. □

簡單來說 Theorem 6.3.1 告訴我們從 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformations 和 $m\times n$ matrices 之間有一個一對一的對應關係 (注意矩陣階數與定義域, 對應域之間的關係). 由於一個 linear transformation 和其對應的 $m\times n$ matrix 關係特別密切, 我們有以下的定義.

Definition 6.3.2. 假設 $T:\mathbb{F}^n\to\mathbb{F}^m$ 為 linear transformation 且 $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ 為 $\mathbb{F}^n$ 的 standard basis. 則對於 $i=1,\dots,n$, 其 i-th column 為 $T(\mathbf{e}_i)$ 的 $m\times n$ matrix 稱為 T 的 *standard matrix representation.*

由於 T 的 standard matrix representation 是唯一的且和 T 有關, 以後我們都用 $[T]$ 來表示 T 的 standard matrix representation. 也就是說對任意 $\mathbf{v}\in\mathbb{F}^n$, 我們有 $T(\mathbf{v})=[T]\mathbf{v}$.

Example 6.3.3. 我們探討 Example 6.2.5 中的 linear transformation 其 standard matrix representation.

(1) 考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix})=\begin{bmatrix} x_1+x_2\\ x_1-x_3\end{bmatrix}$. 由於

$$T(\mathbf{e}_1)=T(\begin{bmatrix}1\\0\\0\end{bmatrix})=\begin{bmatrix}1\\1\end{bmatrix},T(\mathbf{e}_2)=T(\begin{bmatrix}0\\1\\0\end{bmatrix})=\begin{bmatrix}1\\0\end{bmatrix},T(\mathbf{e}_3)=T(\begin{bmatrix}0\\0\\1\end{bmatrix})=\begin{bmatrix}0\\-1\end{bmatrix},$$

故得

$$[T]=\begin{bmatrix} | & | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \\ | & | & | \end{bmatrix}=\begin{bmatrix} 1 & 1 & 0\\ 1 & 0 & -1\end{bmatrix}.$$

事實上我們有

$$[T]\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix} 1 & 1 & 0\\ 1 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix} x_1+x_2\\ x_1-x_3\end{bmatrix}=T(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}).$$

(2) 考慮 $T:\mathbb{R}^2\to\mathbb{R}^3$ 定義為 $T(\begin{bmatrix} x_1\\ x_2\end{bmatrix})=\begin{bmatrix} x_1\\ x_1+x_2\\ x_1-x_2\end{bmatrix}$. 由於

$$T(\mathbf{e}_1)=T(\begin{bmatrix}1\\0\end{bmatrix})=\begin{bmatrix}1\\1\\1\end{bmatrix},T(\mathbf{e}_2)=T(\begin{bmatrix}0\\1\end{bmatrix})=\begin{bmatrix}0\\1\\-1\end{bmatrix},$$

故得

$$[T]=\begin{bmatrix} | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) \\ | & | \end{bmatrix}=\begin{bmatrix} 1 & 0\\ 1 & 1\\ 1 & -1\end{bmatrix}.$$

事實上我們有

$$[T]\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix} 1 & 0\\ 1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix} x_1\\ x_1+x_2\\ x_1-x_2\end{bmatrix}=T(\begin{bmatrix}x_1\\x_2\end{bmatrix}).$$

有了 standard matrix representation, 我們就可以利用以下的定理幫助我們找出它的的 range 和 null space.

Proposition 6.3.4. 假設 $T:\mathbb{F}^n\to\mathbb{F}^m$ 為 *linear transformation* 且令 $[T]\in M_{m\times n}(\mathbb{F})$ 為其 *standard matrix representation.* 則 T 的 *range* 等於 $[T]$ 的 *column space,* 而 T 的 *null space* 等於 $[T]$ 的 *null space.*

Proof. 由於 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 為 $\mathbb{F}^n$ 的一組 basis, 由 Proposition 6.2.4 我們知 T 的 range, 即 $\mathsf{R}(T)=\mathrm{Span}(T(\mathbf{e}_1),\dots,T(\mathbf{e}_n))$. 然而 $T(\mathbf{e}_1),\dots,T(\mathbf{e}_n)$ 剛好就是 $[T]$ 的 n 個 column, 故由定義 $\mathrm{Span}(T(\mathbf{e}_1),\dots,T(\mathbf{e}_n))$ 就是 $[T]$ 的 column space. 得證 T 的 range 就是 $[T]$ 的 column space.

另一方面, 若 $\mathbf{v} \in \mathsf{N}(T)$, 表示 $T(\mathbf{v}) = \mathbf{0}$. 然而依 standard matrix representation 之定義 $T(\mathbf{v}) = [T]\mathbf{v}$, 故得 $[T]\mathbf{v} = \mathbf{0}$, 亦即 $\mathbf{v}$ 屬於 $[T]$ 的 null space. 得證 $\mathsf{N}(T)$ 包含於 $[T]$ 的 null space. 反之, 若 $\mathbf{v}$ 屬於 $[T]$ 的 null space, 表示 $[T]\mathbf{v} = \mathbf{0}$, 亦即 $T(\mathbf{v}) = \mathbf{0}$, 得證 $\mathbf{v} \in \mathsf{N}(T)$. 證明了 $[T]$ 的 null space 包含於 $\mathsf{N}(T)$, 因此 T 的 null space 等於 $[T]$ 的 null space. □

回顧一個矩陣 A 的 column space 的維度, 我們稱為 A 的 rank, 用 $\mathrm{rank}(A)$ 來表示. 而 A 的 null space 的維度稱為 A 的 nullity, 用 $\mathrm{nullity}(A)$ 來表示 (參見 Definition 3.7.13). 由 Proposition 6.3.4, 我們知道 T 的 range 的維度等於 $\mathrm{rank}([T])$, 而 T 的 null space 的維度等於 $\mathrm{nullity}([T])$, 也就是說我們有以下的結果.

Corollary 6.3.5. 假設 $T : \mathbb{F}^n \to \mathbb{F}^m$ 為 *linear transformation* 且令 $[T] \in M_{m\times n}(\mathbb{F})$ 為其 *standard matrix representation.* 則

$$\mathrm{rank}(T) = \dim(\mathsf{R}(T)) = \mathrm{rank}([T]) \quad \text{and} \quad \mathrm{nullity}(T) = \dim(\mathsf{N}(T)) = \mathrm{nullity}([T]).$$

因為這個原因一般我們也稱 T 的 range 的維度為 T 的 *rank*, 而 T 的 null space 的維度稱為 T 的 *nullity*. 這更進一步的強調了 linear transformation 以及 matrix 之間的關係. 例如我們也很容易利用 linear transformation 的 Dimension Theorem (Theorem 6.2.9) 推得矩陣的 Dimension Theorem (Theorem 3.7.14).

Example 6.3.6. 我們利用 Example 6.3.3 中的 linear transformation 及其 standard matrix representation 探討其 range 和 null space.

(1) 考慮 $T : \mathbb{R}^3 \to \mathbb{R}^2$ 定義為 $T(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_3 \end{bmatrix}$, 以及其 standard matrix representation $[T] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$. 由於 $[T]$ 的 column space 為 $\mathrm{Span}(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}) = \mathbb{R}^2$. 因此由 Proposition 6.3.4 我們有 $\mathsf{R}(T) = \mathbb{R}^2$ (此與 Example 6.2.5(1) 一致). 另一方面, $[T]$ 的 null space 為聯立方程組 $[T]\mathbf{x} = \mathbf{0}$, 即

$$\left\{\begin{array}{rrrcl} x_1 & +x_2 & & = & 0 \\ x_1 & & -x_3 & = & 0 \end{array}\right. \quad \sim \quad \left\{\begin{array}{rrrcl} x_1 & +x_2 & & = & 0 \\ & -x_2 & -x_3 & = & 0 \end{array}\right.$$

的解集合. 因此由 Proposition 6.3.4 我們有 $\mathsf{N}(T) = \mathrm{Span}(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix})$ (此與 Example 6.2.7(1) 一致).

(2) 考慮 $T : \mathbb{R}^2 \to \mathbb{R}^3$ 定義為 $T(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}$, 以及其 standard matrix representation $[T] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$. 由於 $[T]$ 的 column space 為 $\mathrm{Span}(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix})$. 因此由 Proposition 6.3.4 我們有 $\mathsf{R}(T) = \mathrm{Span}(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix})$ (此與 Example 6.2.5(2) 一致). 另一方面, $[T]$ 的 null

space 為聯立方程組 $[T]\mathbf{x}=\mathbf{0}$, 即

$$\begin{cases} x_1 & & = & 0 \\ x_1 & +x_2 & = & 0 \\ x_1 & -x_2 & = & 0 \end{cases} \sim \begin{cases} x_1 & & = & 0 \\ & x_2 & = & 0 \end{cases}$$

的解集合. 因此由 Proposition 6.3.4 我們有 $\mathsf{N}(T)=\{\begin{bmatrix}0\\0\end{bmatrix}\}=\{\mathbf{0}\}$ (此與 Example 6.2.7(2) 一致).

當 T_1,T_2 皆為 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation 時, 對任意 $c_1,c_2\in\mathbb{F}$, 我們可以利用它們得到一個新的 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation $c_1T_1+c_2T_2$ (參見 Proposition 6.1.6). 我們自然會想知道 $c_1T_1+c_2T_2$ 的 standard matrix representation 和 T_1,T_2 的 standard matrix representation 是否有關. 另外, 若 T 為 $\mathbb{F}^m$ 到 $\mathbb{F}^k$ 的 linear transformation, 我們可得合成函數 $T\circ T_1$ 為 $\mathbb{F}^n$ 到 $\mathbb{F}^k$ 的 linear transformation (參見 Proposition 6.1.7). 同樣的, 我們要探討 $T\circ T_1$ 的 standard matrix representation 和 T_1,T 的 standard matrix representation 是否有關.

Lemma 6.3.7. 設 T_1,T_2 為 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 *linear transformations,* 而 T 為 $\mathbb{F}^m$ 到 $\mathbb{F}^k$ 的 *linear transformation.* 令 $[T_1],[T_2]$ 以及 $[T]$ 分別為 T_1,T_2 和 T 的 *standard matrix representation.*

(1) 對任意 $c_1,c_2\in\mathbb{F}$, 皆有 $c_1T_1+c_2T_2:\mathbb{F}^n\to\mathbb{F}^m$ 的 *standard matrix representation* 為 $c_1[T_1]+c_2[T_2]$, 亦即
$$[c_1T_1+c_2T_2]=c_1[T_1]+c_2[T_2].$$

(2) $T\circ T_1:\mathbb{F}^n\to\mathbb{F}^k$ 的 *standard matrix representation* 為 $[T][T_1]$, 亦即
$$[T\circ T_1]=[T][T_1].$$

Proof. (1) 依定義對任意 $\mathbf{v}\in\mathbb{F}^n$, 我們有 $(c_1T_1+c_2T_2)(\mathbf{v})=c_1T_1(\mathbf{v})+c_2T_2(\mathbf{v})$. 又依 standard matrix representation 的定義 $T_1(\mathbf{v})=[T_1]\mathbf{v},T_2(\mathbf{v})=[T_2]\mathbf{v}$, 故依矩陣乘法的分配律得
$$(c_1T_1+c_2T_2)(\mathbf{v})=c_1[T_1]\mathbf{v}+c_2[T_2]\mathbf{v}=(c_1[T_1]+c_2[T_2])\mathbf{v}.$$
換言之, $c_1[T_1]+c_2[T_2]$ 是一個 $m\times n$ matrix 且滿足 $c_1T_1+c_2T_2:\mathbb{F}^n\to\mathbb{F}^m$ 的 standard matrix representation 之要求, 故由 standard matrix representation 的唯一性 (Theorem 6.3.1) 知 $[c_1T_1+c_2T_2]=c_1[T_1]+c_2[T_2]$.

(2) 依定義對任意 $\mathbf{v}\in\mathbb{F}^n$, 我們有 $(T\circ T_1)(\mathbf{v})=T(T_1(\mathbf{v}))$. 又依 standard matrix representation 的定義 $T_1(\mathbf{v})=[T_1]\mathbf{v}$, 故得 $(T\circ T_1)(\mathbf{v})=T([T_1]\mathbf{v})$. 又依定義, 對任意 $\mathbf{w}\in\mathbb{F}^m$ 皆有 $T(\mathbf{w})=[T]\mathbf{w}$, 故得 $(T\circ T_1)(\mathbf{v})=T([T_1]\mathbf{v})=[T]([T_1]\mathbf{v})$. 再依矩陣乘法的結合律得 $[T]([T_1]\mathbf{v})=([T][T_1])\mathbf{v}$. 換言之, $[T][T_1]$ 是一個 $k\times n$ matrix 且滿足 $T\circ T_1:\mathbb{F}^n\to\mathbb{F}^k$ 的 standard matrix representation 之要求 $(T\circ T_1)(\mathbf{v})=([T][T_1])\mathbf{v}$, 故由 standard matrix representation 的唯一性 (Theorem 6.3.1) 知 $[T\circ T_1]=[T][T_1]$. □

Example 6.3.8. 我們利用 Example 6.3.3 中的 linear transformations 及其 standard matrix representations 探討它們的合成函數.

考慮 $T:\mathbb{R}^3\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix})=\begin{bmatrix} x_1+x_2\\ x_1-x_3\end{bmatrix}$. 我們知 T 的 standard matrix representation 為 $[T]=\begin{bmatrix} 1 & 1 & 0\\ 1 & 0 & -1\end{bmatrix}$. 另外考慮 $T':\mathbb{R}^2\to\mathbb{R}^3$ 定義為 $T'(\begin{bmatrix} x_1\\ x_2\end{bmatrix})=\begin{bmatrix} x_1\\ x_1+x_2\\ x_1-x_2\end{bmatrix}$. 我們知 T' 的 standard matrix representation 為 $[T']=\begin{bmatrix} 1 & 0\\ 1 & 1\\ 1 & -1\end{bmatrix}$. 依合成函數定義 $T'\circ T:\mathbb{R}^3\to\mathbb{R}^3$ 滿足

$$(T'\circ T)(\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix})=T'(\begin{bmatrix} x_1+x_2\\ x_1-x_3\end{bmatrix})=\begin{bmatrix} x_1+x_2\\ (x_1+x_2)+(x_1-x_3)\\ (x_1+x_2)-(x_1-x_3)\end{bmatrix}=\begin{bmatrix} x_1+x_2\\ 2x_1+x_2-x_3\\ x_2+x_3\end{bmatrix}.$$

依此結果, 我們得 $T'\circ T$ 的 standard matrix representation 為 $[T'\circ T]=\begin{bmatrix} 1 & 1 & 0\\ 2 & 1 & -1\\ 0 & 1 & 1\end{bmatrix}$. 另一方面, 考慮矩陣乘法, 我們有 $[T'][T]=\begin{bmatrix} 1 & 0\\ 1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix} 1 & 1 & 0\\ 1 & 0 & -1\end{bmatrix}=\begin{bmatrix} 1 & 1 & 0\\ 2 & 1 & -1\\ 0 & 1 & 1\end{bmatrix}$. 的確得到 $[T'\circ T]=[T'][T]$.

我們利用 matrix 來幫助我們了解 linear transformation. 反過來, 我們也可以利用 linear transformation 來幫助我們了解 matrix 的性質. 例如當 $T:\mathbb{F}^n\to\mathbb{F}^m$, $T':\mathbb{F}^m\to\mathbb{F}^k$ 為 linear transformations. 由於 T 的 range 是 $\mathbb{F}^m$ 的 subspace, 即 $\mathsf{R}(T)=T(\mathbb{R}^n)\subseteq\mathbb{R}^m$, 我們有 $(T'\circ T)(\mathbb{R}^n)=T'(T(\mathbb{R}^n))\subseteq T'(\mathbb{R}^m)$. 也就是說 $T'\circ T:\mathbb{R}^n\to\mathbb{R}^k$ 這一個 linear transformation 的 range 是包含於 T' 的 range, 即 $\mathsf{R}(T'\circ T)\subseteq\mathsf{R}(T')$. 利用 subspace 之間 dimension 的關係 (Proposition 3.6.10(4)), 我們得 $\mathrm{rank}(T'\circ T)\le\mathrm{rank}(T')$. 因此當 $A\in M_{m\times n}(\mathbb{F})$, $B\in M_{k\times m}(\mathbb{F})$, 我們可以推得 $\mathrm{Col}(BA)\subseteq\mathrm{Col}(B)$ 且 $\mathrm{rank}(BA)\le\mathrm{rank}(B)$ (Proposition 3.7.16(1)).

Question 6.12. 假設 $T:\mathbb{F}^n\to\mathbb{F}^m$, $T':\mathbb{F}^m\to\mathbb{F}^k$ 為 *linear transformations.* 證明 $\mathsf{N}(T)\subseteq\mathsf{N}(T'\circ T)$, 並依此證明若 $A\in M_{m\times n}(\mathbb{F})$, $B\in M_{k\times m}(\mathbb{F})$, 則 $\mathsf{N}(A)\subseteq\mathsf{N}(BA)$ 且 $\mathrm{rank}(BA)\le\mathrm{rank}(A)$.

6.3.2. Coordinatization. 我們將介紹一種很重要的 linear transformation, 就是將一個 vector space 裡的元素坐標化. 利用坐標化我們可以將抽象的 vector space 的問題, 化成具體的 $\mathbb{F}^n$ 空間的問題處理.

假設 V 是 finite dimensional vector space, 選定 V 的一組 basis, 我們可以將此組 basis 裡的元素排序, 並固定這個順序不變, 那麼這樣的一組有順序的 basis, 我們稱之為 *ordered basis* (有序基底). 這裡要特別強調, 即使 basis 裡的元素相同但排序不同, 我們也視為相異的 ordered basis. 所以一般在談論 ordered basis 時, 我們會用 $(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 來表示, 以強調其順序. 舉例來說 $\left(\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right)$ 和 $\left(\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix}\right)$ 就是 $\mathbb{R}^2$ 中兩組不同的 ordered basis.

有時為了方便起見, 給了一組 ordered basis 後, 我們會用一個符號來表示這一組 ordered basis. 例如給定 $\beta=(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 為 V 的一組 ordered basis, 我們就會 β 來表示這一組

ordered basis $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$. 對於 $\mathbb{F}^n$ 的 standard basis, 我們通常用 ε 來表示 $(\mathbf{e}_1,\ldots,\mathbf{e}_n)$ 這一組 ordered basis.

有了 vector space V 的一組 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 後, 我們就可以將 V 中的元素 "坐標化" (*coordinatization*). 意思就是對任意 $\mathbf{v}\in V$, 我們利用 β 這一組 ordered basis 將 $\mathbf{v}$ 寫成 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 後, $\begin{bmatrix} c_1\\ \vdots\\ c_n\end{bmatrix}$ 就是利用 β 將 $\mathbf{v}$ 坐標化後所得的坐標表示法. 為了方便, 我們就用 $[\mathbf{v}]_\beta$ 來表示利用 β 將 $\mathbf{v}$ 坐標化後所得的坐標. 坐標化的好處是, 我們可以將 $[\mathbf{v}]_\beta$ 看成是 $\mathbb{F}^n$ 中的一個向量. 這樣我們就可以將較抽象的 vector space 中的元素, 看成是 $\mathbb{F}^n$ 中的向量來處理.

Example 6.3.9. 我們看看前面提過的幾個 vector space 坐標化的情形.

(A) 考慮 $M_{2\times 2}(\mathbb{F})$ 及其 ordered basis

$$\varepsilon=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$$

(通常我們稱這一組 basis 為 $M_{2\times 2}(\mathbb{F})$ 的 standard basis). 對於任意 $M_{2\times 2}(\mathbb{F})$ 中的元素 $\begin{bmatrix}a&b\\c&d\end{bmatrix}$, 由於

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}=a\begin{bmatrix}1&0\\0&0\end{bmatrix}+b\begin{bmatrix}0&1\\0&0\end{bmatrix}+c\begin{bmatrix}0&0\\1&0\end{bmatrix}+d\begin{bmatrix}0&0\\0&1\end{bmatrix},$$

我們得 $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ 利用 ε 所得的坐標表示法為

$$\left[\begin{bmatrix}a&b\\c&d\end{bmatrix}\right]_\varepsilon=\begin{bmatrix}a\\b\\c\\d\end{bmatrix}.$$

例如在 $M_{2\times 2}(\mathbb{R})$, 我們有

$$\left[\begin{bmatrix}1&-2\\-3&4\end{bmatrix}\right]_\varepsilon=\begin{bmatrix}1\\-2\\-3\\4\end{bmatrix}.$$

(B) 在 $P_2(\mathbb{F})$ 中通常我們會稱 $1,x,x^2$ 這組 basis 為 standard basis. 考慮 $\varepsilon=(1,x,x^2)$ 這組 ordered basis. 很容易看出在 $P_2(\mathbb{R})$ 中, $2x^2-3x+4$ 用 ε 所得的坐標為 $\begin{bmatrix}4\\-3\\2\end{bmatrix}$, 所以我們有

$$[2x^2-3x+4]_\varepsilon=\begin{bmatrix}4\\-3\\2\end{bmatrix}.$$

我們也可考慮 ordered basis $\beta=(p_1(x),p_2(x),p_3(x))$ 其中

$$p_1(x)=-(x-1)(x+1),\quad p_2(x)=(1/2)x(x+1)\quad\text{and}\quad p_3(x)=(1/2)x(x-1)$$

(參見 Example 3.6.11). 由於

$p_1(0)=1, p_1(1)=p_1(-1)=0; p_2(1)=1, p_2(0)=p_2(-1)=0; p_3(-1)=1, p_3(0)=p_3(1)=0,$

若 $2x^2-3x+4=c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$, 則分別代 $x=0,1,-1$, 可得 $c_1=4, c_2=3, c_3=9$. 故

$$[2x^2-3x+4]_\beta=\begin{bmatrix}4\\3\\9\end{bmatrix}.$$

(C) 我們也可將 $\mathbb{F}^n$ 中的向量用不同的 ordered basis 坐標化. 例如在 $\mathbb{R}^3$ 中考慮 ordered basis $\beta=(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix})$. 若要求向量 $\begin{bmatrix}1\\2\\3\end{bmatrix}$ 以 β 為 ordered basis 的坐標表示, 我們要求出 $c_1,c_2,c_3\in\mathbb{R}$ 滿足

$$\begin{bmatrix}1\\2\\3\end{bmatrix}=c_1\begin{bmatrix}1\\1\\1\end{bmatrix}+c_2\begin{bmatrix}0\\1\\1\end{bmatrix}+c_3\begin{bmatrix}1\\0\\1\end{bmatrix}=\begin{bmatrix}1&0&1\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}.$$

解聯立方程組得, $c_1=0, c_2=2, c_3=1$, 故得

$$\begin{bmatrix}1\\2\\3\end{bmatrix}_\beta=\begin{bmatrix}0\\2\\1\end{bmatrix}.$$

要注意這裡我們有 $\begin{bmatrix}1\\2\\3\end{bmatrix}_\varepsilon=\begin{bmatrix}1\\2\\3\end{bmatrix}$, 其中 ε 為 $\mathbb{R}^3$ 的 standard ordered basis $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$. 這是因為我們原來就是用 standard ordered basis ε 來將所有 $\mathbb{R}^3$ 的向量的坐標化.

給定 V 的一組 ordered basis $\beta=(\mathbf{v}_1,\dots,\mathbf{v}_n)$, 將 V 中的元素利用 β 坐標化, 其實就定出了一個從 V 到 $\mathbb{F}^n$ 的函數 $T_\beta:V\to\mathbb{F}^n$, 其中對任意 $\mathbf{v}\in V$, 我們有 $T_\beta(\mathbf{v})=[\mathbf{v}]_\beta$. 為什麼這是一個函數呢? 因為 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 V 的一組 basis, 所以由 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 是 V 的 spanning set, 可得任意的 $\mathbf{v}\in V$ 確實都存在 $c_1,\dots,c_n\in\mathbb{F}$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 所以 T_β 確實可以將每個定義域中的元素 $\mathbf{v}$ 對應到對應域 $\mathbb{F}^n$ 中的向量 $\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$. 而且這個對應關係是 well-defined, 也就是說不會有將同一個 $\mathbf{v}$ 對應到 $\mathbb{F}^n$ 中兩個不同向量的情況. 這是因為 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 linearly independent, 所以每個 $\mathbf{v}\in V$, 僅有一組 $c_1,\dots,c_n\in\mathbb{F}$ 會使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$.

既然 $T_\beta:V\to\mathbb{F}^n$ 是一個 well-defined 的函數, 那它會是 linear transformation 嗎? 答案是肯定的. 考慮 $\mathbf{v},\mathbf{w}\in V$ 且假設 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, $\mathbf{w}=d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n$, 其中 $c_1,\dots,c_n$ 與 $d_1,\dots,d_n$ 皆屬於 $\mathbb{F}$. 依定義我們有

$$T_\beta(\mathbf{v})=[\mathbf{v}]_\beta=\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}\quad\text{and}\quad T_\beta(\mathbf{w})=[\mathbf{w}]_\beta=\begin{bmatrix}d_1\\\vdots\\d_n\end{bmatrix}.$$

對於任意 $r\in\mathbb{F}$, 由於

$$\mathbf{v}+r\mathbf{w}=(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)+r(d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n)=(c_1+rd_1)\mathbf{v}_1+\cdots+(c_n+rd_n)\mathbf{v}_n,$$

我們有

$$T_\beta(\mathbf{v}+r\mathbf{w})=[\mathbf{v}+r\mathbf{w}]_\beta=\begin{bmatrix}c_1+rd_1\\ \vdots\\ c_n+rd_n\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}+r\begin{bmatrix}d_1\\ \vdots\\ d_n\end{bmatrix}=T_\beta(\mathbf{v})+rT_\beta(\mathbf{w}).$$

得證 $T_\beta:V\to\mathbb{F}^n$ 為 linear transformation.

$T_\beta:V\to\mathbb{R}^n$ 不只是 linear transformation, 事實上 $T_\beta:V\to\mathbb{R}^n$ 是 one-to-one 且 onto. 要檢查 T_β 為 one-to-one, 我們僅要檢查 $\mathsf{N}(T_\beta)=\{\mathbf{0}\}$ 即可 (參見 Proposition 6.2.6). 然而若 $\mathbf{v}\in\mathsf{N}(T_\beta)$, 表示 $\mathbf{v}$ 用 β 的坐標表示法是 $\mathbb{F}^n$ 中的零向量, 亦即 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 其中 $c_1=\cdots=c_n=0$. 很自然的, 這表示 $\mathbf{v}=\mathbf{0}$, 故知 $\mathsf{N}(T_\beta)=\{\mathbf{0}\}$. 要檢查 T_β 為 onto, 我們可以利用 $T_\beta(\mathbf{v}_i)=\mathbf{e}_i$, $\forall i=1,\dots,n$, 故得證 $\mathsf{R}(T_\beta)=\mathrm{Span}(T_\beta(\mathbf{v}_1),\dots,T_\beta(\mathbf{v}_n))=\mathrm{Span}(\mathbf{e}_1,\dots,\mathbf{e}_n)=\mathbb{F}^n$ (參見 Proposition 6.2.4), 即 T_β 為 onto. 我們證得了以下的定理.

Theorem 6.3.10. 假設 V 為 *vector space over* $\mathbb{F}$, $\dim(V)=n$ 且 β 為 V 的一組 *ordered basis.* 考慮 $T_\beta:V\to\mathbb{F}^n$ 定義為 $T_\beta(\mathbf{v})=[\mathbf{v}]_\beta$, $\forall\mathbf{v}\in V$. 則 T_β 為 *linear transformation* 且是 *one-to-one* 以及 *onto.*

一般來說當一個 linear transformation $T:V\to W$ 是 one-to-one 且 onto 時, 為了方便以及強調其特殊性, 我們會稱 T 為一個 *isomorphism.* 知道 $T_\beta:V\to\mathbb{F}^n$ 為 isomorphism 的好處就是, 以後我們要探討 V 中元素的性質, 我們可以利用 T_β, 將問題轉換成大家熟悉的 $\mathbb{F}^n$ 中的向量的性質. 例如我們要判斷 V 中的元素 $\mathbf{w}_1,\dots,\mathbf{w}_k$ 是否為 linearly independent. 我們可以先找到一組 V 的 ordered basis β, 然後考慮 $[\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_k]_\beta$, 這一組 $\mathbb{F}^n$ 中的向量. 利用我們熟悉的判斷 $\mathbb{F}^n$ 中向量是否為 linearly independent 的方法判斷 $[\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_k]_\beta$ 是否為 linearly independent. 由於對於 $i=1,\dots,k$, $[\mathbf{w}_i]_\beta=T_\beta(\mathbf{w}_i)$, 因此由 T_β 為 isomorphism 以後我們會知道 $\mathbf{w}_1,\dots,\mathbf{w}_k$ 為 linearly independent 若且唯若 $[\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_k]_\beta$ 為 linearly independent (參見 Proposition 6.4.4). 因此我們可以由 $[\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_k]_\beta$ 是否為 linearly independent, 來決定 $\mathbf{w}_1,\dots,\mathbf{w}_k$ 是否為 linearly independent. 我們看以下的例子.

Example 6.3.11. 我們看利用坐標化來處理一般 vector space 是否 linear independent 的問題.

(A) 考慮 $P_2(\mathbb{R})$ 中 3 個非零多項式 $f_2(x),f_1(x),f_0(x)$, 其次數分別為為 2, 1, 0 的多項式. 假設 $f_2(x)=ax^2+bx+c$, $f_1(x)=dx+e$, $f_0(x)=r$ 其中 a,d,r 皆不等於 0. 我們利用 $P_2(\mathbb{R})$ 的 standard ordered basis $\varepsilon=(1,x,x^2)$, 可得

$$[f_2(x)]_\varepsilon=\begin{bmatrix}c\\ b\\ a\end{bmatrix},\quad [f_1(x)]_\varepsilon=\begin{bmatrix}e\\ d\\ 0\end{bmatrix}\quad\text{and}\quad [f_0(x)]_\varepsilon=\begin{bmatrix}r\\ 0\\ 0\end{bmatrix}.$$

由於 a,d,r 皆不等於 0, 很容易看出矩陣

$$\begin{bmatrix}c&e&r\\ b&d&0\\ a&0&0\end{bmatrix}$$

的 rank 為 3, 亦即 $[f_2(x)]_\varepsilon,[f_1(x)]_\varepsilon,[f_0(x)]_\varepsilon$ 為 linearly independet. 因此得證 $f_2(x)$, $f_1(x)$, $f_0(x)$ 為 linearly independent. 再由 $\dim(P_2(\mathbb{R}))=3$, 得證 $f_0(x),f_1(x),f_2(x)$ 為 $P_2(\mathbb{R})$ 的一組 basis.

我們可以將這個結果推廣到 $P_n(\mathbb{R})$. 也就是說考慮 $P_n(\mathbb{R})$ 中 $n+1$ 個非零多項式 $f_0(x),\dots,f_n(x)$, 其中對於 $i=0,\dots,n$, $f_i(x)$ 是次數為 i 的多項式. 利用對 standard ordered basis $\varepsilon=(1,x,\dots,x^n)$ 坐標化, 我們可得 $f_0(x),\dots,f_n(x)$ 為 $P_n(\mathbb{R})$ 的一組 basis.

(B) 假設 V 為 vector space over $\mathbb{R}$, 且 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4\in V$ 為 linearly independent. 令

$$\begin{array}{rrrrr}
\mathbf{w}_1= & \mathbf{v}_1 & -\mathbf{v}_2 & +2\mathbf{v}_3 & +\mathbf{v}_4\\
\mathbf{w}_2= & 2\mathbf{v}_1 & -2\mathbf{v}_2 & +4\mathbf{v}_3 & +2\mathbf{v}_4\\
\mathbf{w}_3= & \mathbf{v}_1 & & +\mathbf{v}_3 & +\mathbf{v}_4\\
\mathbf{w}_4= & & \mathbf{v}_2 & & +\mathbf{v}_4\\
\mathbf{w}_5= & -\mathbf{v}_1 & +\mathbf{v}_2 & -\mathbf{v}_3 &
\end{array}.$$

我們要找出 $W=\operatorname{Span}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4,\mathbf{w}_5)$ 的一組 basis.

考慮 $U=\operatorname{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4)$, 因為 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ 為 linearly independent, 我們知 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ 為 U 的一組 basis. 由於 $\mathbf{w}_1,\dots,\mathbf{w}_5\in U$, 我們知 W 為 U 的 subspace. 我們的想法是利用 $\beta=(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4)$ 這組 U 的 ordered basis 將 U 的元素坐標化. 由於 $\mathbf{w}_1,\dots,\mathbf{w}_5\in U$, 我們可以將 $\mathbf{w}_1,\dots,\mathbf{w}_5$ 坐標化, 得 $[\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_5]_\beta$ 這 5 個 $\mathbb{R}^4$ 中的向量. 利用過去我們知道求 $\mathbb{R}^4$ 中 $\operatorname{Span}([\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_5]_\beta)$ 的 basis 的方法求出一組 basis. 再將它們還原成 U 中的元素, 就得到 W 的一組 basis.

現由於

$$[\mathbf{w}_1]_\beta=\begin{bmatrix}1\\-1\\2\\1\end{bmatrix},[\mathbf{w}_2]_\beta=\begin{bmatrix}2\\-2\\4\\2\end{bmatrix},[\mathbf{w}_3]_\beta=\begin{bmatrix}1\\0\\1\\1\end{bmatrix},[\mathbf{w}_4]_\beta=\begin{bmatrix}0\\1\\0\\1\end{bmatrix},[\mathbf{w}_5]_\beta=\begin{bmatrix}-1\\1\\-1\\0\end{bmatrix}$$

我們考慮以它們為 column 的 4×5 matrix 並利用 elementary row operations 將之化為 echelon form 得

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

由於 echelon form 的 1-st, 3-rd, 4-th column 為 pivot 所在位置, 故知 $[\mathbf{w}_1]_\beta,[\mathbf{w}_3]_\beta,[\mathbf{w}_4]_\beta$ 為 $\operatorname{Span}([\mathbf{w}_1]_\beta,\dots,[\mathbf{w}_5]_\beta)$ 的一組 basis (參見 Proposition 3.7.8). 由於 T_β 為 isomprphism 保持 spanning set 以及 linearly independent 的性質, 我們得 $\mathbf{w}_1,\mathbf{w}_3,\mathbf{w}_4$ 為 W 的一組 basis.

Question 6.13. 考慮 *Example 6.3.11 (B)* 中的 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ 以及 $\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4,\mathbf{w}_5$. 試問 $\dim(\operatorname{Span}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4,\mathbf{w}_5))$ 為何? 並將 $\mathbf{w}_2,\mathbf{w}_5$ 寫成 $\mathbf{w}_1,\mathbf{w}_3,\mathbf{w}_4$ 的 *linear combination.*

6.3.3. Matrix Representation of general linear transformation. 當 V,W 分別為 dimension 為 n,m 的 vector space over $\mathbb{F}$. 我們可以透過 V,W 的 ordered basis, 將 V,W 的元素轉換成 $\mathbb{F}^n$ 和 $\mathbb{F}^m$ 的 vector. 因此我們可以將 V 到 W 的 linear transformation T 視為 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation, 而談論 T 的 matrix representation.

分別給定 V,W 的一組 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 以及 $\gamma=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$. 令 $T_\beta:V\to\mathbb{F}^n$ 與 $T_\gamma:W\to\mathbb{F}^m$ 分別為利用 β 以及 γ 將 V, W 的元素坐標化的 linear transformation. 回顧一下, 這裡 T_β,T_γ 皆為 isomorphism. 由於 isomorphism 是 one-to-one 且 onto, 故其反函數是存在的, 且將來我們會證明此反函數仍為 linear transformation (參見 Theorem 6.4.3). 現對於任意 V 到 W 的 linear transformation T, 我們考慮合成函數 $T_\gamma\circ T\circ T_\beta^{-1}:\mathbb{F}^n\to\mathbb{F}^m$. 由於 T_γ, T 以及 T_β^{-1} 皆為 linear transformation, 所以 $T_\gamma\circ T\circ T_\beta^{-1}$ 是一個 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation (Proposition 6.1.7). 因此 $T_\gamma\circ T\circ T_\beta^{-1}$ 有一個 standard matrix representation, 我們定義這個 standard matrix representation 為 T 相對於 β, γ 這兩組 ordered basis 所得的 *matrix representation*, 並用 $[T]_\beta^\gamma$ 來表示. 到底 $[T]_\beta^\gamma$ 是怎樣的矩陣呢? 依定義它是一個 $m\times n$ matrix, 且對於 $i=1,\ldots,n$, $[T]_\beta^\gamma$ 的 i-th column 應為 $T_\gamma\circ T\circ T_\beta^{-1}(\mathbf{e}_i)$. 由於 $T_\beta(\mathbf{v}_i)=\mathbf{e}_i$, 所以 $T_\beta^{-1}(\mathbf{e}_i)=\mathbf{v}_i$. 因此得

$$T_\gamma\circ T\circ T_\beta^{-1}(\mathbf{e}_i)=T_\gamma(T(T_\beta^{-1}(\mathbf{e}_i)))=T_\gamma(T(\mathbf{v}_i))=[T(\mathbf{v}_i)]_\gamma.$$

也就是說 $[T]_\beta^\gamma$ 的 i-th column 就是將 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 的第 i 個元素 $\mathbf{v}_i$ 代入 T 中所得的 $T(\mathbf{v}_i)\in W$, 再利用 γ 將其坐標化所得的 $\mathbb{R}^m$ 中的向量. 我們大致上有以下的表示法

$$[T]_\beta^\gamma=\begin{bmatrix} | & | & & | \\ [T(\mathbf{v}_1)]_\gamma & [T(\mathbf{v}_2)]_\gamma & \cdots & [T(\mathbf{v}_n)]_\gamma \\ | & | & & | \end{bmatrix}.$$

Example 6.3.12. 考慮假設 V 為 finite dimensional vector space over $\mathbb{F}$. 考慮 V 上的 identity map $\mathrm{id}:V\to V$, 亦即對任意 $\mathbf{v}\in V$, 我們定義 $\mathrm{id}(\mathbf{v})=\mathbf{v}$. 很容易看出 id 是一個 linear transformation. 現對任意 V 的 ordered basis $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 我們想知道 $\mathrm{id}:V\to V$ 對定義域和對應域都用 β 這個 ordered basis 所得的 matrix representation $[\mathrm{id}]_\beta^\beta$ 為何? 依照前面的討論, 我們知道 $[\mathrm{id}]_\beta^\beta$ 的 1-st column 就是 $\mathrm{id}(\mathbf{v}_1)$ 利用 β 坐標化所得的 $\mathbb{F}^n$ 中的向量. 由於 $\mathrm{id}(\mathbf{v}_1)=\mathbf{v}_1$, 而 $\mathbf{v}_1$ 又是 β 中第一個向量, 故知 $[\mathrm{id}(\mathbf{v}_1)]_\beta=[\mathbf{v}_1]_\beta=\mathbf{e}_1$, 也就是說 $[\mathrm{id}]_\beta^\beta$ 的 1-st column 就是 $\mathbf{e}_1$. 同理 $[\mathrm{id}]_\beta^\beta$ 的 i-th column 就是 $\mathbf{e}_i$. 也就是說 $[\mathrm{id}]_\beta^\beta$ 就是 $n\times n$ 的 identity matrix I_n.

Example 6.3.13. 考慮 $P_2(\mathbb{R})$ 上的 standard ordered basis $\varepsilon_2=(1,x,x^2)$ 以及 $P_3(\mathbb{R})$ 上的 standard ordered basis $\varepsilon_3=(1,x,x^2,x^3)$. 考慮函數 $T:P_2(\mathbb{R})\to P_3(\mathbb{R})$ 定義為, 對任意 $p(x)\in P_2(\mathbb{R})$, $T(p(x))=(x+1)p(x-1)$. 我們先驗證 T 為 linear transformation. 對任意 $p(x),q(x)\in P_2(\mathbb{R})$ 以及 $r\in\mathbb{R}$, 我們有

$$\begin{aligned}&T(p(x)+rq(x))=\\ &\quad(x+1)(p(x-1)+rq(x-1))=(x+1)p(x-1)+r(x+1)q(x-1)=T(p(x))+rT(q(x)).\end{aligned}$$

得證 T 為 linear transformation. 接下來我們要求 T 對於 $\varepsilon_2,\varepsilon_3$ 的 matrix representation $[T]_{\varepsilon_2}^{\varepsilon_3}$. 依照前面的探討, 我們知 $[T]_{\varepsilon_2}^{\varepsilon_3}$ 的 1-st column 應該就是將 1 代入 T, 得 $T(1)=(x+1)\cdot 1$ 再利用 ε_3 將 $x+1$ 坐標化寫成 $\mathbb{R}^4$ 的向量. 由於 $x+1=1+x+0x^2+0x^3$, 所以得 $[T]_{\varepsilon_2}^{\varepsilon_3}$ 的

1-st column 為

$$[T(1)]_{\varepsilon_3} = [x+1]_{\varepsilon_3} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

同理我們有 $[T]_{\varepsilon_2}^{\varepsilon_3}$ 的 2-nd, 3-rd column 分別為

$$[T(x)]_{\varepsilon_3} = [(x+1)(x-1)]_{\varepsilon_3} = [x^2-1] = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$[T(x^2)]_{\varepsilon_3} = [(x+1)(x-1)^2]_{\varepsilon_3} = [x^3-x^2-x+1]_{\varepsilon_3} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

因此得

$$[T]_{\varepsilon_2}^{\varepsilon_3} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

我們更換 $P_2(\mathbb{R})$ 以及 $P_3(\mathbb{R})$ 的 ordered basis. 回顧在 Example 3.6.11 中利用 Lagrange interpolation polynomial 在 $-1,0,1$ 的情形, 我們可以考慮 $P_2(\mathbb{R})$ 的一組 basis $p_1(x), p_2(x), p_3(x)$, 其中

$$\begin{array}{lll} p_1(-1)=1 & p_1(0)=0 & p_1(1)=0 \\ p_2(-1)=0 & p_2(0)=1 & p_2(1)=0 \\ p_3(-1)=0 & p_3(0)=0 & p_3(1)=1 \end{array}$$

令 $\beta = (p_1(x), p_2(x), p_3(x))$ 為 $P_2(\mathbb{R})$ 的 ordered basis. 同樣的利用 Lagrange interpolation polynomial 在 $-1,0,1,2$ 的情形我們考慮 $P_3(\mathbb{R})$ 的一組 basis $q_1(x), q_2(x), q_3(x), q_4(x)$, 其中

$$\begin{array}{llll} q_1(-1)=1 & q_1(0)=0 & q_1(1)=0 & q_1(2)=0 \\ q_2(-1)=0 & q_2(0)=1 & q_2(1)=0 & q_2(2)=0 \\ q_3(-1)=0 & q_3(0)=0 & q_3(1)=1 & q_3(2)=0 \\ q_4(-1)=0 & q_4(0)=0 & q_4(1)=0 & q_4(2)=1. \end{array}$$

令 $\gamma = (q_1(x), q_2(x), q_3(x), q_4(x))$ 為 $P_3(\mathbb{R})$ 的 ordered basis. 我們要得到 T 對於 β, γ 的 matrix representation $[T]_\beta^\gamma$. 首先 $[T]_\beta^\gamma$ 的 1-st column 為 $T(p_1(x)) = (x+1)p_1(x-1)$ 利用 γ 坐標化所得 $\mathbb{R}^4$ 的向量. 現若 $(x+1)p_1(x-1) = c_1q_1(x) + c_2q_2(x) + c_3q_3(x) + c_4q_4(x)$. 將 x 分別代 $-1,0,1,2$, 我們得到

$$c_1 = (-1+1)p_1(-2) = 0, c_2 = (0+1)p_1(-1) = 1, c_3 = (1+1)p_1(0) = 0, c_4 = (2+1)p_1(1) = 0.$$

也就是說 $[T]_\beta^\gamma$ 的 1-st column 為

$$[T(p_1(x))]_\gamma = [(x+1)p_1(x-1)]_\gamma = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

同樣的方法我們可以得到 $[T]_\beta^\gamma$ 的 2-nd, 3-rd column 分別為

$$[T(p_2(x))]_\gamma=[(x+1)p_2(x-1)]\gamma=\begin{bmatrix}0\\0\\2\\0\end{bmatrix},[T(p_3(x))]_\gamma=[(x+1)p_3(x-1)]_\gamma=\begin{bmatrix}0\\0\\0\\3\end{bmatrix}.$$

因此得

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

由 Example 6.3.13 我們知道同樣的 linear transformation 用不同的 ordered basis 會有不同的 matrix representation. 也因此要注意當要寫下 matrix representation 時一定要表明定義域和對應域的 ordered basis 為何.

到底 matrix representation 有何用處呢? 就如同在 $\mathbb{F}^n$ 上的 standard matrix representation, 利用 matrix representation 可以很快的幫我們求出 linear transformation 在定義域的每個元素的取值. 通常我們會利用圖示來幫助我們了解較複雜的函數合成問題. 給定 V,W 的 ordered basis β,γ, 以及一個 V 到 W 的 linear transformation T, 我們可以圖示如下:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, shift right, "T_\beta"'] & W \arrow[d, shift left, "T_\gamma"] \\ \mathbb{F}^n \arrow[r] \arrow[u, shift right, "T_\beta^{-1}"'] & \mathbb{F}^m \arrow[u, shift left, "T_\gamma^{-1}"] \end{tikzcd}$$

這樣的圖示一般稱為 *commutative diagram*. 它表示底下 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的函數為 $T_\gamma\circ T\circ T_\beta^{-1}$, 亦即先利用 T_β^{-1} 將 $\mathbb{F}^n$ 映射到 V, 再利用 T 將 V 映射到 W, 最後利用 T_γ 將 W 映射到 $\mathbb{F}^m$. Commutative diagram 的好處是幫助我們看出這些函數合成後如何取值. 事實上 commutative diagram 指的就是圖形上任一點到另一點若有不同的可行路徑, 經由這兩種路徑所得的結果會相同. 例如在上圖中, 從 V 到 W 有兩個路徑: 一個是直接利用 T; 另一個是從 V 先經由 T_β 到 $\mathbb{F}^n$, 接著藉由 $T_\gamma\circ T\circ T_\beta^{-1}$ 從 $\mathbb{F}^n$ 到 $\mathbb{F}^m$, 最後從 $\mathbb{F}^m$ 藉由 T_β^{-1} 到達 W. 也就是說對任意 $\mathbf{v}\in V$, 我們可以由 T 得到 $T(\mathbf{v})$. 也可先由 T_β 得到 $T_\beta(\mathbf{v})$, 接著利用 $T_\gamma\circ T\circ T_\beta^{-1}$ 將 $T_\beta(\mathbf{v})$ 送至 $T_\gamma\circ T\circ T_\beta^{-1}(T_\beta(\mathbf{v}))=T_\gamma(T(\mathbf{v}))$, 最後再利用 T_γ^{-1} 將 $T_\gamma(T(\mathbf{v}))$ 送至 $T_\gamma^{-1}(T_\gamma(T(\mathbf{v})))=T(\mathbf{v})$.

利用 $V\to\mathbb{F}^n$ 接 $\mathbb{F}^n\to\mathbb{F}^m$ 再接 $\mathbb{F}^m\to W$ 這樣的路徑來表示原本 T 從 V 到 W 這樣的路徑到底有何好處呢? 主要的原因是 $\mathbb{F}^n\to\mathbb{F}^m$ 這一段的路徑, 有 standard matrix representation, 即 $[T]_\beta^\gamma$. 也就是說對於任意 $\mathbb{F}^n$ 的向量, 我們只要左邊乘上 $[T]_\beta^\gamma$ 就可以知道會被映射到哪一個 $\mathbb{F}^m$ 的向量. 所以對任意 $\mathbf{v}\in V$, 我們可以先利用 β 將 $\mathbf{v}$ 坐標化得 $T_\beta(\mathbf{v})=[\mathbf{v}]_\beta$. 接著由於 $[\mathbf{v}]_\beta\in\mathbb{F}^n$, 故將之代入 $T_\gamma\circ T\circ T_\beta^{-1}$ 就是將 $[\mathbf{v}]_\beta$ 左邊乘上 $[T]_\beta^\gamma$ 這一個 matrix. 也就是說 $T_\gamma\circ T\circ T_\beta^{-1}([\mathbf{v}]_\beta)$ 就是 $[T]_\beta^\gamma[\mathbf{v}]_\beta$. 最後再利用 W 的 ordered basis γ 將 $[T]_\beta^\gamma[\mathbf{v}]_\beta$ 這個 $\mathbb{F}^m$ 中的向量還原回 W 的元素 $T_\gamma^{-1}([T]_\beta^\gamma[\mathbf{v}]_\beta)$, 就是 $T(\mathbf{v})$ 之值. 因此我們有以下之結果.

Proposition 6.3.14. 假設 V,W 為 *vector space over* $\mathbb{F}$ 且 $\beta=(\mathbf{v}_1,\dots,\mathbf{v}_n)$, $\gamma=(\mathbf{w}_1,\dots,\mathbf{w}_m)$ 分別為 V,W 的 *ordered basis.* 設 $T:V\to W$ 為 *linear transformation* 且 $[T]_\beta^\gamma\in M_{m\times n}(\mathbb{F})$ 為

T 相對於 β,γ 的 *matrix representation.* 對於任意 $\mathbf{v}\in V$, 若 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 且

$$[T]_\beta^\gamma\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}=\begin{bmatrix}d_1\\ \vdots\\ d_m\end{bmatrix},$$

則 $T(\mathbf{v})=d_1\mathbf{w}_1+\cdots+d_m\mathbf{w}_m$. 亦即

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

Proof. 因 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 依定義 $\mathbf{v}$ 利用 β 坐標化所得 $\mathbb{F}^n$ 的向量 $T_\beta(\mathbf{v})$ 為 $\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$. 故由前面所述 $T_\gamma\circ T\circ T_\beta^{-1}(T_\beta(\mathbf{v}))=T_\gamma(T(\mathbf{v}))$ 就是將 $\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$ 左邊乘上 $[T]_\beta^\gamma$ 所得的 $\mathbb{F}^m$ 中向量 $\begin{bmatrix}d_1\\ \vdots\\ d_m\end{bmatrix}$. 故由 $T_\gamma(T(\mathbf{v}))=\begin{bmatrix}d_1\\ \vdots\\ d_m\end{bmatrix}$ 可得 $T(\mathbf{v})=d_1\mathbf{w}_1+\cdots+d_m\mathbf{w}_m$. □

Example 6.3.15. 我們考慮 Example 6.3.13 的例子, 即考慮 $T:P_2(\mathbb{R})\to P_3(\mathbb{R})$, 其中對於任意 $p(x)\in P_2(\mathbb{R})$, $T(p(x))=(x+1)p(x-1)$. 考慮 $p(x)=x^2-1$ 的情形. 因 $p(x-1)=(x-1)^2-1$, 依 T 的定義得

$$T(p(x))=(x+1)p(x-1)=(x+1)((x-1)^2-1)=x^3-x^2-2x.$$

當考慮 $P_2(\mathbb{R})$ 的 ordered basis $\varepsilon_2=(1,x,x^2)$ 以及 $P_3(\mathbb{R})$ 的 ordered basis $\varepsilon_3=(1,x,x^2,x^3)$, 我們知道 T 對於 $\varepsilon_2,\varepsilon_3$ 的 matrix representation 為 $[T]_{\varepsilon_2}^{\varepsilon_3}=\begin{bmatrix}1&-1&1\\1&0&-1\\0&1&-1\\0&0&1\end{bmatrix}$ 故由 $[p(x)]_{\varepsilon_2}=[x^2-1]_{\varepsilon_2}=\begin{bmatrix}-1\\0\\1\end{bmatrix}$ 得

$$[T(p(x))]_{\varepsilon_3}=\begin{bmatrix}1&-1&1\\1&0&-1\\0&1&-1\\0&0&1\end{bmatrix}\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0\\-2\\-1\\1\end{bmatrix},$$

即 $T(p(x))=x^3-x^2-2x$.

當然我們也可以利用 Example 6.3.13 中 V 的 ordered basis $\beta=(p_1(x),p_2(x),p_3(x))$ 以及 W 的 ordered basis $\gamma=(q_1(x),q_2(x),q_3(x),q_4(x))$ 求出 $T(x^2-1)$. 此時若 $p(x)=x^2-1=c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$, 則代 $x=-1,0,1$ 得 $c_1=0$, $c_2=-1$, $c_3=0$, 亦即 $[p(x)]_\beta=\begin{bmatrix}0\\-1\\0\end{bmatrix}$.

故將 $[p(x)]_\beta$ 左邊乘上 T 對於 β,γ 的 representation matrix $[T]_\beta^\gamma$ 得

$$[T(p(x))]_\gamma = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \end{bmatrix}.$$

得知 $T(p(x)) = -2q_3(x)$. 由於 $q_3(-1) = q_3(0) = q_3(2) = 0$ 以及 $q_3(1) = 1$, 我們有 $q_3(x) = (x+1)x(x-2)/(-2)$, 故得 $T(p(x)) = (x+1)x(x-2) = x^3 - x^2 - 2x$.

當 $T : \mathbb{F}^n \to \mathbb{F}^m$ 是 linear transformation 時, 我們可以利用 T 的 standard matrix representation $[T]$ 的 null space 來決定 T 的 null space, 也可利用 $[T]$ 的 column space 來決定 T 的 range. 同樣的, 當 $T : V \to W$, 為 linear transformation, 我們也可利用 T 的 matrix representation 來決定 T 的 null space 和 range. 分別選定 V 和 W 的 ordered basis $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 和 $\gamma = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$. 由前面的 commutative diagram, 利用 $V \to W$ 接著 $W \to \mathbb{F}^m$ 的路徑以及 $V \to \mathbb{F}^n$ 接著 $\mathbb{F}^n \to \mathbb{F}^m$ 的路徑, 對於任意 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V$, 我們有

$$[T(\mathbf{v})]_\gamma = [T]_\beta^\gamma \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \tag{6.1}$$

現若 $\mathbf{v} \in \mathsf{N}(T)$, 表示 $T(\mathbf{v}) = \mathbf{0}$, 因此由 $[T(\mathbf{v})]_\gamma = [\mathbf{0}]_\gamma = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, 得 $[T]_\beta^\gamma \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, 亦即 $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = [\mathbf{v}]_\beta$ 為 $[T]_\beta^\gamma$ 的 null space 的向量. 反之, 若 $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$ 為 $[T]_\beta^\gamma$ 的 null space 的向量, 表示 $[T]_\beta^\gamma \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. 故由式子 (6.1) 知, 當 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ 時, 我們有 $[T(\mathbf{v})]_\gamma = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. 此即表示 $T(\mathbf{v}) = \mathbf{0}$, 得證 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in \mathsf{N}(T)$.

另一方面, 若 $\mathbf{w} = d_1\mathbf{w}_1 + \cdots + d_m\mathbf{w}_m \in \mathsf{R}(T)$, 表示存在 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V$, 使得 $T(\mathbf{v}) = \mathbf{w}$. 因此由式子 (6.1) 知, $\begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} = [T(\mathbf{v})]_\gamma = [T]_\beta^\gamma \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, 亦即 $\begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}$ 是 $[T]_\beta^\gamma$ 的 column space 的向量. 反之, 若 $\begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} \in \mathbb{F}^m$ 為 $[T]_\beta^\gamma$ 的 column space 的向量, 表示存在 $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$ 使得 $\begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} = [T]_\beta^\gamma \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$. 故由式子 (6.1) 知, 若令 $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, 我們有 $[T(\mathbf{v})]_\gamma = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}$, 得證 $d_1\mathbf{w}_1 + \cdots + d_m\mathbf{v}_m = T(\mathbf{v}) \in \mathsf{R}(T)$. 我們證得了以下的結果.

Proposition 6.3.16. 假設 V,W 為 *vector space* 且 $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $\gamma = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ 分別為 V,W 的 *ordered basis*. 設 $T : V \to W$ 為 *linear transformation* 且 $[T]_\beta^\gamma \in M_{m\times n}(\mathbb{F})$ 為 T 相

對於 β,γ 的 *matrix representation.* 則 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n\in\mathsf{N}(T)$ 若且唯若 $\begin{bmatrix} c_1\\ \vdots\\ c_n\end{bmatrix}$ 屬於 $[T]_\beta^\gamma$ 的 *null space.* 而 $d_1\mathbf{w}_1+\cdots+d_m\mathbf{w}_m\in\mathsf{R}(T)$ 若且唯若 $\begin{bmatrix} d_1\\ \vdots\\ d_m\end{bmatrix}$ 屬於 $[T]_\beta^\gamma$ 的 *column space.*

Example 6.3.17. 我們考慮 Example 6.3.13 的例子, 當考慮 $P_2(\mathbb{R})$ 的 ordered basis $\varepsilon_2=(1,x,x^2)$ 以及 $P_3(\mathbb{R})$ 的 ordered basis $\varepsilon_3=(1,x,x^2,x^3)$, 我們知道 T 對於 $\varepsilon_2,\varepsilon_3$ 的 matrix representation $T_{\varepsilon_2}^{\varepsilon_3}$ 利用 elementary row operations 化為 echelon form 可得

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

由於 pivot 的個數等於 column 的個數, 我們知 $T_{\varepsilon_2}^{\varepsilon_3}$ 的 null space 為 $\{\begin{bmatrix}0\\0\\0\end{bmatrix}\}$, 故知 $\mathsf{N}(T)=\{\mathbf{0}\}$, 亦即 T 為 one-to-one. 另一方面 $[T]_{\varepsilon_2}^{\varepsilon_3}$ 的 rank 為 3, 故 $\{\begin{bmatrix}1\\1\\0\\0\end{bmatrix},\begin{bmatrix}-1\\0\\1\\0\end{bmatrix},\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}\}$ 為 column space 的一組 basis. 因此得 $\{x+1,x^2-1,x^3-x^2-x+1\}$ 為 $\mathsf{R}(T)$ 的一組 basis. 由於 $\dim(P_3(\mathbb{R}))=4\neq\dim(\mathsf{R}(T))=3$, 我們知 $\mathsf{R}(T)\neq P_3(\mathbb{R})$, 故 T 不是 onto.

Example 6.3.18. 考慮 $M_{2\times2}(\mathbb{R})$ 所形成的 vector space (參見 Example 3.2.2 (A)). 考慮函數 $T:M_{2\times2}(\mathbb{R})\to M_{2\times2}(\mathbb{R})$ 定義為 $T(A)=A-A^{\mathrm{t}}$, $\forall A\in M_{2\times2}(\mathbb{R})$. 我們可得 T 為 linear transformation. 這是因為對任意 $A,B\in M_{2\times2}(\mathbb{R})$ 以及 $r\in\mathbb{R}$, 我們有

$$T(A+rB)=(A+rB)-(A+rB)^{\mathrm{t}}=A+rB-A^{\mathrm{t}}-rB^{\mathrm{t}}=(A-A^{\mathrm{t}})+r(B-B^{\mathrm{t}})=T(A)+rT(B).$$

考慮 $M_{2\times2}(\mathbb{R})$ 的 ordered basis $\varepsilon=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$. 由於

$$T(\begin{bmatrix}1&0\\0&0\end{bmatrix})=\begin{bmatrix}0&0\\0&0\end{bmatrix},\quad T(\begin{bmatrix}0&1\\0&0\end{bmatrix})=\begin{bmatrix}0&1\\-1&0\end{bmatrix},$$

$$T(\begin{bmatrix}0&0\\1&0\end{bmatrix})=\begin{bmatrix}0&-1\\1&0\end{bmatrix},\quad T(\begin{bmatrix}0&0\\0&1\end{bmatrix})=\begin{bmatrix}0&0\\0&0\end{bmatrix},$$

我們得 T 對於 ε,ε 的 matrix representation 為 $[T]_\varepsilon^\varepsilon=\begin{bmatrix}0&0&0&0\\0&1&-1&0\\0&-1&1&0\\0&0&0&0\end{bmatrix}$. 利用 elementary row operation 將 $[T]_\varepsilon^\varepsilon$ 化為 echelon form $\begin{bmatrix}0&1&-1&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$. 得到 $[T]_\varepsilon^\varepsilon$ 的 null space 的一組

basis 為 $\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\}$, 因此得 $\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\}$ 為 $\mathsf{N}(T)$ 的一組 basis. 又 $[T]_\varepsilon^\varepsilon$ 的 column space 的一組 basis 為 $\{\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}\}$, 故得 $\{\begin{bmatrix}0&1\\-1&0\end{bmatrix}\}$ 為 T 的 range $\mathsf{R}(T)$ 的一組 basis. 注意我們有 $\dim(\mathsf{R}(T))+\dim(\mathsf{N}(T))=1+3=4=\dim(M_{2\times 2}(\mathbb{R}))$. 另外若 $A\in\mathsf{N}(T)$ 表示 $T(A)=A-A^{\mathrm{t}}=\mathbf{0}$, 亦即 $A=A^{\mathrm{t}}$. 反之亦然, 也就是說 $A\in\mathsf{N}(T)$ 若且唯若 A 為 symmetric matrix. 因此由 $\dim(\mathsf{N}(T))=3$, 我們知所有 2×2 的 symmetric matrices 所成的 subspace 的維度為 3.

Question 6.14. 試求所有 3×3 的 *symmetric matrices* 所成的 *subspace* 的維度為何?

當 T_1,T_2 皆為 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 linear transformation 時, 對任意 $c_1,c_2\in\mathbb{F}$, 我們知道 $c_1T_1+c_2T_2$ 的 standard matrix representation $[c_1T_1+c_2T_2]$ 和 T_1,T_2 的 standard matrix representations $[T_1],[T_2]$ 的關係為 $[c_1T_1+c_2T_2]=c_1[T_1]+c_2[T_2]$ (參見 Lemma 6.3.7). 這對於一般的 linear transformations $T_1:V\to W$ 以及 $T_2:V\to W$ 的 matrix representations 也是對的. 不過要特別注意, 一般的 linear transformation 的 matrix representation 是和定義域以及對應域的 ordered basis 有關, 所以只有當 T_1, T_2 都考慮對應相同的 ordered basis 所得的 matrix representation, 這樣的矩陣運算才有意義. 也就是說當分別給定 V,W 的 ordered basis, β,γ, 我們會有

$$[c_1T_1+c_2T_2]_\beta^\gamma=c_1[T_1]_\beta^\gamma+c_2[T_2]_\beta^\gamma.$$

對於合成函數也有類似的情況, 若 $T:V\to W$, $T':W\to U$ 為 linear transformations. 若分別給定 V,W,U 的 ordered basis α,β,γ, 我們有以下的圖示

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift right, "T_\alpha"'] & W \arrow[d, shift left, "T_\beta"] & W \arrow[r, "T'"] \arrow[d, shift right, "T_\beta"'] & U \arrow[d, shift left, "T_\gamma"] \\
\mathbb{F}^n \arrow[r] \arrow[u, shift right, "T_\alpha^{-1}"'] & \mathbb{F}^m \arrow[u, shift left, "T_\beta^{-1}"] & \mathbb{F}^m \arrow[r] \arrow[u, shift right, "T_\beta^{-1}"'] & \mathbb{F}^k \arrow[u, shift left, "T_\gamma^{-1}"]
\end{tikzcd}$$

這裡由於 T 的對應域和 T' 的定義域相同, 所以我們可以考慮合成函數 $T'\circ T$. 又由於 W 都用固定的 ordered basis β, 所以兩邊 W 到 $\mathbb{F}^m$ 的之間的函數相同 (皆為 T_β). 因此我們可以將上面兩個 commutative diagrams 合併成一個 commutative diagram 如下:

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift right, "T_\alpha"'] & W \arrow[r, "T'"] \arrow[d, shift left, "T_\beta"] & U \arrow[d, shift left, "T_\gamma"] \\
\mathbb{F}^n \arrow[r] \arrow[u, shift right, "T_\alpha^{-1}"'] & \mathbb{F}^m \arrow[r] \arrow[u, shift left, "T_\beta^{-1}"] & \mathbb{F}^k \arrow[u, shift left, "T_\gamma^{-1}"]
\end{tikzcd}$$

由於底部 $\mathbb{F}^n\to\mathbb{F}^m$ 的 matrix representation 為 $[T]_\alpha^\beta$, 而 $\mathbb{F}^m\to\mathbb{F}^k$ 的 matrix representation 為 $[T']_\beta^\gamma$, 因此由 Lemma 6.3.7 知, 它們的合成所對應的 matrix representation 為 $[T']_\beta^\gamma[T]_\alpha^\beta$. 因此我們有

$$[T'\circ T]_\alpha^\gamma=[T']_\beta^\gamma[T]_\alpha^\beta.$$

綜合以上的討論, 我們有以下有關 Lemma 6.3.7 的推廣.

Theorem 6.3.19. 假設 V,W,U 為 *finite dimensional vector space over* $\mathbb{F}$ 且令 α,β,γ 分別為 V,W,U 的 *ordered basis.*

(1) 假設 T_1,T_2 為 V 到 W 的 *linear transformations.* 則對任意 $c_1,c_2\in\mathbb{F}$, 我們有

$$[c_1T_1+c_2T_2]_\alpha^\beta=c_1[T_1]_\alpha^\beta+c_2[T_2]_\alpha^\beta.$$

(2) 設 $T:V\to W$ 及 $T':W\to U$ 為 *linear transformation.* 則

$$[T'\circ T]_\alpha^\gamma=[T']_\beta^\gamma[T]_\alpha^\beta.$$

假設 V,W 為 finite dimensional vector space over $\mathbb{F}$, 其中 $\dim(V)=n,\dim(W)=m$. 令 $\mathscr{L}(V,W)$ 為所有 V 到 W 的 linear transformations 所成的集合. Proposition 6.1.6 告訴我們 $\mathscr{L}(V,W)$ 有加法和係數積的封閉性. 很容易證明 $\mathscr{L}(V,W)$ 是一個 over $\mathbb{F}$ 的 vector space. 令 β, γ 分別為 V,W 上的 ordered basis, 我們可以訂一出一個由 $\mathscr{L}(V,W)$ 到 $M_{m\times n}(\mathbb{F})$ 的函數 $\mathscr{M}:\mathscr{L}(V,W)\to M_{m\times n}(\mathbb{F})$, 其定義為對任意 $T\in\mathscr{L}(V,W)$, $\mathscr{M}(T)=[T]_\beta^\gamma$. Theorem 6.3.19 告訴我們 $\mathscr{M}:\mathscr{L}(V,W)\to M_{m\times n}(\mathbb{F})$ 是一個 linear transformation. 我們有以下的結果.

Theorem 6.3.20. 假設 V,W 為 *finite dimensional vector space over* $\mathbb{F}$, 其中 $\dim(V)=n,\dim(W)=m$. 給定 β, γ 分別為 V,W 上的 *ordered basis,* 定義函數 $\mathscr{M}:\mathscr{L}(V,W)\to M_{m\times n}(\mathbb{F})$, 其中 $\forall T\in\mathscr{L}(V,W)$, $\mathscr{M}(T)=[T]_\beta^\gamma$. 則 $\mathscr{M}:\mathscr{L}(V,W)\to M_{m\times n}(\mathbb{F})$ 是一個 *one-to-one* 且 *onto* 的 *linear transformation.* 並可得 $\dim_{\mathbb{F}}(\mathscr{L}(V,W))=mn$.

Proof. 對任意 $T_1,T_2\in\mathscr{L}(V,W)$ 以及 $c\in\mathbb{F}$, 我們有 $\mathscr{M}(T_1+cT_2)=[T_1+cT_2]_\beta^\gamma$, 而 $\mathscr{M}(T_1)+c\mathscr{M}[T_2]=[T_1]_\beta^\gamma+c[T_2]_\beta^\gamma$, 故由 Theorem 6.3.19(1) 知 $\mathscr{M}(T_1+cT_2)=\mathscr{M}(T_1)+c\mathscr{M}(T_2)$.

假設 $\beta=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 以及 $\gamma=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$, 對任意 $A=[a_{ij}]\in M_{m\times n}(\mathbb{F})$, 由於 A 的 i-th column 為 $\begin{bmatrix}a_{1i}\\ \vdots\\ a_{mi}\end{bmatrix}$, 對任意 $i=1,\ldots,n$ 我們考慮 $T\in\mathscr{L}(V,W)$ 為唯一的 linear transformation 滿足 $T(\mathbf{v}_i)=a_{1i}\mathbf{w}_1+\cdots+a_{mi}\mathbf{w}_m$ (參見 Theorem 6.1.8). 依定義 $[T]_\beta^\gamma$ 的 i-th column 為 $[T(\mathbf{v}_i)]_\gamma$ 與 A 的 i-th column 相同, 故證得 $\mathscr{M}(T)=[T]_\beta^\gamma=A$. 這證得了 $\mathscr{M}:\mathscr{L}(V,W)\to M_{m\times n}(\mathbb{F})$ 是 onto, 也證得它是 one-to-one, 因為這樣的 $T\in\mathscr{L}(V,W)$ 是唯一的.

最後利用 Dimension Theorem (Theorem 6.2.9), 我們知道 $\operatorname{rank}(\mathscr{M})+\operatorname{nullity}(\mathscr{M})=\dim_{\mathbb{F}}(\mathscr{L}(V,W))$. 由於 $\mathscr{M}$ 是 one-to-one, $\operatorname{nullity}(\mathscr{M})=0$. 又由於 $\mathscr{M}$ 是 onto, 我們知 $\operatorname{rank}(\mathscr{M})=\dim_{\mathbb{F}}(M_{m\times n}(\mathbb{F}))=mn$. 故得證 $\dim_{\mathbb{F}}(\mathscr{L}(V,W))=mn$. □

Theorem 6.3.20, 告訴我們線性映射和矩陣間的對應關係. 也就是說我們可以利用表現矩陣來了解線性映射, 也可以用線性映射來了解矩陣. 兩者之間互相的關係大家應充分體會.

6.4. Invertible Linear Transformation

一般來說, 當一個函數是 invertible (即 one-to-one 且 onto) 時, 並不容易將其 inverse (反函數) 具體的寫下來. 這一節中我們將學得, 對於 invertible linear transformation, 利用 matrix representation 我們可以很容易的將其 inverse 寫下.

首先我們來探討何時一個 linear transformation 會是 invertible. 假設 $T : V \to W$ 為 invertible linear transformation. 由於 T 為 onto, 我們需要 $\mathrm{rank}(T) = \dim(W)$ (Proposition 6.2.3). 然而 T 為 one-to-one, 故知 $\dim(\mathsf{N}(T)) = \mathrm{nullity}(T) = 0$ (Proposition 6.2.6). 利用 Dimension Theorem for linear transformation (Theorem 6.2.9) 我們得

$$\dim(V) = \mathrm{rank}(T) + \mathrm{nullity}(T) = \dim(W) + 0 = \dim(W).$$

這告訴我們只有在 $\dim(V) = \dim(W)$ 時, T 才有可能為 invertible. 現假設 $\dim(V) = \dim(W)$ 且 $T : V \to W$ 為 one-to-one, 由 $\mathrm{nullity}(T) = 0$, 我們得 $\mathrm{rank}(T) = \dim(V) = \dim(W)$, 亦即 T 為 onto. 同樣的, 若 T 為 onto, 則由 $\mathrm{rank}(T) = \dim(W) = \dim(V)$ 得 $\mathrm{nullity}(T) = 0$, 亦即 T 為 one-to-one. 這告訴我們當 $\dim(V) = \dim(W)$ 時, 對於 linear transformation $T : V \to W$, T 為 one-to-one 和 T 為 onto 是等價的. 因而只要其中一個是對的, 就可以得到 T 為 invertible.

Lemma 6.4.1. 假設 V, W 為 *finite dimensional vector spaces over* $\mathbb{F}$ 且 $T : V \to W$ 為 *linear transformation.* 則僅有當 $\dim(V) = \dim(W)$ 時, T 才有可能是 *invertible.* 又當 $\dim(V) = \dim(W)$ 時, T 為 *invertible* 和 T 為 *onto* 是等價的也和 T 為 *one-to-one* 等價.

由 Lemma 6.4.1 我們知道若 $T : V \to W$ 為 invertible, 則 $\dim(V) = \dim(W)$. 換句話說若 $\dim(V) \neq \dim(W)$, 則不可能存在一個 V 到 W 的 invertible linear transformation. 這個性質的反向也是對的, 我們有以下的定理.

Proposition 6.4.2. 設 V, W 為 *finite dimensional vector spaces over* $\mathbb{F}$. 則存在 $T : V \to W$ 為 *invertible linear transformation* 若且唯若 $\dim(V) = \dim(W)$.

Proof. 由 Lemma 6.4.1, 我們僅要證明若 $\dim(V) = \dim(W)$ 則存在 linear transformation $T : V \to W$ 為 invertible. 令 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 和 $\mathbf{w}_1, \ldots, \mathbf{w}_n$ 分別為 V 的一組 basis 和 W 的一組 basis. 考慮一 linear transformation $T : V \to W$ 滿足 $T(\mathbf{v}_i) = \mathbf{w}_i$, $\forall i = 1, \ldots, n$ (參見 Theorem 6.1.8). 我們要說明 T 為 invertible. 然而由 Proposition 6.2.4 知 $\mathsf{R}(T) = \mathrm{Span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)) = \mathrm{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_n)$. 又由 $\mathbf{w}_1, \ldots, \mathbf{w}_n$ 為 W 的一組 basis 知 $\mathrm{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_n) = W$. 故得證 $\mathsf{R}(T) = W$, 即 T 為 onto. 再由 Lemma 6.4.1 得證 T 為 invertible. □

當 $T : V \to W$ 是 one-to-one 且 onto 時, 我們知道它是 invertible, 亦即存在 $T^{-1} : W \to V$, 滿足對所有 $\mathbf{v} \in V$ 皆有 $T^{-1} \circ T(\mathbf{v}) = T^{-1}(T(\mathbf{v})) = \mathbf{v}$ 以及對任意 $\mathbf{w} \in W$ 皆有 $T \circ T^{-1}(\mathbf{w}) = T(T^{-1}(\mathbf{w})) = \mathbf{w}$. 一般來說我們稱 T^{-1} 為 T 的 *inverse* (反函數). 我們都知道 $T^{-1} : W \to V$ 仍為 one-to-one 且 onto, 不過令人好奇的是它是否仍為 linear transformation? 我們有以下之結果.

Theorem 6.4.3. 假設 V,W 為 *vector spaces over* $\mathbb{F}$ 且 $T:V\to W$ 為 *linear transformation.* 若 T 為 *invertible,* 則 T 的 *inverse* $T^{-1}:W\to V$ 亦為 *linear transformation.*

Proof. 對於任意 $\mathbf{w}\in W$, 由於 T 是 one-to-one 且 onto, 故存在唯一的 $\mathbf{v}\in V$ 滿足 $T(\mathbf{v})=\mathbf{w}$. 依反函數定義此時 $T^{-1}(\mathbf{w})=\mathbf{v}$. 所以 T^{-1} 確實定出一個從 W 到 V 的函數. 我們要證明 $T^{-1}:W\to V$ 是一個 linear transformation. 也就是說任取 $\mathbf{w}_1,\mathbf{w}_2\in W$, $r\in\mathbb{F}$, 我們要證明 $T^{-1}(\mathbf{w}_1+r\mathbf{w}_2)=T^{-1}(\mathbf{w}_1)+rT^{-1}(\mathbf{w}_2)$. 現假設 $T^{-1}(\mathbf{w}_1)=\mathbf{v}_1$, $T^{-1}(\mathbf{w}_2)=\mathbf{v}_2$. 亦即 $T(\mathbf{v}_1)=\mathbf{w}_1$ 且 $T(\mathbf{v}_2)=\mathbf{w}_2$. 要檢查 $T^{-1}(\mathbf{w}_1+r\mathbf{w}_2)$ 是否等於 $T^{-1}(\mathbf{w}_1)+rT^{-1}(\mathbf{w}_2)$ 就等同於檢查是否 $T^{-1}(\mathbf{w}_1+r\mathbf{w}_2)=\mathbf{v}_1+r\mathbf{v}_2$, 也就說是否 $T(\mathbf{v}_1+r\mathbf{v}_2)=\mathbf{w}_1+r\mathbf{w}_2$. 然而已知 T 為 linear transformation, 我們有 $T(\mathbf{v}_1+r\mathbf{v}_2)=T(\mathbf{v}_1)+rT(\mathbf{v}_2)=\mathbf{w}_1+r\mathbf{w}_2$. 故得證 $T^{-1}(\mathbf{w}_1+r\mathbf{w}_2)=\mathbf{v}_1+r\mathbf{v}_2=T^{-1}(\mathbf{w}_1)+rT^{-1}(\mathbf{w}_2)$, 亦即 T^{-1} 是一個 linear transformation. □

在 Proposition 6.2.4 我們知道當 $T:V\to W$ 是 onto 時會保持 spanning set. 也就是說若 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 是 V 的一組 spanning set, 則 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)$ 會是 W 的一組 spanning set. 然而怎樣的 linear transformation 會保持 linearly independent 的關係呢? 我們有以下之定理.

Proposition 6.4.4. 假設 V,W 為 *vector spaces over* $\mathbb{F}$ 且假設 $T:V\to W$ 為 *one-to-one linear transformation.* 則 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 *linearly independent* 若且唯若 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 *linearly independent.*

Proof. ($\Rightarrow$) 假設 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 linearly independent, 我們希望說明 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 linearly independent. 用反證法, 假設 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 linearly dependent, 亦即存在 $c_1,\dots,c_n\in\mathbb{F}$ 不全為 0 使得 $c_1T(\mathbf{v}_1)+\cdots+c_nT(\mathbf{v}_n)=\mathbf{0}$. 現因 T 為 linear, 我們有 $\mathbf{0}=c_1T(\mathbf{v}_1)+\cdots+c_nT(\mathbf{v}_n)=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)$. 因此利用 T 為 one-to-one, 知僅有在 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$ 時才有可能使得 $T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=\mathbf{0}$. 故由 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 linearly independent 得知 $c_1=\cdots=c_n=0$. 此和 $c_1,\dots,c_n\in\mathbb{F}$ 不全為 0 的假設相矛盾, 故知 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 linearly independent.

($\Leftarrow$) 這個方向的證明不需要 T 為 one-to-one, 僅需要 T 為 linear transformation. 現假設 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 linearly independent. 我們用反證法假設 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 linearly dependent, 亦即存在 $c_1,\dots,c_n\in\mathbb{F}$ 不全為 0 使得 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{0}$. 現因 T 為 linear, 我們有 $c_1T(\mathbf{v}_1)+\cdots+c_nT(\mathbf{v}_n)=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=T(\mathbf{0})=\mathbf{0}$. 但由於 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)\in W$ 為 linearly independent, 我們得到 $c_1=\cdots=c_n=0$ 之矛盾. 故得證 $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ 為 linearly independent. □

當 $T:V\to W$ 是 invertible 的 linear transformation 時, 我們稱 T 為一個 *isomorphism.* 意思是此時 V 和 W 看成 vector space 時有相似的結構而 T 就是保持這個結構的函數. 事實上由 T 為 onto, 我們知道 T 會保持 V 和 W 的 spanning set, 而由 T 為 one-to-one 我們知道 T 會保持 linearly independent 的關係, 所以我們有以下的結果.

Theorem 6.4.5. 設 V,W 為 *vector spaces over* $\mathbb{F}$ 且 $T:V\to W$ 為 *isomorphism.* 則 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 為 V 的一組 *basis* 若且唯若 $T(\mathbf{v}_1),\dots,T(\mathbf{v}_n)$ 為 W 的一組 *basis.*

Proof. ($\Rightarrow$) 假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis. 因為 T 為 onto, 我們有 $\mathsf{R}(T)=W$. 又因 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 spanning set, 利用 Proposition 6.2.4 知 $\mathrm{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))=\mathsf{R}(T)=W$, 亦即 $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)\}$ 為 W 的 spanning set. 又因為 T 為 one-to-one 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$ 為 linearly independent, 利用 Proposition 6.4.4 知 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)\in W$ 為 linearly independent. 故得 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)$ 為 W 的一組 basis.

($\Leftarrow$) 假設 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)$ 為 W 的一組 basis. 由於 $T^{-1}:W\to V$ 為 linear transformation (Theorem 6.4.3), 且 $T:V\to W$ 為其 inverse, 故知 T^{-1} 為 one-to-one 且 onto, 也就是說 $T^{-1}:W\to V$ 為 isomorphism. 故由 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)$ 為 W 的一組 basis, 套用前面所證可得 $T^{-1}(T(\mathbf{v}_1)),\ldots,T^{-1}(T(\mathbf{v}_n))$ 為 V 的一組 basis. 由於對於 $i=1,\ldots,n$ 皆有 $T^{-1}(T(\mathbf{v}_i))=\mathbf{v}_i$, 故知 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis. $\square$

假設 V 為 vector space over $\mathbb{F}$ 且 β 為 V 的一組 ordered basis. 令 $T_\beta:V\to\mathbb{R}^n$ 定義為 $T_\beta(\mathbf{v})=[\mathbf{v}]_\beta$, $\forall\mathbf{v}\in V$, 在 Theorem 6.3.10 我們知道 T_β 是 isomorphism. 而在 Proposition 6.3.16 中我們知道若 V,W 為 vector space over $\mathbb{F}$ 且 β,γ 分別為 V,W 的一組 ordered basis, 則對於 linear transformation $T:V\to W$, $\mathbf{v}\in\mathsf{N}(T)$ 若且唯若 $T_\beta(\mathbf{v})\in\mathsf{N}([T]_\beta^\gamma)$, 而 $\mathbf{w}\in\mathsf{R}(T)$ 若且唯若 $T_\gamma(\mathbf{w})\in\mathrm{Col}([T]_\beta^\gamma)$. 因此 T_β 定義出了一個從 $\mathsf{N}(T)$ 到 $\mathsf{N}([T]_\beta^\gamma)$ 的 isomorphism 且 T_γ 定義出了一個從 $\mathsf{R}(T)$ 到 $\mathrm{Col}([T]_\beta^\gamma)$ 的 isomorphism. 利用 Theorem 6.4.5, 我們得到以下的結果.

Corollary 6.4.6. V,W 為 *vector space over* $\mathbb{F}$ 且 β,γ 分別為 V,W 的一組 *ordered basis.* 則 $\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ 為 $\mathsf{N}(T)$ 的一組 *basis* 若且唯若 $\{[\mathbf{v}_1]_\beta,\ldots,[\mathbf{v}_r]_\beta\}$ 為矩陣 $[T]_\beta^\gamma$ 的 *null space* $\mathsf{N}([T]_\beta^\gamma)$ 的一組 *basis.* 而 $\{\mathbf{w}_1,\ldots,\mathbf{w}_s\}$ 為 $\mathsf{R}(T)$ 的一組 *basis* 若且唯若 $\{[\mathbf{w}_1]_\gamma,\ldots,[\mathbf{w}_s]_\gamma\}$ 為矩陣 $[T]_\beta^\gamma$ 的 *column space* $\mathrm{Col}([T]_\beta^\gamma)$ 的一組 *basis.* 也因此得到

$$\mathrm{nullity}(T)=\mathrm{nullity}([T]_\beta^\gamma),\quad \mathrm{rank}(T)=\mathrm{rank}([T]_\beta^\gamma).$$

當 $T:V\to W$ 為 isomorphism 時, 我們已知 T 的 inverse $T^{-1}:W\to V$ 亦為 linear transformation 且 $\dim(V)=\dim(W)$. 此時當 V 和 W 分別選定固定的 ordered basis β,γ 後, 我們有以下的 commutative diagram.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift right, "T_\beta"'] & W \arrow[r, "T^{-1}"] \arrow[d, shift left, "T_\gamma"] & V \arrow[d, shift left, "T_\beta"] \\
\mathbb{F}^n \arrow[r] \arrow[u, shift right, "T_\beta^{-1}"'] & \mathbb{F}^n \arrow[r] \arrow[u, shift left, "T_\gamma^{-1}"] & \mathbb{F}^n \arrow[u, shift left, "T_\beta^{-1}"]
\end{tikzcd}$$

由於兩端 V 所用的 ordered basis 是一致的, 所以由 $T^{-1}\circ T$ 是 V 到 V 的 identity map $\mathrm{id}:V\to V$. 在 Example 6.3.12, 我們知道 $[\mathrm{id}]_\beta^\beta=I_n$, 故利用 Theorem 6.3.19 推得

$$[T^{-1}]_\gamma^\beta[T]_\beta^\gamma=[T^{-1}\circ T]_\beta^\beta=[\mathrm{id}]_\beta^\beta=I_n.$$

同理, 由於 $T\circ T^{-1}$ 為 W 到 W 的 identity map, 我們得 $[T]_\beta^\gamma[T^{-1}]_\gamma^\beta=I_n$. 得證 $[T^{-1}]_\gamma^\beta$ 為 $[T]_\beta^\gamma$ 的反矩陣.

綜合以上的討論, 我們有以下的結論.

Theorem 6.4.7. 假設 V,W 為 *finite dimensional vector spaced over* $\mathbb{F}$ 且令 β,γ 分別為 V,W 的 *ordered basis.* 設 $T:V\to W$ 為 *linear transformation.* 則 T 為 *isomorphism* 若且唯若 $[T]_\beta^\gamma$ 為 *invertible matrix.* 又此時 $T^{-1}:W\to V$ 對應於 β,γ 的 *matrix representation* 為

$$[T^{-1}]_\gamma^\beta=([T]_\beta^\gamma)^{-1}.$$

Proof. 由前面的討論已知當 T 為 isomorphism 時其對應於 ordered basis β,γ 的 matrix representation $[T]_\beta^\gamma$ 為 invertible matrix. 我們僅須證明當 $[T]_\beta^\gamma$ 為 invertible matrix 時, $T:V\to W$ 為 isomorphism, 亦即證明 T 為 one-to-one 且 onto. 然而由 $[T]_\beta^\gamma$ 為 invertible matrix, 我們知 $[T]_\beta^\gamma$ 為方陣, 亦即 $\dim(V)=\dim(W)$. 因為 invertible matrix 的 null space 是 $\{\mathbf{0}\}$, 故由 Proposition 6.3.16 知 $\mathsf{N}(T)=\{\mathbf{0}\}$, 亦即 T 為 one-to-one. 最後由 $\dim(V)=\dim(W)$ 以及 Lemma 6.4.1 我們得證 T 亦為 onto. $\square$

Example 6.4.8. 考慮 $T:\mathbb{R}^3\to\mathbb{R}^3$ 定義為 $T(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=\begin{bmatrix}2x_2\\x_1-x_2\\2x_2+x_3\end{bmatrix}$. 我們可得 T 的 standard matrix representation 為 $[T]=\begin{bmatrix}0&2&0\\1&-1&0\\0&2&1\end{bmatrix}$. 在 Example 2.5.8 中我們算出 $[T]^{-1}=\begin{bmatrix}\frac{1}{2}&1&0\\\frac{1}{2}&0&0\\-1&0&1\end{bmatrix}$. 故得 $T^{-1}:\mathbb{R}^3\to\mathbb{R}^3$ 的定義為 $T^{-1}(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=\begin{bmatrix}\frac{1}{2}x_1+x_2\\\frac{1}{2}x_1\\-x_1+x_3\end{bmatrix}$. 我們驗證

$$(T^{-1}\circ T)(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=T'(\begin{bmatrix}2x_2\\x_1-x_2\\2x_2+x_3\end{bmatrix})=\begin{bmatrix}\frac{1}{2}(2x_2)+(x_1-x_2)\\\frac{1}{2}(2x_2)\\-(2x_2)+(2x_2+x_3)\end{bmatrix}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},$$

$$(T\circ T^{-1})(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix})=T(\begin{bmatrix}\frac{1}{2}x_1+x_2\\\frac{1}{2}x_1\\-x_1+x_3\end{bmatrix})=\begin{bmatrix}2(\frac{1}{2}x_1)\\(\frac{1}{2}x_1+x_2)-(\frac{1}{2}x_1)\\2(\frac{1}{2}x_1)+(-x_1+x_3)\end{bmatrix}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},$$

得知 T^{-1} 確為 T 的 inverse.

6.5. 結論

我們學習了 $\mathbb{R}^n$ 上重要的函數, linear transformation. 一個定義在 $\mathbb{R}^n$ 的 linear transformation 可以由一組 $\mathbb{R}^n$ 的 basis 所映得的向量唯一確定, 所以我們得到所謂的 standard matrix representation. 利用 standard matrix representation, 可以幫助我們了解 linear transformation. 因此我們可以利用矩陣的性質推得許多有關 linear transformation 的性質.

由於 linear transformation 可以由一組 basis 所決定. 因此我們可以依 linear transformation 的特性選定一組較易掌握的 basis 來決定此 linear transformation. 利用這樣的概念, 我們也可以由一組 basis 經由此 linear transformation 的變化情形將一個較複雜的 linear transformation 拆解成一些較容易掌握的 linear transformations 的合成.

我們介紹了一般的 finite dimensional vector spaces 以及它們之間的 linear transformations. 由於 finite dimensional vector space 也會有 basis 存在, 所以我們可以加它們坐標化以至於將它們如我們熟悉的 $\mathbb{R}^n$ 來處理. 也因此這樣的 linear transformation 都可以用矩陣

來表示. 所以我們可以利用矩陣的理論來更進一步了解這些 linear transformations. 不過要注意一個 linear transformation 的 matrix representation 和 ordered basis 的選取有關. 有時選取好的 ordered basis 可以讓我們得到更好的 matrix representation 以至於更容易了解這個 linear transformation. 因此一個 linear transformation 選用不同 order basis 所得的 matrix representation 之間的關係分外重要. 希望大家能好好了解它們之間的關係.