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.

Linear Transformations of $\mathbb{R}^n$

在本章中我們介紹在 $\mathbb{R}^n$ 中重要的函數, 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.

5.1. 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.

5.1.1. Function. 給定 $\mathbb{R}^n$, $\mathbb{R}^m$. 若有一個從 $\mathbb{R}^n$ 的向量對應到 $\mathbb{R}^m$ 的向量的對應關係, 即對任意 $\mathbf{v}\in\mathbb{R}^n$, 此對應會將 $\mathbf{v}$ 對應到 $\mathbb{R}^m$ 中一個向量 $\mathbf{w}$. 若對每一個 $\mathbf{v}\in\mathbb{R}^n$, 所對應到的 $\mathbf{w}$ 我們用 $T(\mathbf{v})$ 來表示, 我們就用 $T:\mathbb{R}^n\to\mathbb{R}^m$, $T(\mathbf{v})=\mathbf{w}\in\mathbb{R}^m$, $\forall\mathbf{v}\in\mathbb{R}^n$ 來表示這一個對應關係, 而稱 T 是一個從 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *function* (函數). 這裡 $\mathbb{R}^n$ 稱為 T 的 *domain* (定義域), 而 $\mathbb{R}^m$ 就稱為 T 的 *codomain* (對應域). 注意依函數的定義, 若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是一個函數, 則對任意 $\mathbf{v}\in\mathbb{R}^n$, $T(\mathbf{v})$ 一定要是 $\mathbb{R}^m$ 中的一個確定的向量. 也就是說 $T(\mathbf{v})\in\mathbb{R}^m$, 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 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是一個函數, 是容許在 $\mathbb{R}^n$ 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 $\mathbb{R}^n$ 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 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是一個函數, 是容許在 $\mathbb{R}^m$ 中

有向量 $\mathbf{w}$, 沒有任何 $\mathbb{R}^n$ 的向量會對應到 $\mathbf{w}$ (即不存在 $\mathbf{v}\in\mathbb{R}^n$ 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 $\mathbb{R}^m$ 中任意向量 $\mathbf{w}$, 皆可找到 $\mathbf{v}\in\mathbb{R}^n$ 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).

5.1.2. Linear Transformation. 要了解 $\mathbb{R}^n$ 中的向量, 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 $\mathbb{R}^n$ 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 5.1.1. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為一個函數, 若 T 滿足對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k\in\mathbb{R}^n$ 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 $\mathbb{R}^n$ 中向量的線性組合, 而 $c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$ 是 $\mathbb{R}^m$ 中向量的線性組合, 要區分清楚. 尤其在 $n\neq m$ 時要特別注意. 特別是當 $\mathbf{O}\in\mathbb{R}^n$ 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 $\mathbb{R}^m$ 中的零向量. 也就是說一個 linear transformation $T:\mathbb{R}^n\to\mathbb{R}^m$, 會將 $\mathbb{R}^n$ 中的零向量映射到 $\mathbb{R}^m$ 中的零向量. 雖然當 $n\neq m$ 時, $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 的零向量是不同的, 不過一般我們都用 $\mathbf{O}$ 來表示, 而不區分它. 所以我們仍用 $T(\mathbf{O})=\mathbf{O}$ 來表示 linear transformation 會將 $\mathbb{R}^n$ 中的零向量映射到 $\mathbb{R}^m$ 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 5.1.2. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為一個 *linear transformation.* 則 T 會將 $\mathbb{R}^n$ 中的零向量映射到 $\mathbb{R}^m$ 中的零向量, 亦即 $T(\mathbf{O})=\mathbf{O}$.

再次提醒, 這裡兩個 $\mathbf{O}$ 哪一個是 $\mathbb{R}^n$ 的零向量, 哪一個是 $\mathbb{R}^m$ 的零向量, 一定要區分清楚.

依定義要檢查 $T:\mathbb{R}^n\to\mathbb{R}^m$ 是否為 linear transformation, 我們必須考慮 $\mathbb{R}^n$ 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 Proposition 4.1.2), 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 5.1.3. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為一個函數, 則 T 為 *linear transformation* 若且唯若對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$, $r\in\mathbb{R}$ 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 $\mathbf{u},\mathbf{v}\in V$, $r\in\mathbb{R}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$.

($\Leftarrow$) : 我們要利用對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$, $r\in\mathbb{R}$ 皆有 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$ 這個性質來證明對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k\in\mathbb{R}^n$ 以及 $c_1,\ldots,c_k\in\mathbb{R}$ 皆有 $T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$. 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 $\mathbf{u}=\mathbf{O}$, $\mathbf{v}=\mathbf{v}_1$, $r=c_1$, 依 Lemma 5.1.2, 我們有 $T(c_1\mathbf{v}_1)=T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})=\mathbf{O}+rT(\mathbf{v})=c_1T(\mathbf{v}_1)$. 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 $\mathbf{v}_1,\ldots,\mathbf{v}_k\in\mathbb{R}^n$ 以及 $c_1,\ldots,c_k\in\mathbb{R}$ 皆有 $T(c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)=c_1T(\mathbf{v}_1)+\cdots+c_kT(\mathbf{v}_k)$. 我們要證明對任意 $\mathbf{v}_1,\ldots,\mathbf{v}_k,\mathbf{v}_{k+1}\in\mathbb{R}^n$ 以及 $c_1,\ldots,c_k,c_{k+1}\in\mathbb{R}$ 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 5.1.4. (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 $T(\mathbf{O})=\begin{bmatrix}1\\0\end{bmatrix}\neq\mathbf{O}$, 故由 Lemma 5.1.2 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 $T(\mathbf{O})=\mathbf{O}$, 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 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation 都是這樣的形式.

Lemma 5.1.5. 令 A 為一個 $m\times n$ *matrix.* 考慮 $T:\mathbb{R}^n\to\mathbb{R}^m$ 定義為*:* $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$*.* 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 $\mathbb{R}^n$ 映到 $\mathbb{R}^m$ 的函數. 任取 $\mathbf{v}\in\mathbb{R}^n$, 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 $\mathbf{v}\in\mathbb{R}^n$ 為 $n\times 1$ matrix), 故 $A\mathbf{v}\in\mathbb{R}^m$. 得 T 確實是一個從 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 function.

現要證明 T 為 linear, 亦即對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 以及 $r\in\mathbb{R}$, 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 3.1.8 和 Proposition 3.1.9) 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 5.1.5 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 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 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 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的函數. 對於任意 $\mathbf{v}\in\mathbb{R}^n$, 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 $\mathbb{R}^n$ 的向量映射到 $\mathbb{R}^m$ 中. 這是因為依假設, 對任意 $\mathbf{v}\in\mathbb{R}^n$, 我們有 $T_1(\mathbf{v})\in\mathbb{R}^m$ 以及 $T_2(\mathbf{v})\in\mathbb{R}^m$, 所以自然有 $(T_1+T_2)(\mathbf{v})=T_1(\mathbf{v})+T_2(\mathbf{v})\in\mathbb{R}^m$. 所以 $T_1+T_2:\mathbb{R}^n\to\mathbb{R}^m$ 確實是 well-defined function. 接下來我們要說明若 $T_1:\mathbb{R}^n\to\mathbb{R}^m$, $T_2:\mathbb{R}^n\to\mathbb{R}^m$ 皆為 linear transformation, 則 $T_1+T_2:\mathbb{R}^n\to\mathbb{R}^m$ 亦為 linear transformation. 也就是說對於任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 以及 $r\in\mathbb{R}$, 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 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation.

Question 5.1. 若 A_1,A_2 皆為 $m\times n$ *matrix.* 考慮 $T_1:\mathbb{R}^n\to\mathbb{R}^m$, $T_2:\mathbb{R}^n\to\mathbb{R}^m$, 分別定義為 $T_1(\mathbf{v})=A_1\mathbf{v}$, $T_2(\mathbf{v})=A_2\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$. 則 T_1+T_2 是怎樣的函數?

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

Question 5.2. 設 A 為 $m\times n$ *matrix.* 考慮 $T:\mathbb{R}^n\to\mathbb{R}^m$ 定義為 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$. 則對於 $c\in\mathbb{R}$, cT 是怎樣的函數?

設 $T_1,\cdots,T_k$ 為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformations. $c_1,\dots,c_k\in\mathbb{R}$, 則由前知 $c_1T_1,\dots,c_kT_k$ 皆為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 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 5.1.6. 設 $T_1,\cdots,T_k$ 為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *linear transformations,* $c_1,\dots,c_k\in\mathbb{R}$ 則 $c_1T_1+\cdots+c_kT_k$ 為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *linear transformation.*

另一個產生 linear transformation 的方法就是利用 "合成函數". 若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 和 $T':\mathbb{R}^m\to\mathbb{R}^k$ 為函數, 由於對任意 $\mathbf{v}\in\mathbb{R}^n$, 依定義 $T(\mathbf{v})\in\mathbb{R}^m$, 也就是說 $T(\mathbf{v})$ 會落在 T' 的定義域中. 所以我們可以將 $T(\mathbf{v})$ 代入 T' 中, 亦即得 $T'(T(\mathbf{v}))\in\mathbb{R}^k$. 這樣的方法幫我們定義出一個從 $\mathbb{R}^n$ 到 $\mathbb{R}^k$ 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 $T'\circ T:\mathbb{R}^n\to\mathbb{R}^k$ 的定義為 $T'\circ T(\mathbf{v})=T'(T(\mathbf{v}))$, $\forall\mathbf{v}\in\mathbb{R}^n$. 我們有下面之結果.

Proposition 5.1.7. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 和 $T':\mathbb{R}^m\to\mathbb{R}^k$ 為 *linear transformation,* 則 $T'\circ T:\mathbb{R}^n\to\mathbb{R}^k$ 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 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 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 5.3. 設 A 為 $m\times n$ *matrix,* B 為 $k\times m$ *matrix.* 考慮 $T:\mathbb{R}^n\to\mathbb{R}^m$, $T':\mathbb{R}^m\to\mathbb{R}^k$, 分別定義為 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$ 且 $T'(\mathbf{w})=B\mathbf{w}$, $\forall\mathbf{w}\in\mathbb{R}^m$. 則 $T'\circ T$ 是怎樣的函數?

當 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 $\mathbb{R}^n$ 的一組 basis 時, 對任意 $\mathbf{v}\in\mathbb{R}^n$, 皆存在唯一的一組 $c_1,\ldots,c_n\in\mathbb{R}$, 使得 $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$. 現若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 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 $\mathbb{R}^m$ 中的哪些向量, 則對於任意 $\mathbf{v}\in\mathbb{R}^n$, 我們都可以知道 $T(\mathbf{v})$ 為何. 因此我們有以下的定理.

Theorem 5.1.8. 假設 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in\mathbb{R}^n$, 為 $\mathbb{R}^n$ 的一組 *basis.* 令 $\mathbf{w}_1,\ldots,\mathbf{w}_n\in\mathbb{R}^m$, 則存在唯一的 *linear transformation* $T:\mathbb{R}^n\to\mathbb{R}^m$, 滿足 $T(\mathbf{v}_1)=w_1,\ldots,T(\mathbf{v}_i)=w_i,\ldots,T(\mathbf{v}_n)=w_n$.

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

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

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 $T':\mathbb{R}^n\to\mathbb{R}^m$ 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 $\mathbf{v}\in\mathbb{R}^n$ 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 5.1.8 中的 $\mathbf{w}_1,\ldots,\mathbf{w}_n\in\mathbb{R}^m$ 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 $\mathbb{R}^n$ 的一組 basis 後, 我們可以將這組 basis 裡的向量對應到 $\mathbb{R}^m$ 中任意的向量, 就會得到一個 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 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 5.1.8 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.

5.2. Kernel and Range

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, “kernel” 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 $T:\mathbb{R}^n\to\mathbb{R}^m$, 任取定義域 $\mathbb{R}^n$ 中的子集合 S, 我們很自然的會考慮 $T(S)=\{T(\mathbf{v})\in\mathbb{R}^m\mid\mathbf{v}\in S\}$ 這一個集合, 它就是將所有 S 中的元素利用 T 映射到 $\mathbb{R}^m$ 後的元素收集起來所得的集合. 同樣的, 對於對應域 $\mathbb{R}^m$ 中的子集合 S', 我們也會考慮 $T^{-1}(S')=\{\mathbf{v}\in\mathbb{R}^n\mid T(\mathbf{v})\in S'\}$ 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 $\mathbb{R}^m$ 中的子集合, 而 $T^{-1}(S')$ 會是定義域 $\mathbb{R}^n$ 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 $T(S)=\{\mathbf{w}\in\mathbb{R}^m\mid\mathbf{w}=T(\mathbf{v}),\text{ for some }\mathbf{v}\in S\}$, 有時為了強調 $T(S)$ 為 $\mathbb{R}^m$ 的子集合, 我們也會用這種表示法.

由於 linear transformation 的特色, 當 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation 時, 我們專注於考慮 V 為 $\mathbb{R}^n$ 的 subspace 時的情況. 也就是說我們要了解

$$T(V)=\{T(\mathbf{v})\in\mathbb{R}^m\mid\mathbf{v}\in V\}=\{\mathbf{w}\in\mathbb{R}^m\mid\mathbf{w}=T(\mathbf{v}),\text{ for some }\mathbf{v}\in V\}$$

的特性. 同樣的若 W 為 $\mathbb{R}^m$ 的 subspace, 我們也要了解

$$T^{-1}(W)=\{\mathbf{v}\in\mathbb{R}^n\mid T(\mathbf{v})\in W\}$$

的特性. 事實上, 我們有以下之結果.

Proposition 5.2.1. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 若 V 為 $\mathbb{R}^n$ 的 *subspaces,* 則 $T(V)$ 是 $\mathbb{R}^m$ 的 *subspace.* 另外*,* 若 W 為 $\mathbb{R}^m$ 的 *subspaces,* 則 $T^{-1}(W)$ 是 $\mathbb{R}^n$ 的 *subspace.*

Proof. 依定義我們知 $T(V)$ 會是 $\mathbb{R}^m$ 的子集合, 而 $T^{-1}(W)$ 會是 $\mathbb{R}^n$ 的子集合. 故現僅需證明它們為 subspaces, 即利用 Proposition 4.1.2 我們要證明, 若 $\mathbf{w},\mathbf{w}'\in T(V)$ 且 $r\in\mathbb{R}$, 則 $\mathbf{w}+r\mathbf{w}'\in T(V)$ 以及若 $\mathbf{v},\mathbf{v}'\in T^{-1}(W)$ 且 $r\in\mathbb{R}$, 則 $\mathbf{v}+r\mathbf{v}'\in T^{-1}(W)$.

首先再強調一次, 當我們說 $\mathbb{R}^m$ 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 $r\in\mathbb{R}$, 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 $\mathbb{R}^n$ 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 $r\in\mathbb{R}$, 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 $\mathbb{R}^m$ 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 $V=\mathbb{R}^n$ 和 $W=\{\mathbf{O}\}$ 這兩個特殊情況時, 即

$$T(\mathbb{R}^n)=\{\mathbf{w}\in\mathbb{R}^m \mid \mathbf{w}=T(\mathbf{v}) \text{ for some } \mathbf{v}\in\mathbb{R}^n\} \quad \text{and} \quad T^{-1}(\{\mathbf{O}\})=\{\mathbf{v}\in\mathbb{R}^n \mid T(\mathbf{v})=\mathbf{O}\}$$

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

Definition 5.2.2. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation. 我們稱 $\mathbb{R}^m$ 的 subspace $T(\mathbb{R}^n)$ 為 T 的 *range* (有時也稱為 *image*). 我們稱 $\mathbb{R}^n$ 的 subspace $T^{-1}(\{\mathbf{O}\})$ 為 T 的 *kernel*, 通常用 $\ker(T)$ 來表示.

首先我們來看 linear transformation 的 range. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation. 由 Proposition 5.2.1 我們知 T 的 range $T(\mathbb{R}^n)$ 是 $\mathbb{R}^m$ 的 subspace, 故知 $\dim(T(\mathbb{R}^n))\leq\dim(\mathbb{R}^m)=m$. 若 $\dim(T(\mathbb{R}^n))=m$, 則依 Corollary 4.3.6 知此時 $T(\mathbb{R}^n)=\mathbb{R}^m$. 也就是說對於任意的 $\mathbf{w}\in\mathbb{R}^m$, 由於 $\mathbf{w}\in T(\mathbb{R}^n)$ 依定義存在 $\mathbf{v}\in\mathbb{R}^n$ 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 $\mathbf{w}\in\mathbb{R}^m$, 存在 $\mathbf{v}\in\mathbb{R}^n$ 使得 $T(\mathbf{v})=\mathbf{w}$ 故得 $\mathbf{w}\in T(\mathbb{R}^n)$, 得證 $\mathbb{R}^m\subseteq T(\mathbb{R}^n)$. 再利用已知 $T(\mathbb{R}^n)\subseteq\mathbb{R}^m$, 得證 $T(\mathbb{R}^n)=\mathbb{R}^m$. 我們有以下的性質.

Proposition 5.2.3. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 則 T 為 *onto* 若且唯若 $\dim(T(\mathbb{R}^n))=m$.

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 5.2.3 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 5.2.4. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation* 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in\mathbb{R}^n$ 為 $\mathbb{R}^n$ 的一組 *spanning vectors.* 則

$$T(\mathbb{R}^n)=\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n)).$$

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

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

得證 $T(\mathbb{R}^n)\subseteq\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))$.

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

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

得證 $\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))\subseteq T(\mathbb{R}^n)$. 因此證明了 $T(\mathbb{R}^n)=\operatorname{Span}(T(\mathbf{v}_1),\ldots,T(\mathbf{v}_n))$. $\square$

Example 5.2.5. (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 5.1.4 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 5.2.4 我們有 $T(\mathbb{R}^3)=\mathrm{Span}(\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\-1\end{bmatrix})=\mathbb{R}^2$. 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 5.2.4 我們有 $T(\mathbb{R}^2)=\mathrm{Span}(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix})$. 由於 $\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix}$ 為 linearly independent, 故得 $\dim(T(\mathbb{R}^2))=2$. 由 Proposition 5.2.3 知 T 不是 onto.

Question 5.4. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 若 $m>n$, 則 T 有可能是 *onto* 嗎?

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

Question 5.5. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation* 且 $\mathbf{v}_1,\dots,\mathbf{v}_n\in\mathbb{R}^n$ 為 $\mathbb{R}^n$ 的一組 *basis.* 試證明 T 為 *zero mapping* 若且唯若 $T(\mathbf{v}_1)=\cdots=T(\mathbf{v}_n)=\mathbf{O}$.

接下來我們看 kernel 與 linear transformation 的關係. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation. 若 T 為 one-to-one, 由於已知 $T(\mathbf{O})=\mathbf{O}$, 故知不可能有非零的向量 $\mathbf{v}$ 使得 $T(\mathbf{v})=\mathbf{O}$. 因此可得 $\ker(T)=\{\mathbf{O}\}$. 其實反過來 $\ker(T)=\{\mathbf{O}\}$, 也會使得 T 為 one-to-one, 我們有以下的結果.

Proposition 5.2.6. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 則 T 為 *one-to-one* 若且唯若 $\dim(\ker(T))=0$, 亦即 $\ker(T)=\{\mathbf{O}\}$.

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

Example 5.2.7. 我們探討 Example 5.2.5 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 $\mathbf{v}=\begin{bmatrix} a\\ b\\ c\end{bmatrix}\in\ker(T)$, 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 $\begin{bmatrix} 1\\ -1\\ 1\end{bmatrix}\in\ker(T)$, 知 $\ker(T)\neq\{\mathbf{O}\}$ (事實上 $\ker(T)=\mathrm{Span}(\begin{bmatrix} 1\\ -1\\ 1\end{bmatrix})$). 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 $\mathbf{v}=\begin{bmatrix} a\\ b\end{bmatrix}\in\ker(T)$, 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 $\ker(T)=\{\mathbf{O}\}$, 所以由 Proposition 5.2.6 知 T 是 one-to-one.

要注意 Proposition 5.2.6 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 5.6. *Proposition 5.2.6* 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 $\ker(T)=\{\mathbf{O}\}$ 還是由 $\ker(T)=\{\mathbf{O}\}$ 推得 T 為 *one-to-one?*

Question 5.7. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 在 *Question 5.5* 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 *kernel* 的等價關係嗎?

對於一般的函數, 要判斷其是否為 one-to-one 並不容易, 但當 T 為 linear transformation 時, Proposition 5.2.6 提供我們一個簡便的方法判斷 T 是否為 one-to-one. 也就是僅要確認是否有非零向量會映射到 $\mathbf{O}$ 即可. 所以求一個 linear transformation 的 kernel 是一個重要的課題. 在下一節中, 我們將介紹求得 kernel 的方法.

5.3. Matrix Representation

給定一個 $m\times n$ matrix A, 前面我們已知可以定義出一個 linear transformation $T:\mathbb{R}^n\to\mathbb{R}^m$, 其定義為 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$. 在這一節中, 我們要說明所有的 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformations 都可以寫成這樣的形式. 並利用此將 linear transformation 和 matrix 相連結, 來推得一些重要的性質.

前面 Theorem 5.1.8 告訴我們給定一個 linear transformation, 只要知道此 linear transformation 將一組 $\mathbb{R}^n$ 的 basis 對應到哪些向量, 就可以唯一確定這一個 linear transformation. 在 $\mathbb{R}^n$ 中, 我們有一個最簡單的 basis, 即 standard basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$. 若 $T:\mathbb{R}^n\to\mathbb{R}^m$

是一個 linear transformation, 由前面所述, 我們僅要知道 $T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n)$ 是哪些 $\mathbb{R}^m$ 的 vectors, 就可以知道任意 $\mathbf{v}\in\mathbb{R}^n$, $T(\mathbf{v})$ 為何了.

事實上對於每一個 $\mathbf{v}\in\mathbb{R}^n$, 都可以找到 $c_1,\ldots,c_n\in\mathbb{R}$ 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 $\mathbf{v}\in\mathbb{R}^n$, 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 5.3.1. *給定一個 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 function T. 則 T 為 linear transformation 若且唯若存在一個 $m\times n$ matrix A 使得 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$. 此 $m\times n$ matrix A 是唯一的, 事實上對任意 $i=1,\ldots,n$, A 的 i-th column 為 $T(\mathbf{e}_i)$, 其中 $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ 為 $\mathbb{R}^n$ 的 standard basis.*

Proof. 由 Lemma 5.1.5 我們知道, 若 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$, 則 T 為 linear transformation. 反之, 若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 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 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$.

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 $i=1,\ldots,n$, $B\mathbf{e}_i$ 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 5.3.1 告訴我們從 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 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 5.3.2. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation 且 $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ 為 $\mathbb{R}^n$ 的 standard basis. 則對於 $i=1,\ldots,n$, 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 $\mathbf{v}\in\mathbb{R}^n$, 我們有 $T(\mathbf{v})=[T]\mathbf{v}$.

Example 5.3.3. 我們探討 Example 5.2.5 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一個 linear transformation 的 range 和 kernel.

Proposition 5.3.4. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation* 且令 $[T]\in\mathscr{M}_{m\times n}$ 為其 *standard matrix representation.* 則 T 的 *range* 等於 $[T]$ 的 *column space,* 而 T 的 *kernel* 等於 $[T]$ 的 *nullspace.*

Proof. 由於 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 為 $\mathbb{R}^n$ 的一組 basis, 由 Proposition 5.2.4 我們知 T 的 range, 即 $T(\mathbb{R}^n)=\mathrm{Span}(T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n))$. 然而 $T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n)$ 剛好就是 $[T]$ 的 n 個 column, 故由定義 $\mathrm{Span}(T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n))$ 就是 $[T]$ 的 column space. 得證 T 的 range 就是 $[T]$ 的 column space.

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

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 nullspace 的維度稱為 A 的 nullity, 用 $\mathrm{null}(A)$ 來表示 (參見 Definition 2.3.1). 由 Proposition 5.3.4, 我們知道 T 的 range 的維度等於 $\mathrm{rank}([T])$, 而 T 的 kernel 的維度等於 $\mathrm{null}([T])$, 也就是說我們有以下的結果.

Corollary 5.3.5. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation* 且令 $[T]\in\mathscr{M}_{m\times n}$ 為其 *standard matrix representation.* 則

$$\dim(T(\mathbb{R}^n))=\mathrm{rank}([T])\quad\text{and}\quad\dim(\ker(T))=\mathrm{null}([T]).$$

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 kernel 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, 因為這一層關係我們也很容易推得 linear transformation 的 Dimension Theorem.

Theorem 5.3.6 (Dimension Theorem for Linear Transformation)**.** 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 則

$$\dim(T(\mathbb{R}^n))+\dim(\ker(T))=n.$$

Proof. 由於 T 的 standard matrix representation $[T]$ 為 $m\times n$ matrix, 由 Theorem 4.4.13 我們知 $\mathrm{rank}([T])+\mathrm{null}([T])=n$. 故套用 Corollary 5.3.5, 得證本定理. □

Example 5.3.7. 我們利用 Example 5.3.3 中的 linear transformation 及其 standard matrix representation 探討其 range 和 kernel.

(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 5.3.4 我們有 $T(\mathbb{R}^3)=\mathbb{R}^2$ (此與 Example 5.2.5(1) 一致). 另一方面, $[T]$ 的 nullspace 為聯立方程組 $[T]\mathbf{x}=\mathbf{O}$, 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 5.3.4 我們有 $\ker(T)=\mathrm{Span}(\begin{bmatrix}1\\-1\\1\end{bmatrix})$ (此與 Example 5.2.7(1) 一致). 檢查 Dimension Theorem, 我們確實有 $\dim(T(\mathbb{R}^3))+\dim(\ker(T))=2+1=3$.

(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 5.3.4 我們有 $T(\mathbb{R}^3)=\mathrm{Span}(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix})$ (此與 Example 5.2.5(2) 一致). 另一方面, $[T]$ 的

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

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

的解集合. 因此由 Proposition 5.3.4 我們有 $\ker(T)=\{\begin{bmatrix}0\\0\end{bmatrix}\}=\{\mathbf{O}\}$ (此與 Example 5.2.7(2) 一致). 檢查 Dimension Theorem, 我們確實有 $\dim(T(\mathbb{R}^2))+\dim(\ker(T))=2+0=2$.

當 T_1,T_2 皆為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation 時, 對任意 $c_1,c_2\in\mathbb{R}$, 我們可以利用它們得到一個新的 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 linear transformation $c_1T_1+c_2T_2$ (參見 Proposition 5.1.6). 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 $\mathbb{R}^m$ 到 $\mathbb{R}^k$ 的 linear transformation, 我們可得合成函數 $T\circ T_1$ 為 $\mathbb{R}^n$ 到 $\mathbb{R}^k$ 的 linear transformation (參見 Proposition 5.1.7). 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 5.3.8. 設 T_1,T_2 為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *linear transformations,* 而 T 為 $\mathbb{R}^m$ 到 $\mathbb{R}^k$ 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 $c_1,c_2\in\mathbb{R}$, 皆有 $c_1T_1+c_2T_2:\mathbb{R}^n\to\mathbb{R}^m$ 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) $T\circ T_1:\mathbb{R}^n\to\mathbb{R}^k$ 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 $\mathbf{v}\in\mathbb{R}^n$, 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 $c_1T_1+c_2T_2:\mathbb{R}^n\to\mathbb{R}^m$ 的 standard matrix representation 之要求, 故由 standard matrix representation 的唯一性 (Theorem 5.3.1) 知 $[c_1T_1+c_2T_2]=c_1[T_1]+c_2[T_2]$.

(2) 依定義對任意 $\mathbf{v}\in\mathbb{R}^n$, 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 $\mathbf{w}\in\mathbb{R}^m$ 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 $T\circ T_1:\mathbb{R}^n\to\mathbb{R}^k$ 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 5.3.1) 知 $[T\circ T_1]=[T][T_1]$. □

Example 5.3.9. 我們利用 Example 5.3.3 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 $T:\mathbb{R}^n\to\mathbb{R}^m$, $T':\mathbb{R}^m\to\mathbb{R}^k$ 為 linear transformations. 由於 T 的 range 是 $\mathbb{R}^m$ 的 subspace, 即 $T(\mathbb{R}^n)\subseteq\mathbb{R}^m$, 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 的 subspace. 利用 subspace 之間 dimension 的關係 (Corollary 4.3.6), 我們得 $\dim((T'\circ T)(\mathbb{R}^n))\leq\dim(T'(\mathbb{R}^m))$. 利用此我們可以推得以下關於矩陣的性質.

Proposition 5.3.10. *假設 A 為 $m\times n$ matrix, B 為 $k\times m$ matrix, 則* $\operatorname{rank}(BA)\leq\operatorname{rank}(A)$ *且* $\operatorname{rank}(BA)\leq\operatorname{rank}(B)$.

Proof. 考慮 $T:\mathbb{R}^n\to\mathbb{R}^m$ 定義為 $T(\mathbf{v})=A\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$, 以及 $T':\mathbb{R}^m\to\mathbb{R}^k$ 定義為 $T'(\mathbf{w})=B\mathbf{w}$, $\forall\mathbf{w}\in\mathbb{R}^m$. 依此定義, 我們有 $[T]=A,[T']=B$ 以及由 Lemma 5.3.8 知 $[T'\circ T]=[T'][T]=BA$. 由前面的討論我們又知 $\dim((T'\circ T)(\mathbb{R}^n))\leq\dim(T'(\mathbb{R}^m))$, 因此利用 Corollary 5.3.5 知 $\dim((T'\circ T)(\mathbb{R}^n))=\operatorname{rank}([T'\circ T])=\operatorname{rank}(BA)$ 以及 $\dim(T'(\mathbb{R}^m))=\operatorname{rank}([T'])=\operatorname{rank}(B)$, 得證 $\operatorname{rank}(BA)\leq\operatorname{rank}(B)$.

另一方面, 套用剛剛的結果, 我們有 $\operatorname{rank}(A^{\mathrm{T}}B^{\mathrm{T}})\leq\operatorname{rank}(A^{\mathrm{T}})$. 由於 $\operatorname{rank}(A^{\mathrm{T}})=\operatorname{rank}(A)$ 以及 $\operatorname{rank}((BA)^{\mathrm{T}})=\operatorname{rank}(BA)$ (Proposition 4.4.14), 我們得

$$\operatorname{rank}(BA)=\operatorname{rank}((BA)^{\mathrm{T}})=\operatorname{rank}(A^{\mathrm{T}}B^{\mathrm{T}})\leq\operatorname{rank}(A^{\mathrm{T}})=\operatorname{rank}(A).$$

□

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. 不過對於 invertible linear transformation, 利用 standard 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 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 invertible linear transformation. 由於 T 為 onto, 我們需要 $\dim(T(\mathbb{R}^n))=m$ (Proposition 5.2.3). 然而 T 為 one-to-one, 故知 $\dim(\ker(T))=0$ (Proposition 5.2.6). 利用 Dimension Theorem for linear transformation (Theorem 5.3.5) 我們得 $m=\dim(T(\mathbb{R}^n))=\dim(T(\mathbb{R}^n))+\dim(\ker(T))=n$. 這告訴我們只有在 $m=n$ 時, T 才有可能為 invertible. 現假設 $T:\mathbb{R}^n\to\mathbb{R}^n$ 為 one-to-one, 由 $\dim(\ker(T))=0$, 我們得 $\dim(T(\mathbb{R}^n))=\dim(T(\mathbb{R}^n))+\dim(\ker(T))=n$, 亦即 T 為 onto. 同樣的, 若 T 為 onto, 則由 $\dim(T(\mathbb{R}^n))=n$ 得 $\dim(\ker(T))=n-\dim(T(\mathbb{R}^n))=0$, 亦即 T 為 one-to-one. 這告訴我們當 T 是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的 linear transformation 時, 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 5.3.11. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation.* 則僅有當 $m=n$ 時, T 才有可能是 *invertible.* 又當 $m=n$ 時, T 為 *invertible* 和 T 為 *onto* 是等價的也和 T 為 *one-to-one* 等價.

由 Lemma 5.3.11 我們知道一個 linear transformation T 的 standard matrix representation $[T]$ 必須是 square matrix, T 才有可能是 invertible. 而且此時, 若 $[T]$ 為 $n\times n$ matrix (即 T 為 $\mathbb{R}^n$ 到 $\mathbb{R}^n$), 則 T 是 invertible 若且唯若 $\dim(T(\mathbb{R}^n))=n$, 亦即 $\operatorname{rank}([T])=n$. 另外這也等價於 $\dim(\ker(T))=0$, 亦即 $\operatorname{null}([T])=0$. 這都表示 $[T]$ 為 invertible matrix (參見 Theorem 3.5.2). 因此我們知 T 為 invertible 若且唯若 $[T]$ 為 invertible. 現若 T 為 invertible, 則因 $[T]$ 為 invertible matrix, 我們知 $[T]$ 的反矩陣 $[T]^{-1}$ 存在. 現考慮 $T':\mathbb{R}^n\to\mathbb{R}^n$ 其定義為 $T'(\mathbf{v})=[T]^{-1}\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^n$. 則我們有 $T'\circ T:\mathbb{R}^n\to\mathbb{R}^n$ 的 standard matrix representation 為 $[T'\circ T]=[T'][T]=[T]^{-1}[T]=I_n$. 同理 $T\circ T':\mathbb{R}^n\to\mathbb{R}^n$ 的 standard matrix representation 為 $[T\circ T']=[T][T']=[T][T]^{-1}=I_n$. $T'\circ T$ 和 $T\circ T'$ 的 standard matrix representation 皆等於 I_n 表示 $T'\circ T=T\circ T'$ 而且對任意 $\mathbf{v}\in\mathbb{R}^n$ 皆有 $(T'\circ T)(\mathbf{v})=(T\circ T')(\mathbf{v})=I_n\mathbf{v}=\mathbf{v}$, 得知 T' 為 T 的 inverse (反函數). 因此我們證得了以下之定理.

Theorem 5.3.12. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 *linear transformation* 且令 $[T]\in\mathscr{M}_{n\times n}$ 為其 *standard matrix representation.* 則 T 為 *invertible function* 若且唯若 $[T]$ 為 *invertible matrix.* 又若 T 為 *invertible,* 則 T 的 *inverse,* 亦為 *linear transformation* 且其 *standard matrix representation* 為 $[T]^{-1}$.

一般來講, 我們會將 invertible function f 的 inverse 用 f^{-1} 來表示. 所以當 $T:\mathbb{R}^n\to\mathbb{R}^n$ 為 invertible linear transformation, 我們也用 T^{-1} 來表示其 inverse. Theorem 5.3.12 告訴我們, 當 $T:\mathbb{R}^n\to\mathbb{R}^n$ 為 invertible linear transformation 時, $T^{-1}:\mathbb{R}^n\to\mathbb{R}^n$ 亦為 invertible linear transformation, 而且其 standard matrix representation 會是 $[T^{-1}]=[T]^{-1}$.

Example 5.3.13. 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 3.5.8 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.

5.4. 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.