

随机矩阵的一些性质

杨载朴 朱庆国

(盐城工学院基础科学部, 盐城, 224003)

摘 要 论述并证明了带行和 S 的既约广义随机矩阵的性质, 双随机矩阵积的性质和 P 类矩阵的一些性质。

关键词 随机矩阵 既约广义随机矩阵 双随机矩阵

分类号 O151

引 言

随机矩阵(stochastic matrix)是概率论中研究 Markov 链的有力工具, 所谓齐次 Markov 链是指这样一种物理体系 S , 它在每一时刻处于 $B_1, B_2, \dots, B_n$ 这 n 个状态之一, 且只在 $t_1 < t_2 < \dots < t_m < \dots$ 诸时刻改变其状态, 同时它在时刻 $t (t_i \leq t < t_{i+1})$ 时状态为 $B_j (j=1, 2, \dots, n)$ 的概率 p_{ij} 只与该体系在时刻 $t (t_{i-1} \leq t < t_i)$ 的状态 B_i 有关, 换句话说, 在前一个时刻处于状态 B_i 的条件下, 后一时刻处于状态 B_j 的概率为 p_{ij} , 如果再记体系 S 在初始时刻 $t_0 (t_0 < t_1)$ 为状态 B_i 之概率为 $p_i^{(0)}$, 且记

$$P^{(0)} = (p_1^{(0)}, \dots, p_n^{(0)}), \quad P = (p_{ij})_{n \times n}$$

由概率论中的基本常识知

$$0 \leq p_j^{(0)} \leq 1, (j=1, \dots, n) \quad \sum_{j=1}^n p_j^{(0)} = 1 \quad (1)$$

$$0 \leq p_{ij} \leq 1, (i, j=1, \dots, n) \quad \sum_{j=1}^n p_{ij} = 1, (i=1, \dots, n) \quad (2)$$

因体系 S 在 t_0 时刻处于状态 B_i 的概率为 $p_i^{(0)}$, 由全概率公式, 它在 t_1 时刻处于 B_j 的概率为 $\sum_{i=1}^n p_i^{(0)} p_{ij}$, 这恰为 n 元向量

$$P^{(0)}P = (p_1^{(0)}, p_2^{(0)}, \dots, p_n^{(0)}) \begin{bmatrix} p_{11} & \dots & p_{1n} \\ \dots & \dots & \dots \\ p_{n1} & \dots & p_{nn} \end{bmatrix}$$

的第 j 个分量, $(j=1, \dots, n)$, 一般地 $P^{(0)}P^m$ 之第 j 个分量即表示体系 S 在时刻 t_m 处于状态 B_j 的概率 $(j=1, \dots, n)$, 因此有了 $P^{(0)}$ 与 P , 这个体系 S 在任何时刻所处的各状态 $B_1, \dots, B_n$ 的概率就知道了, 称 $P^{(0)}$ 为体系 S 的初始分布向量, 且给出:

定义 1 称满足(2)的矩阵 $P=(p_{ij})_{n \times n}$ 为随机矩阵, 更一般地, 若 $A=(a_{ij})_{n \times n}$, 满足条件

$$0 \leq a_{ij} \leq S, (i, j=1, 2, \dots, n) \quad \sum_{j=1}^n a_{ij} = S, (i=1, \dots, n)$$

则称 A 为行和是 S 的广义随机矩阵。

定义 2 设 $A = (a_{ij})_{n \times n}$, 若 A, A^T 均为随机矩阵, 则称 A 为双随机矩阵(doubly stochastic matrix)。

由此可见, 给出随机矩阵或双随机矩阵 A , 研究其幂 A^n 的性质, 无论在理论上或实践上都有一定的意义, 本文在一定条件下给出双随机矩阵 A 的乘积或幂的若干性质, 以及既约广义随机矩阵的性质。

1 既约广义随机矩阵的性质

先引进

引理 1^{[1][2][3]} 设 $A = (a_{ij})_{n \times n}$ 为带行和 S 的既约广义随机矩阵, 则

(1) $\lambda = S$ 为 A 的一个特征值, $e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ 为 A 的对应于 $\lambda = S$ 的一个特征向量。

(2) $\rho(A) = S$, 这里 $\rho(A)$ 为 A 的谱半径, 即特征值的模的最大值。

(3) S 为 A 的单重特征值。

(4) 若 $a_{ii} = \min_{1 \leq j \leq n} a_{jj}$, 则 A 的谱

$$\sigma(A) \subset D_i \equiv \{ |Z - a_{ii}| \leq S - a_{ii} \}.$$

即 A 的全部特征值包含在以 a_{ii} 为圆心, $S - a_{ii}$ 为半径的闭圆域内, 且 A 之 n 个 Γ epwzopuH 圆均内切于大圆 Γ 于 S 点 (图 1, 含重合的情形)。

因既约广义随机矩阵为 β 类矩阵的特例, 因而有关 β 类矩阵的结论对既约广义随机矩阵都适用, 本文不再赘述。

定理 1 设 $A = (a_{ij})_{n \times n}$ 为行和 S 的既约广义随机矩阵, 且 $a_{jj} \neq 0, j = 1, 2, \dots, n$, 则 $\lambda = S$ 为 A 的模为 S 的唯一特征值。

证 用反证法, 设 A 有另一特征值 α 的模为 S , 则 α 在大圆 Γ 上, $\Gamma: |Z| = S$, 又由引理 1, α 在某一个 Γ epwzopuH 圆盘:

$$D_i: |Z - a_{ii}| \leq S - a_{ii}$$

上, 但由于 D_i 的边界 Γ_i 内切(重合)于大圆 Γ (图 1), 所以 α 既在 Γ_i 上又在 Γ 上, 即 α, S 均为 Γ_i 与 Γ 的切点, 因而 Γ_i 与 Γ 重合, 所以它们的圆心亦重合, 因而 $a_{ii} = 0$, 这与条件矛盾, 证毕。

推论 1 设 $A = (a_{ij})_{n \times n}$ 为既约随机矩阵, 且 $a_{jj} \neq 0, j = 1, 2, \dots, n$, 则 $A, A^2, \dots, A^n, \dots$ 是收敛序列。

证 因 A 为既约随机矩阵, 所以 $\rho(A) = 1$, 且 $\lambda = 1$ 为单重特征值, 且为 A 的模为 1 的唯一特征值, 由 [1] 中第一章定理 2 知, $A, A^2, \dots, A^n, \dots$ 是收敛序列。

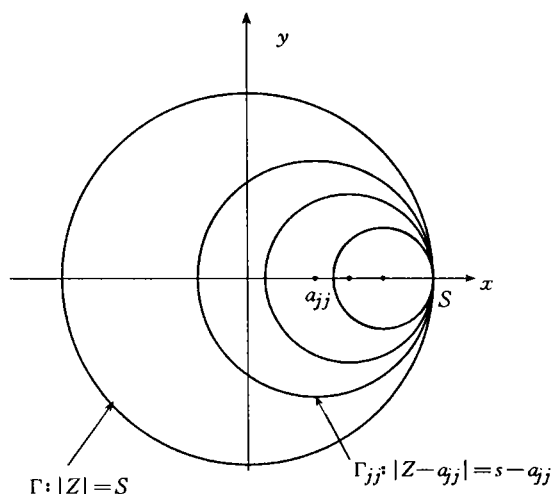


图 1

2 双随机矩阵积的性质

引理 2 设 $A=(a_{ij})_{n \times n}, B=(b_{ij})_{n \times n}$ 均为双随机矩阵, 则 AB 亦为双随机矩阵。

证 记 $AB=C=(c_{ij})_{n \times n}$, 考察 C 之第 i 行元素之和, 并注意到 A, B 均为双随机矩阵, 则有

$$\sum_{j=1}^n c_{ij} = \sum_{j=1}^n \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n a_{ik} \left(\sum_{j=1}^n b_{kj} \right) = \sum_{k=1}^n a_{ik} \cdot 1 = 1 \quad (i=1, 2, \dots, n)$$

且 c 之第 j 列元素之和

$$\sum_{i=1}^n c_{ij} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n b_{kj} \left(\sum_{i=1}^n a_{ik} \right) = \sum_{k=1}^n b_{kj} \cdot 1 = 1 \quad (j=1, 2, \dots, n)$$

因而 AB 为双随机矩阵。

推论 2 设 A 为双随机矩阵, m 为任一自然数, 则 A^m 为双随机矩阵。

推论 2 由引理 2 并用数学归纳法即得。

在本文中, 若 $A=(a_{ij})_{n \times n}$ 为既约双随机矩阵, 且 $a_{jj} > 0, j=1, 2, \dots, n$, 则记为 $A \in \mathcal{D}$ 。

定理 2 若 $A=(a_{ij})_{n \times n} \in \mathcal{D}, B=(b_{ij})_{n \times n} \in \mathcal{D}$, 则 $AB \in \mathcal{D}$ 。于是全体 n 阶 $\mathcal{D}$ 类矩阵成一半群。

证 由引理 2, AB 为双随机矩阵, 下面来证 AB 为既约矩阵, 用反证法。

设 AB 可约, 则存在置换矩阵 Q (注意到 $Q^{-1}=Q^T$), 使

$$Q(AB)Q^{-1} = \begin{pmatrix} C & M \\ 0 & D \end{pmatrix} \quad (3)$$

其中 C 为 $r \times r$ 矩阵, D 为 $(n-r) \times (n-r)$ 矩阵。

记 $e=(1, 1, \dots, 1)^T$, 因 A 为双随机矩阵, 所以 $\sum_{j=1}^n a_{ij} \cdot 1 = 1, i=1, 2, \dots, n$, 即 $Ae=1 \cdot e=e$, 同理 $Be=e$, 于是 $(AB)e=A(Be)=Ae=e$, 上式两边左乘 Q 得: $Q(AB)e=Qe$, 即 $(QABQ^{-1})Qe=Qe$, 所以有 (注意到 $Qe=e$)

$$(QABQ^{-1})e=e \quad (4)$$

记 $e = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$, 其中 $e_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, $e_2 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ 分别为 r 元和 $n-r$ 元列向量, 由 (3) (4) 得:

$\begin{pmatrix} C & M \\ 0 & D \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$, 所以 $De_2=e_2=1 \cdot e_2$, 因而 $\lambda=1$ 是 D 的一个特征值。

另一方面, 因 A^T, B^T 亦为双随机矩阵, 所以 $(B^T A^T)e=B^T(A^T e)=B^T e=e$, 所以 $Q(B^T A^T)e=Qe$, 即 $Q(B^T A^T Q^{-1})(Qe)=Qe$, 即 $(QB^T A^T Q^{-1})e=e$, 又因 $(QABQ^{-1})^T=QB^T A^T Q^{-1}=\begin{pmatrix} C^T & 0 \\ M^T & D^T \end{pmatrix}$, 将此式代入上式得 $\begin{pmatrix} C^T & 0 \\ M^T & D^T \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ 所以 $C^T e_1=e_1=1 \cdot e_1$, 从而 $\lambda=1$ 是 C^T

的一个特征值, 也是 C 的特征值, 因而 $\lambda=1$ 是 $QABQ^{-1}=\begin{pmatrix} C & M \\ 0 & D \end{pmatrix}$ 的重特征值, 也是 AB 的重特征值, 与引理 1 矛盾 (这里 $S=1$)。

下面再来证明 AB 之对角线元素大于零。

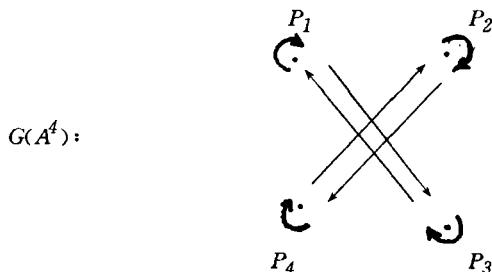
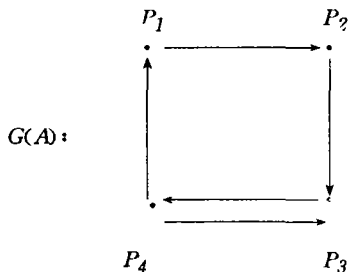
记 $AB=(c_{ij})_{n \times n}, c_{ii}=\sum_{k=1}^n a_{ik} b_{ki} \geq a_{ii} b_{ii} > 0$, 综上所述有 $AB \in \mathcal{D}$ 。

引理 3^[1] 若 A 为可约矩阵, m 为任一自然数, 则 A^m 为可约矩阵。

由引理 3 知, 若 A^m 为既约矩阵, 则 A 必为既约矩阵, 但其逆命题不真, 例如

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad A^4 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

但它们的方向图:



易见, A 为既约矩阵, 而 A^4 为可约矩阵, 但我们有以下结果:

定理 3 设 $A = (a_{ij})_{n \times n}$ 为双随机矩阵, 且 $a_{jj} \neq 0, j = 1, \dots, n$ 。 m 为任一自然数, 则 A^m 为既约矩阵的充要条件是 A 为既约矩阵。

证 必要性由引理 3 得。

充分性。设 A 为既约矩阵, 则 $A \in \mathcal{D}$, 由定理 2 并用数学归纳法得 $A^m \in \mathcal{D}$, 即 A^m 为既约矩阵。

例:

$$A = \begin{bmatrix} 0.1 & 0.2 & 0 & 0.7 \\ 0.5 & 0.3 & 0.2 & 0 \\ 0 & 0.4 & 0.5 & 0.1 \\ 0.4 & 0.1 & 0.3 & 0.2 \end{bmatrix}$$

容易验证 A 为双随机矩阵, 且其对角线元素不为 0, 又因为 A 的方向图 $G(A)$ 为强连接的, 所以 A 既约, 因而 $A \in \mathcal{D}$, 由定理 3 知对任意自然数 m , 有 $A^m \in \mathcal{D}$, 事实上

$$A^2 = \begin{bmatrix} 0.39 & 0.15 & 0.25 & 0.21 \\ 0.20 & 0.27 & 0.16 & 0.37 \\ 0.24 & 0.33 & 0.36 & 0.07 \\ 0.17 & 0.25 & 0.23 & 0.35 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0.198 & 0.244 & 0.218 & 0.340 \\ 0.303 & 0.222 & 0.245 & 0.230 \\ 0.217 & 0.298 & 0.267 & 0.218 \\ 0.282 & 0.236 & 0.270 & 0.212 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 0.2778 & 0.2340 & 0.2598 & 0.2284 \\ 0.2333 & 0.2482 & 0.2359 & 0.2826 \\ 0.2579 & 0.2614 & 0.2585 & 0.2222 \\ 0.2310 & 0.2564 & 0.2458 & 0.2668 \end{bmatrix}$$

易见 A^2, A^3, A^4 均为既约双随机矩阵。

推论 4 设 $A \in \mathcal{D}$, m 为任一自然数, 则

(1) $\lambda = 1$ 为 A^m 的单重特征值, 也是模为 1 的唯一特征值。 $e = (1, \dots, 1)^T$ 为 A^m 的对应于 $\lambda = 1$ 的一个特征向量。

(2) $\rho(A^m)=1$, 这里 $\rho(A^m)$ 为 A^m 的谱半径.

证 由定理 3, A^m 为双随机矩阵, 且其对角线元素大于 0. 再由引理 1、定理 1 知 (这里 $S=1$) $\lambda=1$ 为 A^m 的单重特征值, 也是模为 1 的唯一特征值, $e=(1,1,\dots,1)^T$ 为 A^m 的对应于 $\lambda=1$ 的一个特征向量, 且 $\rho(A^m)=1$.

3 $\mathcal{D}$ 类矩阵的其它一些性质

定理 4 设 $A=(a_{ij})_{n \times n} \in \mathcal{D}$, $0 \leq a_l \leq 1$, ($l=1,2,\dots,k$), $\sum_{l=1}^k a_l=1$, 则 $\sum_{l=1}^k a_l A^l \in \mathcal{D}$

证 记 $A^l=(a_{ij}^{(l)})_{n \times n}$, 由定理 3 知 $A^l \in \mathcal{D}$, 所以 A^l 为双随机矩阵, 因而 $\sum_{j=1}^n a_{ij}^{(l)}=1$, ($i=1,2,\dots,n$); $\sum_{l=1}^k a_l A^l=(\sum_{l=1}^k a_l a_{ij}^{(l)})_{n \times n}$, 其第 i 行元素之和为:

$$\sum_{j=1}^n (\sum_{l=1}^k a_l a_{ij}^{(l)}) = \sum_{l=1}^k a_l (\sum_{j=1}^n a_{ij}^{(l)}) = \sum_{l=1}^k a_l \cdot 1 = 1$$

同理 $\sum_{l=1}^k a_l A^l$ 的第 j 列元素之和为

$$\sum_{i=1}^n (\sum_{l=1}^k a_l a_{ij}^{(l)}) = \sum_{l=1}^k a_l (\sum_{i=1}^n a_{ij}^{(l)}) = \sum_{l=1}^k a_l \cdot 1 = 1$$

故 $\sum_{l=1}^k a_l A^l$ 为双随机矩阵.

又由条件 $a_1, \dots, a_k$ 中至少有一数大于 0, 不妨设 $a_m > 0$, 因 $A^m \in \mathcal{D}$, 因此 $\sum_{l=1}^k a_l A^l$ 任一元素 $\sum_{l=1}^k a_l a_{ij}^{(l)} \geq a_m a_{ij}^{(m)}$, 故 $\sum_{l=1}^k a_l A^l$ 既约, 且其对角线元素 $\sum_{l=1}^k a_l a_{ii}^{(l)} \geq a_m a_{ii}^{(m)} > 0$, $i=1,2,\dots,n$, 所以 $\sum_{l=1}^k a_l A^l \in \mathcal{D}$.

定理 5 记 $S=\{A^1, A^2, \dots, A^k | k \in N\} \subset \mathcal{D}$, $L(S)=\{\sum_{l=1}^k a_l A^l | A^l \in S, 0 \leq a_l \leq 1, \sum_{l=1}^k a_l=1\}$

若 $(\sum_{l=1}^k a_l A^l) \in L(S)$, $(\sum_{l=1}^m b_l A^l) \in L(S)$, 则 $(\sum_{l=1}^k a_l A^l) \cdot (\sum_{l=1}^m b_l A^l) \in L(S)$, 因此 $(L(S) \cdot)$ 是 $(\mathcal{D} \cdot)$ 的子半群.

证 $(\sum_{l=1}^k a_l A^l) \cdot (\sum_{l=1}^m b_l A^l) = \sum_{l=1}^k \sum_{l'=1}^m a_l b_{l'} A^{l+l'} = a_1 b_1 A^2 + (a_1 b_2 + a_2 b_1) A^3 + \dots + a_k b_m A^{k+m}$
其系数之和为

$$a_1 b_1 + (a_1 b_2 + a_2 b_1) + \dots + a_k b_m = (\sum_{l=1}^k a_l) \cdot (\sum_{l=1}^m b_l) = 1 \times 1 = 1$$

故

$$(\sum_{l=1}^k a_l A^l) \cdot (\sum_{l=1}^m b_l A^l) \in L(S)$$

参考文献

- 1 Pullman, N. J., Matrix Theory and its Applications, Marcel Dekker Inc, New York and Basel, 1976, P209~227
- 2 Franklin, J. N., Matrix Theory, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1968