

現在我們假設 $A \in M_{n \times n}(\mathbb{R})$ 為 regular stochastic matrix 且為 diagonalizable. 令 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 為 \mathbb{R}^n 的一組 basis 且為 A 的 eigenvectors. 假設 \mathbf{v}_i 所對應的 eigenvalue 為 λ_i , 其中 $\lambda_1 = 1$. 由 Proposition 8.5.8 我們可設 \mathbf{v}_1 為 probability vector, 又對於 $i = 2, \dots, n$,

$|\lambda_i| < 1$ 故若 $\mathbf{v}_i = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 我們有 $c_1 + \dots + c_n = 0$. 現令 $P = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}$, 我們得 $A =$

$$P \begin{bmatrix} 1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} P^{-1}. \text{ 由於矩陣的乘法僅牽涉到每個 entry 之間的加法與乘法, 所以當取}$$

極限時可以分別取極限再乘在一起. 因此我們有

$$\lim_{k \rightarrow \infty} A^k = \lim_{k \rightarrow \infty} P \begin{bmatrix} 1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}^k P^{-1} = \lim_{k \rightarrow \infty} P \begin{bmatrix} 1 & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} P^{-1} = P \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} P^{-1}.$$

注意對於 $i = 2, \dots, n$, \mathbf{v}_i 每一個 entry 之和為 0, 亦即 $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$. 我們知 $\mathbf{u} \in \text{Span}(\mathbf{v}_2, \dots, \mathbf{v}_n)^\perp$, 再因 $\mathbf{v}_2, \dots, \mathbf{v}_n$ 為 linearly independent, 知 $\dim(\text{Span}(\mathbf{v}_2, \dots, \mathbf{v}_n)^\perp) = n - (n-1) = 1$. 因此得 $\text{Span}(\mathbf{v}_2, \dots, \mathbf{v}_n)^\perp = \text{Span}(\mathbf{u})$. 另一方面因 \mathbf{v}_1 為 probability vector, 故 $\langle \mathbf{u}, \mathbf{v}_1 \rangle = 1$, 因此利用 $P^{-1}P = I_n$, 推得 P^{-1} 的 1-st row 為 \mathbf{u}^t , 即每一個 entry 皆為 1. 故得

$$P \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} P^{-1} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ - & \mathbf{0} & - \\ \vdots & & \\ - & \mathbf{0} & - \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_1 & \cdots & \mathbf{v}_1 \\ | & | & & | \end{bmatrix}.$$

也就是說當 A 為 regular stochastic matrix 且為 diagonalizable, 則當 k 越來越大時 A^k 會趨近一個每個 column 皆為 \mathbf{v}_1 的矩陣, 其中 \mathbf{v}_1 為 \mathbb{R}^n 中唯一的 probability vector 滿足 $A\mathbf{v}_1 = \mathbf{v}_1$. 這個結果其實對於一般的 regular stochastic matrix (不需 diagonalizable 之假設) 皆成立. 它的證明並沒有用到太多線性代數的技巧, 而著重於數列極限的估計. 不過為了完整性, 我們仍簡略的證明一下. 同學們可以忽略它的證明, 不過我們希望大家仍能充分了解此定理的敘述以及其衍伸的結果 (Corollary 8.5.10).

Theorem 8.5.9. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 regular stochastic matrix 且 $\mathbf{v} \in \mathbb{R}^n$ 為唯一的 probability vector 滿足 $A\mathbf{v} = \mathbf{v}$. 則

$$\lim_{k \rightarrow \infty} A^k = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v} & \mathbf{v} & \cdots & \mathbf{v} \\ | & | & & | \end{bmatrix}.$$

Proof. 當 $n = 1$ 時 $A = [1]$ 是唯一的 stochastic matrix, 此時本定理自然成立, 所以我們考慮 $n \geq 2$ 的情形. 此時我們要證明 $\lim_{k \rightarrow \infty} A^k$ 的每一個 column 是相同的 vector.

要如何證明 $\lim_{k \rightarrow \infty} A^k$ 的每一個 column 是相同的 vector 呢? 給定 $i \in \{1, \dots, n\}$, 我們令 \mathbf{v} 的 i -th entry c_i , 以及對所有 $k \in \mathbb{N}$ 考慮 A^k 的 i -th row, 且令 M_k, m_k 分別為 A^k 的 i -th row 的 entry 中的最大值與最小值. 我們要證明 $m_k \leq c_i \leq M_k$ 且 $\lim_{k \rightarrow \infty} (M_k - m_k) = 0$, 因此

得證 $\lim_{k \rightarrow \infty} A^k$ 的 i -th row 的每個 entry 皆為同一個數 c_i . 因為這是對所有 $i = 1, \dots, n$ 皆成立, 也因此得證 $\lim_{k \rightarrow \infty} A^k$ 的每一個 column 皆為相同的 vector \mathbf{v} .

首先, 令 a_{ij} 為 A 的 (i, j) -th entry 且當 $k > 1$ 時令 $a_{ij}^{(k)}$ 為 A^k 的 (i, j) -th entry. 因 $A\mathbf{v} = \mathbf{v}$, 我們有 $A^k\mathbf{v} = \mathbf{v}, \forall k \in \mathbb{N}$. 若 $\mathbf{v} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, 則考慮 $A^k\mathbf{v}$ 的 i -th entry, 我們有

$$c_i = c_1 a_{i1}^{(k)} + \dots + c_n a_{in}^{(k)} = \sum_{j=1}^n c_j a_{ij}^{(k)} \quad (8.20)$$

因 \mathbf{v} 為 probability vector, 我們有 $c_1 + \dots + c_n = 1$. 又依假設 $m_k \leq a_{ij}^{(k)}, \forall j = 1, \dots, n$, 故由式子 (8.20) 知 $c_i \geq m_k(c_1 + \dots + c_n) = m_k$. 同理, 因 $a_{ij}^{(k)} \leq M_k$, 故證得 $m_k \leq c_i \leq M_k$.

接下來, 我們證明 $\lim_{k \rightarrow \infty} (M_k - m_k) = 0$. 依矩陣乘法的定義以及 $A^{k+1} = A^k A$, 我們有 $a_{ij}^{(k+1)} = \sum_{q=1}^n a_{iq}^{(k)} a_{qj}$. 假設 A^k 的 i -th row 的最大值和最小值分別發生在 A^k 的 (i, s) -th entry 以及 (i, t) -th entry (即 $M_k = a_{is}^{(k)}, m_k = a_{it}^{(k)}$), 又假設 A^{k+1} 的 i -th row 的最大值和最小值分別發生在 A^{k+1} 的 (i, j) -th entry 以及 (i, l) -th entry (即 $M_{k+1} = a_{ij}^{(k+1)}, m_{k+1} = a_{il}^{(k+1)}$). 此時我們有

$$M_{k+1} = a_{ij}^{(k+1)} = a_{it}^{(k)} a_{tj} + \sum_{q \neq t} a_{iq}^{(k)} a_{qj} \leq m_k a_{tj} + M_k \sum_{q \neq t} a_{qj} = m_k a_{tj} + M_k (1 - a_{tj}). \quad (8.21)$$

$$m_{k+1} = a_{il}^{(k+1)} = a_{is}^{(k)} a_{sl} + \sum_{q \neq s} a_{iq}^{(k)} a_{ql} \geq M_k a_{sl} + m_k \sum_{q \neq s} a_{ql} = M_k a_{sl} + m_k (1 - a_{sl}). \quad (8.22)$$

注意式子 (8.21), (8.22) 用到了 A 為 stochastic matrix 之假設, 即 $\sum_{q=1}^n a_{qj} = \sum_{q=1}^n a_{ql} = 1$. 現在將式子 (8.21) 減去式子 (8.22) 得

$$M_{k+1} - m_{k+1} \leq M_k (1 - a_{tj} - a_{sl}) + m_k (a_{tj} - (1 - a_{sl})) = (M_k - m_k) (1 - a_{tj} - a_{sl}). \quad (8.23)$$

由於 $a_{tj} \geq 0, a_{sl} \geq 0$, 我們有 $1 - a_{tj} - a_{sl} \leq 1$, 故由式子 (8.23) 得 $M_{k+1} - m_{k+1} \leq M_k - m_k$. 也就是說數列 $(M_k - m_k)_{k=1}^{\infty}$ 是一個遞減的數列.

如何證明 $(M_k - m_k)_{k=1}^{\infty}$ 這個遞減的數列其極限為 0 呢? 我們只要找到它的一個 subsequence 趨近於 0 即可. 這裡就需要用到 A 為 regular 的假設. 也就是說存在 $r \in \mathbb{N}$ 使得 A^r 的每個 entry 皆為正實數. 為了方便起見, 我們令 $B = A^r$ 且令 b_{ij} 為 B 的 (i, j) -th entry 且當 $k > 1$ 時令 $b_{ij}^{(k)}$ 為 $B^k = A^{rk}$ 的 (i, j) -th entry. 注意依假設我們有 $b_{ij} > 0$ 且 $b_{ij}^{(k)} > 0$. 又給定 $i \in \{1, \dots, n\}$, 我們有 M_{rk}, m_{rk} 分別為 $B^k = A^{rk}$ 的 i -th row 的 entry 中的最大值與最小值. 因此用相同的方法, 套用式子 (8.21), (8.22) 在 B^k 和 B^{k+1} 的情形, 可以推得式子 (8.23) 在 B^k 與 B^{k+1} 相對應的結果

$$M_{r(k+1)} - m_{r(k+1)} \leq M_{rk} (1 - b_{tj} - b_{sl}) + m_{rk} (b_{tj} - (1 - b_{sl})) = (M_{rk} - m_{rk}) (1 - b_{tj} - b_{sl}). \quad (8.24)$$

要注意式子 (8.24) 與式子 (8.23) 最大的不同處在於 $b_{tj} > 0$ 且 $b_{sl} > 0$. 現令 β 為 B 的所有 entry 中的最小值. 依 β 的定義, 我們有 $b_{tj} \geq \beta, b_{sl} \geq \beta$, 故得 $1 - b_{tj} - b_{sl} \leq 1 - 2\beta$. 因此由式子 (8.23) 得

$$M_{r(k+1)} - m_{r(k+1)} \leq (M_{rk} - m_{rk}) (1 - 2\beta). \quad (8.25)$$

因 β 所在的那個 column 所有的 entry 之和等於 1 且大於等於 $n\beta$, 故得 $0 < \beta \leq \frac{1}{n} \leq \frac{1}{2}$ (注意我們有 $n \geq 2$ 之假設). 所以 $0 \leq 1 - 2\beta < 1$, 利用式子 (8.25) 以及數學歸納法得

$$M_{r(k+1)} - m_{r(k+1)} \leq (M_r - m_r)(1 - 2\beta)^k,$$

且因此得證 $(M_{rk} - m_{rk})_{k=1}^{\infty}$ 這個 subsequence 滿足 $\lim_{k \rightarrow \infty} (M_{rk} - m_{rk}) = 0$, 故得原 sequence $(M_k - m_k)_{k=1}^{\infty}$ 滿足 $\lim_{k \rightarrow \infty} (M_k - m_k) = 0$. 得證本定理. \square

再次重申 Theorem 8.5.9 對於一般的 stochastic matrix 未必會成立, 須加上 regular 的條件才一定成立. 這個定理的最重要的結果就是在 Markov process 中不管初始的 probability vector 為何, 經過 regular stochastic matrix 多次的作用之下, 最後會趨近於一個穩定的狀態. 我們將此結果敘述如下.

Corollary 8.5.10. 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 regular stochastic matrix 且 $\mathbf{v} \in \mathbb{R}^n$ 為唯一的 probability vector 滿足 $A\mathbf{v} = \mathbf{v}$. 則對於任意 \mathbb{R}^n 的 probability vector \mathbf{w} 皆有

$$\lim_{k \rightarrow \infty} (A^k \mathbf{w}) = \mathbf{v}.$$

Proof. 令 $\mathbf{w} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$. 依 \mathbf{w} 為 probability vector 之假設, 我們知 $c_1 + \cdots + c_n = 1$. 故由 Theorem 8.5.9 知

$$\lim_{k \rightarrow \infty} (A^k \mathbf{w}) = \left(\lim_{k \rightarrow \infty} A^k \right) \mathbf{w} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v} & \mathbf{v} & \cdots & \mathbf{v} \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c_1 \mathbf{v} + c_2 \mathbf{v} + \cdots + c_n \mathbf{v} = \mathbf{v}.$$

\square

Question 8.8. Corollary 8.5.10 中 \mathbf{w} 需要假設是 probability vector 才會成立嗎?

Exercise 8.28. 參考 Exercise 8.26，試回答以下 stochastic matrices 相關問題：

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

- (1) 說明 $\lim_{k \rightarrow \infty} A_i^k$ 是否存在。若存在，求其極限。
- (2) 令 $\mathbf{p} = \begin{bmatrix} 1/6 \\ 1/3 \\ 1/2 \end{bmatrix}$ ，說明 $\lim_{k \rightarrow \infty} A_i^k \mathbf{p}$ 是否存在。若存在，求其極限。
- (3) 對任意的 probability vector \mathbf{v} 。請回答以下問題並說明之。
 - (a) 那個 A_i 會使得 $\lim_{k \rightarrow \infty} A_i^k \mathbf{v}$ 一定存在且不管 \mathbf{v} 為何都會趨近到同一個向量；
 - (b) 那個 A_i 會使得 $\lim_{k \rightarrow \infty} A_i^k \mathbf{v}$ 一定存在但 \mathbf{v} 改變會趨近於不同的向量；
 - (c) 那個 A_i 會使得 $\lim_{k \rightarrow \infty} A_i^k \mathbf{v}$ 不一定存在。
 - (d) 那個 A_i 會使得 $\lim_{k \rightarrow \infty} A_i^k \mathbf{v}$ 一定不存在。

Exercise 8.29.

已知 A 為 regular stochastic matrix.

- (1) 證明若 \mathbf{v} 為 A 的 eigenvalue 為 1 的 eigenvector，則 \mathbf{v} 的每一個 entry 皆不等於 0。
- (2) 假設 $A \in M_{n \times n}(\mathbb{R})$ 為 diagonalizable 且 Q 為 invertible matrix 使得 $Q^{-1}AQ = D$ ，其中 D 為 diagonal matrix 且其 (1,1)-entry 為 1。證明 Q^{-1} 的 1-st row 為 $[r \cdots r]$ 其中 r 為非零實數。特別的若 Q 的 1-st column 為 $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ ，則 $r = (c_1 + \cdots + c_n)^{-1}$ 。

Exercise 8.30. 考慮函數 $S: \mathbb{R}^n \rightarrow \mathbb{R}$ 定義為 $S(a_1, \dots, a_n) = \sum_{i=1}^n a_i$ 。對 $A \in M_{n \times n}(\mathbb{R})$ 令 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ，定義為 $T_A(\mathbf{v}) = A\mathbf{v}$ 。令 $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ 分別為 \mathbb{R}^n 和 \mathbb{R} 的 standard ordered basis.

- (1) 試證明 S 為 linear map 且寫下 S 對 $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ 的表現矩陣 $[S]_{\boldsymbol{\varepsilon}'}^{\boldsymbol{\varepsilon}}$ 。
- (2) 考慮合成函數 $S \circ T_A: \mathbb{R}^n \rightarrow \mathbb{R}$ 。試說明 $[S \circ T_A]_{\boldsymbol{\varepsilon}'}^{\boldsymbol{\varepsilon}}$ 為何。
- (3) 假設 A 的每個 entry 皆為非負實數。證明 A 為 stochastic matrix 若且唯若 $S \circ T_A = S$ 。
- (4) 假設 A 為 stochastic matrix 且 $B \in M_{n \times n}(\mathbb{R})$ 的每個 entry 皆為非負實數。利用函數合成的結合律 (即 $S \circ (T_A \circ T_B) = (S \circ T_A) \circ T_B$) 以及函數合成與矩陣乘法的關係證明 B 為 stochastic matrix 若且唯若 AB 為 stochastic matrix。
- (5) 試找到 $A, B \in M_{2 \times 2}(\mathbb{R})$ 其中 A 和 BA 為 stochastic matrix 但是 B 並不是 stochastic matrix。