大學線性代數初步

李華介

國立台灣師範大學數學系

前言

本講義主要目的是針對數學系大一學生介紹有關線性代數基本的理論, 大致上僅談論實係數的向量空間.

一般經驗上大一學生會覺得線性代數學習起來會比微積分吃力, 主要是因素可能是在高中時期學習有關線性代數部分主要著重於操作而較少論證. 這一點在大學線性代數中就較不同了. 不過經驗上也告訴我們當學生到大三大四時 (因準備研究所考試) 再重新溫習線性代數時, 會覺得它不再那麼難以親近. 由此可知大學時期的線性代數其理論並不難懂, 較大的障礙是要開始學習數學的論證. 這個障礙或許到大三大四一些數學思維較成熟時稍可解除, 不過對一些同學來說可能為時已晚, 畢竟線性代數的理論或概念與其他課程都脫不了關係. 基於這個原因寫下這份講義, 希望藉由較平易近人的方式介紹線性代數也慢慢引導熟悉數學的論證方式. 本講義希望以淺顯易懂為主旨, 而不是生動有趣. 畢竟有些事情要說明清楚就會顯得囉唆, 當然就不有趣了.

研讀本講義的同學要有心理準備, 本講義是針對數學系學生而寫, 自然偏重於整個線性代數的理論架構. 對於線性代數在其他領域的應用幾乎沒有著墨. 我們依循一貫的原則就是理論清楚了, 接下來的應用或推廣就不難了. 所以對應用有興趣的同學應再參考其他的參考書籍. 另外本講義並未提供習題, 不過在某些概念講述之後有時會提供一些問題 (Question). 這些問題幾乎是檢視觀念是否正確或是對內容是否了解, 大部分問題若觀念已清楚應可以立即回答. 所以這些問題的份量仍不及一般習題, 對熟習線性代數所給予的訓練. 針對這一點, 請欲學習好線性代數的同學務必參閱一般線性代數書籍的習題, 自行磨練.

本講義雖然主要以中文撰寫, 不過當涉及定義或專有名詞時, 為免翻譯的困擾將以英文取代. 因此將以中英夾雜較不傳統的方式顯現, 若有不便請見諒.

本講義編寫費時, 編寫完後並未經過嚴謹的校對. 疏漏在所難免, 雖不至於有理論性上嚴重的錯誤, 但讀者仍應注意不宜概括全收. 若發現錯誤, 歡迎提出寶貴的意見.

Chapter 6

Linear Transformation of General Vector Space

在這一章中, 我們將推廣前兩章的結果到一般的 vector space. 我們將介紹 vector space 並探討 vector space 上的 linear transformation. 由於前兩章的結果, 僅利用 vector space 的特性便可推得, 而不必侷限在 $\mathbb{R}^n$ 的情況. 所以我們很容易推得和前面一致的結果.

6.1. Vector Space and Subspace

我們曾經提過像 $\mathbb{R}^m$ 這樣, 裡面任意有限多個向量的線性組合仍在 $\mathbb{R}^m$ 中, 而向量的運算又符合 Proposition 1.2.3 的 8 項規則, 我們便稱之為 vector space. 在這一節中我們將正式定義 vector space 並探討 vector space 以及其 subspaces 的相關性質.

給定一非空集合 V, 我們說 V 中有加法運算 $+$, 表示對於任意 V 中兩個元素 $\mathbf{u},\mathbf{v} \in V$, 經由這個運算所得的結果 $\mathbf{u}+\mathbf{v}$ 仍然是 V 中的元素 (此為加法封閉性). 又我們說 $\mathbb{R}$ 對 V 有係數積表示對任意 $r \in \mathbb{R}$ 以及 $\mathbf{v} \in V$, 皆有 r 對 $\mathbf{v}$ 運算所得 $r\mathbf{v}$ 仍然在 V 中 (此為係數積封閉性).

Definition 6.1.1. 假設非空集合 V 中有加法運算 $+$, 以及 $\mathbb{R}$ 對 V 的係數積. 若這兩種運算符合以下 8 項性質, 則稱 V 為一個 *vector space.*

(1) 對任意 $\mathbf{u},\mathbf{v} \in V$, 皆有 $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$.

(2) 對任意 $\mathbf{u},\mathbf{v},\mathbf{w} \in V$, 皆有 $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(3) 存在一向量 $\mathbf{O} \in V$ 滿足對任意 $\mathbf{u} \in V$ 皆有 $\mathbf{O}+\mathbf{u} = \mathbf{u}$.

(4) 對任意 $\mathbf{u} \in V$ 皆可找到 $\mathbf{u}' \in V$ 滿足 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$.

(5) 對任意 $r,s \in \mathbb{R}$ 以及 $\mathbf{u} \in V$, 皆有 $r(s\mathbf{u}) = (rs)\mathbf{u}$.

(6) 對任意 $r,s \in \mathbb{R}$ 以及 $\mathbf{u} \in V$, 皆有 $(r+s)\mathbf{u} = r\mathbf{u}+s\mathbf{u}$.

(7) 對任意 $r \in \mathbb{R}$ 以及 $\mathbf{u}, \mathbf{v} \in V$ 皆有 $r(\mathbf{u}+\mathbf{v}) = r\mathbf{u}+r\mathbf{v}$.

(8) 對任意 $\mathbf{u} \in V$, 皆有 $1\mathbf{u} = \mathbf{u}$.

在此要說明一下, 一般來說我們不能說一個集合是 vector space, 一定要附帶說明它的加法及係數積為何. 不過當我們談到一般抽象的 vector space 時, 我們說 V 是一個 vector space 時就隱含其中有加法運算且直接用 + 表示, 同時也隱含其中有實數的係數積, 而不再去強調其中有加法及係數積. 同樣的對於常見的 vector space, 例如 $\mathbb{R}^n$, 由於我們已經有常用的加法及係數積, 所以不會再次強調其加法及係數積為何. 不過當我們要介紹一個新的具體的 vector space 時, 就一定要說明如何定出其加法及係數積. 另外 vector space 不一定要具有實數的係數積, 具一般的 "field" 的係數積也可. 所以一般來說我們必須說是 over 哪一個 field 的 vector space (因為 over 不同的 field 影響很大). 不過由於許多同學可能不知道甚麼是 field 而且我們這裡談論的都是實數 $\mathbb{R}$ 的情形, 所以即使依上面的定義我們應該說是 over $\mathbb{R}$ 的 vector space, 不過我們都省略僅說是一個 vector space 即可.

在 Corollary 1.2.4 我們證明了在 $\mathbb{R}^n$ 的情況, 上述性質 (3) 中 $\mathbf{O}$ 是唯一的而 (4) 中若給定 $\mathbf{u}$, 則 $\mathbf{u}'$ 也是唯一的. 由於當時我們的證明故意避開用 $\mathbb{R}^n$ 的加法與係數積的定義處理, 而僅利用上述 8 項性質來證明, 所以知道這個結果對一般的 vector space 也成立.

Proposition 6.1.2. 假設 V 為 *vector space,* 則在 V 中存在唯一的向量 $\mathbf{O}$ 滿足對任意 $\mathbf{u} \in V$ 皆有 $\mathbf{O}+\mathbf{u} = \mathbf{u}$. 另外, 給定 $\mathbf{u} \in V$, 存在唯一的 $\mathbf{u}' \in V$ 滿足 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$.

既然 $\mathbf{O}$ 是唯一的, 以後就用 $\mathbf{O}$ 這個專屬的符號來表示 V 中唯一符合 $\mathbf{O}+\mathbf{u} = \mathbf{u}, \forall \mathbf{u} \in V$ 的這個元素, 且稱之為 V 的 *additive identity* 或依慣例稱之為 *zero element.* 又給定 $\mathbf{u} \in V$, 存在唯一的 $\mathbf{u}'$ 使得 $\mathbf{u}+\mathbf{u}' = \mathbf{O}$, 依慣例我們以後就用 $-\mathbf{u}$ 來表示這一個唯一的 $\mathbf{u}'$, 且稱之為 $\mathbf{u}$ 的 *additive inverse.*

接著我們要再強調的是, 雖然在性質 (3) 中提到 $\mathbf{O}$ 是必須滿足對所有 $\mathbf{u} \in V$ 皆有 $\mathbf{O}+\mathbf{u} = \mathbf{u}$ 才可以 (事實上 Proposition 6.1.2 的證明需用到對所有 $\mathbf{u}$ 皆對才可以), 也就是說依定義要驗證對所有 $\mathbf{u} \in V$ 皆有 $\mathbf{w}+\mathbf{u} = \mathbf{u}$, 才能確定 $\mathbf{w} = \mathbf{O}$. 不過有了 Proposition 6.1.2 後, 我們便可推得只要有一個 $\mathbf{u} \in V$, 會使得 $\mathbf{w}+\mathbf{u} = \mathbf{u}$, 就可以認定 $\mathbf{w} = \mathbf{O}$ 了. 利用這個性質我們便可推得許多有關於 $\mathbf{O}$ 的性質. 例如對任意實數 r, 我們皆會有 $r\mathbf{O} = \mathbf{O}$. 這是因為利用性質 (7) 可得 $r\mathbf{O}+r\mathbf{O} = r(\mathbf{O}+\mathbf{O}) = r\mathbf{O}$. 也就是說 $r\mathbf{O}$ 加了 $r\mathbf{O}$ 仍為 $r\mathbf{O}$, 所以由前面所提可知 $r\mathbf{O} = \mathbf{O}$. 為了方便起見, 我們列出以下的結果, 至於證明請參見 Corollary 1.2.6 以及 Corollary 1.2.7.

Proposition 6.1.3. 假設 V 為 *vector space,* 我們有以下之結果.

(1) 若對於 $\mathbf{w} \in V$ 存在 $\mathbf{u} \in V$ 使得 $\mathbf{w}+\mathbf{u} = \mathbf{u}$, 則 $\mathbf{w} = \mathbf{O}$.

(2) 對任意 $\mathbf{v} \in V$ 皆有 $0\mathbf{v} = \mathbf{O}$.

(3) 對任意 $r \in \mathbb{R}$ 皆有 $r\mathbf{O} = \mathbf{O}$.

(4) 對任意 $r \in \mathbb{R}, \mathbf{v} \in V$ 皆有 $(-1)(r\mathbf{v}) = -(r\mathbf{v}) = r(-\mathbf{v})$.

由 Proposition 6.1.3 我們知當 $r=0$ 或 $\mathbf{v}=\mathbf{O}$ 時會有 $r\mathbf{v}=\mathbf{O}$, 但若 $r\neq 0$ 且 $\mathbf{v}\neq\mathbf{O}$, 是否有可能 $r\mathbf{v}=\mathbf{O}$ 呢? 答案是不可能. 這是因為若 $r\neq 0$ 且 $\mathbf{v}\neq\mathbf{O}$, 我們可以考慮 $1/r$ 乘上 $r\mathbf{v}$ 由 vector space 運算性質 (5) 得 $\frac{1}{r}(r\mathbf{v})=\frac{r}{r}\mathbf{v}=1\mathbf{v}$, 再由性質 (8) 知 $1\mathbf{v}=\mathbf{v}$, 故得 $\frac{1}{r}(r\mathbf{v})=\mathbf{v}\neq\mathbf{O}$. 現若 $r\mathbf{v}=\mathbf{O}$, 由 Proposition 6.1.3 (3) 會得到 $\frac{1}{r}(r\mathbf{v})=\frac{1}{r}\mathbf{O}=\mathbf{O}$, 造成矛盾. 故知此時 $r\mathbf{v}$ 絕對不會是 $\mathbf{O}$.

Question 6.1. *假設 V 為 vector space, $\mathbf{v}\in V$ 且 $\mathbf{v}\neq\mathbf{O}$, 試證明若 $r,s\in\mathbb{R}$ 且 $r\neq s$, 則 $r\mathbf{v}\neq s\mathbf{v}$, 並利用這個結果證明若 V 是 vector space over $\mathbb{R}$ 且 V 中有非零元素, 則 V 有無窮多個元素.*

總而言之, vector space 中所要求加法及係數積的 8 項性質, 就是要確保一個 vector space 中的元素運算都可像實數一般處理. 例如, 我們可以如實數一樣引用“減法”的符號, 也就是說將 $\mathbf{w}+(-\mathbf{v})$ 寫成 $\mathbf{w}-\mathbf{v}$. 如此一來以後我們在一些等式的推演時就直接沿用大家習慣的「移項」的說法. 例如 $2\mathbf{u}+\mathbf{v}=\mathbf{w}$, 我們就直接移項且乘以 $1/2$ 得 $\mathbf{u}=\frac{1}{2}(\mathbf{w}-\mathbf{v})$.

接下來我們看一些常見的 vector space 的例子.

Example 6.1.4. (A) 考慮 $\mathscr{M}_{m\times n}$ 為所有 $m\times n$ matrices 所成的集合. 利用我們一般定義矩陣的加法與係數積 (參見 Definition 3.1.2), 我們可以證得這兩種運算符合 vector space 運算的 8 項性質 (參見 Proposition 3.1.3). 所以在這個加法運算與係數積之下 $\mathscr{M}_{m\times n}$ 是一個 vector space.

(B) 考慮 $P(\mathbb{R})$ 為所有以 x 為變數的實係數多項式. 考慮一般多項式的加法與係數積, 我們證明 $P(\mathbb{R})$ 為 vector space. 首先回顧若 $f(x)=a_nx^n+\cdots+a_1x+a_0, g(x)=b_mx^m+\cdots+b_1x+b_0\in P(\mathbb{R})$, 其中 $m\leq n$, 則我們可以將 $g(x)$ 寫成 $g(x)=b_nx^n+\cdots+b_mx^m+\cdots+b_1x+b_0$, 其中 $b_n=b_{n-1}=\cdots=b_{m+1}=0$. 為了方便起見雖然多項式的次數可能不同, 以後我們都用這種方式將它們補成相同次數再相加. 所以我們可以將 $f(x)+g(x)$ 的定義寫成 $f(x)+g(x)=\sum_{i=0}^n(a_i+b_i)x^i$. 而對於 $r\in\mathbb{R}$, 係數積 $rf(x)$ 的定義為 $rf(x)=\sum_{i=0}^n(ra_i)x^i$. 利用這個定義, 我們知道兩多項式相加仍為多項式以及實數對多項式的係數積仍為多項式, 所以在此定義之下加法和係數積確為 $P(\mathbb{R})$ 中的運算. 接著我們要一一檢查是否符合 vectors space 的 8 項運算規則.

(1) 對任意 $f(x)=\sum_{i=0}^n a_ix^i$, $g(x)=\sum_{i=0}^n b_ix^i\in P(\mathbb{R})$ 我們有

$$f(x)+g(x)=\sum_{i=0}^n(a_i+b_i)x^i=\sum_{i=0}^n(b_i+a_i)x^i=g(x)+f(x).$$

(2) 對任意 $f(x)=\sum_{i=0}^n a_ix^i$, $g(x)=\sum_{i=0}^n b_ix^i, h(x)=\sum_{i=0}^n c_ix^i\in P(\mathbb{R})$ 我們有

$$(f(x)+g(x))+h(x)=\sum_{i=0}^n(a_i+b_i)x^i+\sum_{i=0}^n c_ix^i=\sum_{i=0}^n((a_i+b_i)+c_i)x^i,$$

$$f(x)+(g(x)+h(x))=\sum_{i=0}^n a_ix^i+\sum_{i=0}^n(b_i+c_i)x^i=\sum_{i=0}^n(a_i+(b_i+c_i))x^i.$$

由於 $(a_i+b_i)+c_i=a_i+(b_i+c_i)$, 故得證 $(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x))$.

(3) 考慮零多項式 $g(x) = 0 = \sum_{i=0}^{n} b_i x^i \in P(\mathbb{R})$, 其中 $b_i = 0$, $\forall i = 0, 1, \ldots, n$. 此時對任意 $f(x) = \sum_{i=0}^{n} a_i x^i \in P(\mathbb{R})$, 可得

$$f(x) + g(x) = \sum_{i=0}^{n} (a_i + b_i) x^i = \sum_{i=0}^{n} a_i x^i = f(x).$$

(4) 給定 $f(x) = \sum_{i=0}^{n} a_i x^i \in P(\mathbb{R})$, 我們考慮 $h(x) = \sum_{i=0}^{n} (-a_i) x^i \in P(\mathbb{R})$, 則

$$f(x) + h(x) = \sum_{i=0}^{n} (a_i - a_i) x^i = \sum_{i=0}^{n} 0 x^i = 0.$$

(5) 對任意 $r, s \in \mathbb{R}$ 以及 $f(x) = \sum_{i=0}^{n} a_i x^i \in P(\mathbb{R})$, 我們有

$$r(sf(x)) = r(\sum_{i=0}^{n} (sa_i) x^i) = \sum_{i=0}^{n} (r(sa_i)) x^i = \sum_{i=0}^{n} ((rs)a_i) x^i = (rs) \sum_{i=0}^{n} a_i x^i = (rs) f(x).$$

(6) 對任意 $r, s \in \mathbb{R}$ 以及 $f(x) = \sum_{i=0}^{n} a_i x^i \in P(\mathbb{R})$, 皆有

$$(r+s) f(x) = \sum_{i=0}^{n} ((r+s)a_i) x^i = \sum_{i=0}^{n} (ra_i + sa_i) x^i = \sum_{i=0}^{n} (ra_i) x^i + \sum_{i=0}^{n} (sa_i) x^i = rf(x) + sf(x).$$

(7) 對任意 $r \in \mathbb{R}$ 以及 $f(x) = \sum_{i=0}^{n} a_i x^i$, $g(x) = \sum_{i=0}^{n} b_i x^i \in P(\mathbb{R})$ 皆有

$$r(f(x) + g(x)) = r(\sum_{i=0}^{n} (a_i + b_i) x^i) = \sum_{i=0}^{n} (r(a_i + b_i)) x^i = \sum_{i=0}^{n} (ra_i + rb_i) x^i = rf(x) + rg(x).$$

(8) 對任意 $f(x) = \sum_{i=0}^{n} a_i x^i \in P(\mathbb{R})$, 皆有

$$1f(x) = \sum_{i=0}^{n} (1a_i) x^i = \sum_{i=0}^{n} a_i x^i = f(x).$$

因為 $P(\mathbb{R})$ 的加法與係數積符合 vector space 的 8 項運算規則, 所以在這個加法與係數積的運算之下 $P(\mathbb{R})$ 是一個 vector space.

(C) 考慮 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 為所有 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformations 所成的集合. 我們定義過 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 的加法與係數積, 並知在加法與係數積之下 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 是封閉的 (參見 Proposition 5.1.6). 由於 linear transformation 的加法及係數積的定義是對定義域裡的每一個元素代入後定義的, 所以由 $\mathbb{R}^m$ 是 vector space, 我們很容易證明 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 的加法與係數積符合 vector space 的 8 項運算規則. 例如對於 (2), 假設 $T_1, T_2, T_3 \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. 對於所有 $\mathbf{v} \in \mathbb{R}^n$, 我們有

$$((T_1 + T_2) + T_3)(\mathbf{v}) = (T_1 + T_2)(\mathbf{v}) + T_3(\mathbf{v}) = (T_1(\mathbf{v}) + T_2(\mathbf{v})) + T_3(\mathbf{v}),$$

而

$$(T_1 + (T_2 + T_3))(\mathbf{v}) = T_1(\mathbf{v}) + (T_2 + T_3)(\mathbf{v}) = T_1(\mathbf{v}) + (T_2(\mathbf{v}) + T_3(\mathbf{v})).$$

故利用 $\mathbb{R}^m$ 中的向量加法符合結合律, 得 $((T_1 + T_2) + T_3)(\mathbf{v}) = (T_1 + (T_2 + T_3))(\mathbf{v})$. 由於這是對所有 $\mathbf{v} \in \mathbb{R}^n$ 皆成立的, 故由函數相等的定義得 $(T_1 + T_2) + T_3 = T_1 + (T_2 + T_3)$. 另外要說明的是 (3) 中的 zero element 指的就是 zero mapping, 即 $T : \mathbb{R}^n \to \mathbb{R}^m$, 滿足 $T(\mathbf{v}) = \mathbf{O}$, $\forall \mathbf{v} \in \mathbb{R}^n$ 這樣的 linear transformation. 由於 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 的加法與係數積符合 vector space 的 8 項運算規則, 因此在這個加法與係數積的運算之下 $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ 是一個 vector space.

在 $\mathbb{R}^n$ 的情況, 我們介紹過 $\mathbb{R}^n$ 的非空子集合若在 $\mathbb{R}^n$ 的加法及係數積運算之下為 vector space, 便稱為 $\mathbb{R}^n$ 的 subspace. 同樣的對於一般 vector space 的非空子集合, 如果在此 vector space 的加法及係數積運算之下這個子集合亦為 vector space, 則稱為此 vector space 的 *subspace*.

Definition 6.1.5. 假設 V 為 vector space, W 為 V 的 nonempty subset. 若在 V 的加法及係數積運算之下 W 亦為 vector space, 則稱 W 為 V 的 *subspace*.

雖然一個 vector space 的 subspace 仍為 vector space, 但要檢查是否為 subspace 不必像檢查 vector space 一樣要去檢查 8 項的運算規則. 這是因為原本 vector space 的 8 項運算規則中除了 (3)(4) 兩項會和所在的集合有關外, 其他各項僅是元素間的運算規則, 和所在的集合無關. 所以我們有和 $\mathbb{R}^n$ 的 subspace 同樣的性質 (參見 Proposition 4.1.2).

Proposition 6.1.6. 假設 V 為 *vector space* 且 W 為 V 的子集合. 則 W 為 V 的 *subspace* 若且唯若 $\mathbf{O} \in W$ 且對任意 $\mathbf{u}, \mathbf{v} \in W$, $r \in \mathbb{R}$ 皆有 $\mathbf{u} + r\mathbf{v} \in W$.

Proof. $(\Rightarrow)$: 依 subspace 的定義, 加法及係數積皆有封閉性, 故對任意 $\mathbf{u}, \mathbf{v} \in W$, $r \in \mathbb{R}$ 皆有 $\mathbf{u} + r\mathbf{v} \in W$. 又依定義 W 為非空集合, 故必存在一向量 $\mathbf{w} \in W$. 現考慮 $0\mathbf{w}$, 依封閉性 $0\mathbf{w} \in W$. 又因 V 為 vector space, 我們知 $0\mathbf{w} = \mathbf{O}$ (Proposition 6.1.3(2)). 故得證 $\mathbf{O} = 0\mathbf{w} \in W$.

$(\Leftarrow)$: 由 $\mathbf{O} \in W$, 我們知 W 為 V 的非空子集合. 現由對任意 $\mathbf{u}, \mathbf{v} \in W$, $r \in \mathbb{R}$ 皆有 $\mathbf{u} + r\mathbf{v} \in W$, 表示 V 的加法和係數積對 W 的元素有封閉性, 亦即這兩個運算也可視為 W 上的加法與係數積. 所以要證明 W 為 V 的 subspace, 我們須證明此加法與係數積在 W 上符合 vector space 的 8 項的運算規則. 然而除了 (3)(4) 兩項以外, 其他各項對於所有 V 的元素都成立. 而 W 為 V 的子集合, W 的元素皆為 V 的元素, 因此知這幾項對於所有 W 的元素皆成立. 又已知 $\mathbf{O} \in W$, 所以 (3) 成立. 而對於任意 $\mathbf{w} \in W$, 由係數積封閉性知 $(-1)\mathbf{w} \in W$, 又因 $\mathbf{w} + (-1)\mathbf{w} = 0\mathbf{w} = \mathbf{O}$, 得證 (4) 也成立. $\square$

由 Proposition 6.1.6, 我們知道要檢查一個 vector space V 中的子集合 W 是否為 V 的 subspace, 我們僅要檢查

(1) $\mathbf{O} \in V$

(2) $\mathbf{u}, \mathbf{v} \in V$, $r \in \mathbb{R} \Rightarrow \mathbf{u} + r\mathbf{v} \in V$.

是否成立即可. 我們看以下的例子.

Example 6.1.7. (A) 考慮 $\mathscr{M}_{n\times n}$, 即所有 $n \times n$ 方陣所成的 vector space. 我們想知道 $\mathscr{M}_{n\times n}$ 中所有的 symmetric matrices (對稱矩陣) 所成的集合是否為 $\mathscr{M}_{n\times n}$ 的 subspace. 首先回顧 $A \in \mathscr{M}_{n\times n}$ 為 symmetric matrix, 表示 $A^{\mathrm{T}} = A$. 很明顯的 $n \times n$ 階零矩陣 $\mathbf{O}$ 為 symmetric matrix. 而若 $A, B \in \mathscr{M}_{n\times n}$ 滿足 $A^{\mathrm{T}} = A, B^{\mathrm{T}} = B$, 則對任意 $r \in \mathbb{R}$, 我們有 $(A + rB)^{\mathrm{T}} = A^{\mathrm{T}} + (rB)^{\mathrm{T}} = A + rB$ (參見 Proposition 3.2.4). 亦即 $A + rB$ 亦為 symmetric matrix, 得證 $\mathscr{M}_{n\times n}$ 中所有的 symmetric matrices 所成的集合為 $\mathscr{M}_{n\times n}$ 的 subspace.

$\mathscr{M}_{n\times n}$ 中所有的 invertible matrices (可逆矩陣) 所成的集合是否為 $\mathscr{M}_{n\times n}$ 的 subspace 呢? 答案是否定的. 很明顯的零矩陣 $\mathbf{O}$ 就不是 invertible, 所以由 $\mathbf{O}$ 不在其中就可得 $\mathscr{M}_{n\times n}$ 中所有的 invertible matrices 所成的集合不是 $\mathscr{M}_{n\times n}$ 的 subspace. 其實即使我們考慮 invertible matrices 所成的集合與 $\{\mathbf{O}\}$ 的聯集, 仍不會是 $\mathscr{M}_{n\times n}$ 的 subspace. 因為即使此時 $\mathbf{O}$ 在其中, 但仍有可能兩個 invertible matrices 相加後就不是 invertible. 例如在 2×2 的情形, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 和 $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 皆為 invertible, 但是 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}+\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 不是 invertible.

(B) 考慮 $P(\mathbb{R})$, 即所有以 x 為變數的實係數多項式所成的 vector space. 給定一自然數 $n\in\mathbb{N}$, 我們說明所有次數小於等於 n 的多項式所成的集合 $P_n(\mathbb{R})$ 是 $P(\mathbb{R})$ 的 subspace. 首先我們可以將 $P_n(\mathbb{R})$ 寫成 $P_n(\mathbb{R})=\{\sum_{i=0}^{n} a_i x^i \mid a_i\in\mathbb{R}\}$. 很明顯的零多項式屬於 $P_n(\mathbb{R})$ (注意一般數學上定義零多項式的次數為 $-\infty$, 而不是 0. 這個部分以後代數課程會去談論, 這裡就不多談). 又若 $f(x)=\sum_{i=0}^{n} a_i x^i, g(x)=\sum_{i=0}^{n} b_i x^i\in P_n(\mathbb{R})$, 則對任意 $r\in\mathbb{R}$, 我們有 $f(x)+rg(x)=\sum_{i=0}^{n}(a_i+rb_i)x^i\in P_n(\mathbb{R})$. 故知 $P_n(\mathbb{R})$ 為 $P(\mathbb{R})$ 的 subspace. 要注意, 若僅考慮次數等於 n 的多項式所成的集合, 那麼就不會是 $P(\mathbb{R})$ 的 subspace 了. 很明顯的零多項式就不會在裡面. 又即使加入零多項式, 但仍有可能兩個次數為 n 的多項式相加之後其次數變小了, 例如 $(x^2+x+1)+(-x^2+x+1)=2x+2$. 也就是說在這情況之下加法是不封閉的, 所以無法成為一個 vector space.

(C) 考慮 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 為所有 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformations 所成的 vector space. 給定 $\mathbb{R}^n$ 的一個 subspace V, 我們考慮

$$N(V)=\{T\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)\mid V\subseteq\ker(T)\}.$$

亦即 $N(V)$ 裡的元素皆為滿足對任意 $\mathbf{v}\in V$, $T(\mathbf{v})=\mathbf{O}$ 的 linear transformation T. 很明顯的, $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 裡的 additive identity, 即 zero mapping, 因其 kernel 為 $\mathbb{R}^n$ 故由 $V\subseteq\mathbb{R}^n$ 知 zero mapping 屬於 $N(V)$. 現若 $T_1,T_2\in N(V)$, 則對於任意 $r\in\mathbb{R}$, 由於對所有 $\mathbf{v}\in V$ 皆有

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

得證 $V\subseteq\ker(T_1+rT_2)$, 因此 $T_1+rT_2\in N(V)$. 得證 $N(V)$ 為 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 的 subspace.

Question 6.2. 考慮 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 為所有 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *linear transformations* 所成的 *vector space*. 給定 $\mathbb{R}^n$ 的一個 *subspace* V, 我們考慮

$$S(V)=\{T\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)\mid\ker(T)\subseteq V\}.$$

是否 $S(V)$ 為 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 的 *subspace?*

給定 $\mathbb{R}^m$ 的一個 *subspace* W, 我們考慮

$$R(W)=\{T\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)\mid T(\mathbb{R}^n)\subseteq W\}.$$

是否 $R(W)$ 為 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 的 *subspace?* 另外若考慮

$$I(W)=\{T\in\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)\mid W\subseteq T(\mathbb{R}^n)\}.$$

是否 $I(W)$ 為 $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ 的 *subspace?*

6.2. Basis and Dimension

在這一節中, 我們將介紹 vector space 中 basis 的概念, 並介紹 finitely generated vector space 中 dimension 之概念及性質.

我們在介紹 $\mathbb{R}^n$ 的 subspace 時, 引進了 basis 的概念. 有了 basis, 表示此 subspace 中的元素, 都可以用這一組 basis 的線性組合唯一表示. 這又等同於這組 basis 是此 subspace 的 spanning vectors 且為 linearly independent (參見 Proposition 4.1.8). 我們可將此概念推廣到一般的 vector space.

Definition 6.2.1. 假設 V 為 vector space 且 $\mathbf{v}_1,\dots,\mathbf{v}_n \in V$. 對於任意 $c_1,\dots,c_n \in \mathbb{R}$, 我們稱 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 為 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 的 *linear combination*. 所有 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 的 linear combination 所成的集合, 我們用 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 來表示, 亦即 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=\{\sum_{i=1}^n c_i\mathbf{v}_i \mid c_1,\dots,c_n \in \mathbb{R}\}$.

我們知道 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 會是 V 的 subspace (參見 Proposition 1.3.2). 事實上它是包含 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 最小的 subspace. 這是因為若 W 是 V 的 subspace 且 $\mathbf{v}_1,\dots,\mathbf{v}_n \in W$, 則由 subspace 的加法與係數積的封閉性得 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)\subseteq W$. 特別的, 當 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=V$ 時, 我們給它一個特別的名稱.

Definition 6.2.2. 假設 V 為 vector space. 若 $\mathbf{v}_1,\dots,\mathbf{v}_n \in V$ 滿足 $\mathrm{Span}(\mathbf{v}_1,\dots,\mathbf{v}_n)=V$, 則稱 $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ 為 V 的一組 *spanning set*. 此時我們稱 V 為 *finitely generated vector space*.

為何特別稱之為 finitely generated vector space 呢? 這表示此 vector space 中所有的元素都可表示成固定有限多個元素的線性組合. 例如 $\mathbb{R}^n$ 就是 finitely generated (可由 $\mathbf{e}_1,\dots\mathbf{e}_n$ 所展成). 由於當初我們僅談 $\mathbb{R}^n$ 及其 subspaces, 它們都是 finitely generated vector space 所以不特別去強調. 現在我們談論一般的 vector space, 有可能不是 finitely generated, 所以要區分出來, 我們看以下的例子.

Example 6.2.3. 我們討論前面提的一些 vector space 哪些是 finitely generated vector space.

(A) $\mathscr{M}_{m\times n}$ 是 finitely generated. 因為考慮 $M_{i,j}\in\mathscr{M}_{m\times n}$, 表示 (i,j)-th entry 為 1, 其他 entry 為 0 的 $m\times n$ matrix. 很容易看出所有的 $m\times n$ matrix 皆可寫成 $M_{i,j}$ 其中 $1\leq i\leq m$, $1\leq j\leq n$ 的 linear combination. 所以 $\{M_{i,j} \mid 1\leq i\leq m, 1\leq j\leq n\}$ 是 $\mathscr{M}_{m\times n}$ 的 spanning set, 也因此 $\mathscr{M}_{m\times n}$ 是 finitely generated vector space.

(B) $P(\mathbb{R})$ 不是 finitely generated vector space. 這是因為如果 $\{f_1(x),\dots,f_n(x)\}$ 為 $P(\mathbb{R})$ 的 spanning set, 假設 $f_1(x),\dots,f_n(x)$ 的最高次為 m, 則任何 $f_1(x),\dots,f_n(x)$ 的 linear combination $c_1f_1(x)+\cdots+c_nf_n(x)$ 的次數皆不可能大於 m. 也就是說 $\mathrm{Span}(f_1(x),\dots,f_n(x))$ 不可能包含次數大於 m 的多項式. 此與 $\{f_1(x),\dots,f_n(x)\}$ 為 $P(\mathbb{R})$ 的 spanning set 明顯不合, 故知 $P(\mathbb{R})$ 不可能是 finitely generated. 不過次數小於等於 n 的多項式所成的集合 $P_n(\mathbb{R})$ 就是 finitely generated vector space. 很容易看出 $\{x^n,\dots,x,1\}$ 就是 $P_n(\mathbb{R})$ 的 spanning set.

大家或許直覺會認為 finite generated vector space 的 subspace 一定也是 finitely generated. 這是對的, 不過證明卻不是如直覺那麼簡單 (大家不妨現在試著證明看看). 它的證明等到介紹完 linearly independence 的概念, 我們就可以處理.

Spanning set 的概念是處理 linear combination 的存在性, 而 linear independence 的概念便是處理 linear combination 的唯一性. 我們可推廣 $\mathbb{R}^n$ 中向量 linearly independent 的定義到一般的 vector space.

Definition 6.2.4. 假設 V 為 vector space 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$. 若存在 $c_1,\ldots,c_n \in \mathbb{R}$ 不全為 0 使得 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{O}$, 則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly dependent*. 反之, 若只有在 $c_1,\ldots,c_n$ 皆為 0 的情況才會使得 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{O}$, 則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 *linearly independent*.

如同在 $\mathbb{R}^n$ 的情形, $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent, 可以確保選取不同的 $c_1,\ldots,c_n$ 就會得到不同的 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 也就是說每一個不同的 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination 都會有不同的結果. 反之, 如果 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly dependent, 假設 $d_1,\ldots,d_n$ 不全為 0 使得 $d_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{O}$. 則對任意 $c_1,\ldots,c_n\in\mathbb{R}$, $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 和 $(c_1+d_1)\mathbf{v}_1+\cdots+(c_n+d_n)\mathbf{v}_n$ 是 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 不同的 linear combination, 但它們是表示同樣的元素. 這是因為我們有

$$\begin{aligned}&(c_1+d_1)\mathbf{v}_1+\cdots+(c_n+d_n)\mathbf{v}_n=\\ &\quad(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)+(d_1\mathbf{v}_1+\cdots+d_n\mathbf{v}_n)=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n+\mathbf{O}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n.\end{aligned}$$

Example 6.2.5. 很容易看出在 $P(\mathbb{R})$ 中 $x^n,x^{n-1},\ldots,x,1$ 為 linearly independent. 這是因為如果 $c_n,\ldots,c_1,c_0\in\mathbb{R}$ 使得 $c_nx^n+\cdots+c_1x+c_01$ 為零多項式, 則依零多項式的定義, $c_n,\ldots,c_1,c_0$ 必全為 0. 現在我們介紹 $P(\mathbb{R})$ 中另一種重要的 linearly independent 的多項式的建構方法, 稱為 *Lagrange interpolation polynomials*. 我們僅舉出一個例子, 一般狀況請大家自行推廣.

給定 a,b,c 三相異實數, 我們希望找到三個二次多項式 $p_1(x),p_2(x),p_3(x)$ 滿足

$$p_1(a)=1,p_1(b)=p_1(c)=0,\quad p_2(b)=1,p_2(a)=p_2(c)=0\quad\text{and}\quad p_3(c)=1,p_3(a)=p_3(b)=0.$$

由於 $p_1(b)=p_1(c)=0$, 我們知 $p_1(x)$ 應為 $(x-b)(x-c)$ 的倍式, 也就是存在實數 r 使得 $p_1(x)=r(x-b)(x-c)$. 但又要求 $p_1(a)=1$, 故代入 $x=a$ 得 $r=1/(a-b)(a-c)$. 同理可求出 $p_2(x),p_3(x)$ 因此我們有

$$p_1(x)=\frac{(x-b)(x-c)}{(a-b)(a-c)},\quad p_2(x)=\frac{(x-a)(x-c)}{(b-a)(b-c)}\quad\text{and}\quad p_3(x)=\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

我們要說明 $p_1(x),p_2(x),p_3(x)$ 為 linearly independent. 首先觀察, 若 $f(x)=c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$, 則代入 $x=a$ 時可由 $p_1(a)=1,p_2(a)=p_3(a)=0$, 得 $f(a)=c_1$. 同理知 $f(b)=c_2,f(c)=c_3$. 因此現若 $f(x)$ 為零多項式, 由 $f(a)=f(b)=f(c)=0$, 可得 $c_1=c_2=c_3=0$. 也就是說只有當 $c_1=c_2=c_3=0$ 時才會使得 $c_1p_1(x)+c_2p_2(x)+c_3p_3(x)$ 為零多項式, 得證 $p_1(x),p_2(x),p_3(x)$ 為 linearly independent.

我們可以將上面的討論推廣. 對於任意 n 個相異實數 $a_1,\ldots,a_n$, 我們都可以得到 n 個 $n-1$ 次的多項式 $p_1(x),\ldots,p_n(x)$ 滿足 $p_i(a_i)=1$ 且當 $j\neq i$ 時, $p_i(a_j)=0$. 利用前面的論述, 我們可以得到 $p_1(x),\ldots,p_n(x)$ 為 linearly independent.

Spanning set 和 linearly independent 性質看似無關, 不過它們之間卻存有某種程度的互補關係. 我們將它們的關係列出如下.

Lemma 6.2.6. 假設 V 為 *vector space* 且令 $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$.

(1) $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}) \neq \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_{n-1},\mathbf{v}_n)$ 若且唯若 $\mathbf{v}_{n+1} \notin \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$.

(2) 若已知 $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}$ 為 *linearly independent,* 則 $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1},\mathbf{v}_n$ 為 *linearly independent* 若且唯若 $\mathbf{v}_n \notin \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_{n-1})$.

(3) 假設 $\mathbf{w}_1,\ldots,\mathbf{w}_m \in V$. 若 $\mathbf{w}_1,\ldots,\mathbf{w}_m \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 且 $m>n$, 則 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 *linearly dependent.*

Proof. 其實 (1), (2), (3) 分別就是 Lemma 4.2.1, Lemma 4.2.4 和 Lemma 4.2.5 在一般 vector space 的情形. 由於 (1), (2) 的證明和 Lemma 4.2.1, Lemma 4.2.4 一致. 所以我們僅證明 (3) 即可. 在 Lemma 4.2.5 的證明方式用到了 $\mathbb{R}^m$ 向量的坐標表示法, 我們現在談論的是一般 vector space, 所以要稍加修改.

由於 $\mathbf{w}_1,\ldots,\mathbf{w}_m \in \mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 因此對任意 $j=1,\ldots,m$, $\mathbf{w}_j$ 都可以寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination. 也就是說, 存在 $a_{1,j},\ldots,a_{i,j},\ldots,a_{n,j} \in \mathbb{R}$ 使得

$$\mathbf{w}_j = a_{1,j}\mathbf{v}_1+\cdots+a_{i,j}\mathbf{v}_i+\cdots+a_{n,j}\mathbf{v}_n.$$

現在我們要找到 $c_1,\ldots,c_m \in \mathbb{R}$ 不全為 0 使得 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m = \mathbf{O}$, 便證得 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 linearly dependent. 現將 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m$ 中每一個 $\mathbf{w}_j$ 換成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination 後會等於

$$(c_1a_{1,1}+\cdots+c_ma_{1,m})\mathbf{v}_1+\cdots+(c_1a_{i,1}+\cdots+c_ma_{i,m})\mathbf{v}_i+\cdots+(c_1a_{m,1}+\cdots+c_ma_{m,m})\mathbf{v}_m. \quad (6.1)$$

因此若我們能找到 $c_1,\ldots,c_m \in \mathbb{R}$ 使得式子 (6.1) 中每個 $\mathbf{v}_i$ 的係數等於 0, 便可得到 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m = \mathbf{O}$. 因此我們只要找到聯立方程組

$$\begin{cases} a_{1,1}x_1+\cdots+a_{1,m}x_m &= 0 \\ \qquad\vdots & \\ a_{i,1}x_1+\cdots+a_{i,m}x_m &= 0 \\ \qquad\vdots & \\ a_{n,1}x_1+\cdots+a_{n,m}x_m &= 0 \end{cases}$$

的一組解 $x_1=c_1,\ldots,x_m=c_m$, 就可以使得 $c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m=\mathbf{O}$. 然而這個 homogeneous linear system 的方程式個數 n 少於未知數個數 m, 由 Corollary 3.4.7 知必存在不全為 0 的 $c_1,\ldots,c_m \in \mathbb{R}$ 使得 $x_1=c_1,\ldots,x_m=c_m$ 為其一組解. 故得證 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 為 linearly dependent. □

現在我們可以回答 finitely generated vector space 的 subspace 也是 finitely generated.

Proposition 6.2.7. 假設 V 為 *finitely generated vector space.* 若 W 為 V 的 *subspace,* 則 W 為 *finitely generated vector space.*

Proof. 依 V 為 finitely generated 的假設, 存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$ 滿足 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)=V$. 由於 $\{\mathbf{O}\}=\mathrm{Span}(\mathbf{O})$ 為 finitely generated, 我們僅需要考慮 $W\neq\{\mathbf{O}\}$ 的情況. 我們用反證法, 假設 W 不是 finitely generated. 現任取 $\mathbf{w}_1\in W$ 其中 $\mathbf{w}_1\neq\mathbf{O}$. 由於 W 不是 finitely generated, 我們知 $\mathrm{Span}(\mathbf{w}_1)\neq W$, 亦即存在 $\mathbf{w}_2\in W$ 且 $\mathbf{w}_2\notin\mathrm{Span}(\mathbf{w}_1)$. 由 Lemma 6.2.6 (2) 知 $\mathbf{w}_1,\mathbf{w}_2$ 為 linearly independent. 同理, 因 W 不是 finitely generated, 我們知 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)\neq W$ 亦即存在 $\mathbf{w}_3\in W$ 且 $\mathbf{w}_3\notin\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)$. 由 Lemma 6.2.6 (2) 知 $\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3$ 為 linearly independent. 這樣一直下去, 利用數學歸納法假設我們得 $\mathbf{w}_1,\ldots,\mathbf{w}_k\in W$ 為 linearly independent. 由於 $\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k)\neq W$, 存在 $\mathbf{w}_{k+1}\in W$ 且 $\mathbf{w}_{k+1}\notin\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k)$. 因此再由 Lemma 6.2.6 (2) 知 $\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+}$ 為 linearly independent. 我們利用數學歸納法證明了, 若 W 不是 finitely generated, 則對任意 $m\in\mathbb{N}$, 皆存在 $\mathbf{w}_1,\ldots,\mathbf{w}_m\in W$ 為 linearly independent. 然而這在 $m>n$ 是會造成矛盾的. 因為此時由於 $\mathbf{w}_1,\ldots,\mathbf{w}_m\in W\subseteq V=\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, Lemma 6.2.6 (3) 告訴我們 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 必為 linearly dependent. 因此知 W 為 finitely generated vector space. □

Question 6.3. 利用在 $P(\mathbb{R})$ 中對於任意 $n\in\mathbb{N}$, $x^n,x^{n-1},\ldots,x,1$ 為 *linearly independent,* 證明 $P(\mathbb{R})$ 不是 *finitely generated vector space.*

有了 spanning set 的存在性以及 linearly independent 的唯一性, 我們知道若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 vector space V 的 spanning set 且為 linearly independent, 則所有 V 中的元素都可以唯一寫成 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的 linear combination, 也因此我們有以下的定義.

Definition 6.2.8. 假設 V 為 vector space. 若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 spanning set 且為 linearly independent, 則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 *basis.*

接下來我們要說明非零的 finitely generated vector space 都會有 basis. 其實這個性質對於不是 finitely generated 的 vector space 也對, 不過由於這會牽涉到較抽象的邏輯概念而且我們以後談論的 vector space 都是 finitely generated, 所以我們不去談論它.

Theorem 6.2.9. 假設 $V\neq\{\mathbf{O}\}$ 為 *finitely generated vector space.* 則存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$ 為 V 的一組 *basis.* 又若 $\mathbf{w}_1,\ldots,\mathbf{w}_m\in V$ 亦為 V 的一組 *basis,* 則 $m=n$.

Proof. 這個定理是 Theorem 4.3.1 和 Theorem 4.3.2 在一般 vector space 的推廣. 我們可以利用 Theorem 4.3.1 的方法處理 basis 的存在性. 不過這裡我們利用 finitely generated 的性質, 直接用數學歸納法處理.

我們對 vector space 的 spanning set 的元素個數做數學歸納法. 首先假設 V 可由一個元素展成. 也就是說 $V=\mathrm{Span}(\mathbf{u})$. 此時由於 $V\neq\{\mathbf{O}\}$, 知 $\mathbf{u}\neq\mathbf{O}$. 故 $\mathbf{u}$ 本身是 linearly independent 且 $\{\mathbf{u}\}$ 是 V 的 spanning set 故 $\mathbf{u}$ 是 V 的 basis. 利用納法假設當一個 vector space 可由 k 個元素展成時, basis 是存在的. 現假設 V 是一個可由 $k+1$ 個元素展成的 vector space, 我們假設 $V=\mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_k,\mathbf{u}_{k+1})$. 現考慮 $W=\mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_k)$. 由於 W 是一個可由 k 個元素展成的 vector space, 依歸納法假設存在 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in W$ 為 W 的一組 basis. 亦即 $\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_n)=W$ 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent. 現若 $W=V$, 當然 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 就是 V 的一組 basis. 而若 $W\neq V$, 表示 $\mathbf{u}_{k+1}\notin W$ (否則會造成 $V=\mathrm{Span}(\mathbf{u}_1,\ldots,\mathbf{u}_k,\mathbf{u}_{k+1})\subseteq W$ 之矛

盾). 因此由 $\mathbf{u}_{k+1} \notin W = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 以及 Lemma 6.2.6 (2) 知 $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1}$ 為 linearly independent. 又因為 $V = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1})$ 而 $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = W = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 故得 $V = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1})$. 得證 $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_{k+1}$ 為 V 的一組 basis.

對於組成一組 basis 的元素個數是唯一的證明, 由於 Lemma 4.2.5 以及 Theorem 4.3.2 的證明只需用到 vector space 的假設, 並不需用到 $\mathbb{R}^m$ 的假設. 因此可以直接套用其證明, 我們就不再重複了. □

Theorem 6.2.9 告訴我們組成 V 的一組 basis 的元素個數是固定的. 也就是說若找到 n 個元素形成 V 的 basis, 則 V 其他的 basis 一定也會是由 n 個元素所組成. 由於這個重要的結果我們有以下的定義.

Definition 6.2.10. 假設 V 是一個 finitely generated 的 vector space. 組成 V 的一組 basis 的元素個數稱為 V 的 *dimension* (維度), 用 $\dim(V)$ 來表示.

由於組成 finitely generated vector space 的一組 basis 的元素個數是有限的, 所以一般我們也稱 finitely generated vector space 為 *finite dimensional vector space.*

Example 6.2.11. 我們探討在 Example 6.2.3 中的 finite dimensional vector space 的維度為多少.

(A) 在 Example 6.2.3 (A) 我們知道 $\{M_{i,j} \in \mathscr{M}_{m\times n} \mid 1 \le i \le m, 1 \le j \le n\}$ 是 $\mathscr{M}_{m\times n}$ 的 spanning set. 很容易檢查它們為 linearly independent, 所以 $\{M_{i,j} \mid 1 \le i \le m, 1 \le j \le n\}$ 是 $\mathscr{M}_{m\times n}$ 的一組 basis. 因此 $\dim(\mathscr{M}_{m\times n}) = m \times n$.

(B) 我們知道 $\{x^n, \ldots, x, 1\}$ 是 $P_n(\mathbb{R})$ 的 spanning set 且為 linearly independent. 故知 $x^n, \ldots, x, 1$ 為 $P_n(\mathbb{R})$ 的一組 basis, 因此 $\dim(P_n(\mathbb{R})) = n+1$.

對於 finite dimensional vector space 有關於 dimension 的性質, 我們匯集如下. 相關的證明請參考 Proposition 4.3.4, Corollary 4.3.6.

Proposition 6.2.12. 假設 V 為 *finite dimensional vector space.*

(1) 若 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 V 的 *spanning set,* 則 $\dim(V) \le n$. 特別的, 若此時 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 為 *linearly dependent,* 則 $\dim(V) < n$.

(2) 若 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ 為 *linearly independent,* 則 $\dim(V) \ge n$. 特別的, 若此時 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 不是 V 的 *spanning set,* 則 $\dim(V) > n$.

(3) 若 W 為 V 的 *subspace,* 則 $\dim(W) \le \dim(V)$. 此時 $\dim(W) = \dim(V)$ 若且唯若 $V = W$.

(4) 假設 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$. 下列的敘述為等價.

(a) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 為 V 的一組 *basis.*

(b) $\dim(V) = n$ 且 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 V 的 *spanning set.*

(c) $\dim(V) = n$ 且 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 為 *linearly independent.*

強調一下, Proposition 6.2.12 (4) 告訴我們當要檢查 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是否為 V 的一組 basis 時, 若我們知道 $\dim(V)$ 恰好是 n, 則僅要檢查它們是否為 spanning set 或 linearly independent 其中一項就可.

Example 6.2.13. 在 Example 6.2.5 中給定 a,b,c 三相異實數, 我們找到三個二次多項式

$$p_1(x)=\frac{(x-b)(x-c)}{(a-b)(a-c)},\quad p_2(x)=\frac{(x-a)(x-c)}{(b-a)(b-c)}\quad\text{and}\quad p_3(x)=\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

滿足

$$p_1(a)=1,p_1(b)=p_1(c)=0;\quad p_2(b)=1,p_2(a)=p_2(c)=0;\quad p_3(c)=1,p_3(a)=p_3(b)=0.$$

我們知道 $p_1(x),p_2(x),p_3(x)\in P_2(\mathbb{R})$ 為 linearly independent 且 $\dim(P_2(\mathbb{R}))=3$, 故由 Proposition 6.2.12 (4) 知 $p_1(x),p_2(x),p_3(x)$ 為 $P_2(\mathbb{R})$ 的一組 basis.

同理對於任意 n 個相異實數 $a_1,\ldots,a_n$, 我們有 n 個 $n-1$ 次的多項式 $p_1(x),\ldots,p_n(x)$ 滿足 $p_i(a_i)=1$ 且當 $j\neq i$ 時, $p_i(a_j)=0$. 由於 $p_1(x),\ldots,p_n(x)\in P_{n-1}(\mathbb{R})$ 且為 linearly independent, 故由 $\dim(P_{n-1}(\mathbb{R}))=n$ 知 $p_1(x),\ldots,p_n(x)$ 為 $P_{n-1}(\mathbb{R})$ 的一組 basis.

6.3. Linear Transformation

我們將介紹在一般 vector space 之間的 linear transformation 及其相關性質. 如同在 $\mathbb{R}^n$ 的情形, 要探討兩個 vector spaces 之間的關係, 我們需要的函數是能保持 vector space 的加法及係數積運算的, 所以有以下之定義.

Definition 6.3.1. 假設 V,W 為 vector spaces, $T:V\to W$ 為一個函數. 若 T 滿足對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k\in V$ 以及 $c_1,\ldots,c_k\in\mathbb{R}$ 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 中的 linear combination, 而 $c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$ 是 W 中的 linear combination, 要區分清楚. 特別是當 $\mathbf{O}\in V$ 為 V 的 zero element 時, 依 linear transformation 的定義, 我們有 $T(\mathbf{O})=T(\mathbf{O}+\mathbf{O})=T(\mathbf{O})+T(\mathbf{O})$. 此時兩邊加上 $T(\mathbf{O})$ 的加法反元素, 得 $T(\mathbf{O})$ 應為 W 中的 zero element. 也就是說一個 linear transformation $T:V\to W$, 會將 V 中的 zero element 映射到 W 中的 zero element. 雖然當 $V\neq W$ 時, 它們的 zero element 是不同的, 不過一般我們都還是用 $\mathbf{O}$ 來表示, 而不區分它. 所以我們仍用 $T(\mathbf{O})=\mathbf{O}$ 來表示.

依定義要檢查 $T:V\to W$ 是否為 linear transformation, 我們必須考慮 V 中任意的 linear combination 代入 T 中是否符合 linear transformation 的要求. 不過就如同 $\mathbb{R}^n$ 的情況 (參見 Proposition 5.1.3), 我們只要檢查任兩個元素的 linear combination 即可.

Proposition 6.3.2. 假設 V,W 為 *vector spaces* 且 $T:V\to W$ 為一個函數. 則 T 為 *linear transformation* 若且唯若對任意 $\mathbf{u},\mathbf{v}\in V$, $r\in\mathbb{R}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$.

當 T_1, T_2 皆為 V 到 W 的 linear transformation 時, 我們可以定義 T_1, T_2 之間的加法 $T_1 + T_2 : V \to W$, 為 $(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v})$, $\forall \mathbf{v} \in V$. 我們也可定義 linear transformation $T : V \to W$ 的係數積. 即對於 $r \in \mathbb{R}$, 定義 $rT : V \to W$ 為 $(rT)(\mathbf{v}) = r(T(\mathbf{v}))$, $\forall \mathbf{v} \in V$. 在此定義之下 $T_1 + T_2$ 和 rT 仍為 V 到 W 的 linear transformation (參見 Proposition 5.1.6). 事實上這樣的加法及係數積運算會符合 vector space 的 8 項運算規則, 其中 (3) 的 additive identity 就是 V 到 W 的 zero mapping (就是所有 V 中的元素映射至 W 的 $\mathbf{O}$). 所以我們有以下的結論 (證明從略).

Proposition 6.3.3. *假設 V, W 為 vector spaces, 令 $\mathscr{L}(V, W)$ 為所有 V 到 W 的 linear transformation 所成的集合. 則利用函數的加法及係數積的定義, $\mathscr{L}(V, W)$ 是一個 vector space.*

另一個產生 linear transformation 的方法就是利用“合成函數”. 假設 U, V, W 為 vector spaces, 若 $T : V \to W$ 和 $T' : W \to U$ 為 linear transformations, 我們定義 $T' \circ T : V \to U$ 為 $T' \circ T(\mathbf{v}) = T'(T(\mathbf{v}))$, $\forall \mathbf{v} \in V$. 我們有下面之結果 (證明請參見 Proposition 5.1.7).

Proposition 6.3.4. *假設 U, V, W 為 vector spaces. 若 $T : V \to W$ 和 $T' : W \to U$ 為 linear transformations, 則 $T' \circ T : V \to U$ 亦為 linear transformation.*

對於 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation 會有 standard matrix representation 一個主要的關鍵是 linear transformation 可以由定義域的一組 basis 唯一確定. 這對於一般的 linear transformation 也是對的 (證明請參見 Theorem 5.1.8).

Theorem 6.3.5. *假設 V, W 為 vector spaces 且 $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, 為 V 的一組 basis. 任取 $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W$, 則存在唯一的 linear transformation $T : V \to W$, 滿足對於任意 $i = 1, \ldots, n$ 皆有 $T(\mathbf{v}_i) = w_i$.*

由於 linear transformation 保持加法及係數積的特性, 它會保持 subspace 之間的關係. 也就是說當 $T : V \to W$ 為 linear transformation 時, 對於 V 的 subspace V' 我們考慮

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

同樣的若 W' 為 W 的 subspace, 我們也考慮

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

我們會有以下之結果 (證明請參見 Proposition 5.2.1).

Proposition 6.3.6. *假設 V, W 為 vector spaces 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.*

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

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

這兩個 subspaces, 對我們了解 T 這個 linear transformation 非常有幫助. 如同在 $\mathbb{R}^n$ 的情形我們有以下之定義.

Definition 6.3.7. 假設 V,W 為 vector spaces 且 $T:V\to W$ 為 linear transformation. 我們稱 W 的 subspace $T(V)$ 為 T 的 *range* (有時也稱為 *image*). 我們稱 V 的 subspace $T^{-1}(\{\mathbf{O}\})$ 為 T 的 *kernel*, 通常用 $\ker(T)$ 來表示.

依定義, 一個 linear transformation $T:V\to W$ 是 onto 就表示其 range $T(V)$ 是整個 W. 要如何得知 T 的 range 呢? 我們有以下的結果 (證明請參見 Proposition 5.2.4).

Proposition 6.3.8. 假設 V,W 為 *vector spaces* 且 $T:V\to W$ 為 *linear transformation.* 若 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 為 V 的 *spanning set.* 則

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

特別的, 我們有 T 為 *onto* 若且唯若 $W=\mathrm{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))$.

至於 kernel 就和是否為 one-to-one 有關了. 同樣的我們有和 Proposition 5.2.6 一致的結果.

Proposition 6.3.9. 假設 V,W 為 *vector spaces* 且 $T:V\to W$ 為 *linear transformation.* 則 T 為 *one-to-one* 若且唯若 $\ker(T)=\{\mathbf{O}\}$.

在 Proposition 6.3.8 我們知道了 $T:V\to W$ 是 onto 時會保持 spanning set. 也就是說 T 會將 V 的一組 spanning set 送到 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 6.3.10. 假設 V,W 為 *vector spaces* 且假設 $T:V\to W$ 為 *one-to-one linear transformation.* 則 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V$ 為 *linearly independent* 若且唯若 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)\in W$ 為 *linearly independent.*

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

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*. 在 Theorem 5.3.12 中我們知道 $\mathbb{R}^n$ 上的 linear transformation 如果是 invertible, 則其 inverse 亦為 linear transformation. 這在一般的情形也對, 不過由於在 Theorem 5.3.12 我們是用 standard matrix representation 來證明. 這裡我們尚未知道有沒有 matrix representation, 所以必須有另外的證明.

Theorem 6.3.11. 假設 V,W 為 *vector spaces* 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 $\mathbf{w}_1, \mathbf{w}_2 \in V$, $r \in R$, 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 $T^{-1}(\mathbf{w}_1) + rT^{-1}(\mathbf{w}_2) = \mathbf{v}_1 + r\mathbf{v}_2$ 就等同於檢查是否 $T(\mathbf{v}_1 + r\mathbf{v}_2) = \mathbf{w}_1 + r\mathbf{w}_2$. 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. □

當 $T: V \to W$ 是 invertible 的 linear transformation 時, 我們也時常稱 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 6.3.12. 設 V,W 為 *vector spaces* 且 $T: V \to W$ 為 *isomorphism.* 則 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 為 V 的一組 *basis* 若且唯若 $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ 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 且 $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 為 V 的 spanning set, 利用 Proposition 6.3.8 知 $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ 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 6.3.10 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 6.3.11), 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. □

當 V,W 為 finite dimensional vector space 時, 由 Theorem 6.3.12, 我們知若 $T: V \to W$ 為 isomorphism, 則 V 和 W 的 basis 有相同個數的組成元素, 也就是說 $\dim(V) = \dim(W)$. 換句話說若 $\dim(V) \neq \dim(W)$, 則不可能存在一個 V 到 W 的 isomorphism. 這個性質的反向也是對的, 我們有以下的定理.

Corollary 6.3.13. 設 V,W 為 *finite dimensional vector spaces.* 則存在 $T:V\to W$ 為 *isomorphism* 若且唯若 $\dim(V)=\dim(W)$.

Proof. $(\Rightarrow)$ 假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis, 即 $\dim(V)=n$. 此時若 $T:V\to W$ 為 isomorphism 由 Theorem 6.3.12 知 $T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)$ 為 W 的一組 basis. 故得 $\dim(W)=n=\dim(V)$.

$(\Leftarrow)$ 假設 $\dim(V)=\dim(W)$, 且令 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 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 6.3.5). 我們要說明 T 為 isomorphism. 首先由 6.3.8 知 $T(V)=\mathrm{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))=\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_n)$. 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 $T(V)=W$, 即 T 為 onto. 現假設 $\mathbf{v}\in\ker(T)$. 由於 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis, 故存在 $c_1,\ldots,c_n\in\mathbb{R}$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 依假設 $T(\mathbf{v})=\mathbf{O}$, 亦即

$$\mathbf{O}=T(c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n)=c_1T(\mathbf{v}_1)+\cdots+c_nT(\mathbf{v}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n.$$

然而因 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 為 linearly independent, 知 $c_1=\cdots=c_n=0$. 故得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n=\mathbf{O}$. 我們證得 $\ker(T)=\{\mathbf{O}\}$, 故由 Proposition 6.3.9 知 T 為 one-to-one, 得證 T 為 isomorphism. □

6.4. Coordinatization

這一節中, 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 $\mathbb{R}^n$ 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.* 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 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 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 $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 為 V 的一組 ordered basis, 我們就會 $\mathscr{B}$ 來表示這一組 ordered basis $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$. 對於 $\mathbb{R}^n$ 的 standard basis, 我們通常用 $\mathscr{E}$ 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 $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ 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 $\mathscr{B}$ 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 $\mathscr{B}$ 將 $\mathbf{v}$ 坐標化後所得的坐標表示法. 為了方便, 我們就用 ${}_{\mathscr{B}}[\mathbf{v}]$ 來表示利用 $\mathscr{B}$ 將 $\mathbf{v}$ 坐標化後所得的坐標. 坐標化的好處是, 我們可

以將 ${}_{\mathscr{B}}[\mathbf{v}]$ 看成是 $\mathbb{R}^n$ 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 $\mathbb{R}^n$ 中的向量來處理.

Example 6.4.1. 我們看看前面提過的幾個 vector space 坐標化的情形.

(A) 考慮 $\mathscr{M}_{2\times 2}$ 及其 ordered basis

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

(通常我們稱這一組 basis 為 $\mathscr{M}_{2\times 2}$ 的 standard basis). 對於任意 $\mathscr{M}_{2\times 2}$ 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 $\mathscr{E}$ 所得的坐標表示法為

$${}_{\mathscr{E}}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}a\\b\\c\\d\end{bmatrix}.$$

例如我們有

$${}_{\mathscr{E}}\begin{bmatrix}1&-2\\-3&4\end{bmatrix}=\begin{bmatrix}1\\-2\\-3\\4\end{bmatrix}.$$

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

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

我們也可考慮 ordered basis $\mathscr{B}=(p_1(x),p_2(x),p_3(x))$ 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 6.2.5, Example 6.2.13). 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

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

(C) 我們也可將 $\mathbb{R}^n$ 中的向量用不同的 ordered basis 坐標化. 例如在 $\mathbb{R}^3$ 中考慮 ordered basis $\mathscr{B}=(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix})$. 若要求向量 $\begin{bmatrix}1\\2\\3\end{bmatrix}$ 以 $\mathscr{B}$ 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

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

要注意這裡我們有 ${}_{\mathscr{E}}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}$, 其中 $\mathscr{E}$ 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 $\mathscr{E}$ 來將所有 $\mathbb{R}^3$ 的向量的坐標化.

給定 V 的一組 ordered basis $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, 將 V 中的元素利用 $\mathscr{B}$ 坐標化, 其實就定出了一個從 V 到 $\mathbb{R}^n$ 的函數 $T_{\mathscr{B}}:V\to\mathbb{R}^n$, 其中對任意 $\mathbf{v}\in V$, 我們有 $T_{\mathscr{B}}(\mathbf{v})={}_{\mathscr{B}}[\mathbf{v}]$. 為什麼這是一個函數呢? 因為 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis, 所以由 $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 是 V 的 spanning set, 可得任意的 $\mathbf{v}\in V$ 確實都存在 $c_1,\ldots,c_n\in\mathbb{R}$ 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 所以 $T_{\mathscr{B}}$ 確實可以將每個定義域中的元素 $\mathbf{v}$ 對應到對應域 $\mathbb{R}^n$ 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 $\mathbb{R}^n$ 中兩個不同向量的情況. 這是因為 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 linearly independent, 所以每個 $\mathbf{v}\in V$, 僅有一組 $c_1,\ldots,c_n\in\mathbb{R}$ 會使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$.

既然 $T_{\mathscr{B}}:V\to\mathbb{R}^n$ 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 $c_1,\ldots,c_n$ 與 $d_1,\ldots,d_n$ 皆為實數. 依定義我們有

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

對於任意 $r\in\mathbb{R}$, 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

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

得證 $T_{\mathscr{B}}:V\to\mathbb{R}^n$ 為 linear transformation.

$T_{\mathscr{B}}: V \to \mathbb{R}^n$ 不只是 linear transformation, 事實上它是 V 和 $\mathbb{R}^n$ 之間的 isomorphism. 也就是說 $T_{\mathscr{B}}: V \to \mathbb{R}^n$ 是 one-to-one 且 onto. 要檢查 $T_{\mathscr{B}}$ 為 one-to-one, 我們僅要檢查 $\ker(T_{\mathscr{B}}) = \{\mathbf{O}\}$ 即可 (參見 Proposition 6.3.9). 然而若 $\mathbf{v} \in \ker(T_{\mathscr{B}})$, 表示 $\mathbf{v}$ 用 $\mathscr{B}$ 的坐標表示法是 $\mathbb{R}^n$ 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 $\mathbf{v} = \mathbf{O}$, 故知 $\ker(T_{\mathscr{B}}) = \{\mathbf{O}\}$. 要檢查 $T_{\mathscr{B}}$ 為 onto, 我們可以利用 $T_{\mathscr{B}}(\mathbf{v}_i) = \mathbf{e}_i$, $\forall i = 1, \ldots, n$, 故得證 $T_{\mathscr{B}}(V) = \mathrm{Span}(T_{\mathscr{B}}(\mathbf{v}_1), \ldots, T_{\mathscr{B}}(\mathbf{v}_n)) = \mathrm{Span}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \mathbb{R}^n$ (參見 Proposition 6.3.8), 即 $T_{\mathscr{B}}$ 為 onto. 我們證得了以下的定理.

Theorem 6.4.2. 假設 V 為 *vector space,* $\dim(V) = n$ 且 $\mathscr{B}$ 為 V 的一組 *ordered basis.* 考慮 $T_{\mathscr{B}}: V \to \mathbb{R}^n$ 定義為 $T_{\mathscr{B}}(\mathbf{v}) = {}_{\mathscr{B}}[\mathbf{v}]$, $\forall \mathbf{v} \in V$. 則 $T_{\mathscr{B}}$ 為 V 到 $\mathbb{R}^n$ 的 *isomorphism.*

知道 $T_{\mathscr{B}}: V \to \mathbb{R}^n$ 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 $T_{\mathscr{B}}$, 將問題轉換成大家熟悉的 $\mathbb{R}^n$ 中的向量的性質. 例如我們要判斷 V 中的元素 $\mathbf{w}_1, \ldots, \mathbf{w}_k$ 是否為 linearly independent. 我們可以先找到一組 V 的 ordered basis $\mathscr{B}$, 然後考慮 ${}_{\mathscr{B}}[\mathbf{w}_1], \ldots, {}_{\mathscr{B}}[\mathbf{w}_k]$, 這一組 $\mathbb{R}^n$ 中的向量. 利用我們熟悉的判斷 $\mathbb{R}^n$ 中向量是否為 linearly independent 的方法判斷 ${}_{\mathscr{B}}[\mathbf{w}_1], \ldots, {}_{\mathscr{B}}[\mathbf{w}_k]$ 是否為 linearly independent. 由於對於 $i = 1, \ldots, k$, ${}_{\mathscr{B}}[\mathbf{w}_i] = T_{\mathscr{B}}(\mathbf{w}_i)$, 因此由 $T_{\mathscr{B}}$ 為 isomorphism 知 $\mathbf{w}_1, \ldots, \mathbf{w}_k$ 為 linearly independent 若且唯若 ${}_{\mathscr{B}}[\mathbf{w}_1], \ldots, {}_{\mathscr{B}}[\mathbf{w}_k]$ 為 linearly independent (參見 Proposition 6.3.10). 因此我們可以由 ${}_{\mathscr{B}}[\mathbf{w}_1], \ldots, {}_{\mathscr{B}}[\mathbf{w}_k]$ 是否為 linearly independent, 來決定 $\mathbf{w}_1, \ldots, \mathbf{w}_k$ 是否為 linearly independent. 我們看以下的例子.

Example 6.4.3. 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 $f_2(x), f_1(x), f_0(x)$ 分別為次數為 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 $\mathscr{E} = (x^2, x, 1)$, 可得

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

由於 a, d, r 皆不等於 0, 很容易看出矩陣

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

的 rank 為 3, 亦即 ${}_{\mathscr{E}}[f_2(x)], {}_{\mathscr{E}}[f_1(x)], {}_{\mathscr{E}}[f_0(x)]$ 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 $f_0(x), \ldots, f_n(x)$, 其中對於 $i = 0, \ldots, n$, $f_i(x)$ 是次數為 i 的多項式. 利用對 standard ordered basis $\mathscr{E} = (x^n, \ldots, x, 1)$ 坐標化, 我們可得 $f_0(x), \ldots, f_n(x)$ 為 $P_n(\mathbb{R})$ 的一組 basis.

(B) 假設 V 為 vector space, 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 $W=\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4,\mathbf{w}_5)$ 的一組 basis.

考慮 $U=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4)$, 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 $\mathscr{B}=(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4)$ 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 ${}_{\mathscr{B}}[\mathbf{w}_1],\dots,{}_{\mathscr{B}}[\mathbf{w}_5]$ 這 5 個 $\mathbb{R}^4$ 中的向量. 利用過去我們知道求 $\mathbb{R}^4$ 中 $\mathrm{Span}({}_{\mathscr{B}}[\mathbf{w}_1],\dots,{}_{\mathscr{B}}[\mathbf{w}_5])$ 的 basis 的方法求出 $\mathrm{Span}({}_{\mathscr{B}}[\mathbf{w}_1],\dots,{}_{\mathscr{B}}[\mathbf{w}_5])$ 的一組 basis. 再將它們還原成 U 中的元素, 就得到 W 的一組 basis.

現由於

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

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

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

由於 echelon form 的 1-st, 3-rd, 4-th column 為 pivot 所在位置, 故知 ${}_{\mathscr{B}}[\mathbf{w}_1],{}_{\mathscr{B}}[\mathbf{w}_3],{}_{\mathscr{B}}[\mathbf{w}_4]$ 為 $\mathrm{Span}({}_{\mathscr{B}}[\mathbf{w}_1],\dots,{}_{\mathscr{B}}[\mathbf{w}_5])$ 的一組 basis (參見 Proposition 4.4.8). 由於 $T_{\mathscr{B}}$ 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 6.4. 考慮 *Example 6.4.3 (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 $\dim(\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4,\mathbf{w}_5))$ 為何? 並將 $\mathbf{w}_2,\mathbf{w}_5$ 寫成 $\mathbf{w}_1,\mathbf{w}_3,\mathbf{w}_4$ 的 *linear combination.*

6.5. Matrix Representation

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

分別給定 V,W 的一組 ordered basis $\mathscr{B}=(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 以及 $\beta=(\mathbf{w}_1,\dots,\mathbf{w}_m)$. 令 $T_{\mathscr{B}}:V\to\mathbb{R}^n$ 與 $T_\beta:W\to\mathbb{R}^m$ 分別為利用 $\mathscr{B}$ 以及 β 將 V,W 的元素坐標化的 linear transformation. 回顧一下, 這裡 $T_{\mathscr{B}},T_\beta$ 皆為 isomorphism. 對於任意 V 到 W 的 linear transformation T, 我們考慮合成函數 $T_\beta\circ T\circ T_{\mathscr{B}}^{-1}:\mathbb{R}^n\to\mathbb{R}^m$. 由於 T_β, T 以及 $T_{\mathscr{B}}^{-1}$ 皆為 linear transformation, 所以 $T_\beta\circ T\circ T_{\mathscr{B}}^{-1}$ 是一個 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation (Proposition 6.3.4). 因此 $T_\beta\circ T\circ T_{\mathscr{B}}^{-1}$

有一個 standard matrix representation, 我們定義這個 standard matrix representation 為 T 相對於 $\mathscr{B}$, β 這兩組 ordered basis 所得的 *matrix representation*, 並用 ${}_\beta[T]_\mathscr{B}$ 來表示. 到底 ${}_\beta[T]_\mathscr{B}$ 是怎樣的矩陣呢? 依定義它是一個 $m\times n$ matrix, 且對於 $i=1,\dots,n$, ${}_\beta[T]_\mathscr{B}$ 的 i-th column 應為 $T_\beta\circ T\circ T_\mathscr{B}^{-1}(\mathbf{e}_i)$. 由於 $T_\mathscr{B}(\mathbf{v}_i)=\mathbf{e}_i$, 所以 $T_\mathscr{B}^{-1}(\mathbf{e}_i)=\mathbf{v}_i$. 因此得

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

也就是說 ${}_\beta[T]_\mathscr{B}$ 的 i-th column 就是將 ordered basis $\mathscr{B}=(\mathbf{v}_1,\dots,\mathbf{v}_n)$ 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 β 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

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

Example 6.5.1. 考慮 $P_2(\mathbb{R})$ 上的 standard ordered basis $\mathscr{E}=(x^2,x,1)$ 以及 $P_3(\mathbb{R})$ 上的 standard ordered basis $\varepsilon=(x^3,x^2,x,1)$. 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 $\mathscr{E},\varepsilon$ 的 matrix representation ${}_\varepsilon[T]_\mathscr{E}$. 依照前面的探討, 我們知 ${}_\varepsilon[T]_\mathscr{E}$ 的 1-st column 應該就是將 x^2 代入 T, 得 $T(x^2)=(x+1)(x-1)^2$ 再利用 ε 將 $(x+1)(x-1)^2$ 坐標化寫成 $\mathbb{R}^3$ 的向量. 由於 $(x+1)(x-1)^2=x^3-x^2-x+1$, 所以得 ${}_\varepsilon[T]_\mathscr{E}$ 的 1-st column 為

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

同理我們有 ${}_\varepsilon[T]_\mathscr{E}$ 的 2-nd, 3-rd column 分別為

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

因此得

$${}_\varepsilon[T]_\mathscr{E}=\begin{bmatrix}1&0&0\\-1&1&0\\-1&0&1\\1&-1&1\end{bmatrix}.$$

我們更換 $P_2(\mathbb{R})$ 以及 $P_3(\mathbb{R})$ 的 ordered basis. 回顧在 Example 6.2.5 以及 Example 6.2.13 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 $\mathscr{B}=(p_1(x),p_2(x),p_3(x))$ 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 $\beta=(q_1(x),q_2(x),q_3(x),q_4(x))$ 為 $P_3(\mathbb{R})$ 的 ordered basis. 我們要得到 T 對於 $\mathscr{B},\beta$ 的 matrix representation ${}_\beta[T]_\mathscr{B}$. 首先 ${}_\beta[T]_\mathscr{B}$ 的 1-st column 為 $T(p_1(x))=(x+1)p_1(x-1)$ 利用 β 坐標化所得 $\mathbb{R}^3$ 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 ${}_\beta[T]_\mathscr{B}$ 的 1-st column 為

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

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

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

因此得

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

由 Example 6.5.1 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 $\mathbb{R}^n$ 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 $\mathscr{B},\beta$, 以及一個 V 到 W 的 linear transformation T, 我們可以圖示如下:

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

這樣的圖示一般稱為 *commutative diagram*. 它表示底下 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的函數為 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}$, 亦即先利用 $T_{\mathscr{B}}^{-1}$ 將 $\mathbb{R}^n$ 映射到 V, 再利用 T 將 V 映射到 W, 最後利用 T_β 將 W 映射到 $\mathbb{R}^m$. 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 $T_{\mathscr{B}}$ 到 $\mathbb{R}^n$, 接著藉由 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}$ 從 $\mathbb{R}^n$ 到 $\mathbb{R}^m$, 最後從 $\mathbb{R}^m$ 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 $T_{\mathscr{B}}$ 得到 $T_{\mathscr{B}}(\mathbf{v})$, 接著利用 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}$ 將 $T_{\mathscr{B}}(\mathbf{v})$ 送至 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}(T_{\mathscr{B}}(\mathbf{v})) = T_\beta(T(\mathbf{v}))$, 最後再利用 T_β^{-1} 將 $T_\beta(T(\mathbf{v}))$ 送至 $T_\beta^{-1}(T_\beta(T(\mathbf{v}))) = T(\mathbf{v})$.

利用 $V \to \mathbb{R}^n$ 接 $\mathbb{R}^n \to \mathbb{R}^m$ 再接 $\mathbb{R}^m \to W$ 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 $\mathbb{R}^n \to \mathbb{R}^m$ 這一段的路徑, 有 standard matrix representation, 即 ${}_\beta[T]_{\mathscr{B}}$. 也就是說對於任意 $\mathbb{R}^n$ 的向量, 我們只要左邊乘上 ${}_\beta[T]_{\mathscr{B}}$ 就可以知道會被映射到哪一個 $\mathbb{R}^m$ 的向量. 所以對任意 $\mathbf{v} \in V$, 我們可以先利用 $\mathscr{B}$ 將 $\mathbf{v}$ 坐標化得 $T_{\mathscr{B}}(\mathbf{v}) = {}_{\mathscr{B}}[\mathbf{v}]$. 接著由於 ${}_{\mathscr{B}}[\mathbf{v}] \in \mathbb{R}^n$, 故將之代入 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}$ 就是將 ${}_{\mathscr{B}}[\mathbf{v}]$ 左邊乘上 ${}_\beta[T]_{\mathscr{B}}$ 這一個 matrix. 也就是說 $T_\beta \circ T \circ T_{\mathscr{B}}^{-1}({}_{\mathscr{B}}[\mathbf{v}])$ 就是 ${}_\beta[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathbf{v}]$. 最後再利用 W 的 ordered basis β 將 ${}_\beta[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathbf{v}]$ 這個 $\mathbb{R}^m$ 中的向量還原回 W 的元素 $T_\beta^{-1}({}_\beta[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathbf{v}])$, 就是 $T(\mathbf{v})$ 之值. 因此我們有以下之結果.

Proposition 6.5.2. *假設 V,W 為 vector space 且 $\mathscr{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $\beta = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ 分別為 V,W 的 ordered basis. 設 $T : V \to W$ 為 linear transformation 且 ${}_\beta[T]_{\mathscr{B}} \in \mathscr{M}_{m\times n}$ 為 T 相對於 $\mathscr{B},\beta$ 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*

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

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

$${}_\beta[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathbf{v}] = {}_\beta[T(\mathbf{v})].$$

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

Example 6.5.3. 我們考慮 Example 6.5.1 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 $\mathscr{E} = (x^2, x, 1)$ 以及 $P_3(\mathbb{R})$ 的 ordered basis $\varepsilon = (x^3, x^2, x, 1)$, 我們知道 T 對於 $\mathscr{E}, \varepsilon$ 的 matrix representation 為 ${}_{\varepsilon}[T]_{\mathscr{E}} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ 故由 ${}_{\mathscr{E}}[p(x)] = {}_{\mathscr{E}}[x^2 - 1] = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ 得

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

即 $T(p(x)) = x^3 - x^2 - 2x$.

當然我們也可以利用 Example 6.5.1 中 V 的 ordered basis $\mathscr{B} = (p_1(x), p_2(x), p_3(x))$ 以及 W 的 ordered basis $\beta = (q_1(x), q_2(x), q_3(x), q_4(x))$ 求出 $T(x^2 - 1)$. 此時若 $p(x) = x^2 - 1 = c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x)$, 則代 $x = -1, 0, 1$ 得 $c_1 = 0$, $c_2 = -1$, $c_3 = 0$, 亦即 ${}_{\mathscr{B}}[p(x)] = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$. 故將 ${}_{\mathscr{B}}[p(x)]$ 左邊乘上 T 對於 $\mathscr{B}, \beta$ 的 representation matrix ${}_{\beta}[T]_{\mathscr{B}}$ 得

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

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 $T : \mathbb{R}^n \to \mathbb{R}^m$ 是 linear transformation 時, 我們可以利用 T 的 standard matrix representation $[T]$ 的 nullspace 來決定 T 的 kernel, 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 kernel 和 range. 分別選定 V 和 W 的 ordered basis $\mathscr{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 和 $\beta = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$. 由前面的 commutative diagram, 利用 $V \to W$ 接著 $W \to \mathbb{R}^m$ 的路徑以及 $V \to \mathbb{R}^n$ 接著 $\mathbb{R}^n \to \mathbb{R}^m$ 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

$${}_{\beta}[T(\mathbf{v})] = {}_{\beta}[T]_{\mathscr{B}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \tag{6.2}$$

現若 $\mathbf{v}\in\ker(T)$, 表示 $T(\mathbf{v})=\mathbf{O}$, 因此由 ${}_\beta[T(\mathbf{v})]={}_\beta[\mathbf{O}]=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}$, 得 ${}_\beta[T]_{\mathscr{B}}\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}$, 亦即 $\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}={}_{\mathscr{B}}[\mathbf{v}]$ 為 ${}_\beta[T]_{\mathscr{B}}$ 的 nullspace 的向量. 反之, 若 $\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}\in\mathbb{R}^n$ 為 ${}_\beta[T]_{\mathscr{B}}$ 的 nullspace 的向量, 表示 ${}_\beta[T]_{\mathscr{B}}\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}$. 故由式子 (6.2) 知, 當 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ 時, 我們有 ${}_\beta[T(\mathbf{v})]=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}$. 此即表示 $T(\mathbf{v})=\mathbf{O}$, 得證 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n\in\ker(T)$.

另一方面, 若 $\mathbf{w}=d_1\mathbf{w}_1+\cdots+d_m\mathbf{w}_m\in T(V)$, 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 (6.2) 知, $\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}={}_\beta[T(\mathbf{v})]={}_\beta[T]_{\mathscr{B}}\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$, 亦即 $\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}$ 是 ${}_\beta[T]_{\mathscr{B}}$ 的 column space 的向量. 反之, 若 $\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}\in\mathbb{R}^m$ 為 ${}_\beta[T]_{\mathscr{B}}$ 的 column space 的向量, 表示存在 $\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}\in\mathbb{R}^n$ 使得 $\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}={}_\beta[T]_{\mathscr{B}}\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$. 故由式子 (6.2) 知, 若令 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$, 我們有 ${}_\beta[T(\mathbf{v})]=\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}$, 得證 $d_1\mathbf{w}_1+\cdots+d_m\mathbf{v}_m=T(\mathbf{v})\in T(V)$. 我們證得了以下的結果.

Proposition 6.5.4. 假設 V,W 為 *vector space* 且 $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, $\beta=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$ 分別為 V,W 的 *ordered basis*. 設 $T:V\to W$ 為 *linear transformation* 且 ${}_\beta[T]_{\mathscr{B}}\in\mathscr{M}_{m\times n}$ 為 T 相對於 $\mathscr{B},\beta$ 的 *matrix representation*. 則 $c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n\in\ker(T)$ 若且唯若 $\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$ 屬於 ${}_\beta[T]_{\mathscr{B}}$ 的 *nullspace*. 而 $d_1\mathbf{w}_1+\cdots+d_m\mathbf{w}_m\in T(V)$ 若且唯若 $\begin{bmatrix}d_1\\\vdots\\d_m\end{bmatrix}$ 屬於 ${}_\beta[T]_{\mathscr{B}}$ 的 *column space*.

分別給定 $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$, $\beta=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$ 為 V,W 的 ordered basis, 我們知道 $T_{\mathscr{B}}:V\to\mathbb{R}^n$, $T_\beta:W\to\mathbb{R}^m$ 為 isomorphism, 因此由 Proposition 6.5.4 我們知道 $T_{\mathscr{B}}$ 會將 $\ker(T)$ 的一組 basis 送至 ${}_\beta[T]_{\mathscr{B}}$ 的 nullspace 的一組 basis 且 T_β 會將 T 的 range $T(V)$ 的一組 basis 送至 ${}_\beta[T]_{\mathscr{B}}$ 的 column space 的一組 basis (參見 Theorem 6.3.12). 所以我們有 $\dim(\ker(T))=\mathrm{null}({}_\beta[T]_{\mathscr{B}})$ 以及 $\dim(T(V))=\mathrm{rank}({}_\beta[T]_{\mathscr{B}})$. 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 kernel 的維度稱為 T 的 *nullity*. 利用此一結果, 我們推廣了 Theorem 5.3.6.

Theorem 6.5.5 (Dimension Theorem for Linear Transformation)**.** 假設 V,W 為 *finite dimensional vector space* 且 $T:V\to W$ 為 *linear transformation.* 則

$$\dim(T(V))+\dim(\ker(T))=\dim(V).$$

Proof. 考慮 V,W 的 ordered basis $\mathscr{B}=(\mathbf{v}_1,\ldots,\mathbf{v}_n),\beta=(\mathbf{w}_1,\ldots,\mathbf{w}_m)$. 由於 T 對於 $\mathscr{B},\beta$ 的 matrix representation ${}_\beta[T]_{\mathscr{B}}$ 為 $m\times n$ matrix, 由 Theorem 4.4.13 我們知 $\mathrm{rank}({}_\beta[T]_{\mathscr{B}})+\mathrm{null}({}_\beta[T]_{\mathscr{B}})=n$. 故套用前述 $\dim(\ker(T))=\mathrm{null}({}_\beta[T]_{\mathscr{B}})$ 以及 $\dim(T(V))=\mathrm{rank}({}_\beta[T]_{\mathscr{B}})$, 得證本定理. □

Example 6.5.6. 我們考慮 Example 6.5.1 的例子, 當考慮 $P_2(\mathbb{R})$ 的 ordered basis $\mathscr{E}=(x^2,x,1)$ 以及 $P_3(\mathbb{R})$ 的 ordered basis $\varepsilon=(x^3,x^2,x,1)$, 我們知道 T 對於 $\mathscr{E},\varepsilon$ 的 matrix representation ${}_\varepsilon[T]_{\mathscr{E}}$ 利用 elementary row operations 化為 echelon form 可得

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

由於 pivot 的個數等於 column 的個數, 我們知 ${}_\varepsilon[T]_{\mathscr{E}}$ 的 nullspace 為 $\{\begin{bmatrix}0\\0\\0\end{bmatrix}\}$, 故知 $\ker(T)=\{\mathbf{O}\}$, 亦即 T 為 one-to-one. 另一方面 ${}_\varepsilon[T]_{\mathscr{E}}$ 的 rank 為 3, 故 $\{\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix},\begin{bmatrix}0\\1\\0\\-1\end{bmatrix},\begin{bmatrix}0\\0\\1\\1\end{bmatrix}\}$ 為 column space 的一組 basis. 因此得 $\{x^3-x^2-x+1,x^2-1,x+1\}$ 為 $T(P_2(\mathbb{R}))$ 的一組 basis. 由於 $\dim(P_3(\mathbb{R}))=4\neq\dim(T(P_2(\mathbb{R})))=3$, 我們知 $T(P_2(\mathbb{R}))\neq P_3(\mathbb{R})$, 故 T 不是 onto. 最後驗證 Dimension Theorem, 我們有

$$\dim(T(P_2(\mathbb{R})))+\dim(\ker(T))=3+0=3=\dim(P_2(\mathbb{R})).$$

Example 6.5.7. 考慮 $\mathscr{M}_{2\times2}$ 所形成的 vector space (參見 Example 6.1.4 (A)). 考慮函數 $T:\mathscr{M}_{2\times2}\to\mathscr{M}_{2\times2}$ 定義為 $T(A)=A-A^{\mathrm{T}}$, $\forall A\in\mathscr{M}_{2\times2}$. 我們可得 T 為 linear transformation. 這是因為對任意 $A,B\in\mathscr{M}_{2\times2}$ 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 $\mathscr{M}_{2\times2}$ 的 ordered basis $\mathscr{E}=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$. 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 $\mathscr{E},\mathscr{E}$ 的 matrix representation 為 ${}_{\mathscr{E}}[T]_{\mathscr{E}}=\begin{bmatrix}0&0&0&0\\0&1&-1&0\\0&-1&1&0\\0&0&0&0\end{bmatrix}$. 利用 elementary row operation 將 ${}_{\mathscr{E}}[T]_{\mathscr{E}}$ 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 ${}_{\mathscr{E}}[T]_{\mathscr{E}}$ 的 nullspace 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 $\ker(T)$ 的一組 basis. 又 ${}_{\mathscr{E}}[T]_{\mathscr{E}}$ 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 $T(\mathscr{M}_{2\times 2})$ 的一組 basis. 注意我們有 $\dim(T(\mathscr{M}_{2\times 2}))+\dim(\ker(T))=1+3=4=\dim(\mathscr{M}_{2\times 2})$. 另外若 $A\in\ker(T)$ 表示 $T(A)=A-A^{\mathrm{T}}=\mathbf{O}$, 亦即 $A=A^{\mathrm{T}}$. 反之亦然, 也就是說 $A\in\ker(T)$ 若且唯若 A 為 symmetric matrix. 因此由 $\dim(\ker(T))=3$, 我們知所有 2×2 的 symmetric matrices 所成的 subspace 的維度為 3.

Question 6.5. 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 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation 時, 對任意 $c_1,c_2\in\mathbb{R}$, 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 5.3.8). 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, $\mathscr{B},\beta$, 我們會有

$$ {}_{\beta}[c_1T_1+c_2T_2]_{\mathscr{B}}=c_1\,{}_{\beta}[T_1]_{\mathscr{B}}+c_2\,{}_{\beta}[T_2]_{\mathscr{B}}. $$

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 $\mathscr{B},\beta,\gamma$, 我們有以下的圖示

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

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 $\mathbb{R}^m$ 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:

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

由於底部 $\mathbb{R}^n \to \mathbb{R}^m$ 的 matrix representation 為 ${}_{\beta}[T]_{\mathscr{B}}$, 而 $\mathbb{R}^m \to \mathbb{R}^k$ 的 matrix representation 為 ${}_{\gamma}[T']_{\beta}$, 因此由 Lemma 5.3.8 知, 它們的合成所對應的 matrix representation 為 ${}_{\gamma}[T']_{\beta}\,{}_{\beta}[T]_{\mathscr{B}}$. 因此我們有

$$ {}_{\gamma}[T' \circ T]_{\mathscr{B}} = {}_{\gamma}[T']_{\beta}\,{}_{\beta}[T]_{\mathscr{B}}. $$

特別的, 當 $T : V \to W$ 為 isomorphism 時, 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 $\mathscr{B}, \beta$ 後, 我們有以下的 commutative diagram.

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

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 知, 底部 $\mathbb{R}^n \to \mathbb{R}^n$ 再接 $\mathbb{R}^n \to \mathbb{R}^n$ 這個路徑是將 $\mathbb{R}^n$ 的向量固定不動, 也就是說它們的合成所對應的 matrix representation 為 identity matrix. 亦即 ${}_{\mathscr{B}}[T^{-1}]_{\beta}\,{}_{\beta}[T]_{\mathscr{B}} = I_n$. 同理, 由於 $T \circ T^{-1}$ 為 W 到 W 的 identity map, 我們得 ${}_{\beta}[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[T^{-1}]_{\beta} = I_n$. 得證 ${}_{\mathscr{B}}[T^{-1}]_{\beta}$ 為 ${}_{\beta}[T]_{\mathscr{B}}$ 的反矩陣. 綜合以上的討論, 我們有以下有關 Lemma 5.3.8 和 Theorem 5.3.12 的推廣.

Theorem 6.5.8. 假設 V, W, U 為 *finite dimensional vector space* 且令 $\mathscr{B}, \beta, \gamma$ 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 $c_1, c_2 \in \mathbb{R}$, 我們有

$$ {}_{\beta}[c_1T_1 + c_2T_2]_{\mathscr{B}} = c_1\,{}_{\beta}[T_1]_{\mathscr{B}} + c_2\,{}_{\beta}[T_2]_{\mathscr{B}}. $$

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

$$ {}_{\gamma}[T' \circ T]_{\mathscr{B}} = {}_{\gamma}[T']_{\beta}\,{}_{\beta}[T]_{\mathscr{B}}. $$

(3) 設 $T : V \to W$ 為 *linear transformation.* 則 T 為 *isomorphism* 若且唯若 ${}_{\beta}[T]_{\mathscr{B}}$ 為 *invertible matrix.* 又此時 $T^{-1} : W \to V$ 對應於 $\beta, \mathscr{B}$ 的 *matrix representation* 為

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

我們知道一個 linear transformation, 當我們用不同的 ordered bases 所得的 matrix representation 會不同. 假設 $\mathscr{B}, \mathscr{B}'$ 為 V 的兩組 ordered bases, 而 β, β' 為 W 的兩組 ordered basis. 對於 linear transformation $T : V \to W$, 其對應於這兩對 ordered bases 的 matrix representations ${}_{\beta}[T]_{\mathscr{B}}$ 和 ${}_{\beta'}[T]_{\mathscr{B}'}$ 之間會有甚麼關係呢? 首先我們考慮 identity map $\mathbf{id} : V \to V$. 注意雖然是 identity map, 但其 matrix representation 未必會是 identity matrix. 事實上, 當我們定義域和對應域都選同一組 ordered basis $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, 則由於 $\mathbf{id}(\mathbf{v}_i) = \mathbf{v}_i$, 故其 matrix representation 是 identity matrix. 但若定義域是使用 $\mathscr{B}$ 這一組

ordered basis, 而對應域選的是 $\mathscr{B}'=\{\mathbf{v}_1',\ldots,\mathbf{v}_n'\}$ 這一組 ordered basis, identity map 對應於 $\mathscr{B},\mathscr{B}'$ 的 matrix representation ${}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}$ 其 i-th column 雖然仍和 $\mathrm{id}(\mathbf{v}_i)=\mathbf{v}_i$ 有關, 不過卻是要將 $\mathbf{v}_i$ 寫成以 $\{\mathbf{v}_1',\ldots,\mathbf{v}_n'\}$ 為 ordered basis 的坐標表示法 ${}_{\mathscr{B}'}[\mathbf{v}_i]$. 所以當 $\mathscr{B}$ 和 $\mathscr{B}'$ 相異時, ${}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}$ 不是 identity matrix. 現對任意 $\mathbf{v}\in V$, 因 $\mathbf{v}$ 對於 $\mathscr{B}$ 的坐標表示法為 ${}_{\mathscr{B}}[\mathbf{v}]$, 依 matrix representation 的性質 (Proposition 6.5.2) 可得

$$ {}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathbf{v}] = {}_{\mathscr{B}'}[\mathrm{id}(\mathbf{v})] = {}_{\mathscr{B}'}[\mathbf{v}]. $$

也就是說, 矩陣 ${}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}$ 可以將 V 中元素對於 $\mathscr{B}$ 的坐標表示轉換成對於 $\mathscr{B}'$ 的坐標表示, 也因此我們稱 ${}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}$ 為 *change-of-basis matrix.*

要注意 $\mathrm{id}:V\to V$ 是 isomorphism, 所以由 Theorem 6.5.8 (3), 我們得 ${}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}}$ 為 invertible 且

$$ ({}_{\mathscr{B}'}[\mathrm{id}]_{\mathscr{B}})^{-1} = {}_{\mathscr{B}}[\mathrm{id}^{-1}]_{\mathscr{B}'} = {}_{\mathscr{B}}[\mathrm{id}]_{\mathscr{B}'} \tag{6.3} $$

也就是說將 $\mathscr{B}$ 的坐標表示轉換成對於 $\mathscr{B}'$ 的坐標表示的 change-of-basis matrix 的 inverse 就是 $\mathscr{B}'$ 的坐標表示轉換成對於 $\mathscr{B}$ 的坐標表示的 change-of-basis matrix.

我們回到原先的問題, 假設 $T:V\to W$ 為 linear transformation 且 $\mathscr{B},\mathscr{B}'$ 為 V 的兩組 ordered bases, 而 β,β' 為 W 的兩組 ordered basis. 我們要探討 ${}_{\beta}[T]_{\mathscr{B}}$ 和 ${}_{\beta'}[T]_{\mathscr{B}'}$ 之間的關係. 由於 $\mathrm{id}_V:V\to V$, $T:V\to W$ 和 $\mathrm{id}_W:W\to W$ 之合成 $\mathrm{id}_W\circ T\circ \mathrm{id}_V:V\to W$ 仍為 $T:V\to W$, 所以由 Theorem 6.5.8 (2) 得

$$ {}_{\beta'}[\mathrm{id}_W]_{\beta}\,{}_{\beta}[T]_{\mathscr{B}}\,{}_{\mathscr{B}}[\mathrm{id}_V]_{\mathscr{B}'} = {}_{\beta'}[T]_{\mathscr{B}'}. $$

這就是所謂的 change-of basis formula, 我們將之完整敘述如下.

Theorem 6.5.9 (Change-of-basis Formula)**.** 假設 $T:V\to W$ 為 *linear transformation* 且 $\mathscr{B},\mathscr{B}'$ 為 V 的兩組 *ordered bases,* 而 β,β' 為 W 的兩組 *ordered basis,* 則存在 *invertible matrix* P,Q 使得 ${}_{\beta'}[T]_{\mathscr{B}'}=Q({}_{\beta}[T]_{\mathscr{B}})P$, 其中 P 為將 $\mathscr{B}'$ 的坐標表示轉換成 $\mathscr{B}$ 的坐標表示的 *change-of-basis matrix* ${}_{\mathscr{B}}[\mathrm{id}_V]_{\mathscr{B}'}$, 而 Q 為將 β 的坐標表示轉換成 β' 的坐標表示的 *change-of-basis matrix* ${}_{\beta'}[\mathrm{id}_W]_{\beta}$.

Example 6.5.10. 在 Example 6.5.1 中我們考慮 linear transformation $T:P_2(\mathbb{R})\to P_3(\mathbb{R})$, 其中 $T(p(x))=(x+1)p(x-1)$, $\forall p(x)\in P_2(\mathbb{R})$. 另外我們考慮 $P_2(\mathbb{R})$ 的兩組 ordered bases $\mathscr{E}=(x^2,x,1)$, $\mathscr{B}=(p_1(x),p_2(x),p_3(x))$ 其中

$$ p_1(x)=\frac{1}{2}(x^2-x),\quad p_2(x)=-x^2+1,\quad p_3(x)=\frac{1}{2}(x^2+x) $$

以及 $P_3(\mathbb{R})$ 的兩組 ordered bases $\varepsilon=(x^3,x^2,x,1)$, $\beta=(q_1(x),q_2(x),q_3(x),q_4(x))$ 其中

$$ q_1(x)=\frac{-x^3+3x^2-2x}{6}, q_2(x)=\frac{x^3-2x^2-x+2}{2}, q_3(x)=\frac{-x^3+x^2+2x}{2}, q_4(x)=\frac{x^3-x}{6}. $$

在 Example 6.5.1 中我們得到

$$ {}_{\varepsilon}[T]_{\mathscr{E}}=\begin{bmatrix} 1 & 0 & 0\\ -1 & 1 & 0\\ -1 & 0 & 1\\ 1 & -1 & 1\end{bmatrix},\quad {}_{\beta}[T]_{\mathscr{B}}=\begin{bmatrix} 0 & 0 & 0\\ 1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}. $$

因 ${}_{\mathscr{E}}[p_1(x)] = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}$, ${}_{\mathscr{E}}[p_2(x)] = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, ${}_{\mathscr{E}}[p_3(x)] = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$ 依定義 $\mathscr{B}$ 到 $\mathscr{E}$ 的 change-of-basis matrix 為 ${}_{\mathscr{E}}[\mathrm{id}_{P_2(\mathbb{R})}]_{\mathscr{B}} = \begin{bmatrix} 1/2 & -1 & 1/2 \\ -1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$. 另外若 $x^3 = c_1q_1(x) + c_2q_2(x) + c_3q_3(x) + c_4q_4(x)$, 則因 $q_1(-1) = 1, q_2(-1) = q_3(-1) = q_4(-1) = 0$, 將 $x = -1$ 代入前式得 $c_1 = -1$, 同理我們可得 $c_2 = 0, c_3 = 1, c_4 = 8$, 亦即 ${}_{\beta}[x^3] = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 8 \end{bmatrix}$. 用同樣方法求 $x^2, x, 1$ 對於 β 的坐標表示法, 我們得 ε 到 β 的 change-of-basis matrix 為 ${}_{\beta}[\mathrm{id}_{P_3(\mathbb{R})}]_{\varepsilon} = \begin{bmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \end{bmatrix}$. 我們也可以先寫下 β 到 ε 的 change-of-basis matrix ${}_{\varepsilon}[\mathrm{id}_{P_3(\mathbb{R})}]_{\beta} = \begin{bmatrix} -1/6 & 1/2 & -1/2 & 1/6 \\ 1/2 & -1 & 1/2 & 0 \\ -1/3 & -1/2 & 1 & -1/6 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ 再取 inverse 得 ${}_{\beta}[\mathrm{id}_{P_3(\mathbb{R})}]_{\varepsilon}$. 最後我們驗算

$$ {}_{\beta}[\mathrm{id}_{P_3(\mathbb{R})}]_{\varepsilon}\, {}_{\varepsilon}[T]_{\mathscr{E}}\, {}_{\mathscr{E}}[\mathrm{id}_{P_2(\mathbb{R})}]_{\mathscr{B}} = \begin{bmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{-1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} = {}_{\beta}[T]_{\mathscr{B}}. $$

在線性代數中, 我們經常談論的一種 linear transformation 是其定義域及對應域為相同的 vector space. 這樣的 linear transformation 我們特別稱之為 *linear operator*. 關於 linear operator 我們通常對於定義域及對應域會選同樣的一組 ordered basis. 此時利用 Theorem 6.5.9, 我們得以下之結果.

Corollary 6.5.11. 假設 $T : V \to V$ 為 *linear transformation* 且 $\mathscr{B}, \mathscr{B}'$ 為 V 的兩組 *ordered bases.* 則存在 *invertible matrix* P 使得 ${}_{\mathscr{B}'}[T]_{\mathscr{B}'} = P^{-1}({}_{\mathscr{B}}[T]_{\mathscr{B}})P$, 其中 P 為將 $\mathscr{B}'$ 的坐標表示轉換成 $\mathscr{B}$ 的坐標表示的 *change-of-basis matrix* ${}_{\mathscr{B}}[\mathrm{id}_V]_{\mathscr{B}'}$.

Proof. 考慮 Theorem 6.5.9 其中 $W = V$, $\beta = \mathscr{B}$ 以及 $\beta' = \mathscr{B}'$ 的情形. 此時 $Q = {}_{\mathscr{B}'}[\mathrm{id}_V]_{\mathscr{B}}$ 由式子 (6.3), 知 $Q = ({}_{\mathscr{B}}[\mathrm{id}_V]_{\mathscr{B}'})^{-1} = P^{-1}$, 得證本定理. □

給定一個 $n \times n$ matrix A 我們知道它可以代表某一個 dimension 為 n 的 vector space V 上的 linear operator $T : V \to V$, 對於 V 的某一組 ordered basis 的 matrix representation. 若 P 為 $n \times n$ invertible matrix, 則我們稱 $B = P^{-1}AP$ 和 A 為 *similar*. 意味著我們也可將 B 視為 $T : V \to V$ 的一個 matrix representation 只是選取 V 不同的 ordered basis 而已. 有時一個 linear operator, 若選取夠好的一組 ordered basis, 我們可以得到更好的 matrix representation 以至於更容易了解這個 linear transformation. 有關於這個課題, 等以後談到對角化時我們再進一步探討. 我們先看一個簡單的例子.

Example 6.5.12. 考慮 linear operator $T:\mathbb{R}^2\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix}x\\y\end{bmatrix})=\frac{1}{25}\begin{bmatrix}9x+12y\\12x+16y\end{bmatrix}$. 若利用標準基底 $\varepsilon=(\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix})$ 我們得 ${}_{\varepsilon}[T]_{\varepsilon}=\frac{1}{25}\begin{bmatrix}9&12\\12&16\end{bmatrix}$. 然而若用 $\beta=(\begin{bmatrix}3\\4\end{bmatrix},\begin{bmatrix}-4\\3\end{bmatrix})$ 這組 ordered basis 可由 $T(\begin{bmatrix}3\\4\end{bmatrix})=\begin{bmatrix}3\\4\end{bmatrix}$, $T(\begin{bmatrix}-4\\3\end{bmatrix})=\begin{bmatrix}0\\0\end{bmatrix}$, 得 ${}_{\beta}[T]_{\beta}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$. 我們很容易由

$${}_{\beta}[T\circ T]_{\beta}=({}_{\beta}[T]_{\beta})({}_{\beta}[T]_{\beta})=\begin{bmatrix}1&0\\0&0\end{bmatrix}^2=\begin{bmatrix}1&0\\0&0\end{bmatrix}={}_{\beta}[T]_{\beta}$$

推得 $T\circ T=T$. 事實上從 β 這組 ordered basis 我們很容易看出 T 就是將 $\mathbb{R}^2$ 上的向量對 $\begin{bmatrix}3\\4\end{bmatrix}$ 的 projection. 另外令 $P={}_{\varepsilon}[\mathrm{id}]_{\beta}=\begin{bmatrix}3&-4\\4&3\end{bmatrix}$, 我們得

$$\begin{bmatrix}1&0\\0&0\end{bmatrix}={}_{\beta}[T]_{\beta}=({}_{\beta}[\mathrm{id}]_{\varepsilon})({}_{\varepsilon}[T]_{\varepsilon})({}_{\varepsilon}[\mathrm{id}]_{\beta})=P^{-1}(\frac{1}{25}\begin{bmatrix}9&12\\12&16\end{bmatrix})P,$$

所以 $\begin{bmatrix}1&0\\0&0\end{bmatrix}$ 和 $\frac{1}{25}\begin{bmatrix}9&12\\12&16\end{bmatrix}$ 為 similar.

6.6. 結論

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.