

8.5. Application: Markov Processes

當 $A \in M_{n \times n}(\mathbb{R})$ 為 diagonalizable 時, 對於 $k \in \mathbb{N}$ 我們可以利用對角矩陣很容易求出 A^k . 進而對於任意 $\mathbf{v} \in \mathbb{R}^n$, 推算出 $A^k \mathbf{v}$. 其實還有一種情況 (即使不是 diagonalizable), 當 k 很大時我們也能“估計” $A^k \mathbf{v}$ 大約為何. 這就是本節要探討的課題.

首先我們看 $A \in M_{n \times n}(\mathbb{R})$ 為 diagonalizable 的情形. 此時由於存在 diagonal matrix $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ 以及 invertible matrix $P \in M_{n \times n}(\mathbb{R})$ 使得 $D = P^{-1}AP$, 因此 $A = PDP^{-1}$. 依此我們可以推得

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1} = P \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix} P^{-1},$$

然後用數學歸納法推得

$$A^k = P \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} P^{-1}.$$

Example 8.5.1. 我們利用 Fibonacci sequence $0, 1, 1, 2, 3, 5, 8, 13, \dots$, 來說明如何利用對角化. Fibonacci sequence 是一組滿足 $F_{k+1} = F_k + F_{k-1}$ 的遞迴數列, 其中 $F_0 = 0, F_1 = 1$. 我們令 $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 且對任意 $k \geq 1$ 令 $\mathbf{v}_k = \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$, 則

$$A\mathbf{v}_k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix} = \begin{bmatrix} F_k + F_{k-1} \\ F_k \end{bmatrix} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \mathbf{v}_{k+1}.$$

因此我們有 $\mathbf{v}_{k+1} = A^k \mathbf{v}_1$. 也就是說對於任意 $k \geq 1$, 我們只要能算出 $A^k \mathbf{v}_1$, 就能求出 F_{k+1} 為何. 然而 A 的 characteristic polynomial 為 $P_A(t) = t^2 - t - 1$, 故得 A 的 eigenvalues 為 $\lambda_1 = (1 - \sqrt{5})/2$, $\lambda_2 = (1 + \sqrt{5})/2$. 因 A 是 2×2 matrix, 所以由 A 兩個相異的 eigenvalues 得 A 為 diagonalizable. 事實上 A 對於 λ_1, λ_2 的 eigenvector 分別為 $\mathbf{v}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$.

因此若令 $P = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}$, 我們有

$$P = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}, \quad P^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & \frac{1+\sqrt{5}}{2} \\ 1 & \frac{-1+\sqrt{5}}{2} \end{bmatrix}.$$

因此將 A 對角化得 $A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$, 也因此求出對任意 $k \in \mathbb{N}$,

$$A^k = P \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} P^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \begin{bmatrix} -1 & \lambda_2 \\ 1 & -\lambda_1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_2^{k+1} - \lambda_1^{k+1} & \lambda_2^k - \lambda_1^k \\ \lambda_2^k - \lambda_1^k & \lambda_2^{k-1} - \lambda_1^{k-1} \end{bmatrix}.$$

所以由 $\mathbf{v}_1 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, 我們得

$$\mathbf{v}_{k+1} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = A^k \mathbf{v}_1 = A^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_2^{k+1} - \lambda_1^{k+1} \\ \lambda_2^k - \lambda_1^k \end{bmatrix},$$

故得 $F_{k+1} = \frac{1}{5}(\lambda_2^{k+1} - \lambda_1^{k+1}) = \frac{1}{5} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{k+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{k+1} \right)$. #

在 Example 8.5.1 中矩陣 A 的兩個 eigenvalue λ_1, λ_2 所對應 eigenvector $\mathbf{v}_1, \mathbf{v}_2$ 其實有特殊的意義。事實上 \mathbf{v}_1 代表的便是 $a_0 = 1, a_1 = \lambda_1$ 符合 $a_{n+1} = a_n + a_{n-1}$ 的遞迴數列 $\langle a_n \rangle$ ，此時 $a_2 = 1 + \lambda_1 = \lambda_1^2, a_3 = a_2 + a_1 = \lambda_1^2 + \lambda_1 = \lambda_1^3, \dots, a_n = \lambda_1^n, \dots$ 也就是說 $\langle a_n \rangle$ 恰為公比為 λ_1 的等比數列。同理若數列 $\langle b_n \rangle$ 符合同樣的遞迴關係式 $b_{n+1} = b_n + b_{n-1}$ 且 $b_0 = 1, b_1 = \lambda_2$ ，則 $\langle b_n \rangle$ 便會是公比為 λ_2 的等比數列。這對於更一般的遞迴數列，也成立。

給定實數 c, d ，考慮符合 $a_{n+1} = ca_n + da_{n-1}$ 的遞迴數列 $\langle a_n \rangle$ 。令 $A = \begin{bmatrix} c & d \\ 1 & 0 \end{bmatrix}$ 且對任意 $k \geq 1$ 令 $\mathbf{v}_k = \begin{bmatrix} a_k \\ a_{k-1} \end{bmatrix}$ ，則

$$A\mathbf{v}_k = \begin{bmatrix} c & d \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_k \\ a_{k-1} \end{bmatrix} = \begin{bmatrix} ca_k + da_{k-1} \\ a_k \end{bmatrix} = \begin{bmatrix} a_{k+1} \\ a_k \end{bmatrix} = \mathbf{v}_{k+1}.$$

此時 A 的 characteristic polynomial $p_A(t) = t^2 - ct - d$ 。因此若 r 為 A 的 eigenvalue，則因 $p_A(r) = r^2 - cr - d = 0$ ，可得 $\begin{bmatrix} c & d \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r \\ 1 \end{bmatrix} = \begin{bmatrix} cr + d \\ r \end{bmatrix} = r \begin{bmatrix} r \\ 1 \end{bmatrix}$ ，亦即 $\begin{bmatrix} r \\ 1 \end{bmatrix}$ 是 A 的 eigenvalue r 的 eigenvector。這表示，前兩項為 $a_0 = 1, a_1 = r$ 符合遞迴關係 $a_{n+1} = ca_n + da_{n-1}$ 的數列，會是等比數列。由於等比數列的一般項很好掌握，因此若 A 有兩個相異的 eigenvalue r_1, r_2 （亦即 A 可以對角化），此時對任意符合 $a_{n+1} = ca_n + da_{n-1}$ 的遞迴數列 $\langle a_n \rangle$ ，我們只要解 $x, y \in \mathbb{R}$ 滿足 $x + y = a_0$ 以及 $r_1x + r_2y = a_1$ ，就可寫下 a_n 的一般項為 $a_n = xr_1^n + yr_2^n$ 。

接下來我們要探討的是，有時即使 A 不是 diagonalizable，但我們仍能估計 $A^k \mathbf{v}$ 。這裡要探討的情況是所謂 *Markov Processes*，是機率統計上的課題。由於我們僅專注於線性代數的部分，在這裡就不多談它的由來和例子，直接切入主題。

Definition 8.5.2. 對於一 \mathbb{R}^n 上的 vector $\mathbf{v} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ ，若 $c_1 + \dots + c_n = 1$ 且對於所有 $i =$

$1, \dots, n$ ，皆有 $c_i \geq 0$ ，則稱 \mathbf{v} 為 *probability vector*。若 $A \in M_{n \times n}(\mathbb{R})$ 且其每一個 column vector 皆為 probability vector，則稱 A 為 *stochastic matrix*。另外，一個 stochastic matrix A 若存在 $r \in \mathbb{N}$ 使得 A^r 的每個 entry 皆為正實數，則稱 A 為 *regular*。

Example 8.5.3. $A = \begin{bmatrix} 1/2 & 1 \\ 1/2 & 0 \end{bmatrix}$ 和 $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 皆為 stochastic matrix。而且 A 為 regular，因為 $A^2 = \begin{bmatrix} 3/4 & 1/2 \\ 1/4 & 1/2 \end{bmatrix}$ ，每個 entry 皆為正。然而 I_2 不是 regular，因為對於任意 $r \in \mathbb{N}$ 皆有 $I_2^r = I_2$ （除了對角線，其他位置的 entry 皆為 0）。

由於 probability vector 和 stochastic matrix 的定義的牽涉到 column vectors 各 entry 加總，我們可以利用向量內積來幫忙處理加總運算。也就是利用每個 entry 皆為 1 的 \mathbb{R}^n 向量 $\mathbf{u} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ 與 $\mathbf{v} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 的內積 $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^t \mathbf{v} = [1 \ \dots \ 1] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1 + \dots + c_n$ ，就可求得 \mathbf{v}

每個 entry 之和。因此，若 $\mathbf{v} \in \mathbb{R}^n$ 以及 $A \in M_{n \times n}(\mathbb{R})$ 的每個 entry 皆為非負實數，則分別由 $\langle \mathbf{u}, \mathbf{v} \rangle = 1$ 以及 $\mathbf{u}^t A = \mathbf{u}$ 就可以確認 \mathbf{v} 為 probability vector 以及 A 為 stochastic matrix。至於向量 \mathbf{v} 的每個 entry 是否為非負實數，我們也可利用標準基底的向量 \mathbf{e}_i 與 \mathbf{v} 的內積 $\langle \mathbf{e}_i, \mathbf{v} \rangle$ 來檢驗 \mathbf{v} 的 i -th entry 是否為非負實數。接下來 \mathbf{u} 表示這個 \mathbb{R}^n 中每個 entry 為 1 的向量，就不再贅述。我們利用 \mathbf{u} 看幾個有關 stochastic matrix 的性質。

Lemma 8.5.4. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 stochastic matrix 且 $\mathbf{v} \in \mathbb{R}^n$ 為 probability vector。則 $A\mathbf{v}$ 亦為 probability vector。另外若 A 的每一個 entry 皆為正實數，則 $A\mathbf{v}$ 的每個 entry 亦皆為正實數。

Proof. 令 $A = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, 則 $A\mathbf{v} = c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n$ 。因此 $A\mathbf{v}$ 所有 entries 之和就是 $\langle \mathbf{u}, c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n \rangle = c_1\langle \mathbf{u}, \mathbf{a}_1 \rangle + \cdots + c_n\langle \mathbf{u}, \mathbf{a}_n \rangle$ 。然而因每個 \mathbf{a}_j 為 probability vector, 因此 $\langle \mathbf{u}, \mathbf{a}_j \rangle = 1$, 所以再利用 \mathbf{v} 為 probability vector 得 $\langle \mathbf{u}, A\mathbf{v} \rangle = c_1 + \cdots + c_n = 1$ 。因此 $A\mathbf{v}$ 的所有 entry 之和為 1。

接著對每個 i 利用向量 \mathbf{e}_i 檢查 $A\mathbf{v}$ 的 i -th entry 是否為非負實數。因為

$$\langle \mathbf{e}_i, A\mathbf{v} \rangle = \langle \mathbf{e}_i, c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n \rangle = c_1\langle \mathbf{e}_i, \mathbf{a}_1 \rangle + \cdots + c_n\langle \mathbf{e}_i, \mathbf{a}_n \rangle \quad (8.15)$$

因為 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 中的每個 entry 皆為非負實數，故每個 $\langle \mathbf{e}_i, \mathbf{a}_j \rangle \geq 0$ 。又因為 c_1, \dots, c_n 皆為非負實數，所以 $\langle \mathbf{e}_i, A\mathbf{v} \rangle \geq 0$ 。亦即 $A\mathbf{v}$ 每個 entry 皆為非負實數，得證 $A\mathbf{v}$ 為 probability vector。

另外若 A 的每一個 entry 皆為正實數，即式子 (8.15) 中每個 $\langle \mathbf{e}_i, \mathbf{a}_j \rangle > 0$ ，此時由於 c_1, \dots, c_n 為非負實數，故對任意 $j = 1, \dots, n$ 皆有 $\langle \mathbf{e}_i, A\mathbf{v} \rangle \geq c_j\langle \mathbf{e}_i, \mathbf{a}_j \rangle$ 。又因 c_1, \dots, c_n 不全為 0，故若 $c_j > 0$ ，則 $c_j\langle \mathbf{e}_i, \mathbf{a}_j \rangle > 0$ ，因此 $\langle \mathbf{e}_i, A\mathbf{v} \rangle \geq c_j\langle \mathbf{e}_i, \mathbf{a}_j \rangle > 0$ 。得證 $A\mathbf{v}$ 的每個 entry 皆為正實數。□

現若 $A = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix}$ 為 stochastic matrix，則依矩陣乘法定義 A^2 的 i -th column 為 $A\mathbf{a}_i$ ，故由 Lemma 8.5.4 知， A^2 的每個 column 皆為 probability vector，亦即 A^2 亦為 stochastic matrix。同理對任意 $k \geq 2$ ， A^k 的 i -th column 為 $A^{k-1}\mathbf{a}_i$ ，因此利用數學歸納法以及 Lemma 8.5.4，我們得證 A^k 亦為 stochastic matrix。同樣的利用數學歸納法以及 Lemma 8.5.4，我們可以證明若 A^r 的每一個 entry 皆為正實數，則對於所有 $k \in \mathbb{N}$ ， $A^{r+k} = A^{r+k-1}A$ 的每個 entry 亦皆為正實數。因此有以下的定理。

Proposition 8.5.5. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 stochastic matrix，則對所有 $k \in \mathbb{N}$ ， A^k 亦為 stochastic matrix。又若 A 為 regular 且 A^r 的每個 entry 皆為正實數，則對所有 $k \in \mathbb{N}$ ， A^{r+k} 的每個 entry 亦皆為正實數。

接下來我們要談論 stochastic matrix 的 eigenvalues 以及 eigenvectors. 不像前面的情況, 由於我們探討的是一般的 stochastic matrix 而不是具體的矩陣, 所以我們無法從它的 characteristic polynomial 來處理. 這裡我們需要特定的技巧, 首先我們從轉置矩陣出發.

Lemma 8.5.6. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 stochastic matrix. 則 1 為 A^t 的一個 eigenvalue 且

$\mathbf{v} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ 為其 eigenvector. 另外若 A 的每個 entry 皆為正實數, 則對於 A^t , 其 eigenvalue 1 的 geometric multiplicity 為 1.

Proof. 由於 A 為 stochastic matrix, A 每一個 column vector \mathbf{a}_i 皆為 probability vector,

亦即 $\langle \mathbf{a}_i, \mathbf{v} \rangle = 1$. 因此我們有 $A^t \mathbf{v} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{a}_n, \mathbf{v} \rangle \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{v}$. 得證 \mathbf{v} 為 A^t 的 eigenvector 且其 eigenvalue 為 1.

現考慮 A 的每個 entry 皆為正實數且假設 $\mathbf{w} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 為 A^t 的 eigenvalue 為 1 的 eigenvector 且 c_j 為 c_1, \dots, c_n 的最大值. 考慮 $A^t \mathbf{w}$ 的 j -th entry, 依定義其值為 $\langle \mathbf{a}_j, \mathbf{w} \rangle = a_{1j}c_1 + \dots + a_{nj}c_n$. 由於 $A^t \mathbf{w} = \mathbf{w}$, 所以 $A^t \mathbf{w}$ 的 j -th entry 應為 c_j , 因此有

$$c_j - (a_{1j}c_1 + \dots + a_{nj}c_n) = 0.$$

此時利用 $a_{1j} + \dots + a_{nj} = 1$ 可得

$$a_{1j}(c_j - c_1) + \dots + a_{nj}(c_j - c_n) = 0. \quad (8.16)$$

因為 a_{1j}, \dots, a_{nj} 皆為正實數且對每一個 $i = 1, \dots, n$ 皆有 $(c_j - c_i) \geq 0$, 故式子 (8.16) 中的每一項 $a_{ij}(c_j - c_i) \geq 0$, 但相加等於 0, 故得 $a_{ij}(c_i - c_j) = 0$. 再由 $a_{ij} \neq 0$ 的假設證得了 $c_1 = \dots = c_j = \dots = c_n = r$. 這說明了 $\mathbf{w} = r\mathbf{v}$, 亦即所有 A^t 的 eigenvalue 為 1 的 eigenvector 皆在 $\text{Span}(\mathbf{v})$ 中. 因此得證 A^t 其 eigenvalue 1 的 geometric multiplicity 為 1. \square

回顧 Proposition 7.3.6 和 Proposition 8.1.6 告訴我們 A 和 A^t 有相同的 eigenvalues 而且每個 eigenvalue 對於 A 和 A^t 的 geometric multiplicity 相同. 因此我們有以下的結果.

Proposition 8.5.7. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 stochastic matrix. 則 1 為 A 的一個 eigenvalue. 另外若 A 為 regular, 則對於 A , 其 eigenvalue 1 的 geometric multiplicity 為 1.

Proof. 因 A 為 stochastic matrix, 由 Lemma 8.5.6 知 1 為 A^t 的一個 eigenvalue. 故由 Proposition 7.3.6 知 1 亦為 A 的一個 eigenvalue. 另外, 若 A 為 regular 且假設 $r \in \mathbb{N}$ 使得 A^r 的每個 entry 皆為正實數, 則由 Lemma 8.5.6 知 $(A^r)^t$ 的 eigenvalue 1 其 geometric multiplicity 為 1. 也因此由 Proposition 8.1.6 知 A^r 的 eigenvalue 1, 其 geometric multiplicity 亦為 1, 也就是說 $\dim(E_{A^r}(1)) = 1$. 現對於任意 $\mathbf{v} \in E_A(1)$, 由於 $A\mathbf{v} = \mathbf{v}$, 我們得 $A^r \mathbf{v} = \mathbf{v}$, 亦即 $\mathbf{v} \in E_{A^r}(1)$. 因此得 $E_A(1) \subseteq E_{A^r}(1)$, 所以 $\dim(E_A(1)) \leq \dim(E_{A^r}(1)) = 1$. 然而前面已知 1 為 A

的一個 eigenvalue, 因此 $\dim(E_A(1)) > 0$. 得證 $\dim(E_A(1)) = 1$, 亦即 1 對於 A 的 geometric multiplicity 為 1. \square

現在我們來探討一個 stochastic matrix $A = [a_{ij}] \in M_{n \times n}(\mathbb{R})$ 其 eigenvector 有何特性. 假設 $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ 為 eigenvalue 為 λ 的 eigenvector. 此時由矩陣乘法與內積關係知 $\langle \mathbf{u}, A\mathbf{x} \rangle = \langle A^t \mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{x} \rangle$; 另一方面因 \mathbf{x} 為 eigenvalue 為 λ 的 eigenvector, 我們有 $\langle \mathbf{u}, A\mathbf{x} \rangle = \langle \mathbf{u}, \lambda \mathbf{x} \rangle = \lambda \langle \mathbf{u}, \mathbf{x} \rangle$. 亦即 $\langle \mathbf{u}, \mathbf{x} \rangle = \lambda \langle \mathbf{u}, \mathbf{x} \rangle$, 故知當 $\lambda \neq 1$ 時 $\langle \mathbf{u}, \mathbf{x} \rangle = x_1 + \cdots + x_n = 0$.

另一方面, 因 $A\mathbf{x} = \lambda \mathbf{x}$. 我們有 $\lambda x_i = \sum_{j=1}^n a_{ij} x_j$ 且因為對於任意 $i, j \in \{1, \dots, n\}$ 皆有 $a_{ij} \geq 0$, 我們得

$$\begin{aligned} |\lambda| |x_1| &= |a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n| \leq a_{11}|x_1| + a_{12}|x_2| + \cdots + a_{1n}|x_n| \\ &\vdots \\ |\lambda| |x_i| &= |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \leq a_{i1}|x_1| + a_{i2}|x_2| + \cdots + a_{in}|x_n| \\ &\vdots \\ |\lambda| |x_n| &= |a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n| \leq a_{n1}|x_1| + a_{n2}|x_2| + \cdots + a_{nn}|x_n| \end{aligned} \quad (8.17)$$

將式子 (8.17) 由上往下加起來且將右式同樣的 $|x_j|$ 加在一起, 由於 A 為 stochastic matrix, 對於 $j = 1, \dots, n$, 我們有 $a_{1j} + \cdots + a_{ij} + \cdots + a_{nj} = 1$, 故得

$$|\lambda| (|x_1| + |x_2| + \cdots + |x_n|) \leq |x_1| + |x_2| + \cdots + |x_n|.$$

由於 x_1, \dots, x_n 不全為 0, 得證 $|\lambda| \leq 1$. 也就是說當 A 為 stochastic matrix, 它的 eigenvalue λ 必須滿足 $|\lambda| \leq 1$.

現在我們回到探討 A 為 regular 的情形. 首先考慮 eigenvalue 為 -1 的情形. 假設 $\mathbf{w} \in \mathbb{R}^n$ 為 A 的 eigenvalue 為 -1 的 eigenvector. 此時由於 $A\mathbf{w} = -\mathbf{w}$, 我們得 $A^k \mathbf{w} = (-1)^k \mathbf{w}$. 也就是說當 $k > 0$ 為偶數時 \mathbf{w} 就會是 A^k 的 eigenvalue 為 1 的 eigenvector. 又 A 必有 eigenvalue 為 1 的 eigenvector \mathbf{v} , 且當 A 為 regular 時我們知必存在 $k > 0$ 為偶數使得 A^k 的每個 entry 皆為正實數 (Proposition 8.5.5), 故由 Proposition 8.5.7 的證明我們知此時 $E_A(1) = E_{A^k}(1) = \text{Span}(\mathbf{v})$. 然而 $\mathbf{w} \in E_{A^k}(1) = \text{Span}(\mathbf{v})$, 此與 \mathbf{v}, \mathbf{w} 為 linearly independent (因 \mathbf{v}, \mathbf{w} 分別為 A 對應到 1, -1 的 eigenvector, 由 Proposition 8.1.3 知它們為 linearly independent) 相矛盾. 故知當 A 為 regular stochastic matrix 時 -1 不可能會是 A 的 eigenvalue.

最後我們考慮 A 為 regular 且 eigenvalue 為 1 的情形. 要注意由 Proposition 8.5.7 的證明我們知道, 若 A^r 的每個 entry 皆為正實數, 則 eigenspace $E_A(1) = E_{A^r}(1)$. 所以我們只要考慮 A 為 stochastic matrix 且 $A = (a_{ij})$ 的每個 entry a_{ij} 皆為正實數的情形即可. 假設 \mathbf{x} 為 A 的 eigenvalue 為 1 的 eigenvector. 此時由於 $A\mathbf{x} = \mathbf{x}$, 我們有 $x_i = \sum_{j=1}^n a_{ij} x_j$ 且因為對於任意 $i, j \in \{1, \dots, n\}$ 皆有 $a_{ij} > 0$, 我們得

$$\begin{aligned}
|x_1| &= |a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n| \leq a_{11}|x_1| + a_{12}|x_2| + \cdots + a_{1n}|x_n| \\
&\vdots \\
|x_i| &= |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \leq a_{i1}|x_1| + a_{i2}|x_2| + \cdots + a_{in}|x_n| \quad (8.18) \\
&\vdots \\
|x_n| &= |a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n| \leq a_{n1}|x_1| + a_{n2}|x_2| + \cdots + a_{nn}|x_n|
\end{aligned}$$

同樣的，將式子 (8.18) 由上往下加起來，由於 A 為 stochastic matrix, 可得

$$|x_1| + |x_2| + \cdots + |x_n| \leq |x_1| + |x_2| + \cdots + |x_n|. \quad (8.19)$$

也就是說式子 (8.19) 的等式必須成立. 然而式子 (8.19) 的等式成立若且唯若式子 (8.18) 的每一項等式皆成立. 而又對於任意 $i, j \in \{1, \dots, n\}$ 皆有 $a_{ij} > 0$ 故式子 (8.18) 的每一項等式皆成立若且唯若 x_1, \dots, x_n 皆同時大於等於 0 或同時小於等於 0. 由於 x_1, \dots, x_n 不全為 0, 我們有 $x_1 + \cdots + x_n \neq 0$, 故令 $\mathbf{v} = \frac{1}{x_1 + \cdots + x_n} \mathbf{x}$, 我們得 \mathbf{v} 為 probability vector 且為 A 的 eigenvalue 為 1 的 eigenvector. 因為 $\dim(E_A(1)) = 1$, 我們知道 \mathbf{v} 會是 \mathbb{R}^n 中唯一同時符合這兩個要求的 vector. 綜合以上的討論, 我們有以下的結論.

Proposition 8.5.8. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 stochastic matrix, 則 A 的任一 eigenvalue λ 皆需滿足 $|\lambda| \leq 1$. 若 $\mathbf{w} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n$ 為 A 的 eigenvector 且其 eigenvalue 不等於 1, 則 $c_1 + \cdots + c_n = 0$.

另外若 A 為 regular, 則 -1 不可能會是 A 的 eigenvalue 且在 \mathbb{R}^n 中存在唯一的 probability vector 會是 A 的 eigenvalue 為 1 的 eigenvector.

Question 8.7. 假設 A 為 regular stochastic matrix 且 \mathbf{v} 為 A 的 eigenvalue 為 1 的 eigenvector. 試證明 \mathbf{v} 每一個 entry 皆同時為正或同時為負 (即沒有一個 entry 會是 0).

Exercise 8.25. 利用找到 shift function (數列往後移一位) 的表現矩陣及其 eigenvalues, eigenvectors 的方式處理以下問題。

- (1) 考慮遞迴數列 $\langle z_n \rangle_{n=1}^{\infty}$ 滿足

$$z_{n+1} = 3z_n - 2z_{n-1}, \text{ for } n \geq 2.$$

請找到符合此遞迴關係式的兩個不同公比的等比數列。並寫下 $\langle z_n \rangle$ 首二項為 $z_1 = 2, z_2 = 3$ 的數列其 z_n 的通式。

- (2) 考慮遞迴數列 $\langle z_n \rangle_{n=1}^{\infty}$ 滿足

$$z_{n+1} = 2z_n + z_{n-1} - 2z_{n-2}, \text{ for } n \geq 3.$$

請找到符合此遞迴關係式的三個不同公比的等比數列。並寫下 $\langle z_n \rangle$ 首三項為 $z_1 = 0, z_2 = z_3 = 1$ 其 z_n 的通式。

- (3) 給定 $r \in \mathbb{R}$ 滿足 $0 < |r-1| < 1$ 。考慮遞迴數列 $\langle z_n \rangle_{n=1}^{\infty}$ 滿足 $z_1 = 0, z_2 = 1$ 以及

$$z_{n+1} = rz_n + (1-r)z_{n-1}, \text{ for } n \geq 2.$$

試說明此數列為 *neutrally stable* (意指不存在 $N \in \mathbb{N}$ 使得 $z_m = z_N, \forall m > N$, 但 $\lim_{k \rightarrow \infty} z_k$ 存在), 並求其極限。

Exercise 8.26. 考慮 stochastic matrices

$$A_1 = \begin{bmatrix} 0 & 0.5 & 0 \\ 0.5 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0.5 \\ 0 & 1 & 0.5 \end{bmatrix}; \quad A_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0.5 \\ 0 & 0 & 0.5 \end{bmatrix}; \quad A_4 = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0 \\ 1 & 0 & 0.5 \end{bmatrix}.$$

- (1) 請分別利用 eigenvalues 和 eigenvectors 說明哪些不是 regular stochastic matrix.
- (2) 驗證可能為 regular stochastic matrix 的矩陣。(利用連續平方, 以及矩陣乘法用 column, row 來處理會比較快)
- (3) 說明哪些為 diagonalizable.

Exercise 8.27. 假設 A 為 invertible stochastic matrix.

- (1) 利用 $\mathbf{u} = \begin{bmatrix} 1 \\ \cdots \\ 1 \end{bmatrix}$ 為 A^t 的一個 eigenvalue 為 1 的 eigenvector, 證明 A^{-1} 的每一個 column vector 其 entry 之和為 1.
- (2) 試證明若 A 為 regular, 則 A^{-1} 一定不是 stochastic matrix.