6.3. Change of Basis

在這一節中, 我們介紹 change of basis 的概念, 了解到一個 linear operator 換了 ordered basis 後其表現矩陣的關係. 這個概念能幫助我們以後處理矩陣對角化的問題.

我們知道一個 linear transformation, 當我們用不同的 ordered bases 所得的 matrix representation 會不同. 假設 β,β' 為 V 的兩組 ordered bases, 而 γ,γ' 為 W 的兩組 ordered basis. 對於 linear transformation $T:V\to W$, 其對應於這兩對 ordered bases 的 matrix representations $[T]_\beta^\gamma$ 和 $[T]_{\beta'}^{\gamma'}$ 之間會有甚麼關係呢? 首先我們考慮 identity map $\mathrm{id}:V\to V$. 注意雖然是 identity map, 但其 matrix representation 未必會是 identity matrix. 事實上, 當我們定義域和對應域都選同一組 ordered basis $\beta=\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$, 則由於 $\mathrm{id}(\mathbf{v}_i)=\mathbf{v}_i$, 故其 matrix representation 是 identity matrix. 但若定義域是使用 β 這一組 ordered basis, 而對應域選的是 $\beta'=\{\mathbf{v}'_1,\dots,\mathbf{v}'_n\}$ 這一組 ordered basis, identity map 對應於 β,β' 的 matrix representation $[\mathrm{id}]_\beta^{\beta'}$ 其 i-th column 雖然仍和 $\mathrm{id}(\mathbf{v}_i)=\mathbf{v}_i$ 有關, 不過卻是要將 $\mathbf{v}_i$ 寫成以 $\{\mathbf{v}'_1,\dots,\mathbf{v}'_n\}$ 為 ordered basis 的坐標表示法 $[\mathbf{v}_i]_{\beta'}$. 所以當 β 和 β' 相異時, $[\mathrm{id}]_\beta^{\beta'}$ 不是 identity matrix. 現對任意 $\mathbf{v}\in V$, 因 $\mathbf{v}$ 對於 β 的坐標表示法為 $[\mathbf{v}]_\beta$, 依 matrix representation 的性質 (Proposition 4.3.14) 可得

$$[\mathrm{id}]_\beta^{\beta'}[\mathbf{v}]_\beta=[\mathrm{id}(\mathbf{v})]_{\beta'}=[\mathbf{v}]_{\beta'}.$$

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

要注意 $\mathrm{id}:V\to V$ 是 isomorphism, 所以由 Theorem 4.3.19 (3), 我們得 $[\mathrm{id}]_\beta^{\beta'}$ 為 invertible 且

$$([\mathrm{id}]_\beta^{\beta'})^{-1}=[\mathrm{id}^{-1}]_{\beta'}^{\beta}=[\mathrm{id}]_{\beta'}^{\beta} \tag{6.1}$$

也就是說將 β 的坐標表示轉換成對於 β' 的坐標表示的 change-of-basis matrix 的 inverse 就是 β' 的坐標表示轉換成對於 β 的坐標表示的 change-of-basis matrix.

我們回到原先的問題, 假設 $T:V\to W$ 為 linear transformation 且 β,β' 為 V 的兩組 ordered bases, 而 γ,γ' 為 W 的兩組 ordered basis. 我們要探討 $[T]_\beta^\gamma$ 和 $[T]_{\beta'}^{\gamma'}$ 之間的關係. 由於 $\mathrm{id}_V:V\to V$, $T:V\to W$ 和 $\mathrm{id}_W:W\to W$ 之合成 $\mathrm{id}_W\circ T\circ\mathrm{id}_V:V\to W$ 仍為 $T:V\to W$, 所以由 Theorem 4.3.19 (2) 得

$$[\mathrm{id}_W]_\beta^{\gamma'}[T]_\beta^\gamma[\mathrm{id}_V]_{\beta'}^\beta=[T]_{\beta'}^{\gamma'}.$$

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

Theorem 6.3.1 (Change-of-basis Formula)**.** 假設 $T:V\to W$ 為 *linear transformation* 且 β,β' 為 V 的兩組 *ordered bases,* 而 γ,γ' 為 W 的兩組 *ordered basis,* 則存在 *invertible matrix* P,Q 使得 $[T]_{\beta'}^{\gamma'}=Q([T]_\beta^\gamma)P$, 其中 P 為將 β' 的坐標表示轉換成 β 的坐標表示的 *change-of-basis matrix* $[\mathrm{id}_V]_{\beta'}^\beta$, 而 Q 為將 γ 的坐標表示轉換成 γ' 的坐標表示的 *change-of-basis matrix* $[\mathrm{id}_W]_\gamma^{\gamma'}$.

Example 6.3.2. 在 Example 4.3.13 中我們考慮 linear transformation $T: P_2(\mathbb{R}) \to P_3(\mathbb{R})$, 其中 $T(p(x)) = (x+1)p(x-1)$, $\forall p(x) \in P_2(\mathbb{R})$. 另外我們考慮 $P_2(\mathbb{R})$ 的兩組 ordered bases $\varepsilon = (x^2, x, 1)$, $\beta = (p_1(x), p_2(x), p_3(x))$ 其中

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

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

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

在 Example 4.3.13 中我們得到

$$[T]_{\varepsilon}^{\varepsilon'} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}, \quad [T]_{\beta}^{\beta'} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

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

$$[\mathrm{id}_{P_3(\mathbb{R})}]_{\varepsilon'}^{\beta'} [T]_{\varepsilon}^{\varepsilon'} [\mathrm{id}_{P_2(\mathbb{R})}]_{\beta}^{\varepsilon} = \begin{bmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{-1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} = [T]_{\beta}^{\beta'}.$$

回顧一下, 我們經常談論的一種 linear transformation 是其定義域及對應域為相同的 vector space. 這樣的 linear transformation 我們特別稱之為 *linear operator*. 關於 linear operator 我們通常對於定義域及對應域會選同樣的一組 ordered basis. 此時利用 Theorem 6.3.1, 我們得以下之結果.

Corollary 6.3.3. 假設 $T:V\to V$ 為 *linear transformation* 且 β,β' 為 V 的兩組 *ordered bases.* 則存在 *invertible matrix* P 使得 $[T]_{\beta'}^{\beta'}=P^{-1}([T]_{\beta}^{\beta})P$, 其中 P 為將 β' 的坐標表示轉換成 β 的坐標表示的 *change-of-basis matrix* $[\mathrm{id}_V]_{\beta'}^{\beta}$.

Proof. 考慮 Theorem 6.3.1 其中 $W=V$, $\gamma=\beta$ 以及 $\gamma'=\beta'$ 的情形. 此時 $Q=[\mathrm{id}_V]_{\beta}^{\beta'}$ 由式子 (6.1), 知 $Q=([\mathrm{id}_V]_{\beta'}^{\beta})^{-1}=P^{-1}$, 得證本定理. □

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

Example 6.3.4. 考慮 linear operator $T:\mathbb{R}^2\to\mathbb{R}^2$ 定義為 $T(\begin{bmatrix}x\\y\end{bmatrix})=\frac{1}{25}\begin{bmatrix}9x+12y\\12x+16y\end{bmatrix}$. 若利用標準基底 $\varepsilon=(\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix})$ 我們得 $[T]_{\varepsilon}^{\varepsilon}=\frac{1}{25}\begin{bmatrix}9&12\\12&16\end{bmatrix}$. 然而若用 $\beta=(\begin{bmatrix}3\\4\end{bmatrix},\begin{bmatrix}-4\\3\end{bmatrix})$ 這組 ordered basis 可由 $T(\begin{bmatrix}3\\4\end{bmatrix})=\begin{bmatrix}3\\4\end{bmatrix}$, $T(\begin{bmatrix}-4\\3\end{bmatrix})=\begin{bmatrix}0\\0\end{bmatrix}$, 得 $[T]_{\beta}^{\beta}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$. 我們很容易由

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

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

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

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