7.3. Characteristic Polynomial

我們先介紹如何找到一個方陣的 eigenvalues，再利用它來處理一個 linear operator 的 eigenvalues.

要如何找到一個 $n\times n$ matrix 的 eigenvector 及其對應的 eigenvalue 呢? 其實一般的找法是先找到 eigenvalue, 然後再找出與其對應的 eigenvector. 首先觀察若 $\lambda\in\mathbb{F}$ 是 A 的 eigenvalue, 表示存在一個非零向量 $\mathbf{v}\in\mathbb{F}^n$ 使得 $A\mathbf{v}=\lambda\mathbf{v}$. 由於 $I_n\mathbf{v}=\mathbf{v}$, 所以看成矩陣的運算 $\lambda\mathbf{v}=(\lambda I_n)\mathbf{v}$. 因此 $A\mathbf{v}=\lambda\mathbf{v}$ 就等同於 $(A-\lambda I_n)\mathbf{v}=\mathbf{0}$. 換言之, λ 是 A 的 eigenvalue 等同於由 $n\times n$ matrix $A-\lambda I_n$ 所對應的 linear system $(A-\lambda I_n)\mathbf{x}=\mathbf{0}$ 有 nontrivial solution $\mathbf{x}=\mathbf{v}$. 由 Theorem 2.6.9, 這又等同於 $A-\lambda I_n$ 不是 invertible, 再由 Theorem 5.2.6(1) 知這也等同於 $\det(A-\lambda I_n)=0$. 總言之, 要找到 A 的 eigenvalue λ 就是要找到 λ 滿足 $\det(A-\lambda I_n)=0$.

要怎樣找到 λ 滿足 $\det(A-\lambda I_n)=0$ 呢? 假設 $A=[a_{ij}]$, 若我們將 t 視為變數, 考慮 $\det(A-tI_n)$. 由於

$$A-tI_n=\begin{bmatrix} a_{11}-t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-t \end{bmatrix}$$

利用數學歸納法, 我們可以證明 $\det(A-tI_n)$ 會是一個以 t 為變數的 n 次實係數多項式. 而若 $t=\lambda$ 為此多項式的一實數根, 則 λ 就會滿足 $\det(A-\lambda I_n)=0$, 也就是說 λ 就會是 A 的一個 eigenvalue. 反之, 若 λ 就會是 A 的一個 eigenvalue, 就表示 $t=\lambda$ 會是多項式 $\det(A-tI_n)$ 的一個根. 由此可知多項式 $\det(A-tI_n)$ 可以讓我們完全掌握 A 的 eigenvalue, 我們因而給它一個特別的定義.

Definition 7.3.1. 假設 $A\in\mathsf{M}_{n\times n}(\mathbb{F})$, 考慮以 t 為變數的多項式 $p_A(t)=\det(A-tI_n)$. 我們稱 $p_A(t)$ 為 A 的 *characteristic polynomial* (特徵多項式)..

從上面的討論我們知道 $\lambda\in\mathbb{F}$ 為 characteristic polynomial $p_A(t)$ 的一個根若且唯若 λ 為 A 的 eigenvalue. 這裡要注意要談論 eigenvalue 是必須強調在哪一個 field 的 eigenvalue. 例如當 $A\in\mathsf{M}_{n\times n}(\mathbb{R})$, 其 characteristic polynomial $p_A(t)$ 是一個實係數多項式, 不過 $p_A(t)$ 有可能有非實數的虛根. 此時這個虛根不會是 A 在 $\mathbb{R}^n$ 中的 eigenvector 所對應的 eigenvalue. 事實上如果 $\lambda\in\mathbb{C}\setminus\mathbb{R}$ 是 $p_A(t)$ 的一個虛根, 此時假設存在 $\mathbf{v}\in\mathbb{R}^n$ 使得 $A\mathbf{v}=\lambda\mathbf{v}$. 由於 $A\mathbf{v}\in\mathbb{R}^n$, 但 $\lambda\mathbf{v}\notin\mathbb{R}^n$, 所以 $A\mathbf{v}=\lambda\mathbf{v}$ 不可能成立. 不過依前面的探討我們知道一定會有 $\mathbf{w}\in\mathbb{C}^n$ 滿足 $A\mathbf{w}=\lambda\mathbf{w}$. 在這個課程裡, 當我們探討矩陣 $A\in\mathsf{M}_{n\times n}(\mathbb{F})$ 的 eigenvalue 時, 若沒有特別說明, 都僅討論在 $\mathbb{F}$ 的 eigenvalue. 例如當我們討論實矩陣時, 我們考慮 eigenvalue 僅考慮 characteristic polynomial 的實根.

Example 7.3.2. 考慮 $B=\begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$, 此時 A 的 characteristic polynomial 為

$$p_B(t)=\det(B-tI_3)=\det\begin{bmatrix} -1-t & 4 & 2 \\ 1 & 3-t & 1 \\ -1 & 2 & 2-t \end{bmatrix}.$$

對第一個 row 降階求行列式得

$$p_A(t)=(-1-t)\det\begin{bmatrix}3-t & 1\\ 2 & 2-t\end{bmatrix}-4\det\begin{bmatrix}-1 & 1\\ -1 & 2-t\end{bmatrix}+2\det\begin{bmatrix}-1 & 3-t\\ -1 & 2\end{bmatrix}.$$

化簡可得 $p_A(t)=-(t-1)^2(t-2)$. 也因此 $t=1$ 和 $t=2$ 為 A 的 characteristic polynomial 的二實根, 也因此得 A 有兩個 eigenvalues $1,2$.

接下來我們說明當 $A\in\mathsf{M}_{n\times n}(\mathbb{F})$ 時, 其 characteristic polynomial $\det(A-tI_n)$ 確實是 t 的多項式. 首先觀察當我們在利用降階求 determinant 時, 其實是一些乘積之和. 利用數學歸納法可得這些乘積是由每一個 column 中的某個元素相乘而得而且它們都不會在同一個 row. 例如當我們計算 2×2 matrix $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ 的 characteristic polynomial $\det\begin{bmatrix}a-t & b\\ c & d-t\end{bmatrix}$ 時不難發現會貢獻 t 的最高次項乘積的是 $(a-t)(d-t)$ 而另一個乘積 bc 就僅影響到常數項, 因此其最高次項 t^2 與次高次項 t 的係數就完全由 $(a-t)(d-t)$ 的 t^2 與 t 的係數即 $at^2-(a+d)t$ 所決定. 現考慮 3×3 matrix $A=\begin{bmatrix}a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}$ 的 characteristic polynomial. 利用對第一個 row 降階的方式我們有

$$\det\begin{bmatrix}a-t & b & c\\ d & e-t & f\\ g & h & i-t\end{bmatrix}=(a-t)\det\begin{bmatrix}e-t & f\\ h & i-t\end{bmatrix}-b\det\begin{bmatrix}d & f\\ g & i-t\end{bmatrix}+c\det\begin{bmatrix}d & e-t\\ g & h\end{bmatrix}.$$

從前面 2×2 的情形我們看出 $\det\begin{bmatrix}e-t & f\\ h & i-t\end{bmatrix}$ 的 t^2 與次高次項 t 的係數就完全由 $(e-t)(i-t)$ 的 t^2 與 t 的係數所決定, 因此 $(a-t)(e-t)(i-t)$ 貢獻出 t^3 和 t^2 的係數. 而 $\det\begin{bmatrix}d & f\\ g & i-t\end{bmatrix}$ 和 $\det\begin{bmatrix}d & e-t\\ g & h\end{bmatrix}$ 最多僅有 t 的一次出現, 因此得 $\det(A-tI_3)$ 的 t^3 和 t^2 的係數完全由 $(a-t)(e-t)(i-t)$ 所決定. 也就是說 A 的 chacteristic polynomial $p_A(t)$ 為 3 次多項式且其最高次的兩項為 $(-1)^3t^3+(-1)^2(a+e+i)t^2$. 這裡 a,e,i 為 A 的 diagonal entries, 它們之和 $a+e+i$ 我們稱為 A 的 trace, 用 $\operatorname{tr}(A)$ 來表示. 利用數學歸納法, 我們可得當 $A=[a_{ij}]$ 為 $n\times n$ matrix 時, A 的 characteristic polynomial $p_A(t)=\det(A-tI_n)$ 為 t 的 n 次實係數多項式, 且其最高次的兩項是由 $(a_{11}-t)(a_{22}-t)\cdots(a_{nn}-t)$ 所貢獻因此為 $(-1)^nt^n+(-1)^{n-1}(a_{11}+\cdots+a_{nn})t^{n-1}$. 由於 A 的 diagonal entries 之和 $a_{11}+\cdots+a_{nn}$ 我們定為 $\operatorname{tr}(A)$, 因此有以下之結論.

Proposition 7.3.3. 假設 $A\in\mathsf{M}_{n\times n}(\mathbb{F})$. 則 A 的 *characteristic polynomial* 為 t 的 n 次實係數多項式. 其 t^n 項係數為 $(-1)^n$, t^{n-1} 項係數為 $(-1)^{n-1}\operatorname{tr}(A)$ 而常數項係數為 $\det(A)$.

Proof. 令 $p_A(t)=\det(A-tI_n)$, 由前面的討論我們僅剩討論 $p_A(t)$ 的常數項. 由於 $p_A(t)$ 是多項式所以它的常數項是 $p_A(0)=\det(A-0I_n)=\det(A)$. □

Question 7.4. 假設 $A\in\mathsf{M}_{n\times n}(\mathbb{F})$. 試問 A 最多會有幾個相異的 *eigenvalues?*

Example 7.3.4. 考慮 Example 7.2.5 中 $A=\begin{bmatrix}1 & 3\\ 4 & 2\end{bmatrix}$ 的 characteristic polynomial $p_A(t)$. 由於 $\operatorname{tr}(A)=1+2=3$ 以及 $\det(A)=2-12=-10$, 利用 Proposition 7.3.3 可得

$$p_A(t)=(-1)^2t^2+(-1)3t+(-10)=t^2-3t-10.$$

事實上利用 characteristic polynomial 的定義直接計算可得

$$p_A(t) = \det\begin{bmatrix} 1-t & 3 \\ 4 & 2-t \end{bmatrix} = (1-t)(2-t) - 12 = t^2 - 3t - 10.$$

分解後可得 $-2, 5$ 為 A 的 eigenvalues.

接下來我們介紹一個和 eigenvalue 有關的定義. 若 $\lambda \in \mathbb{F}$ 是 A 的 eigenvalue. 由於 $t = \lambda$ 會是 A 的 characteristic polynomial $p_A(t) = \det(A - tI_n)$ 的一個根. 由因式定理知 $t - \lambda$ 會整除 $p_A(t)$. 若 $(t-\lambda)^m$ 可整除 $p_A(t)$, 但 $(t-\lambda)^{m+1}$ 不能整除 $p_A(t)$, 則我們稱 eigenvalue λ 的 *algebraic multiplicity* (代數重根數) 為 m. 當然了當 $t = \lambda$ 是 $p_A(t)$ 的一個單根, 我們就說 λ 的 algebraic multiplicity 為 1. 例如 Example 7.3.2 中 B 有兩個 eigenvalue 1 和 2, 其中 eigenvalue 1 的 algebraic multiplicity 為 2, 而 eigenvalue 2 的 algebraic multiplicity 為 1. 而 Example 7.3.4 中 A 的兩個 eigenvalue $-2, 5$ 其 algebraic multiplicity 皆為 1. 有關 algebraic multiplicity 的性質, 以後我們還會進一步討論.

Question 7.5. *Identity matrix I_n 的 eigenvalue 有哪些? 其 algebraic multiplicity 為何?*

最後我們介紹一些和 characteristic polynomial 有關的性質. 一般來說兩個 $n \times n$ matrices 的 characteristic polynomial 可能不相同. 不過在一種特殊情況之下, 它們的 characteristic polynomial 會一樣. 前面提過當 A, B 為 $n \times n$ matrices, 若存在 $n \times n$ 的 invertible matrix U, 使得 $B = U^{-1}AU$, 則我們稱 A, B 為 *similar* (關於這個定義的原因我們以後會再詳述). 此時我們可得 A 和 B 的 characteristic polynomial 是相同的.

Proposition 7.3.5. *假設 A, B 為 $n \times n$ matrices 且存在 $n \times n$ 的 invertible matrix U 滿足 $B = U^{-1}AU$. 則 A 和 B 有相同的 characteristic polynomial.*

Proof. 依假設 B 的 characteristic polynomial 為 $\det(B - tI_n) = \det(U^{-1}AU - tI_n)$. 然而依矩陣乘法性質

$$U^{-1}(A - tI_n)U = U^{-1}AU - U^{-1}(tI_n)U = U^{-1}AU - tU^{-1}I_nU = U^{-1}AU - tI_n.$$

因此再由 determinant 的性質 (Theorem 5.2.6) 得

$$\det(B - tI_n) = \det(U^{-1}(A - tI_n)U) = \det(U^{-1})\det(A - tI_n)\det(U) = \det(A - tI_n).$$

得證 A 和 B 有相同的 characteristic polynomial. □

另一個會有相同的 characteristic polynomial 的情況就是 A 和 A^t 有相同的 characteristic polynomial.

Proposition 7.3.6. *假設 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$, 則 A 和 A^t 有相同的 characteristic polynomial*

Proof. 利用 trnaspose 的性質 $(A - tI_n)^t = A^t - tI_n^t = A^t - tI_n$ (Proposition 2.2.4), 故利用 Theorem 5.2.6 (3), 我們有

$$P_{A^t}(t) = \det(A^t - tI_n) = \det((A - tI_n)^t) = \det(A - tI_n) = P_A(t).$$

□

Question 7.6. 試說明 A 和 A^{t} 有相同的 *eigenvalues* 且對每個 *eigenvalue* 其在 A 和 A^{t} 的 *algebraic multiplicity* 也相同.

讓我們回到 linear operator. 要怎樣找一個 linear operator $T: V \to V$ 的 eigenvalue 呢？從上一節的說明可以知道, 任取 V 的一組 ordered basis β, 只要考慮其 matrix representation $[T]_\beta$ 的 eigenvalue 就是 T 的 eigenvalue. 所以我們可以定義一個 linear operator $T: V \to V$ 的 characteristic polynomial $P_T(t)$. 其定義的方法就是任取一個 V 的 ordered basis β, 若 $A = [T]_\beta$, 則定義 $P_T(t) = P_A(t)$. 注意這樣定義出來的 characteristic polynomial 和 β 的選取無關. 主要的原因是若取 V 的另一組 ordered basis, 其表現矩陣會和 A 是 similar. 故由 Proposition 7.3.5 即 similar matrix 的 characteristic polynomial 是一樣的, 得知 $P_T(t)$ 不會因選取的 ordered basis 不同而有所不同. 注意：因為 T 的 eigenvalue 就是其任一有序基底表現矩陣 A 的 characteristic polynomial 的根, 所以依此定義我們也可以說 T 的 eigenvalue 就是 T 的 characteristic polynomial 的根.

Example 7.3.7. 考慮 Example 7.2.2 中的 linear operator $T: P_2(\mathbb{R}) \to P_2(\mathbb{R})$, 以及 $P_2(\mathbb{R})$ 的 standard basis $\varepsilon = (1, x, x^2)$. 由於依定義 $T(1) = 1$, $T(x) = 2x+1$, $T(x^2) = 3x^2 + 2x$, 我們得 $[T]_\varepsilon = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$. 因為 $[T]_\varepsilon$ 是上三角矩陣, 很容易求得其 characteristic polynomial 為 $(1-t)(2-t)(3-t)$. 得知 $[T]_\varepsilon$ 的 eigenvalue 為 $1,2,3$。事實上我們在 Example 7.2.2 已知 $1,3$ 是 T 的 eigenvalue. 而若考慮 $x+1 \in P_2(\mathbb{R})$ 我們有 $T(x+1) = 2(x+1)$，所以 2 確實也是 T 的 eigenvalue.

因為 $\{1, x+1, x^2+2x+1\}$ 是 T 的 eigenvectors（所對應的 eigenvalue 分別為 $1,2,3$）且是 $P_2(\mathbb{R})$ 的一組 basis. 若考慮 ordered basis $\beta = (1, x+1, x^2+2x+1)$, 則 $[T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. 也就是說 $[T]_\beta$ 的 characteristic polynomial 也是 $(1-t)(2-t)(3-t)$.

7.4. Eigenspace 和 Eigenvector

我們了解了如何找到一個 $n \times n$ matrix 的 eigenvalue 之後, 接下來便是要找出這些 eigenvalue 所對應的 eigenvectors.

假設 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ 且 $\lambda \in \mathbb{F}$ 為 A 的一個 eigenvalue. 由於 $\det(A - \lambda I_n) = 0$, 我們知聯立方程組 $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ 存在非零的 nontrivial solution. 現假設 $\mathbf{v} \in \mathbb{F}^n$ 為非零向量且 $\mathbf{x} = \mathbf{v}$ 為 $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ 的一組解. 此即表示 $\mathbf{v}$ 滿足 $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$, 亦即 $A\mathbf{v} = \lambda\mathbf{v}$. 故此時 $\mathbf{v}$ 為 A 的一個以 λ 為 eigenvalue 的 eigenvector. 反之, 若 $\mathbf{v}$ 為 A 的一個以 λ 為 eigenvalue 的 eigenvector, 則 $\mathbf{x} = \mathbf{v}$ 必為 $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ 的一組 nontrivial solution. 因此我們只要掌握 $n \times n$ matrix $A - \lambda I_n$ 的 nullspace (即 $\{\mathbf{v} \in \mathbb{F}^n \mid (A - \lambda I_n)\mathbf{v} = \mathbf{0}\}$) 中的非零向量就會是 A 相對於 λ 的 eigenvector. 由於 nullspace 是 vector space, 因此我們有以下的定義.

Definition 7.4.1. 假設 $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ 且 $\lambda \in \mathbb{F}$ 為 A 的一個 eigenvalue. 則 $A - \lambda I_n$ 的 nullspace 稱為 A 對於 eigenvalue λ 的 *eigenspace*. 我們用 $E_A(\lambda)$ 來表示.

要注意對於 λ 的 eigenspace 並不是由以 λ 為 eigenvalue 的 eigenvectors 所組成. 這是因為零向量 $\mathbf{0}$ 不是 eigenvector, 但 vector space 必須包含 $\mathbf{0}$. 所以對於 λ 的 eigenspace 應該是由所有以 λ 為 eigenvalue 的 eigenvectors 和 $\mathbf{0}$ 所組成. 那為什麼要讓它形成 vector space 呢? 因為 vector space 有其方便性, 例如有了 vector space 我們就可以利用 dimension 來知道它的大小. 因此我們定義 $E_A(\lambda)$ 的 dimension 為 eigenvalue λ 的 *geometric multiplicity* (幾何重根數). 要注意 eigenvalue λ 的 algebraic multiplicity 無法讓我們知道 λ 所對應的 eigenvectors 的多寡, 而是 λ 的 geometric multiplicity 可以提供這一個訊息.

Example 7.4.2. 考慮 $A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$. 由前面 Example 7.3.2, Example 7.3.4 我們已計算出 A 和 B 的 characteristic polynomial 分別為 $p_A(t) = (x+2)(x-5)$, $p_B(t) = -(t-1)^2(t-2)$. 接下來我們分別計算 A 和 B 的 eigenspace.

首先考慮 A 對於 eigenvalue -2 的 eigenspace, 亦即找出 $A-(-2I_2) = \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. 可得 $E_A(1) = \mathrm{Span}(\begin{bmatrix} 1 \\ -1 \end{bmatrix})$. 也就是說 A 對於 eigenvalue 為 -2 的 eigenvector 就是那些和 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 平行的 nonzero vector. 由於 $\dim(E_A(-2)) = 1$, 我們也得到 A 對於 eigenvalue -2 的 geometric multiplicity 為 1. 至於 A 對於 eigenvalue 5 的 eigenspace, 亦即找出 $A - 5I_2 = \begin{bmatrix} -4 & 3 \\ 4 & -3 \end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix} -4 & 3 \\ 0 & 0 \end{bmatrix}$. 因此得 $E_A(5) = \mathrm{Span}(\begin{bmatrix} 3 \\ 4 \end{bmatrix})$. 也就是說 A 對於 eigenvalue 為 5 的 eigenvector 就是那些和 $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ 平行的 nonzero vector, 我們也得到 A 對於 eigenvalue 5 的 geometric multiplicity 為 1. 在 Example 7.2.5 中我們舉出 A 的 eigenvector 的例子其實是這樣得到的.

接著考慮 B 對於 eigenvalue 1 的 eigenspace, 亦即找出 $B - I_3 = \begin{bmatrix} -2 & 4 & 2 \\ -1 & 2 & 1 \\ -1 & 2 & 1 \end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix} 1 & -2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. 可得 $E_B(1) = \mathrm{Span}(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix})$. 也就是說 B 對於 eigenvalue 為 1 的 eigenvector 就是那些由 $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ 的 linear combination 所得的 nonzero vector. 例如 $\mathbf{v} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ 就滿足

$$B\mathbf{v} = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \mathbf{v}.$$

由於 $\dim(E_B(1)) = 2$, 我們也得到 B 對於 eigenvalue 1 的 geometric multiplicity 為 2. 至於 B 對於 eigenvalue 2 的 eigenspace, 亦即找出 $B - 2I_3 = \begin{bmatrix} -3 & 4 & 2 \\ -1 & 1 & 1 \\ -1 & 2 & 0 \end{bmatrix}$ 的 null space. 經由 elementary row operations, 可化為 echelon form $\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. 因此得 $E_B(2) = \mathrm{Span}(\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix})$. 也就是說 B 對於 eigenvalue 為 2 的 eigenvector 就是那些和 $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ 平行的 nonzero vector, 我們也得到 B 對於 eigenvalue 2 的 geometric multiplicity 為 1.

在 Proposition 7.2.3 中我們知道兩個 eigenvectors 若其 eigenvalue 相同, 則其線性組合只要不是 $\mathbf{0}$, 就會是有同樣 eigenvalue 的 eigenvector. 所以一般在探討一個 $n \times n$ 矩陣的 eigenvector 時, 我們只要寫下其 eigenspace 的一組基底即可.

知道如何得到矩陣的 eigenvectors 後，我們就可以探討如何找到 linear operator $T : V \to V$ 的 eigenvectors 了！從前面一開始的說明可以知道, 任取 V 的一組 ordered basis β, 只要考慮其 matrix representation $[T]_\beta$ 的 eigenvalue 和 eigenvector，接著將 eigenvectors 利用 β 還原成 V 中的向量，就會是 T 的 eigenvector 了, 我們看以下的例子.

Example 7.4.3. 考慮 Example 7.2.2 中的 linear operator $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$, 以及 $P_2(\mathbb{R})$ 的 standard basis $\varepsilon = (1, x, x^2)$. 在 Example 7.3.7 我們有 $[T]_\varepsilon = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ 且得 T 的 characteristic polynomial $p_T(t) = (1-t)(2-t)(3-t)$ 以及 T 的 eigenvalue 為 $1, 2, 3$. 接下來我們利用解 $[T]_\varepsilon$ 的 eigenspace 得 $[T]_\varepsilon$ 的 eigenvectors. 對於 eigenvalue 1 所得的 eigenspace 就是 $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$ 的 null space, 即 $\mathrm{Span}(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix})$. 然而 $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ 是 1 在 $P_2(\mathbb{R})$ 利用 ε 所得的坐標表示. 故知 $\mathrm{Span}(1)$ 中的非 0 元素是 T 的 eigenvector 且其 eigenvalue 為 1. 對於 $[T]_\varepsilon$ 的 eigenvalue 2 所得的 eigenspace 就是 $\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ 的 null space, 即 $\mathrm{Span}(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix})$. 然而 $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ 是 $x+1$ 在 $P_2(\mathbb{R})$ 利用 ε 所得的坐標表示. 故知 $\mathrm{Span}(x+1)$ 中的非 0 元素是 T 的 eigenvector 且其 eigenvalue 為 2. 對於 $[T]_\varepsilon$ 的 eigenvalue 3 所得的 eigenspace 就是 $\begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ 的 null space, 即 $\mathrm{Span}(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix})$. 然而 $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ 是 $x^2 + 2x + 1$ 在 $P_2(\mathbb{R})$ 利用 ε 所得的坐標表示. 故知 $\mathrm{Span}(x^2 + 2x + 1)$ 中的非 0 元素是 T 的 eigenvector 且其 eigenvalue 為 3. 事實上在 Example 7.3.7 中我們驗證過這些確為 T 的 eigenvectors。 $\sharp$

接著我們探討一個方陣或 linear operator 的 eigenvalue 其 algebraic multiplicity 和 geometric multiplicity 之間的關係。雖然前面關於 eigenvalue、eigenvector 都是現利用矩陣處理，有趣的是這裡我們反而利用 linear operator 來處理較容易。

在探討函數的理論時, 通常當定義域很大時, 我們可以透過所謂的 restriction 將函數限制在較小的範圍來了解該函數. 給定一個函數 $f: X \to Y$, 以及 X 中的子集合 S, 所謂 f 的 restriction on S, 用 $f|_S$ 表示, 就是將 f 的定義域縮小到 S, 其他對於 f 的映射方式都沒有改變. 也就是說 $f|_S$ 是一個定義域為 S 的函數 $f|_S: S \to Y$, 且對於任意 $s \in S$, $f|_S(s) = f(s)$, 不過若 $x \in X$ 但 $x \notin S$, 則 $f|_S(x)$ 是無定義的. 現若 $T: V \to V$ 是 linear operator, W 為 V 的 subspace, 則 $T|_W$ 依然會是一個 linear transformation (只是定義域在 W 上). 不過 $T|_W$ 未必會是一個 linear operator, 因為 T 未必會將 W 中的元素映射到 W. 如此一來, 我們就不能將過去探討 linear operator 的理論運用在 $T|_W$ 上了. 為了達到 $T|_W$ 仍為 linear operator 的目的, 我們必須選有特殊性質的 W（即 T 會將 W 的元素映射到 W）, 這樣就能套用 linear operator 的理論了. 因此我們有以下的定義.

Definition 7.4.4. 假設 V 是一個 vector space over $\mathbb{F}$ 且 $T: V \to V$ 是 linear operator. 若 W 是 V 的 subspace 且滿足 $T(W) \subseteq W$（即 $T(\mathbf{w}) \in W, \forall \mathbf{w} \in W$）, 則稱 W 為一個 T-invariant subspace.

要注意當 $T: V \to V$ 是 linear operator, Definition 7.4.4 告訴我們 W 是 T-invariant, 表示 $T(W) \subseteq W$, 並不是說 $T(W) = W$, 也不是說 $T(\mathbf{w}) = \mathbf{w}$, $\forall \mathbf{w} \in W$. 請大家不要誤解. 也就是說要檢查 V 的 subspace W 是否為 T-invariant subspace, 我們僅要檢查是否所有 W 的元素 $\mathbf{w}$ 經由 T 的映射（即 $T(\mathbf{w})$）依然在 W 中. 當然了, 因 T 為 linear operator, 對任意 $\mathbf{v} \in V$, 皆有 $T(\mathbf{v}) \in V$, 故 V 本身是 T-invariant. 還有因為 T 是 linear transformation, 我們知道 $T(\mathbf{0}) = \mathbf{0}$, 所以 zero space $\{\mathbf{0}\}$ 也是 T-invariant. 另外若 $\lambda \in \mathbb{F}$ 是 T 的 eigenvalue, 則 λ 所對應的 eigenspace $E_T(\lambda) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda \mathbf{v}\}$ 也會是 T-invariant subspace. 這是因為若 $\mathbf{v} \in E_T(\lambda)$, 則 $T(\mathbf{v}) = \lambda \mathbf{v}$. 由於 $E_T(\lambda)$ 是 V 的 subspace 且 $\mathbf{v} \in E_T(\lambda)$, 自然有 $\lambda \mathbf{v} \in E_T(\lambda)$, 亦即 $T(\mathbf{v}) \in E_T(\lambda)$. 故 $E_T(\lambda)$ 亦為 T-invariant. 還有許多有趣的 T-invariant subspace 可以幫我們探討有關於 linear operator T 的性質。目前我們僅需考慮 eigenspace 這類的 T-invariant subspace.

Linear operator 限制在較小的 T-invariant subspace 基本上和原來的 operator 是相同的, 因此它們的 characteristic polynomial 之間應該有關係. 接下來, 我們便是要探討它們之間的關係. 假設 $T: V \to V$ 是 linear operator, 且 W 為 T-invariant subspace. 要討論 T 和 $T|_W$ 的 characteristic polynomial, 我們需找 W 和 V 的 ordered basis, 然後得到相對應的表現矩陣, 再得到它們的 characteristic polynomial. 因為 W 是 V 的 subspace, 我們又期待它們之間的 characteristic polynomial 相關. 自然的, 我們可以先找 W 的一組 ordered basis $\beta = (\mathbf{w}_1, \ldots, \mathbf{w}_k)$ 再將 β 擴大成 V 的一組 ordered basis $\gamma = (\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n)$. 現假設 $A = [T|_W]_\beta$, 注意 A 的 i-th column 就是 $T(\mathbf{w}_i)$ 用 β 所得的坐標表示 (即 $\mathbb{F}^k$ 中的向量). 現考慮 T 用 ordered basis γ 所得的矩陣表示 $[T]_\gamma$. 要注意當 $1 \leq i \leq k$ 時, $[T]_\gamma$ 的 i-th column 和 A 的 i-th column 一樣是 $T(\mathbf{w}_i)$, 不同的是它應該是 $T(\mathbf{v}_i)$ 用 γ 的坐標表示 (即 $\mathbb{F}^n$ 中的向量). 由於 $T(\mathbf{w}_i) \in W$, 其用 γ 寫下的線性組合, 僅需要用到前面 k 個向量, 即 β 的線性組合, 所以 $T(\mathbf{w}_i)$ 用 γ 的坐標表示基本上和用 β 坐標表示相同, 只是後面 $k+1, \ldots, m$ 這 $m-k$ 個 entry 須補上 0. 至於 $[T]_\gamma$ 在 k_1-th column 之後的 columns 我們就不知道會是怎

樣, 不過這並不會影響我們要探討的問題. 總而言之, 若 $[T|_W]_\beta = A$, 我們可以將 $[T]_\gamma$ 寫成 $\left[\begin{array}{c|c} A & M_1 \\ \hline \mathbf{0} & M_2 \end{array}\right]$ 這樣的形式. 注意這裡 A 是 $k\times k$ matrix, $\mathbf{0}$ 是 $(n-k)\times k$ matrix, 而 M_1, M_2 分別為 $k\times(n-k)$ 和 $(n-k)\times(n-k)$ matrix. 因此 $[T]_\gamma$ 的 characteristic polynomial 應為

$$\det\left[\begin{array}{c|c} A-tI_k & M_1 \\ \hline \mathbf{0} & M_2-tI_{n-k} \end{array}\right].$$

回顧前面提過在計算形如 $\left[\begin{array}{c|c} A & B \\ \hline \mathbf{0} & C \end{array}\right]$ 這樣的矩陣的 determinant 時, 當 A, C 分別為 $k\times k$ 和 $(n-k)\times(n-k)$ matrix 時我們可以先將前 k 個 row 用 elementary row operations 將 A 化成 echelon form, 再將後面 $n-k$ 個 row 用 elementary row operations 將 C 化為 echelon form. 因為這樣最後是一個 upper triangular matrix, 其 determinant 就是對角線相乘, 因此知 $\det\left[\begin{array}{c|c} A & B \\ \hline \mathbf{0} & C \end{array}\right] = (\det A)(\det C)$. 因此前面算 T 的 characteristic polynomial 便會是 $\det(A-tI_k)\det(M_2-tI_{n-k})$. 因為 $\det(A-tI_k)$ 就是 $T|_W$ 的 characteristic polynomial, 因此我們有以下的結果.

Proposition 7.4.5. 假設 V 為 *vector space over* $\mathbb{F}$, $T: V\to V$ 為 *linear operator* 且 W 為 T-*invariant subspace,* 則 $T|_W$ 的 *characteristic polynomial* 會是 T 的 *characteristic polynomial* 的因式. 也就是說, 若 $f(t)$ 是 T 的 *characteristic polynomial* 且 $g(t)$ 是 $T|_W$ 的 *characteristic polynomial,* 則存在係數在 $\mathbb{F}$ 的多項式 $h(t)$ 滿足 $f(t) = h(t)g(t)$.

Exercise 7.7. 考慮 $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -6 & 4 \\ -2 & -4 & 5 \\ -2 & -6 & 7 \end{bmatrix}$.

(1) 試分別求 A, B 的 characteristic polynomial.

(2) 試分別說明 A, B 的 eigenvalues 有哪些, 並計算每個 eigenvalue 的 algebraic multiplicity 和 geometric multiplicity.

Exercise 7.8. 考慮 $T: \mathbb{R}^4 \to \mathbb{R}^4$, 定義為 $T(x,y,z,w) = (x+y+2z-w, y+w, 2z-w, z+w)$. 令 $W = \{(x,y,0,0) : x,y\in\mathbb{R}\}$, $U = \mathrm{Span}((1,1,0,1),(2,1,1,2),(1,-1,2,1))$.

(1) 證明 W 和 U 皆為 $\mathbb{R}^4$ 中的 T-invariant subspace.

(2) 分別找出 W, U 中的 ordered basis $\beta = (\mathbf{w}_1, \mathbf{w}_2)$, $\gamma = (\mathbf{u}_1, \mathbf{u}_2)$。分別寫下 $T|_W$ 以 β 所得的表現矩陣，以及 $T|_U$ 以 γ 所得的表現矩陣.

(3) 說明 $\alpha = (\mathbf{w}_1, \mathbf{w}_2, \mathbf{u}_1, \mathbf{u}_2)$ 可形成 $\mathbb{R}^4$ 的一組 ordered basis, 並寫下 T 以 α 所得的表現矩陣. 計算 $T|_W$, $T|_U$ 以及 T 的 characteristic polynomials. 並說明它們之間關係。

(4) 分別找出 $T|_W$, $T|_U$ 以及 T 的 eigenvalues，並說明它們的代數重根數與幾何重根數。