Chapter 4

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.

4.1. 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像 $\mathbb{R}^n$ 這樣所謂的 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.

4.1.1. 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).

4.1.2. 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 4.1.1. 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 4.1.2. 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 5.1.2), 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 4.1.3. 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 4.1.2, 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

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

故由數學歸納法知 T 為 linear transformation. $\square$

Example 4.1.4. (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 4.1.2 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 4.1.5. 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 4.1.5 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 4.1. 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 $cT:\mathbb{R}^n\to\mathbb{R}^m$ 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 4.2. 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. $c_1,\ldots,c_k\in\mathbb{R}$, 則由前知 $c_1T_1,\ldots,c_kT_k$ 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 4.1.6. 設 $T_1,\cdots,T_k$ 為 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 *linear transformations,* $c_1,\ldots,c_k\in\mathbb{R}$ 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 4.1.7. 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. $\square$

Question 4.3. 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?

回顧在 $\mathbb{R}^n$ 中, 我們有一組所謂的 standard basis $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$. 對於任意 $\mathbb{R}^n$ 上的向量 $\mathbf{v}=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}$, 我們有 $\mathbf{v}=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$. 因此若 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear, 則 $T(\mathbf{v})=T(c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n)=c_1T(\mathbf{e}_1)+\cdots+c_nT(\mathbf{e}_n)$. 也就是說, 只要我們知道 $T(\mathbf{e}_1),\dots,T(\mathbf{e}_n)$ 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 4.1.8. 任意給定 $\mathbf{w}_1,\dots,\mathbf{w}_n\in\mathbb{R}^m$, 則存在唯一的 *linear transformation* $T:\mathbb{R}^n\to\mathbb{R}^m$, 滿足 $T(\mathbf{e}_1)=\mathbf{w}_1,\dots,T(\mathbf{e}_i)=\mathbf{w}_i,\dots,T(\mathbf{e}_n)=\mathbf{w}_n$.

Proof. 首先證明存在性. 定義 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 $T(c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$, $\forall c_1,\dots,c_n\in\mathbb{R}$. 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 $T(\mathbf{v})\in\mathbb{R}^m$. 然而因對任意 $\mathbf{v}\in\mathbb{R}^n$ 皆存在唯一的一組 $c_1,\dots,c_n\in\mathbb{R}$, 使得 $\mathbf{v}=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$. 故此時得 $T(\mathbf{v})=T(c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n)=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n\in\mathbb{R}^m$. 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 $T(\mathbf{u}+r\mathbf{v})=T(\mathbf{u})+rT(\mathbf{v})$. 由於存在 $c_1,\dots,c_n$ 以及 $d_1,\dots,d_n\in\mathbb{R}$ 使得 $\mathbf{u}=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$ 且 $\mathbf{v}=d_1\mathbf{e}_1+\cdots+d_n\mathbf{e}_n$. 故此時 $\mathbf{u}+r\mathbf{v}=(c_1+rd_1)\mathbf{e}_1+\cdots+(c_n+rd_n)\mathbf{e}_n$. 依 T 的定義得

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

另一方面

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

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

接著證明唯一性, 我們用反證法. 也就是說若 $T':\mathbb{R}^n\to\mathbb{R}^m$ 是另一個 linear transformation 滿足 $T'(\mathbf{e}_1)=\mathbf{w}_1,\dots,T'(\mathbf{e}_n)=\mathbf{w}_n$, 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 $c_1,\dots,c_n$, 使得 $\mathbf{v}=c_1\mathbf{e}_1+\cdots+c_n\mathbf{e}_n$, 故依 T,T' 皆為 linear 的假設, 我們有

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

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

要注意 Theorem 4.1.8 中的 $\mathbf{w}_1,\dots,\mathbf{w}_n\in\mathbb{R}^m$ 是可以任意選取的 (甚至不需相異). 這個定理告訴我們可以將 $\mathbb{R}^n$ 的 standard 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 4.1.8 告訴我們僅要檢查兩個 linear transformations 在 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 皆是一致的, 那麼這兩個 linear transformation 事實上就會是相同的函數.

4.2. 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 4.1.8 告訴我們若 $T : \mathbb{R}^n \to \mathbb{R}^m$ 是一個 linear transformation, 我們僅要知道 $T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)$ 是哪些 $\mathbb{R}^m$ 的 vectors, 就可以知道對於任意 $\mathbf{v} \in \mathbb{R}^n$, $T(\mathbf{v})$ 為何了. 事實上對於每一個 $\mathbf{v} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n$ 由於 $\mathbf{v} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$. 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 4.2.1. 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 $i = 1, \dots, n$, A 的 *i-th column* 為 $T(\mathbf{e}_i)$, 其中 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ 為 $\mathbb{R}^n$ 的 *standard basis.*

Proof. 由 Lemma 4.1.5 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 $i = 1, \dots, n$, $B\mathbf{e}_i$ 為 B 的 i-th column. 但由假設 $B\mathbf{e}_i = T(\mathbf{e}_i)$, 亦即 B 的 i-th column 為 $T(\mathbf{e}_i)$. 因此 B 的所有 column 皆與前述 A 的 column 相一致, 證得唯一性. $\square$

簡單來說 Theorem 4.2.1 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 4.2.2. 假設 $T : \mathbb{R}^n \to \mathbb{R}^m$ 為 linear transformation 且 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ 為 $\mathbb{R}^n$ 的 standard basis. 則對於 $i = 1, \dots, n$, 其 i-th column 為 $T(\mathbf{e}_i)$ 的 $m \times n$ matrix 稱為 T 的 *standard matrix representation.*

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

Example 4.2.3. 我們探討幾個 linear transformations 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

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

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 4.1.6). 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 4.1.7). 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 4.2.4. 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 4.2.1) 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 4.2.1) 知 $[T \circ T_1] = [T][T_1]$. □

Example 4.2.5. 我們利用 Example 4.2.3 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 $[T' \circ T] = [T'][T]$.

我們可以利用 standard matrix representation 來幫助我們了解 linear transformation. 假設 $T:\mathbb{R}^n\to\mathbb{R}^m$ 為 linear transformation, 則 T 的 standard matrix representation $[T]$ 為 $m\times n$ matrix. 由於對任意 $\mathbf{v}\in\mathbb{R}^n$, $T(\mathbf{v})=[T]\mathbf{v}$, 因此當 T 為 onto, 表示對任意 $\mathbf{w}\in\mathbb{R}^m$, 存在 $\mathbf{v}\in\mathbb{R}^n$ 使得 $T(\mathbf{v})=[T]\mathbf{v}=\mathbf{w}$. 亦即對於任意 $\mathbf{w}\in\mathbb{R}^m$, 聯立方程組 $[T]\mathbf{x}=\mathbf{w}$ 一定有解. 因此由 Theorem 3.4.2 知 $\mathrm{rank}([T])=m$. 另一方面由於 $T(\mathbf{O})=\mathbf{O}$, 因此若 T 為 one-to-one, 表示 $\mathbf{x}=\mathbf{O}$ 是在 $\mathbb{R}^n$ 中唯一的向量滿足 $T(\mathbf{x})=[T]\mathbf{x}=\mathbf{O}$. 換言之, 這表示聯立方程組 $[T]\mathbf{x}=\mathbf{O}$ 沒有 non-trivial solution. 因此由 Theorem 3.4.6 知 $\mathrm{rank}([T])=n$. 我們將這些結果歸納如下.

Proposition 4.2.6. 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.*

(1) T 為 *onto* 若且唯若 $\mathrm{rank}([T])=m$.

(2) T 為 *one-to-one* 若且唯若 $\mathrm{rank}([T])=n$.

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, 利用 Proposition 4.2.6(1) 知 $\mathrm{rank}([T])=m\leq n$. 然而 T 為 one-to-one, 同樣由 Proposition 4.2.6(2) 知 $\mathrm{rank}([T])=n\leq m$. 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 $\mathrm{rank}([T])=n$, 我們得 T 為 onto. 同理, 若 T 為 onto, 則 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 4.2.7. 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 4.2.7 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 $\mathrm{rank}([T])=n$. 然而這也等價於 $[T]$ 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 4.2.8. 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 4.2.8 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 4.2.9. 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.

4.3. Linear transformations of $\mathbb{R}^2$

在這節中我們將專注於探討 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的 linear transformations. 我們介紹三種特別的 linear transformations: projection (投影), reflection (鏡射) 以及 rotation (旋轉).

所有 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的 linear transformations 的 standard matrix representation 都是 2×2 matrix. 所以我們要探討這樣的 linear transformations 自然是由 2×2 matrices 的分類開始. 我們利用 rank 來分類. 首先最簡單的就是 rank 0 的情形. 此時表示 linear transformation 的 range 就是 $\{\mathbf{O}\}$, 也就是說此 linear transformation 就是將所有 $\mathbb{R}^2$ 的 vectors 送至 $\mathbf{O}$ 的 zero mapping. 簡單來說若 $T:\mathbb{R}^2\to\mathbb{R}^2$ 為 linear transformation 則其 standard matrix representation $[T]$ 的 rank 為 0 若且唯若 $T(\mathbf{v})=\mathbf{O}$, $\forall\mathbf{v}\in\mathbb{R}^2$. 接下來我們要探討 rank 大於 0 的情形.

4.3.1. Rank 1. 當一個 linear transformation 是 rank 1, 表示此 linear transformation 將 $\mathbb{R}^n$ 的向量皆送到與某個特定非零向量平行的向量. 大家應該會想到投影就有這個特點, 因為對某個非零向量的投影就是將所有的向量投射到與此向量平行的向量. 我們將簡單的回

顧一下投影的性質, 然後說明投影確實是 linear transformation. 再進一步談論所有 rank 1 的 linear transformation.

回顧在 Proposition 1.4.9, 我們證明了給定 $\mathbb{R}^n$ 中的非零向量 $\mathbf{w}$, 對任意 $\mathbf{v}\in\mathbb{R}^n$, 我們都可以將 $\mathbf{v}$ 拆寫成兩個向量之和, 其中一個和 $\mathbf{w}$ 平行, 另一個和 $\mathbf{w}$ 垂直. 亦即 $\mathbf{v}=\mathbf{w}'+\mathbf{v}'$, 其中 $\mathbf{w}'=r\mathbf{w}$, 且 $\mathbf{v}'\cdot\mathbf{w}=0$. 重要的是這個寫法是唯一的, 也就是說此時

$$\mathbf{w}'=\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}\quad\text{and}\quad\mathbf{v}'=\mathbf{v}-\mathbf{w}'.$$

我們稱此 $\mathbf{w}'$ 為 $\mathbf{v}$ 在 $\mathbf{w}$ 的 *projection* (投影). 因此我們可以定義所謂的 projection function $\mathrm{proj}_{\mathbf{w}}:\mathbb{R}^n\to\mathbb{R}^n$, 其定義為對所有 $\mathbf{v}\in\mathbb{R}^n$, $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})$ 為 $\mathbf{v}$ 在 $\mathbf{w}$ 的 projection, 亦即

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{v})=\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w},\ \forall\mathbf{v}\in\mathbb{R}^n.\tag{4.1}$$

利用 inner product 的性質, 我們很容易證明 $\mathrm{proj}_{\mathbf{w}}:\mathbb{R}^n\to\mathbb{R}^n$ 為 linear transformation. 事實上對任意 $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ 以及 $r\in\mathbb{R}$, 我們有

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{u}+r\mathbf{v})=\frac{(\mathbf{u}+r\mathbf{v})\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}=\frac{\mathbf{u}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}+r\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}=\mathrm{proj}_{\mathbf{w}}(\mathbf{u})+r\,\mathrm{proj}_{\mathbf{w}}(\mathbf{v}).$$

Question 4.4. 設 $\mathbf{w}$ 為 $\mathbb{R}^n$ 中的非零向量. 試利用 "對任意 $\mathbf{v}\in\mathbb{R}^n$, 存在唯一的 $\mathbf{w}',\mathbf{v}'\in\mathbb{R}^n$ 滿足 $\mathbf{v}=\mathbf{w}'+\mathbf{v}'$, $\mathbf{w}'=r\mathbf{w}$ 且 $\mathbf{v}'\cdot\mathbf{w}=0$" 這個事實證明 $\mathrm{proj}_{\mathbf{w}}$ 是 *linear transformation.*

現在我們回到 $\mathbb{R}^2$ 的情況. 設 $\mathbf{w}\in\mathbb{R}^2$ 且 $\mathbf{w}\neq\mathbf{O}$, 既然 $\mathrm{proj}_{\mathbf{w}}:\mathbb{R}^2\to\mathbb{R}^2$ 為 linear transformation, 我們自然想知道其 standard matrix representation. 依定義 $[\mathrm{proj}_{\mathbf{w}}]$ 會是一個 2×2 matrix, 其 1-st column 為 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)$, 2-nd column 為 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)$. 所以設 $\mathbf{w}=\begin{bmatrix}a\\b\end{bmatrix}$, 則由於 $\mathbf{e}_1\cdot\mathbf{w}=a$, $\mathbf{e}_2\cdot\mathbf{w}=b$ 得 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)=\frac{a}{a^2+b^2}\mathbf{w}=\frac{a}{a^2+b^2}\begin{bmatrix}a\\b\end{bmatrix}$, $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)=\frac{b}{a^2+b^2}\mathbf{w}=\frac{b}{a^2+b^2}\begin{bmatrix}a\\b\end{bmatrix}$. 故知

$$[\mathrm{proj}_{\mathbf{w}}]=\begin{bmatrix}\frac{a^2}{a^2+b^2}&\frac{ab}{a^2+b^2}\\\frac{ab}{a^2+b^2}&\frac{b^2}{a^2+b^2}\end{bmatrix}=\frac{1}{a^2+b^2}\begin{bmatrix}a^2&ab\\ab&b^2\end{bmatrix}.\tag{4.2}$$

Example 4.3.1. 考慮 $\mathbf{w}=\begin{bmatrix}1\\2\end{bmatrix}$, 我們要求出 $\mathrm{proj}_{\mathbf{w}}$ 的 standard matrix representation.

方法一: 利用式子 (4.1) 我們有

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)=\frac{\mathbf{e}_1\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}=\frac{1}{5}\begin{bmatrix}1\\2\end{bmatrix}\quad\text{and}\quad\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)=\frac{\mathbf{e}_2\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}=\frac{2}{5}\begin{bmatrix}1\\2\end{bmatrix}.$$

故得

$$[\mathrm{proj}_{\mathbf{w}}]=\begin{bmatrix}\frac{1}{5}&\frac{2}{5}\\\frac{2}{5}&\frac{4}{5}\end{bmatrix}=\frac{1}{5}\begin{bmatrix}1&2\\2&4\end{bmatrix}.$$

方法二: 考慮 $\mathbf{u}=\begin{bmatrix}-2\\1\end{bmatrix}$, 我們有 $\mathbf{u}\cdot\mathbf{w}=0$, 亦即 $\mathbf{u}$ 和 $\mathbf{w}$ 為垂直. 由於是投影, 我們有 $\mathrm{proj}_{\mathbf{w}}(\mathbf{w})=\mathbf{w}$ 以及 $\mathrm{proj}_{\mathbf{w}}(\mathbf{u})=\mathbf{O}$. 現由 Proposition 1.4.9, 我們知道所有 $\mathbf{v}\in\mathbb{R}^2$ 皆可寫成 $\mathbf{w},\mathbf{u}$ 的 linear combination. 特別的我們可找到 $c_1,c_2\in\mathbb{R}$ 使得 $\begin{bmatrix}1\\0\end{bmatrix}=c_1\begin{bmatrix}1\\2\end{bmatrix}+c_2\begin{bmatrix}-2\\1\end{bmatrix}$, 同理

存在 $d_1,d_2\in\mathbb{R}$ 使得 $\begin{bmatrix}0\\1\end{bmatrix}=d_1\begin{bmatrix}1\\2\end{bmatrix}+d_2\begin{bmatrix}-2\\1\end{bmatrix}$. 也就是說, 我們要解聯立方程組

$$\begin{cases} c_1 & -2c_2 & = & 1\\ 2c_1 & +c_2 & = & 0\end{cases} \quad\text{and}\quad \begin{cases} d_1 & -2d_2 & = & 0\\ 2d_1 & +d_2 & = & 1\end{cases}.$$

解得 $c_1=1/5,c_2=-2/5$ 以及 $d_1=2/5,d_2=1/5$, 亦即 $\mathbf{e}_1=(1/5)\mathbf{w}+(-2/5)\mathbf{u}$ 以及 $\mathbf{e}_2=(2/5)\mathbf{w}+(1/5)\mathbf{u}$. 故知

$$\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)=\mathrm{proj}_{\mathbf{w}}(\frac{1}{5}\mathbf{w}+\frac{-2}{5}\mathbf{u})=\frac{1}{5}\mathrm{proj}_{\mathbf{w}}(\mathbf{w})+\frac{-2}{5}\mathrm{proj}_{\mathbf{w}}(\mathbf{u})=\frac{1}{5}\mathbf{w},$$
$$\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)=\mathrm{proj}_{\mathbf{w}}(\frac{2}{5}\mathbf{w}+\frac{1}{5}\mathbf{u})=\frac{2}{5}\mathrm{proj}_{\mathbf{w}}(\mathbf{w})+\frac{1}{5}\mathrm{proj}_{\mathbf{w}}(\mathbf{u})=\frac{2}{5}\mathbf{w}.$$

因此求出 $[\mathrm{proj}_{\mathbf{w}}]=\frac{1}{5}\begin{bmatrix}1&2\\2&4\end{bmatrix}$.

我們還是要再強調一次, 雖然在 Example 4.3.1 的方法一, 看似較簡單, 但這是因為我們早已求出 projection 的通式 (式子 4.1). 在一般情形當我們不知一個 linear transformation 的通式時, 我們會用類似方法二, 也就是利用此 linear transformation 的特性來得到 standard matrix representation.

Question 4.5. 考慮 $\mathbf{w}'=-2\mathbf{w}$, 其中 $\mathbf{w}$ 就是 *Example 4.3.1* 中的 $\mathbf{w}$. 試求出 $\mathrm{proj}_{\mathbf{w}'}:\mathbb{R}^2\to\mathbb{R}^2$ 的 *standard matrix representation.*

依照 projection 的定義, 若 $\mathbf{w}'=r\mathbf{w}$, 其中 $r\neq 0$, 則 $\mathrm{proj}_{\mathbf{w}}$ 和 $\mathrm{proj}_{\mathbf{w}'}$ 是相同的. 所以一般我們會考慮 $\mathbf{w}$ 為 unit vector 的情形, 即 $\|\mathbf{w}\|=1$ 的情況. 此時若 $\mathbf{w}=\begin{bmatrix}a\\b\end{bmatrix}$, 則有 $a^2+b^2=1$, 所以 $[\mathrm{proj}_{\mathbf{w}}]=\begin{bmatrix}a^2&ab\\ab&b^2\end{bmatrix}$. 不管怎樣我們可以看出 $\mathrm{proj}_{\mathbf{w}}$ 的 rank 為 1. 另外我們注意 $[\mathrm{proj}_{\mathbf{w}}]$ 會是 symmetric matrix (對稱矩陣). 所以並不是所有的 rank 1 linear transformation 都會是 projection.

Question 4.6. (1) 我們知道所有 $\mathbb{R}^2$ 上的 *unit vector* 都可以寫成 $\mathbf{w}=\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$ 的形式, 其中 θ 為 $\mathbf{w}$ 與 $\mathbf{e}_1$ 的夾角. 試寫下 $[\mathrm{proj}_{\mathbf{w}}]$.

(2) 現若 $\mathbf{v}\in\mathbb{R}^2$ 亦為 *unit vector,* 且 $\mathbf{v}$ 與 $\mathbf{e}_1$ 的夾角為 δ, 試求 $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})$.

(3) 由於 $\mathbf{w},\mathbf{v}$ 的夾角為 $\delta-\theta$, 依照投影的定義 $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})=\cos(\delta-\theta)\mathbf{w}$. 利用以上求出的結果推出餘弦函數的差角公式.

既然不是所有 rank 1 的 linear transformation 都是 projection, 那要如何確認那些是 projection? 那些不是呢? 其實這很簡單, 我們只要運用 Example 4.3.1 的方法二, 就可以知道了. 假設 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是 rank 1 的 linear transformation. 我們知道必存在 $\mathbf{w}\in\mathbb{R}^2$ 滿足 $[T]$ 的 column vectors 皆與 $\mathbf{w}$ 平行. 另外我們再找一個非零向量 $\mathbf{u}$ 滿足 $\mathbf{u}\cdot\mathbf{w}=0$. 現若 T 是一個 projection, 則 T 必為對 $\mathbf{w}$ 的投影 (即 $T=\mathrm{proj}_{\mathbf{w}}$). 也就是說 T 會固定 $\mathbf{w}$ (即 $T(\mathbf{w})=\mathbf{w}$), 且將 $\mathbf{u}$ 送到 $\mathbf{O}$ (即 $T(\mathbf{u})=\mathbf{O}$). 事實上我們僅要檢查 $T(\mathbf{v})=\mathbf{v}$ 和 $T(\mathbf{u})=\mathbf{O}$ 是否都成立即可. 因為若有一項不成立, 那 T 當然不會是 projection. 而若兩項都成立, 以後我們會知道此時 T 必為對 $\mathbf{w}$ 的投影. 我們看以下的例子.

Example 4.3.2. 考慮以下三個矩陣, 我們要判斷哪一個是一個 projection 的 standard matrix representation.

$$A=\begin{bmatrix}1&-2\\1&-2\end{bmatrix}, B=\begin{bmatrix}1&1\\1&1\end{bmatrix}, C=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}.$$

可以看出三個矩陣的 column vectors 皆與 $\begin{bmatrix}1\\1\end{bmatrix}$ 平行, 所以我們考慮對 $\mathbf{w}=\begin{bmatrix}1\\1\end{bmatrix}$ 的投影. 另外我們要選 $\mathbf{u}$ 滿足 $\mathbf{u}\cdot\mathbf{w}=0$, 可以考慮 $\mathbf{u}=\begin{bmatrix}-1\\1\end{bmatrix}$. 很容易看出 A 不是 symmetric, 我們知 A 不會是 projection. 事實上我們有 $A\mathbf{u}\neq\mathbf{O}$. 另外雖然 $B\mathbf{u}=\mathbf{O}$, 但 $B\mathbf{w}=2\mathbf{w}$, 所以 B 也不是 projection. 最後我們檢查 $C\mathbf{u}=\mathbf{O}$ 且 $C\mathbf{w}=\mathbf{w}$, 故知 $[\mathrm{proj}_{\mathbf{w}}]=C$.

雖然並不是所有的 rank 1 linear transformation 都是 projection, 不過除了 standard matrix representation 為 $\begin{bmatrix}r&s\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\s&r\end{bmatrix}, s\neq 0$ 這兩類的情況外, 其他的 rank 1 linear transformation 都會是一些正反向伸縮與 projection 的合成. 我們用 Example 4.3.2 的例子來說明.

Example 4.3.3. 在 Example 4.3.2 中我們有 $A=\begin{bmatrix}1&-2\\1&-2\end{bmatrix}$ 為 rank 1, 但不是 projection. 為了方便起見, 我們令 $T_A:\mathbb{R}^2\to\mathbb{R}^2$ 表示 A 所對應的 linear transformation, 即 $T_A(\mathbf{v})=A\mathbf{v}, \forall\mathbf{v}\in\mathbb{R}^2$. 如同 Example 4.3.2 我們令 $\mathbf{w}=\begin{bmatrix}1\\1\end{bmatrix}$.

比較 $T_A(\mathbf{e}_1)=A\mathbf{e}_1=\mathbf{w}$, $T_A(\mathbf{e}_2)=A\mathbf{e}_2=-2\mathbf{w}$ 以及 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)=\frac{1}{2}\mathbf{w}, \mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)=\frac{1}{2}\mathbf{w}$. 我們知道 T_A 會將 $\mathbf{e}_1$ 送到 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)$ 的 2 倍 (即 $T_A(\mathbf{e}_1)=2\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)$). 不過 $\mathrm{proj}_{\mathbf{w}}$ 是 linear transformation, 若我們先將 $\mathbf{e}_1$ 伸長 2 倍後 (即變成 $2\mathbf{e}_1$) 再代入 $\mathrm{proj}_{\mathbf{w}}$, 就會得到 $\mathrm{proj}_{\mathbf{w}}(2\mathbf{e}_1)=2\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_1)=\mathbf{w}=T_A(\mathbf{e}_1)$. 同理 T_A 將 $\mathbf{e}_2$ 送到 $\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)$ 反向的 4 倍 (即 $T_A(\mathbf{e}_2)=-4\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)$). 因此若我們先將 $\mathbf{e}_2$ 反向伸長 4 倍 (即變成 $-4\mathbf{e}_1$) 再代入 $\mathrm{proj}_{\mathbf{w}}$, 就會得到 $\mathrm{proj}_{\mathbf{w}}(-4\mathbf{e}_2)=-4\mathrm{proj}_{\mathbf{w}}(\mathbf{e}_2)=-2\mathbf{w}=T_A(\mathbf{e}_2)$. 也就是說, 如果我們先考慮將 $\mathbf{e}_1$ 送至 $2\mathbf{e}_1$ 且將 $\mathbf{e}_2$ 送至 $-4\mathbf{e}_2$ 的 linear transformation T (即 $T:\mathbb{R}^2\to\mathbb{R}^2$ 滿足 $T(\mathbf{e}_1)=2\mathbf{e}_1$, $T(\mathbf{e}_2)=-4\mathbf{e}_2$), 則將 T 與 $\mathrm{proj}_{\mathbf{w}}$ 合成就會等於 T_A 了 (即 $\mathrm{proj}_{\mathbf{w}}\circ T=T_A$). 由於 T 的 standard matrix representation 為 $[T]=\begin{bmatrix}2&0\\0&-4\end{bmatrix}$, 我們驗證

$$[\mathrm{proj}_{\mathbf{w}}][T]=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}\begin{bmatrix}2&0\\0&-4\end{bmatrix}=\begin{bmatrix}1&-2\\1&-2\end{bmatrix}=A=[T_A].$$

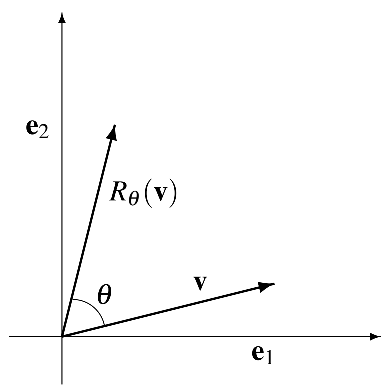
確實得證 $\mathrm{proj}_{\mathbf{w}}\circ T=T_A$.

Question 4.7. 試將 *Example 4.3.2* 的矩陣 B 所對應的 *linear transformation* 寫成一個正反向伸縮與 *projection* 的合成.

這些例子說明了, 我們可以將一個 linear transformation 拆解成一些較簡單的 linear transformations 的合成. 這種方法可以讓我們更容易了解一個 linear transformation, 以後我們還會看到相關的例子.

4.3.2. Rank 2. 由於我們考慮的是 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的 linear transformations. 在這情況下, rank 2 的 linear transformation 都會是 invertible (參見 Theorem 4.2.8). 要注意此時 linear transformation 雖然是 one-to-one and onto, 但它不一定會保持向量的長度. 例如前面所提的正反向伸縮這樣的 linear transformation, 如 Example 4.3.3 中的 T, 它將 $\mathbf{e}_1$ 伸長 2 倍, $\mathbf{e}_2$ 反向伸長 4 倍, 其 standard matrix representation 為 $[T] = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$, 是一個 rank 2 的 invertible matrix.

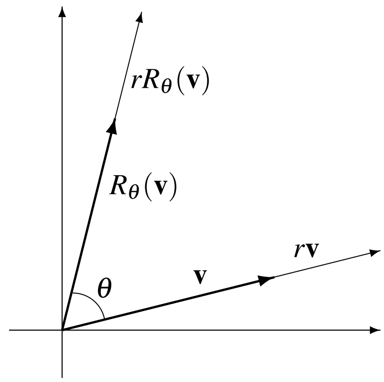
不過 $\mathbb{R}^2$ 中確實有兩種保持長度的 linear transformation. 第一種是所謂的 rotation, 它的作用是將 $\mathbb{R}^2$ 的向量 (將起始點置於原點) 以原點為中心逆時鐘方向繞 θ 角. 若用 $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ 表示這一個函數, 我們有以下的圖示



我們可以說, 若 $\mathbf{v}$ 和 $\mathbf{e}_1$ 的夾角為 δ, 則 $R_\theta(\mathbf{v})$ 是和 $\mathbf{e}_1$ 的夾角為 $\delta+\theta$ 且長度和 $\mathbf{v}$ 相同的向量.

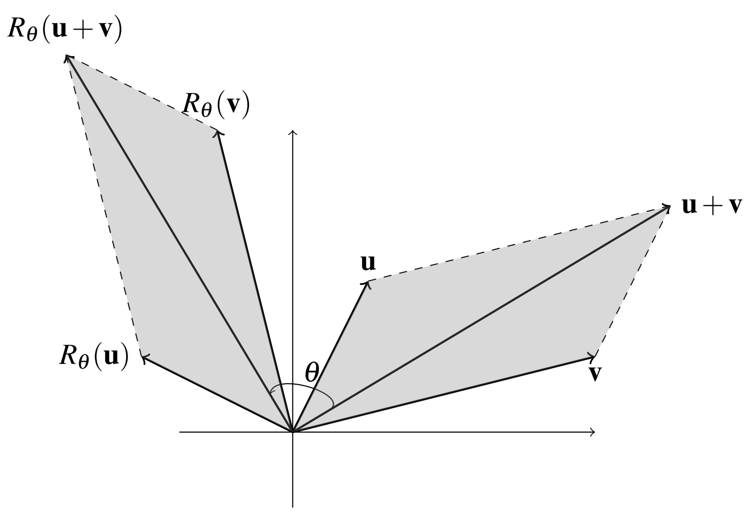
我們必須說明 R_θ 是一個 linear transformation. 由於我們想用幾何的方式來證明, 而一般給定 $\mathbf{u},\mathbf{v} \in \mathbb{R}^n$, $r \in \mathbb{R}$, 我們要表達 $\mathbf{u}+r\mathbf{v}$ 較複雜, 所以我們將 R_θ 為 linear transformation 的證明拆解成分別證明 $R_\theta(r\mathbf{v}) = rR_\theta(\mathbf{v})$ 以及 $R_\theta(\mathbf{u}+\mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$. 很容易看出來, 這就等同於證明 $R_\theta(\mathbf{u}+r\mathbf{v}) = R_\theta(\mathbf{u}) + rR_\theta(\mathbf{v})$.

當 $r \geq 0$ 時 $r\mathbf{v}$ 和 $\mathbf{v}$ 同向只是長度會是 $\mathbf{v}$ 長度的 r 倍. 所以當 $\mathbf{v}$ 經 θ 角旋轉後, $r\mathbf{v}$ 會跟著選轉到與 $R_\theta(\mathbf{v})$ 同向, 且由於 $R_\theta(\mathbf{v})$ 與 $\mathbf{v}$ 等長, 所以 $r\mathbf{v}$ 經旋轉後的長度會是 $R_\theta(\mathbf{v})$ 長度的 r 倍. 換句話說 $R_\theta(r\mathbf{v})$ 和 $R_\theta(\mathbf{v})$ 同向且長度會是 $R_\theta(\mathbf{v})$ 長度的 r 倍, 故得 $R_\theta(r\mathbf{v}) = rR_\theta(\mathbf{v})$, 如下圖所示.

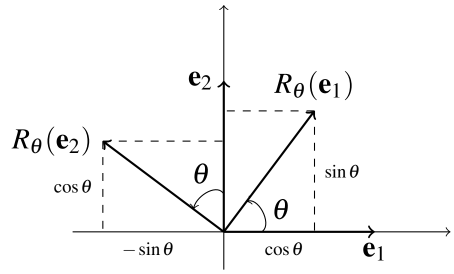


同理當 $r<0$ 時, 由於 $r\mathbf{v}$ 是取 $\mathbf{v}$ 反向的 $-r$ 倍. 所以當 $r\mathbf{v}$ 經 θ 角旋轉後, $r\mathbf{v}$ 會被轉至 $R_\theta(\mathbf{v})$ 的反向的 $-r$ 倍. 也就是 $R_\theta(r\mathbf{v})=rR_\theta(\mathbf{v})$.

現說明 $R_\theta(\mathbf{u}+\mathbf{v})=R_\theta(\mathbf{u})+R_\theta(\mathbf{v})$. 我們考慮 $\mathbf{u},\mathbf{v}$ 所張的平行四邊形. 依定義, $\mathbf{u},\mathbf{v}$ 所夾的對角線所形成的向量就是 $\mathbf{u}+\mathbf{v}$. 現將此平行四邊形經 θ 角旋轉後所得的平行四邊形, 就是 $R_\theta(\mathbf{u}),R_\theta(\mathbf{v})$ 所張的平行四邊形, 如下圖所示. 此時 $\mathbf{u},\mathbf{v}$ 所夾的對角線也被轉至 $R_\theta(\mathbf{u}),R_\theta(\mathbf{v})$ 所夾的對角線. 換言之, $\mathbf{u}+\mathbf{v}$ 會被轉至 $R_\theta(\mathbf{u}),R_\theta(\mathbf{v})$ 所夾的對角線所形成的向量, 即 $R_\theta(\mathbf{u})+R_\theta(\mathbf{v})$. 證得 $R_\theta(\mathbf{u}+\mathbf{v})=R_\theta(\mathbf{u})+R_\theta(\mathbf{v})$.



既然 R_θ 是 linear transformation, 我們自然要找出其 standard matrix representation. 現將 $\mathbf{e}_1$ 逆時鐘方向轉 θ 角後, 所得的向量長度仍為 1, 依正弦餘弦函數定義, 其 x 坐標為 $\cos\theta$, y 坐標為 $\sin\theta$. 同理 $\mathbf{e}_2$ 從原來與 $\mathbf{e}_1$ 夾 $\pi/2$ 轉成夾 $(\pi/2)+\theta$, 所以轉 θ 角後其 x 坐標為 $\cos((\pi/2)+\theta)=-\sin\theta$, y 坐標為 $\sin((\pi/2)+\theta)=\cos\theta$, 如下圖所示.



我們得到 $R_\theta(\begin{bmatrix}1\\0\end{bmatrix})=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$, $R_\theta(\begin{bmatrix}0\\1\end{bmatrix})=\begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$, 故知

$$[R_\theta]=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

Example 4.3.4. 考慮將 $\mathbf{v}=\begin{bmatrix}1\\-2\end{bmatrix}$ 逆時鐘方向轉 $\pi/6$ 角後所得的向量. 我們可將 $\mathbf{v}$ 寫成 $\sqrt{5}\begin{bmatrix}\cos\delta\\ \sin\delta\end{bmatrix}$, 其中 δ 為 $\mathbf{v}$ 和 $\mathbf{e}_1$ 的夾角, 故有 $\cos\delta=1/\sqrt{5},\sin\delta=(-2)/\sqrt{5}$. 經轉 $\pi/6$ 角後得 $T_{\frac{\pi}{6}}(\mathbf{v})=\sqrt{5}\begin{bmatrix}\cos(\delta+(\pi/6))\\ \sin(\delta+(\pi/6))\end{bmatrix}$. 再利用和角公式

$$\cos(\delta+\frac{\pi}{6})=\cos\delta\cos\frac{\pi}{6}-\sin\delta\sin\frac{\pi}{6}=\frac{1}{\sqrt{5}}\frac{\sqrt{3}}{2}-\frac{-2}{\sqrt{5}}\frac{1}{2}=\frac{2+\sqrt{3}}{2\sqrt{5}},$$

$$\sin(\delta+\frac{\pi}{6})=\sin\delta\cos\frac{\pi}{6}+\cos\delta\sin\frac{\pi}{6}=\frac{-2}{\sqrt{5}}\frac{\sqrt{3}}{2}+\frac{1}{\sqrt{5}}\frac{1}{2}=\frac{1-2\sqrt{3}}{2\sqrt{5}},$$

由此得 $R_{\frac{\pi}{6}}(\mathbf{v})=\begin{bmatrix}(2+\sqrt{3})/2\\(1-2\sqrt{3})/2\end{bmatrix}$.

當然了, 我們可以利用 $R_{\frac{\pi}{6}}$ 的 standard matrix representation 輕鬆求出

$$R_{\frac{\pi}{6}}(\mathbf{v})=\begin{bmatrix}\cos\frac{\pi}{6} & -\sin\frac{\pi}{6}\\ \sin\frac{\pi}{6} & \cos\frac{\pi}{6}\end{bmatrix}\begin{bmatrix}1\\-2\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{bmatrix}\begin{bmatrix}1\\-2\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{3}+2}{2}\\ \frac{1-2\sqrt{3}}{2}\end{bmatrix}.$$

從上面的例子, 我們可以利用 R_θ 的 standard matrix representation 幫助我們推導正弦餘弦函數的和角公式. 事實上考慮長度為 1 且和 $\mathbf{e}_1$ 夾 δ 角的向量 $\mathbf{v}=\begin{bmatrix}\cos\delta\\ \sin\delta\end{bmatrix}$, 經逆時鐘方向轉 θ 角後, 我們得 $R_\theta(\mathbf{v})$ 為長度為 1 且和 $\mathbf{e}_1$ 夾 $\delta+\theta$ 角的向量, 即 $R_\theta(\mathbf{v})=\begin{bmatrix}\cos(\delta+\theta)\\ \sin(\delta+\theta)\end{bmatrix}$. 另一方面依 standard matrix representation 的性質,

$$R_\theta(\mathbf{v})=[R_\theta]\mathbf{v}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\delta\\ \sin\delta\end{bmatrix}=\begin{bmatrix}\cos\theta\cos\delta-\sin\theta\sin\delta\\ \sin\theta\cos\delta+\cos\theta\sin\delta\end{bmatrix}.$$

故得證

$$\cos(\theta+\delta)=\cos\theta\cos\delta-\sin\theta\sin\delta,\quad \sin(\theta+\delta)=\sin\theta\cos\delta+\cos\theta\sin\delta.$$

我們也可以用合成函數的觀點來得到和角公式. 由於 R_θ, R_δ 分別為逆時鐘方向轉 θ,δ 角. 所以將它們合成可得逆時鐘方向轉 $\theta+\delta$ 角. 也就是說 $R_\delta\circ R_\theta=R_{\delta+\theta}$ 故由 Lemma 4.2.4,

得 $[R_{\delta+\theta}]=[R_\delta\circ R_\theta]=[R_\delta][R_\theta]$, 因此有

$$\begin{aligned}\begin{bmatrix}\cos(\delta+\theta) & -\sin(\delta+\theta)\\ \sin(\delta+\theta) & \cos(\delta+\theta)\end{bmatrix} &= \begin{bmatrix}\cos\delta & -\sin\delta\\ \sin\delta & \cos\delta\end{bmatrix}\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\\ &= \begin{bmatrix}\cos\delta\cos\theta-\sin\delta\sin\theta & -\cos\delta\sin\theta-\sin\delta\cos\theta\\ \sin\delta\cos\theta+\cos\delta\sin\theta & -\sin\delta\sin\theta+\cos\delta\cos\theta\end{bmatrix}\end{aligned}$$

由此我們也可得到和角公式.

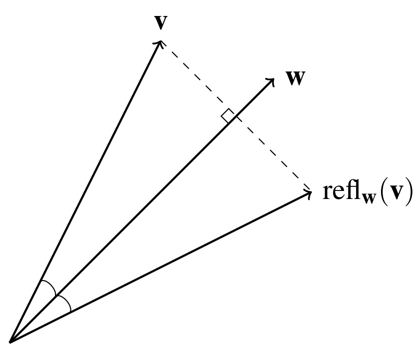
既然旋轉 R_θ 是 rank 2 的 linear transformation, 它自然是 invertible. 甚麼會是 R_θ 的 inverse 呢? 很容易了解逆時鐘轉 θ 角的還原是順時鐘轉 θ 角, 亦即逆時鐘轉 $-\theta$ 角. 因此我們有 $R_\theta^{-1}=R_{-\theta}$. 檢查一下 $R_{-\theta}$ 的 standard matrix representation, 我們有

$$[R_{-\theta}]=\begin{bmatrix}\cos(-\theta) & -\sin(-\theta)\\ \sin(-\theta) & \cos(-\theta)\end{bmatrix}=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix},$$

確實就是 $[R_\theta]$ 的反矩陣 $[R_\theta]^{-1}$.

Question 4.8. 假設 $\mathbf{w}=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ *(參見 Question 4.6)* 我們可以利用旋轉將對 $\mathbf{w}$ 的投影換成對 $\mathbf{e}_1$ 這樣比較簡單的投影. 首先我們利用旋轉 $-\theta$ 角, 即 $R_{-\theta}$ 將 $\mathbf{w}$ 轉至 $\mathbf{e}_1$, 再作 $\mathbf{e}_1$ 上的投影, 即 $\mathrm{proj}_{\mathbf{e}_1}$. 最後再轉 θ 角將 $\mathbf{e}_1$ 轉回 $\mathbf{w}$. 由於所有向量在 $\mathbf{w}$ 的投影所形成的“直角三角形”經由旋轉仍為相等直角三角形. 所以經過上述方法所得的結果確實和在 $\mathbf{w}$ 上的投影一致. 也就是說我們有 $\mathrm{proj}_{\mathbf{w}}=R_\theta\circ\mathrm{proj}_{\mathbf{e}_1}\circ R_{-\theta}$. 試寫下 $[R_\theta],[\mathrm{proj}_{\mathbf{e}_1}],[R_{-\theta}]$ 並利用它們得到 $[\mathrm{proj}_{\mathbf{w}}]$ *(驗證是否和 Question 4.6(1) 的結果相同)*.

第二種 $\mathbb{R}^2$ 中保持長度的 linear transformation 是所謂的 reflection. 它的定義是先給定 $\mathbb{R}^2$ 中的非零向量 $\mathbf{w}$, 以 $\mathbf{w}$ 所在的直線為對稱軸考慮 $\mathbf{v}$ 以此對稱軸所得的對稱向量即為 $\mathbf{v}$ 對 $\mathbf{w}$ 的 reflection, 用 $\mathrm{refl}_{\mathbf{w}}(\mathbf{v})$ 表示. 我們看以下的圖示:



由於 $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})$ 為 $\mathbf{v}$ 在 $\mathbf{w}$ 的投影, 從上面圖示可以看出 $\mathrm{refl}_{\mathbf{w}}(\mathbf{v})-\mathbf{v}=2(\mathrm{proj}_{\mathbf{w}}(\mathbf{v})-\mathbf{v})$. 所以我們可以推得 $\mathrm{refl}_{\mathbf{w}}(\mathbf{v})=2\mathrm{proj}_{\mathbf{w}}(\mathbf{v})-\mathbf{v}$. 令 $\mathrm{id}:\mathbb{R}^2\to\mathbb{R}^2$, 表示 $\mathbb{R}^2$ 上的 identity map, 即 $\mathrm{id}(\mathbf{v})=\mathbf{v}$, $\forall\mathbf{v}\in\mathbb{R}^2$. 則我們有 $\mathrm{refl}_{\mathbf{w}}=2\mathrm{proj}_{\mathbf{w}}-\mathrm{id}$. 因為 $\mathrm{proj}_{\mathbf{w}}:\mathbb{R}^2\to\mathbb{R}^2$ 和 $\mathrm{id}:\mathbb{R}^2\to\mathbb{R}^2$ 皆為 linear transformation, 利用 Proposition 4.1.6, 我們知 $2\mathrm{proj}_{\mathbf{w}}-\mathrm{id}$ 為 linear transformation, 亦即 $\mathrm{refl}_{\mathbf{w}}:\mathbb{R}^2\to\mathbb{R}^2$ 為 linear transformation. 我們也可利用 projection 的公式 (式子 (4.1)) 得到 reflection 的公式

$$\mathrm{refl}_{\mathbf{w}}(\mathbf{v})=2\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}-\mathbf{v},\ \forall\mathbf{v}\in\mathbb{R}^n. \tag{4.3}$$

同樣的, 我們可以利用這個公式證明 $\mathrm{refl}_{\mathbf{w}}$ 為 linear transformation.

Question 4.9. 試利用式子 *(4.3)* 證明 $\mathrm{refl}_{\mathbf{w}}:\mathbb{R}^2\to\mathbb{R}^2$ 為 *linear transformation.*

既然 $\mathrm{refl}_{\mathbf{w}}$ 是 linear transformation, 我們自然要找它的 standard matrix representation. 現在我們已經有許多方法處理這樣的問題了. 首先可以利用 $\mathrm{refl}_{\mathbf{w}}=2\mathrm{proj}_{\mathbf{w}}-\mathrm{id}$, 先求出 $[\mathrm{proj}_{\mathbf{w}}]$ 以及 $[\mathrm{id}]$, 再利用 Lemma 4.2.4 就可得 $[\mathrm{refl}_{\mathbf{w}}]=2[\mathrm{proj}_{\mathbf{w}}]-[\mathrm{id}]$ 了. 現假設 $\mathbf{w}=\begin{bmatrix}a\\b\end{bmatrix}$, 由式子 (4.2), 我們知 $[\mathrm{proj}_{\mathbf{w}}]=\frac{1}{a^2+b^2}\begin{bmatrix}a^2 & ab\\ ab & b^2\end{bmatrix}$. 另外由於 $\mathrm{id}(\mathbf{e}_1)=\mathbf{e}_1,\mathrm{id}(\mathbf{e}_2)=\mathbf{e}_2$, 我們知 $[\mathrm{id}]=I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}$. 故得

$$[\mathrm{refl}_{\mathbf{w}}]=\frac{2}{a^2+b^2}\begin{bmatrix}a^2 & ab\\ ab & b^2\end{bmatrix}-\begin{bmatrix}1&0\\0&1\end{bmatrix}=\frac{1}{a^2+b^2}\begin{bmatrix}a^2-b^2 & 2ab\\ 2ab & b^2-a^2\end{bmatrix}. \tag{4.4}$$

我們也可以利用式子 (4.3), 分別求 $\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1)$ 以及 $\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_2)$. 故當 $\mathbf{w}=\begin{bmatrix}a\\b\end{bmatrix}$ 時, 我們有

$$\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1)=\frac{1}{a^2+b^2}\begin{bmatrix}a^2-b^2\\2ab\end{bmatrix},\quad \mathrm{refl}_{\mathbf{w}}(\mathbf{e}_2)=\frac{1}{a^2+b^2}\begin{bmatrix}2ab\\b^2-a^2\end{bmatrix}.$$

得到和式子 (4.4) 相同的結果.

還有一種方法就是效法 Example 4.3.1 的方法二. 我們令 $\mathbf{u}=\begin{bmatrix}-b\\a\end{bmatrix}$. 此時因為 $\mathbf{w}$ 對 $\mathbf{w}$ 的鏡射是 $\mathbf{w}$ 本身, 我們有 $\mathrm{refl}_{\mathbf{w}}(\mathbf{w})=\mathbf{w}$. 而 $\mathbf{u}$ 和 $\mathbf{w}$ 垂直所以我們有 $\mathbf{u}$ 對 $\mathbf{w}$ 的鏡射是 $\mathbf{u}$ 的反向量, 即 $\mathrm{refl}_{\mathbf{w}}(\mathbf{u})=-\mathbf{u}$, 依此便能確定 $\mathrm{refl}_{\mathbf{w}}$ 了. 我們用下面的例子說明這一個方法.

Example 4.3.5. 如同 Example 4.3.1, 考慮 $\mathbf{w}=\begin{bmatrix}1\\2\end{bmatrix}$, 我們要求出 $\mathrm{refl}_{\mathbf{w}}$ 的 standard matrix representation. 利用類似 Example 4.3.1 的方法二, 我們考慮 $\mathbf{u}=\begin{bmatrix}-2\\1\end{bmatrix}$. 因為 $\mathbf{u}$ 和 $\mathbf{w}$ 為垂直, 我們有 $\mathrm{refl}_{\mathbf{w}}(\mathbf{w})=\mathbf{w}$ 以及 $\mathrm{refl}_{\mathbf{w}}(\mathbf{u})=-\mathbf{u}$. 現因 $\mathbf{w},\mathbf{u}$ 為 $\mathbb{R}^2$ 的一組 basis, 在 Example 4.3.1 我們已求得 $\mathbf{e}_1=(1/5)\mathbf{w}+(-2/5)\mathbf{u}$ 以及 $\mathbf{e}_2=(2/5)\mathbf{w}+(1/5)\mathbf{u}$. 故知

$$\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1)=\mathrm{refl}_{\mathbf{w}}(\frac{1}{5}\mathbf{w}+\frac{-2}{5}\mathbf{u})=\frac{1}{5}\mathrm{refl}_{\mathbf{w}}(\mathbf{w})+\frac{-2}{5}\mathrm{refl}_{\mathbf{w}}(\mathbf{u})=\frac{1}{5}\mathbf{w}+\frac{2}{5}\mathbf{u}=\frac{1}{5}\begin{bmatrix}-3\\4\end{bmatrix},$$

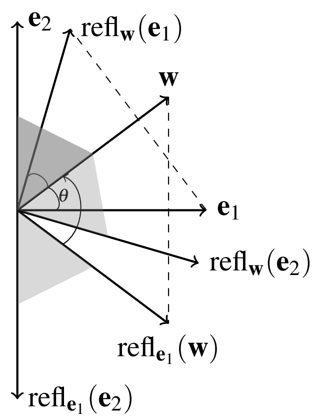
$$\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_2)=\mathrm{refl}_{\mathbf{w}}(\frac{2}{5}\mathbf{w}+\frac{1}{5}\mathbf{u})=\frac{2}{5}\mathrm{refl}_{\mathbf{w}}(\mathbf{w})+\frac{1}{5}\mathrm{refl}_{\mathbf{w}}(\mathbf{u})=\frac{2}{5}\mathbf{w}-\frac{1}{5}\mathbf{u}=\frac{1}{5}\begin{bmatrix}4\\3\end{bmatrix}.$$

因此求出 $[\mathrm{refl}_{\mathbf{w}}]=\frac{1}{5}\begin{bmatrix}-3&4\\4&3\end{bmatrix}$. 與式子 (4.4) 代 $a=1$, $b=2$ 相吻合.

Question 4.10. 試求 $[\mathrm{refl}_{\mathbf{e}_1}]$, $[\mathrm{refl}_{\mathbf{e}_2}]$ 以及 $[\mathrm{refl}_{\mathbf{w}}]$, 其中 $\mathbf{w}=\begin{bmatrix}1\\1\end{bmatrix}$.

和 projection 的情況相同, $\mathrm{refl}_{\mathbf{w}}$ 僅和 $\mathbf{w}$ 的方向有關和 $\mathbf{w}$ 的長度無關. 所以我們可以假設 $\mathbf{w}$ 為長度為 1 且與 $\mathbf{e}_1$ 夾 θ 角的向量, 即 $\mathbf{w}=\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$. 此時利用式子 (4.4) 得 $[\mathrm{refl}_{\mathbf{w}}]=\begin{bmatrix}\cos^2\theta-\sin^2\theta & 2\cos\theta\sin\theta\\ 2\cos\theta\sin\theta & \sin^2\theta-\cos^2\theta\end{bmatrix}$. 再利用倍角公式 $\cos^2\theta-\sin^2\theta=\cos(2\theta)$, $2\cos\theta\sin\theta=\sin(2\theta)$, 我們得 $[\mathrm{refl}_{\mathbf{w}}]=\begin{bmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{bmatrix}$.

為什麼 $\mathrm{refl}_{\mathbf{w}}$ 和 2θ 有關呢? 想想看, 由於 $\mathbf{w}$ 和 $\mathbf{e}_1$ 夾角為 θ, 所以經由對 $\mathbf{w}$ 鏡射, $\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1)$ 會和 $\mathbf{e}_1$ 夾 2θ 角. 所以對 $\mathbf{e}_1$ 來說對 $\mathbf{w}$ 做鏡射就等同於將 $\mathbf{e}_1$ 逆時鐘旋轉 2θ 角. 例如對 $\mathbf{w}$ 就不對了, 因為 $\mathbf{w}$ 對 $\mathbf{w}$ 做鏡射是 $\mathbf{w}$ 本身. 然而如果我們先將 $\mathbf{w}$ 對 $\mathbf{e}_1$ 做鏡射, 此時 $\mathbf{w}$ 會被映射到 $\mathrm{refl}_{\mathbf{e}_1}(\mathbf{w})$ 為與 $\mathbf{e}_1$ 夾 $-\theta$ 角的單位向量, 此時再做 2θ 角的旋轉, 就可將 $\mathbf{w}$ 送回原來的位置. 而對於 $\mathbf{e}_2$, 如果我們先將 $\mathbf{e}_2$ 對 $\mathbf{e}_1$ 做鏡射, $\mathbf{e}_2$ 會被映射到 $\mathrm{refl}_{\mathbf{e}_1}(\mathbf{e}_2) = -\mathbf{e}_2$. 此時 $-\mathbf{e}_2$ 與 $\mathrm{refl}_{\mathbf{e}_1}(\mathbf{w})$ 夾角和 $\mathbf{e}_2$ 與 $\mathbf{w}$ 的夾角相同, 但方向相反. 因此再做 2θ 角的旋轉, 由於 $\mathbf{w}$ 被送回原來的位置, 可得 $\mathbf{e}_2$ 被送至對 $\mathbf{w}$ 的鏡射 $\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_2)$. 另一方面先對 $\mathbf{e}_1$ 做鏡射會將 $\mathbf{e}_1$ 固定, 所以若先做對 $\mathbf{e}_1$ 的鏡射, 再做 2θ 角的旋轉這樣的動作確實將 $\mathbf{e}_1$ 送至 $\mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1)$. 我們有以下的圖示:



由這裡可知, $R_{2\theta}\circ\mathrm{refl}_{\mathbf{e}_1}:\mathbb{R}^2\to\mathbb{R}^2$ 這一個 linear transformation, 會將 $\mathbf{e}_1$ 映至

$$R_{2\theta}(\mathrm{refl}_{\mathbf{e}_1}(\mathbf{e}_1)) = R_{2\theta}(\mathbf{e}_1) = \begin{bmatrix}\cos(2\theta)\\ \sin(2\theta)\end{bmatrix} = \mathrm{refl}_{\mathbf{w}}(\mathbf{e}_1),$$

而將 $\mathbf{e}_2$ 映至

$$R_{2\theta}(\mathrm{refl}_{\mathbf{e}_1}(\mathbf{e}_2) = R_{2\theta}(-\mathbf{e}_2) = -R_{2\theta}(\mathbf{e}_2) = \begin{bmatrix}\sin(2\theta)\\ -\cos(2\theta)\end{bmatrix} = \mathrm{refl}_{\mathbf{w}}(\mathbf{e}_2).$$

由於此時 $R_{2\theta}\circ\mathrm{refl}_{\mathbf{e}_1}:\mathbb{R}^2\to\mathbb{R}^2$ 和 $\mathrm{refl}_{\mathbf{w}}:\mathbb{R}^2\to\mathbb{R}^2$ 在代入 $\mathbf{e}_1,\mathbf{e}_2$ 所得的結果相同, 故由 linear transformation 的唯一性 (Theorem 4.1.8) 得證 $R_{2\theta}\circ\mathrm{refl}_{\mathbf{e}_1} = \mathrm{refl}_{\mathbf{w}}$, 也因此我們會有

$$[\mathrm{refl}_{\mathbf{w}}] = [R_{2\theta}][\mathrm{refl}_{\mathbf{e}_1}] = \begin{bmatrix}\cos(2\theta) & -\sin(2\theta)\\ \sin(2\theta) & \cos(2\theta)\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{bmatrix}$$

之結果.

甚麼是 $\mathrm{refl}_{\mathbf{w}}$ 的 inverse 呢? 依鏡射的定義, 我們很容易看出任意的向量經由對 $\mathbf{w}$ 的鏡射後再做一次對 $\mathbf{w}$ 鏡射就回到原來的向量. 也就是說 $\mathrm{refl}_{\mathbf{w}}\circ\mathrm{refl}_{\mathbf{w}} = \mathrm{id}$. 因此可知 $\mathrm{refl}_{\mathbf{w}}^{-1} = \mathrm{refl}_{\mathbf{w}}$. 我們也可直接驗證 $[\mathrm{refl}_{\mathbf{w}}]^2 = I_2$, 所以得 $[\mathrm{refl}_{\mathbf{w}}]$ 的反矩陣就是 $[\mathrm{refl}_{\mathbf{w}}]$, 即 $[\mathrm{refl}_{\mathbf{w}}]^{-1} = [\mathrm{refl}_{\mathbf{w}}]$.

特別的 $[\mathrm{refl}_{\mathbf{e}_1}]^{-1} = [\mathrm{refl}_{\mathbf{e}_1}]$, 所以當 $\mathbf{w} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ 時, 由 $[\mathrm{refl}_{\mathbf{w}}] = [R_{2\theta}][\mathrm{refl}_{\mathbf{e}_1}]$ 我們推得

$$[R_{2\theta}] = ([R_{2\theta}][\mathrm{refl}_{\mathbf{e}_1}])[\mathrm{refl}_{\mathbf{e}_1}]^{-1} = [\mathrm{refl}_{\mathbf{w}}][\mathrm{refl}_{\mathbf{e}_1}]^{-1} = [\mathrm{refl}_{\mathbf{w}}][\mathrm{refl}_{\mathbf{e}_1}].$$

也就是說當 $\mathbf{w} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ 時, $R_{2\theta} = \mathrm{refl}_{\mathbf{w}} \circ \mathrm{refl}_{\mathbf{e}_1}$. 因此當我們要探討旋轉 θ 角即 $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ 時, 我們可以考慮 $\mathbf{w} = \begin{bmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{bmatrix}$. 此時可得 $R_\theta = \mathrm{refl}_{\mathbf{w}} \circ \mathrm{refl}_{\mathbf{e}_1}$. 也就是說我們可以把任何的旋轉看成是兩次鏡射的合成。

最後我們回到一般 rank 2 的 linear transformation T. 由於此時 T 為 invertible, 即 $[T]$ 為 invertible. 我們知存在 elementary matrices $E_1, \dots, E_n$ 使得 $[T] = E_n \cdots E_1$ (參見 Proposition 3.5.7). 換言之, 只要我們了解每一個 elementary matrix 所對應的 linear transformation, 那麼每一個 rank 2 的 linear transformation 就是這些 elementary matrices 所對應的 linear transformations 的合成.

首先我們來看兩個 row 交換的 elementary matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. 它所對應的 linear transformation 是 $\mathrm{refl}_{\mathbf{w}}$, 其中 $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (請自行套式子 (4.4) 檢驗). 也就是說它所對應的是對直線 $x = y$ 做鏡射.

至於將某個 row 乘上非零實數 r 的 elementary matrix $\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}$ 或 $\begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix}$. 就是前面提過的正反向伸縮. 事實上當 $r > 0$, $\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix}$ 分別表示對 $\mathbf{e}_1$, $\mathbf{e}_2$ 方向伸縮 r 倍. 而當 $r < 0$, 我們可以將它們拆成 $\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -r & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ 以及 $\begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -r \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. 其中 $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 分別表示對 y 軸, x 軸的鏡射. 所以這種 elementary matrix 所對應的 linear transformation, 大致上來說會是對 $\mathbf{e}_1$ 或 $\mathbf{e}_2$ 的伸縮, 或是先對 y 軸或 x 軸的鏡射再合成對 $\mathbf{e}_1$ 或 $\mathbf{e}_2$ 的伸縮.

最後將某個 row 乘上 r 加至另一個 row 的 elementary matrix $\begin{bmatrix} 1 & 0 \\ r & 1 \end{bmatrix}$ 或 $\begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}$. 其中 $\begin{bmatrix} 1 & 0 \\ r & 1 \end{bmatrix}$ 所對應的 linear transformation 將 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 映至 $\begin{bmatrix} 1 \\ r \end{bmatrix}$, 而將 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 固定. 這樣的 linear transformation 由於將垂直 (vertical) 向量 (即 $\mathbf{e}_2$) 固定而將水平 (horizontal) 向量 (即 $\mathbf{e}_1$) 映至一個斜的向量 $\mathbf{e}_1 + r\mathbf{e}_2$, 所以我們稱之為 *vertical shear*. 而 $\begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}$ 所對應的 linear transformation 將 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 固定, 而將 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 映至 $\begin{bmatrix} r \\ 1 \end{bmatrix}$. 這樣的 linear transformation 由於將水平 (horizontal) 向量 (即 $\mathbf{e}_1$) 固定而將垂直 (vertical) 向量 (即 $\mathbf{e}_2$) 映至一個斜的向量 $r\mathbf{e}_1 + \mathbf{e}_2$, 所以我們稱之為 *horizontal shear*. 我們看以下的例子.

Example 4.3.6. 我們來看逆時鐘旋轉 $\pi/3$ 的 linear transformation $R_{\pi/3}$, 可以拆解成哪些 elementary matrices 所對應的 linear transformations 之合成.

首先寫下 $R_{\pi/3}$ 的 standard matrix representation $\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$. 先將 1-st row 乘上 2 得 $\begin{bmatrix} 1 & -\sqrt{3} \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$, 此時左邊乘上的 elementary matrix 為 $E_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. 接著將 1-st row 乘上 $-\sqrt{3}/2$ 加至 2-nd row 得 $\begin{bmatrix} 1 & -\sqrt{3} \\ 0 & 2 \end{bmatrix}$, 此時左邊乘上的 elementary matrix 為 $E_2=\begin{bmatrix}1&0\\-\sqrt{3}/2&1\end{bmatrix}$. 然後將 2-nd row 乘上 1/2 得 $\begin{bmatrix}1&-\sqrt{3}\\0&1\end{bmatrix}$, 此時左邊乘上的 elementary matrix 為 $E_3=\begin{bmatrix}1&0\\0&1/2\end{bmatrix}$. 最後將 2-nd row 乘上 $\sqrt{3}$ 加至 1-st row 得 $\begin{bmatrix}1&0\\0&1\end{bmatrix}$, 此時左邊乘上的 elementary matrix 為 $E_4=\begin{bmatrix}1&\sqrt{3}\\0&1\end{bmatrix}$. 由於 $E_4E_3E_2E_1[R_{\pi/3}]=I_2$, 所以 $[R_{\pi/3}]$ 可分解成

$$\begin{bmatrix}1/2&-\sqrt{3}/2\\\sqrt{3}/2&1/2\end{bmatrix}=E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}=\begin{bmatrix}1/2&0\\0&1\end{bmatrix}\begin{bmatrix}1&0\\\sqrt{3}/2&1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&-\sqrt{3}\\0&1\end{bmatrix}.$$

換言之, 逆時鐘旋轉 $\pi/3$ 的 linear transformation $R_{\pi/3}$ 可以拆解成先做一個 horizontal shear 再將 $\mathbf{e}_2$ 方向伸長 2 倍接著做一個 vertical shear 最後將 $\mathbf{e}_1$ 方向縮成 1/2 倍.

注意利用 row operation 所得 elementary matrices 乘積順序的改變. 其實我們可以考慮 column operation 這樣所寫下的 elementary matrix 的乘積就可以保持順序了. 先將 1-st column 乘上 2 得 $\begin{bmatrix}1&-\sqrt{3}/2\\\sqrt{3}&1/2\end{bmatrix}$, 因為是 column operation 此時是右邊乘上 elementary matrix $E_1=\begin{bmatrix}2&0\\0&1\end{bmatrix}$. 接著將 1-st column 乘上 $\sqrt{3}/2$ 加至 2-nd column 得 $\begin{bmatrix}1&0\\\sqrt{3}&2\end{bmatrix}$, 此時右邊乘上的 elementary matrix 為 $E_2=\begin{bmatrix}1&\sqrt{3}/2\\0&1\end{bmatrix}$. 然後將 2-nd column 乘上 1/2 得 $\begin{bmatrix}1&0\\\sqrt{3}&1\end{bmatrix}$, 此時右邊乘上的 elementary matrix 為 $E_3=\begin{bmatrix}1&0\\0&1/2\end{bmatrix}$. 最後將 2-nd column 乘上 $-\sqrt{3}$ 加至 1-st column 得 $\begin{bmatrix}1&0\\0&1\end{bmatrix}$, 此時右邊乘上的 elementary matrix 為 $E_4=\begin{bmatrix}1&0\\-\sqrt{3}&1\end{bmatrix}$. 我們得到 $[R_{\pi/3}]E_1E_2E_3E_4=I_2$, 所以 $[R_{\pi/3}]=E_4^{-1}E_3^{-1}E_2^{-1}E_1^{-1}$, 亦即

$$\begin{bmatrix}1/2&-\sqrt{3}/2\\\sqrt{3}/2&1/2\end{bmatrix}=E_4^{-1}E_3^{-1}E_2^{-1}E_1^{-1}=\begin{bmatrix}1&0\\\sqrt{3}&1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&-\sqrt{3}/2\\0&1\end{bmatrix}\begin{bmatrix}1/2&0\\0&1\end{bmatrix}.$$

所以這次我們將 $R_{\pi/3}$ 拆解成先將 $\mathbf{e}_1$ 方向縮成 1/2 倍, 再做一個 horizontal shear 然後將 $\mathbf{e}_2$ 方向伸長 2 倍, 最後做一個 vertical shear. 和上一個所拆解的結果不同, 所以我們知這種拆解法是不唯一的.

4.4. 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 standard 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.

我們也探討了一些 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的 linear transformations. 利用一些基本的 linear transformations 的變化情形, 我們學習了如何將一個較複雜的 linear transformation 拆解成一些較容易掌握的 linear transformations 的合成.