

随机马尔可夫乘积矩阵的极限经验谱分布

程慧慧, 宋敏杰

(华北水利水电大学 数学与统计学院, 郑州 450046)

摘要: 设 $\{X_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ 是一列独立同分布的非负实随机变量, 均值为 1, 有限方差 $\sigma^2 > 0$, $M^{(k)} = (X_{ij}^{(k)} / \sum_{j=1}^N X_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$ 是 m 个马尔可夫随机矩阵. 证明了当 N 趋于无穷大时, 实对称矩阵 $\sigma^{-2m} \Pi_m^T \Pi_m$ 的经验谱分布函数 $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}$ 几乎必然弱收敛于 Fuss-Catalan 分布, 其中 $\Pi_m = N^{m/2} M^{(m)} \cdots M^{(2)} M^{(1)}$, $\lambda_1, \lambda_2, \dots, \lambda_N$ 为 $\Pi_m^T \Pi_m$ 的特征值. 并通过数值模拟, 得到模拟结果与理论结论一致.

关键词: 随机马尔可夫矩阵; 经验谱分布; 乘积矩阵; Fuss-Catalan 分布

中图分类号: O211

文献标志码: A

文章编号: 1000-2367(2025)06-0052-06

随机矩阵特征值的极限理论是随机矩阵理论的一个重要分支, 它研究当矩阵的维度趋向于无穷时, 其经验谱分布和极值特征值(最大、最小特征值)的极限问题. 该理论在数学、物理学和统计学等领域有着广泛的应用. 关于随机矩阵的极限经验谱分布问题, 已有很多经典的结果.

在 20 世纪 50 年代, Wigner 提出了著名的半圆律, 即实对称随机矩阵特征值经验谱分布趋向于半圆分布, 该理论用于解释量子力学中的粒子能级分布. 对于样本协方差矩阵, MARČHENKO 和 PASTUR 给出了著名的 Marčenko-Pastur 定理^[1-2]. 假设 $\{X_{ij}, i, j \geq 1\}$ 是一系列独立同分布的随机变量且方差 σ^2 满足 $0 < \sigma^2 < \infty$. 令 $X = (X_{ij})_{N \times N}$, 记 $W = \frac{1}{N} X^T X$, $v_w = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(W)}$, 则 Marčenko-Pastur 定理表明: $v_w \xrightarrow{\ell_b} \mathcal{F}_{MP}$, a. s., 其中“ $\xrightarrow{\ell_b}$ ”表示概率测度弱收敛, $\mathcal{F}_{MP}$ 表示 Marčenko-Pastur 律, 其密度函数为

$$f_{MP}(x) = \begin{cases} \frac{1}{2\pi x \sigma^2} \sqrt{x(4\sigma^2 - x)}, & 0 \leq x \leq 4\sigma^2, \\ 0, & \text{其他.} \end{cases}$$

关于极值特征值的研究, 包括其极限的存在性与对应的极限分布, 如文献[3-4]研究了样本协方差矩阵 $W = \frac{1}{N} X^T X$ 的最大特征值 $\lambda_1(W)$ 的极限性质, 证明 $\lim_{N \rightarrow \infty} \lambda_1(W) = 4\sigma^2$, a. s. 当且仅当 X_{ij} 期望为 0 且四阶矩有限.

随着深度神经网络和深度学习等领域的深入应用, 随机乘积矩阵特征值的研究得到了广泛关注. 假设

收稿日期: 2024-10-14; **修回日期:** 2024-11-18.

基金项目: 河南省高等学校重点科研项目(22B110005).

作者简介: 程慧慧(1983—), 女, 河南清丰人, 华北水利水电大学副教授, 博士, 研究方向为概率论, E-mail: chenghuihui@ncwu.edu.cn.

通信作者: 宋敏杰(2000—), 女, 河南太康人, 华北水利水电大学硕士研究生, 研究方向为概率论, E-mail: songminjie0808@163.com.

引用本文: 程慧慧, 宋敏杰. 随机马尔可夫乘积矩阵的极限经验谱分布[J]. 河南师范大学学报(自然科学版), 2025, 53(6): 52-57. (Cheng Huihui, Song Minjie. Limit of empirical spectral distribution for random Markov product matrices [J]. Journal of Henan Normal University(Natural Science Edition), 2025, 53(6): 52-57. DOI: 10.16366/j.cnki.1000-2367.2024.10.14.0002.)

$\mathbf{X}^{(k)} = (\mathbf{X}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$ 是一列独立同分布的复随机矩阵, 令 $\mathbf{W}_m = \mathbf{X}^{(m)} \cdots \mathbf{X}^{(2)} \mathbf{X}^{(1)}, \mathbf{V}_m = N^{-m} (\mathbf{W}_m)^* (\mathbf{W}_m)$, 则 $\mathbf{V}_m$ 是对称半正定矩阵. 对于有限正整数 m , 文献[5-6] 证明了当 $\mathbf{X}_{11}^{(k)}$ 均值为 0, 方差为 1, 有 $v_{\mathbf{V}_m} \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}, \text{a. s.}$, 其中 $\mathcal{F}_{\text{FC}}$ 为 Fuss-Catalan 分布^[7], 其密度函数为

$$\rho_m(x) = \frac{1}{\sqrt{2\pi}} \frac{m^{m-\frac{3}{2}}}{(m+1)^{m+\frac{1}{2}}} G_{m,m}^{m,0} \left(\begin{matrix} \frac{1}{m}, 0, -\frac{1}{m}, \dots, -\frac{m-2}{m} \\ -\frac{1}{m+1}, -\frac{2}{m+1}, \dots, -\frac{m}{m+1} \end{matrix} \middle| \frac{m^m}{(m+1)^{m+1}} x \right).$$

这里 $G_{m,m}^{m,0}(\cdot | \cdot)$ 是 Meijer G-函数^[8], 并且 $\rho_m(x)$ 的支撑集为 $[0, \frac{(m+1)^{m+1}}{m^m}]$ ^[9]. 其 p 阶矩 $\int_0^\infty x^p \rho_m(x) dx = \frac{1}{m^p + 1} \binom{(m+1)p}{p}$ 对应于 Fuss-Catalan 数^[10].

关于随机马尔可夫矩阵, 记 $\mathbf{M}^{(k)} = (\mathbf{X}_{ij}^{(k)} / \sum_{j=1}^N \mathbf{X}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$. 当 $m=1$ 时, CHAFAÏ^[11] 证明了对独立同分布的指数随机变量 $\{\mathbf{X}_{ij}^{(1)}, 1 \leq i, j \leq N\}$, $N\mathbf{M}^{(1)\text{T}} \mathbf{M}^{(1)}$ 的经验谱分布几乎必然弱收敛为 $\mathcal{F}_{\text{MP}}$. BORDENAVE 等^[12] 进一步推广证明了当 $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N\}$ 为一般的非负随机变量时, $N\mathbf{M}^{(1)\text{T}} \mathbf{M}^{(1)}$ 的经验谱分布依然几乎必然弱收敛为 $\mathcal{F}_{\text{MP}}$. 然而, 本文证明了当 $m > 1$ 时, $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ 是一列独立同分布的非负随机变量, 均值为 1, 有限方差 $\sigma^2 > 0$ 时, $N^m \sigma^{-2m} (\mathbf{M}^{(m)} \cdots \mathbf{M}^{(2)} \mathbf{M}^{(1)})^{\text{T}} (\mathbf{M}^{(m)} \cdots \mathbf{M}^{(2)} \mathbf{M}^{(1)})$ 的经验谱分布几乎必然弱收敛为 Fuss-Catalan 分布.

1 预备知识

假设 $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ 是独立同分布的非负实随机变量且满足

$$E(\mathbf{X}_{11}^{(1)}) = 1, \text{Var}(\mathbf{X}_{11}^{(1)}) = \sigma^2 (0 < \sigma^2 < \infty).$$

定义 $E_N^{(k)} = \{d_{1N}^{(k)} d_{2N}^{(k)} \cdots d_{NN}^{(k)} > 0\}$, 其中 $d_{iN}^{(k)} = \mathbf{X}_{i1}^{(k)} + \mathbf{X}_{i2}^{(k)} + \cdots + \mathbf{X}_{iN}^{(k)}$.

记 $q = P(\mathbf{X}_{11}^{(1)} = 0)$. 由于 $\mathbf{X}_{11}^{(1)}$ 非负且 $E(\mathbf{X}_{11}^{(1)}) = 1$, 故 $q < 1$. 从而有 $\sum_{N=1}^\infty P\{(E_N^{(k)})^c\} = \sum_{N=1}^\infty (1 - (1 - q^N)^N) \leq \sum_{N=1}^\infty Nq^N < \infty$. 根据 Borel-Cantelli 引理, $P((E_N^{(k)})^c \text{ i. o.}) = 0, P(\liminf E_N^{(k)}) = 1$, 当 $N \gg 1$ 时, $d_{1N}^{(k)} d_{2N}^{(k)} \cdots d_{NN}^{(k)} > 0, \text{a. s.}$

定义矩阵 $\mathbf{M}^{(k)} = (\mathbf{M}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$, 这里 $\mathbf{M}_{ij}^{(k)} = \frac{\mathbf{X}_{ij}^{(k)}}{d_{iN}^{(k)}}$. 矩阵 $\mathbf{M}^{(k)}$ 被称为随机马尔可夫矩阵, 其中矩

阵元素满足 $0 \leq \mathbf{M}_{ij}^{(k)} \leq 1$, 行和 $\sum_{j=1}^N \mathbf{M}_{ij}^{(k)} = 1$. 此外, 记 $\mathbf{M}^{(k)} = \mathbf{D}^{(k)} \mathbf{X}^{(k)}$, 这里 $\mathbf{X}^{(k)} = (\mathbf{X}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$,

$\mathbf{D}^{(k)} = \text{diag}(\mathbf{D}_{11}^{(k)}, \mathbf{D}_{22}^{(k)}, \dots, \mathbf{D}_{NN}^{(k)})$ 是对角矩阵, 对角线元素 $\mathbf{D}_{ii}^{(k)} = \frac{1}{d_{iN}^{(k)}}$. 规定当 $N \gg 1$ 时, 若 $d_{iN}^{(k)} = 0$, 则令 $\mathbf{D}_{ii}^{(k)} =$

$1, \mathbf{M}_{ii}^{(k)} = 1, \mathbf{M}_{ij}^{(k)} = 0 (i \neq j)$. 根据定义可知矩阵 $\{\mathbf{M}^{(k)}, 1 \leq k \leq m\}$ 独立同分布, 且矩阵 $\mathbf{M}^{(k)}$ 的所有行向量之间相互独立且同分布.

2 主要结论及证明

本文的主要结论证明了有限多个随机马尔可夫乘积矩阵的极限经验谱分布, 主要结果如下:

定理 1 设 $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ 是独立同分布的非负实随机变量且满足

$$E(\mathbf{X}_{11}^{(1)}) = 1, \text{Var}(\mathbf{X}_{11}^{(1)}) = \sigma^2 (0 < \sigma^2 < \infty).$$

令 $\Pi_m = \sqrt{N}^{-m} \mathbf{M}^{(m)} \cdots \mathbf{M}^{(2)} \mathbf{M}^{(1)}, \{\lambda_i, 1 \leq i \leq N\}$ 是 $\Pi_m^{\text{T}} \Pi_m$ 的特征值, 则当 $N \rightarrow \infty$ 时, $\sigma^{-2m} \Pi_m^{\text{T}} \Pi_m$ 的经验谱分布函数

$$\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}(\sigma^{-2m} \Pi_m^T \Pi_m) \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}, \text{ a. s.} \quad (1)$$

注 1 参考文献 [5–6], 对于一般的情形, 若 $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ 是相互独立的非负实随机变量, 满足 $E(\mathbf{X}_{11}^{(1)}) = 1, \text{Var}(\mathbf{X}_{11}^{(1)}) = \sigma^2 (0 < \sigma^2 < \infty)$, 并且 Lindeberg 条件成立, 即对于任意 $\tau > 0$, 当 $N \rightarrow \infty$ 时, $L_N(\tau) := \max_{1 \leq k \leq m} \frac{1}{N^2} \sum_{i,j=1}^N E |\mathbf{X}_{ij}^{(k)}|^2 I_{\{|\mathbf{X}_{ij}^{(k)}| \geq \tau\sqrt{N}\}} \rightarrow 0$, 则式(1) 成立.

为证明定理 1, 下面给出一些已有的引理结论.

引理 1(一致大数定律) 若 $\{\mathbf{X}_{ij}, 1 \leq i, j < \infty\}$ 是一系列独立同分布的随机变量, 期望为 μ , 方差有限, 记 $S_{iN} = \sum_{j=1}^N \mathbf{X}_{ij}$, 则当 $N \rightarrow \infty$, $\max_{1 \leq i \leq N} \left| \frac{S_{iN}}{N} - \mu \right| \xrightarrow{\text{a. s.}} 0$, 并且, 若 $\mu \neq 0$, 则有 $\max_{1 \leq i \leq N} \left| \frac{N}{S_{iN}} - \frac{1}{\mu} \right| \xrightarrow{\text{a. s.}} 0$.

此引理是文献 [13] 中引理 2 的一个推论. 取 $\alpha = \beta = \mathbf{M} = 1$, 即可得到该结论.

引理 2(参见文献 [2], 定理 A.44) 若 $\mathbf{A}$ 和 $\mathbf{B}$ 是两个 N 阶复矩阵, 则有 $\|v_{\mathbf{A}\mathbf{A}^*} - v_{\mathbf{B}\mathbf{B}^*}\| \leq \frac{1}{N} r(\mathbf{A} - \mathbf{B})$, 其中, $\|f\| = \sup_x |f(x)|$, $r(\mathbf{A})$ 为矩阵 $\mathbf{A}$ 的秩.

该引理阐述了经验谱分布函数 $v_{\mathbf{A}\mathbf{A}^*}$ 和 $v_{\mathbf{B}\mathbf{B}^*}$ 与矩阵秩 $r(\mathbf{A} - \mathbf{B})$ 之间的关系, 当 $N \rightarrow \infty$ 时, 若 $r(\mathbf{A} - \mathbf{B})$ 的大小有限, 则 $\lim_{N \rightarrow \infty} \|v_{\mathbf{A}\mathbf{A}^*} - v_{\mathbf{B}\mathbf{B}^*}\| = 0$, 即 $\mathbf{A}\mathbf{A}^*$ 与 $\mathbf{B}\mathbf{B}^*$ 有相同的极限经验谱分布, 因此可将对前者极限经验谱分布的研究转化为对后者的研究. 这种转化思路在本文主要定理的证明过程中起到了关键作用.

引理 3(奇异值不等式) 若 $\mathbf{A}$ 和 $\mathbf{B}$ 是两个 N 阶复矩阵, 则

$$s_1(\mathbf{A}\mathbf{B}) \leq s_1(\mathbf{A})s_1(\mathbf{B}), s_N(\mathbf{A}\mathbf{B}) \geq s_N(\mathbf{A})s_N(\mathbf{B}).$$

其中 $s_i(\mathbf{A}) = \sqrt{\lambda_i(\mathbf{A}\mathbf{A}^*)}$ ($1 \leq i \leq N$) 为 $\mathbf{A}$ 的奇异值, 并且满足 $s_1(\mathbf{A}) \geq s_2(\mathbf{A}) \geq \cdots \geq s_N(\mathbf{A})$. 此外, 若 $\mathbf{A}$ 是对角矩阵, 则对任意 $1 \leq i \leq N$, 有

$$s_N(\mathbf{A})s_i(\mathbf{B}) \leq s_i(\mathbf{A}\mathbf{B}) \leq s_1(\mathbf{A})s_i(\mathbf{B}).$$

引理 4(参见文献 [12], 引理 C.1) 若 $\{a_{kN}\}_{1 \leq k \leq N}$ 和 $\{b_{kN}\}_{1 \leq k \leq N}$ 为两个复三角序列, 令 v 是 $\mathbb{C}$ 上的概率测度. 若 $\frac{1}{N} \sum_{N=1}^{\infty} \delta_{a_{kN}} \xrightarrow{\ell_b} v$ 且 $\lim_{N \rightarrow \infty} \max_{1 \leq k \leq N} |a_{kN} - b_{kN}| = 0$, 则有 $\frac{1}{N} \sum_{N=1}^{\infty} \delta_{b_{kN}} \xrightarrow{\ell_b} v$.

通过构造 $\{a_{kN}\}_{1 \leq k \leq N}$ 和 $\{b_{kN}\}_{1 \leq k \leq N}$ 的特征函数序列, 根据 Paul Lévy 特征函数准则可证明该结果成立.

定理 1 的证明 因为 $\sqrt{N}\mathbf{M}^{(k)} = N\mathbf{D}^{(k)} N^{-\frac{1}{2}}\mathbf{X}^{(k)}$, 根据引理 1 可以得到

$$\lim_{N \rightarrow \infty} \max_{1 \leq k \leq N} |N\mathbf{D}_{i,i}^{(k)} - 1| = 0, \text{ a. s.} \quad \lim_{N \rightarrow \infty} \max_{1 \leq k \leq N} \left| \frac{1}{N\mathbf{D}_{i,i}^{(k)}} - 1 \right| = 0, \text{ a. s.}$$

故可以得到

$$s_1(N\mathbf{D}^{(k)}) = \max_{1 \leq i \leq N} |N\mathbf{D}_{i,i}^{(k)}| = 1 + o(1), \text{ a. s.}, s_N(N\mathbf{D}^{(k)}) = \min_{1 \leq i \leq N} |N\mathbf{D}_{i,i}^{(k)}| = 1 + o(1), \text{ a. s.} \quad (2)$$

为简便表述, 引入下面记号:

$$\bar{\mathbf{X}}^{(k)} = \mathbf{X}^{(k)} - E(\mathbf{X}^{(k)}), Y_0 = N^{\frac{m}{2}} \mathbf{D}^{(m)} \mathbf{X}^{(m)} \cdots \mathbf{D}^{(2)} \mathbf{X}^{(2)} \mathbf{D}^{(1)} \mathbf{X}^{(1)},$$

$$Y_1 = N^{\frac{m}{2}} \mathbf{D}^{(m)} \mathbf{X}^{(m)} \cdots \mathbf{D}^{(2)} \mathbf{X}^{(2)} \mathbf{D}^{(1)} \bar{\mathbf{X}}^{(1)}, \dots, Y_m = N^{\frac{m}{2}} \mathbf{D}^{(m)} \bar{\mathbf{X}}^{(m)} \cdots \mathbf{D}^{(2)} \bar{\mathbf{X}}^{(2)} \mathbf{D}^{(1)} \bar{\mathbf{X}}^{(1)}.$$

当 $1 \leq k \leq m$ 时, 对于任意 k , 有

$$Y_{k-1} - Y_k = N^{\frac{m}{2}} \mathbf{D}^{(m)} \mathbf{X}^{(m)} \cdots \mathbf{D}^{(k+1)} \mathbf{X}^{(k+1)} [\mathbf{D}^{(k)} (\mathbf{X}^{(k)} - \bar{\mathbf{X}}^{(k)})] \mathbf{D}^{k-1} \bar{\mathbf{X}}^{k-1} \cdots \mathbf{D}^{(1)} \bar{\mathbf{X}}^{(1)}.$$

对于任意 $1 \leq k \leq m$, 根据矩阵秩不等式 $r(\mathbf{A}\mathbf{B}) \leq \min\{r(\mathbf{A}), r(\mathbf{B})\}$, 以及 $E(\mathbf{X}^{(k)})$ 秩为 1, 可知 $r(Y_{k-1} - Y_k) \leq 1$.

由引理 2 可得

$$\begin{aligned} \|v_{\Pi_m^T \Pi_m} - v_{Y_m^T Y_m}\| &= \|v_{Y_0^T Y_0} - v_{Y_m^T Y_m}\| \leq \|v_{Y_0^T Y_0} - v_{Y_1^T Y_1}\| + \|v_{Y_1^T Y_1} - v_{Y_2^T Y_2}\| + \cdots + \\ &\quad \|v_{Y_{m-1}^T Y_{m-1}} - v_{Y_m^T Y_m}\| \leq \frac{m}{N} \rightarrow 0 (N \rightarrow \infty). \end{aligned}$$

这表明 $\Pi_m^T \Pi_m$ 和 $Y_m^T Y_m$ 有相同的极限经验谱分布函数,因此,只需研究 $Y_m^T Y_m$ 的极限经验谱分布函数.

当 $N \gg 1$ 时,由引理 3,可以得到

$$\begin{aligned} s_1(N^m \mathbf{D}^{(m)} \cdots \mathbf{D}^{(2)} \mathbf{D}^{(1)}) &\leq s_1(N \mathbf{D}^{(m)}) \cdots s_1(N \mathbf{D}^{(2)}) s_1(N \mathbf{D}^{(1)}), \\ s_N(N \mathbf{D}^{(m)}) \cdots s_N(N \mathbf{D}^{(2)}) s_N(N \mathbf{D}^{(1)}) &\leq s_N(N^m \mathbf{D}^{(m)} \cdots \mathbf{D}^{(2)} \mathbf{D}^{(1)}). \end{aligned} \quad (3)$$

并且有

$$\begin{aligned} s_N(N \mathbf{D}^{(m)}) \cdots s_N(N \mathbf{D}^{(2)}) s_N(N \mathbf{D}^{(1)}) s_i(N^{-\frac{m}{2}} \bar{\mathbf{X}}^{(m)} \cdots \bar{\mathbf{X}}^{(2)} \bar{\mathbf{X}}^{(1)}) &\leq s_i(N^{\frac{m}{2}} \mathbf{D}^{(m)} \bar{\mathbf{X}}^{(m)} \cdots \mathbf{D}^{(2)} \bar{\mathbf{X}}^{(2)} \mathbf{D}^{(1)} \bar{\mathbf{X}}^{(1)}) \leq \\ s_1(N \mathbf{D}^{(m)}) \cdots s_1(N \mathbf{D}^{(2)}) s_1(N \mathbf{D}^{(1)}) s_i(N^{-\frac{m}{2}} \bar{\mathbf{X}}^{(m)} \cdots \bar{\mathbf{X}}^{(2)} \bar{\mathbf{X}}^{(1)}). \end{aligned} \quad (4)$$

根据文献[14]中最小奇异值的多项式下界可知,对于任意的常数 $a, b > 0$,存在常数 $c > 0$,使得任意确定的复矩阵 $\mathbf{A}$ 满足 $s_1(\mathbf{A}) \leq N^b$,则有

$$P(s_N(\mathbf{X} + \mathbf{A}) \leq N^{-c}) \leq N^{-a}. \quad (5)$$

特别地,根据 Borel-Cantelli 引理,存在可能依赖于 b 的常数 $c > 0$,使得 $\mathbf{X} + \mathbf{A}$ 几乎必然可逆并且当 $N \gg 1$ 时, $s_N(\mathbf{X} + \mathbf{A}) \geq N^{-c}$.

由式(5),当 $N \gg 1$ 时,能得到 $s_N(N^{-\frac{1}{2}} \bar{\mathbf{X}}^{(k)}) > 0, s_N(N \mathbf{D}^{(k)}) > 0, s_N(\sqrt{N} \mathbf{M}^{(k)}) > 0$. a. s. 且当 $N \gg 1$ 时,由式(3)可得

$$s_N(N^m \mathbf{D}^{(m)} \cdots \mathbf{D}^{(2)} \mathbf{D}^{(1)}) > 0, s_N(N^{-\frac{m}{2}} \bar{\mathbf{X}}^{(m)} \cdots \bar{\mathbf{X}}^{(2)} \bar{\mathbf{X}}^{(1)}) > 0. \text{ a. s.} \quad (6)$$

故根据引理 3,有

$$s_N(Y_m) = s_N(N^{\frac{m}{2}} \mathbf{D}^{(m)} \bar{\mathbf{X}}^{(m)} \cdots \mathbf{D}^{(2)} \bar{\mathbf{X}}^{(2)} \mathbf{D}^{(1)} \bar{\mathbf{X}}^{(1)}) > 0. \text{ a. s.} \quad (7)$$

记 $T_m = N^{-\frac{m}{2}} \bar{\mathbf{X}}^{(m)} \cdots \bar{\mathbf{X}}^{(2)} \bar{\mathbf{X}}^{(1)}$, 由引理 3、式(2, 4, 6-7)可以得到 $\max_{1 \leq i \leq N} |\ln(s_i(Y_m)) - \ln(s_i(T_m))| = o(1)$. a. s. 因此有 $\max_{1 \leq i \leq N} |\ln(\lambda_i(Y_m^T Y_m)) - \ln(\lambda_i(T_m^T T_m))| = o(1)$. a. s. 由于 $0 < \sigma^2 < \infty$, 故

$$\max_{1 \leq i \leq N} |\ln(\lambda_i(\sigma^{-2m} Y_m^T Y_m)) - \ln(\lambda_i(\sigma^{-2m} T_m^T T_m))| = o(1). \text{ a. s.} \quad (8)$$

因为 $E(\bar{\mathbf{X}}^{(k)}) = E(\mathbf{X}^{(k)} - E(\mathbf{X}^{(k)})) = 0$, 且根据 V_m 的极限经验谱分布,可得: $v_{\sigma^{-2m} T_m^T T_m} \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}$. a. s.

已知分布函数 $\mathcal{F}_{\text{FC}}$ 的密度函数 $\rho_m(x)$ 的支撑集为 $[0, \frac{(m+1)^{m+1}}{m^m}]$, 故可得:

$$\frac{1}{N} \sum_{i=1}^N \delta_{\ln(\lambda_i(\sigma^{-2m} T_m^T T_m))} \xrightarrow{\ell_b} \ln(\mathcal{F}_{\text{FC}}). \text{ a. s.}$$

根据式(8)和引理 4,可得 $\frac{1}{N} \sum_{i=1}^N \delta_{\ln(\lambda_i(\sigma^{-2m} Y_m^T Y_m))} \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}$. a. s.

因此,进行变量代换 $x \mapsto e^x$, 有 $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(\sigma^{-2m} Y_m^T Y_m)} \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}$. a. s. 故 $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(\sigma^{-2m} \Pi_m^T \Pi_m)} \xrightarrow{\ell_b} \mathcal{F}_{\text{FC}}$. a. s. 即 $\sigma^{-2m} \Pi_m^T \Pi_m$ 的经验谱分布函数几乎必然弱收敛到 Fuss-Catalan 分布. 证毕.

3 数值模拟

为了进一步验证理论结果,用 Mathematica 软件进行如下模拟:首先模拟生成一系列随机数矩阵 $\mathbf{X}^{(k)} = (\mathbf{X}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$;其次,构造马尔可夫矩阵 $\{\mathbf{M}^{(k)}, 1 \leq k \leq m\}$;接着计算 $\Pi_m^T \Pi_m / \sigma^{2m}$ 的特征值,并模拟特征值的概率密度函数图像;最后,将模拟得到的概率密度函数直方图与对应分布函数 $\mathcal{F}_{\text{FC}}$ 的密度函数 $\rho_m(x)$ 进行比较.在整个数值模拟过程中,均不涉及量纲,且令样本量 $N = 400$,图中的纵坐标表示单位长度上的频数,组距为 0.2.

当 $m=1$ 时,假设 $\{\mathbf{X}_{ij}^{(1)}, 1 \leq i, j \leq N\}$ 是独立同分布的指数随机变量,期望 $E(\mathbf{X}_{11}^{(1)}) = 1$,则方差 $\sigma^2 = 1$. 图 1 模拟 $\Pi_1^T \Pi_1$ 特征值的概率密度函数图像.由模拟的结果可知,模拟得到的概率密度结果和理论密度函数曲线 $\rho_1(x)$ 吻合.若 $\{\tilde{\mathbf{X}}_{ij}^{(1)}, 1 \leq i, j \leq N\}$ 是区间 $[0, 2]$ 上独立同分布的均匀随机变量,期望 $E(\tilde{\mathbf{X}}_{11}^{(1)}) = 1$,则

方差 $\sigma^2=1/3$.图2为模拟 $3\tilde{\Pi}_1^T\tilde{\Pi}_1$ 的特征值的概率密度函数图像.同样地,通过观察模拟结果可知,模拟结果与理论上的概率密度函数 $\rho_1(x)$ 吻合.

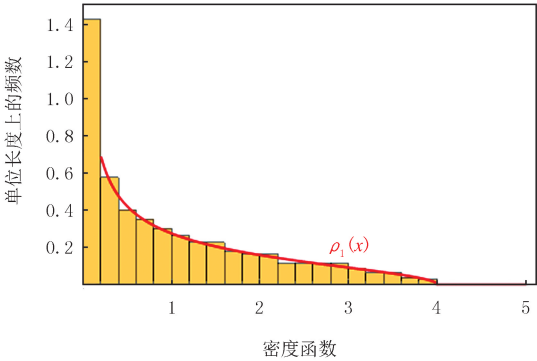


图1 $\Pi_1^T \Pi_1$ 的概率密度函数

Fig.1 Probability density function for $\Pi_1^T \Pi_1$

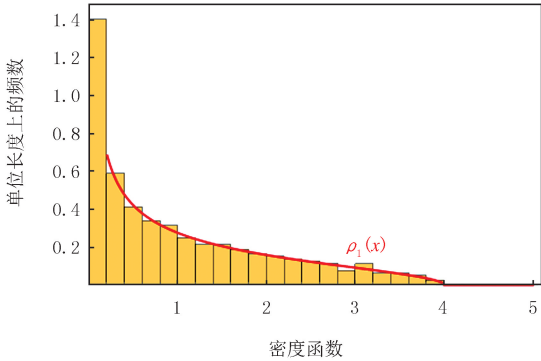


图2 $3\tilde{\Pi}_1^T \tilde{\Pi}_1$ 的概率密度函数

Fig.2 Probability density function for $3\tilde{\Pi}_1^T \tilde{\Pi}_1$

当 $m=2$ 时,假设 $\{\mathbf{X}_{ij}^{(1)}, \mathbf{X}_{ij}^{(2)}, 1 \leq i, j \leq N\}$ 是独立同分布的指数随机变量,期望 $E(\mathbf{X}_{11}^{(1)})=1$,方差 $\sigma^2=1$.图3和图4中分别模拟 $\Pi_2^T \Pi_2$ 和 $3(\Pi_1 \tilde{\Pi}_1)^T \Pi_1 \tilde{\Pi}_1$ 特征值的概率密度函数图像.根据模拟结果可知,模拟结果和理论结果保持一致.

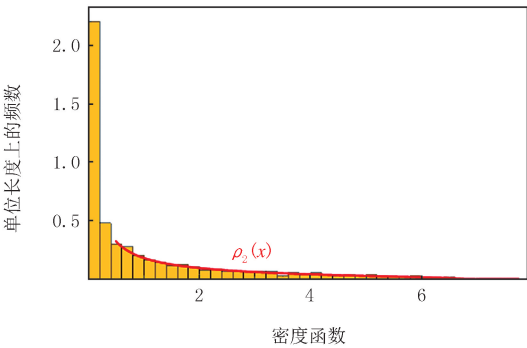


图3 $\Pi_2^T \Pi_2$ 的概率密度函数

Fig.3 Probability density function for $\Pi_2^T \Pi_2$

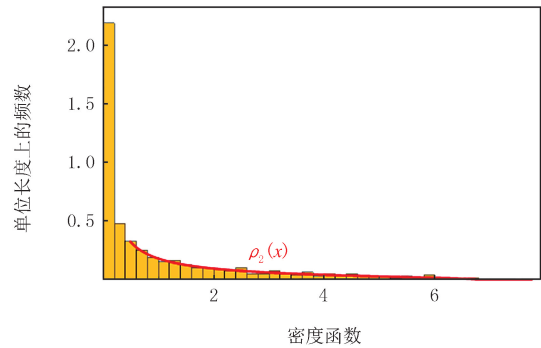


图4 $3(\Pi_1 \tilde{\Pi}_1)^T \Pi_1 \tilde{\Pi}_1$ 的概率密度函数

Fig.4 Probability density function for $3(\Pi_1 \tilde{\Pi}_1)^T \Pi_1 \tilde{\Pi}_1$

类似地,通过模拟有限多个乘积矩阵,可得到 $\Pi_m^T \Pi_m / \sigma^{2m}$ 的概率密度直方图与对应的理论密度函数 $\rho_m(x)$ 保持一致,并且随着样本量 N 的增大,模拟结果更加吻合 $\rho_m(x)$,从而为本文的理论结果提供了强有力的实证支持.

参 考 文 献

[1] MARČHENKO V A,PASTUR L A.Distribution of eigenvalues for some sets of random matrices[J].Mathematics of the USSR-Sbornik, 1967,1(4):457-483.

[2] BAI Z D,SILVERSTEIN J W.Spectral Analysis of Large Dimensional Random Matrices[M].Cham:Springer New York,2010.

[3] BAI Z D,SILVERSTEIN J W,YIN Y Q.A note on the largest eigenvalue of a large dimensional sample covariance matrix[J].Journal of Multivariate Analysis,1988,26(2):166-168.

[4] YIN Y Q,BAI Z D,KRISHNAIAH P R.On the limit of the largest eigenvalue of the large dimensional sample covariance matrix[J].Probability Theory and Related Fields,1988,78(4):509-521.

[5] ALEXEEV N V,GÖTZE F,TIKHOMIROV A N.On the singular spectrum of powers and products of random matrices[J].Doklady Mathematics,2010,82(1):505-507.

[6] ALEXEEV N V,GÖTZE F,TIKHOMIROV A N.On the asymptotic distribution of singular values of powers of random matrices[J].Journal of Mathematical Sciences,2014,199(2):68-87.

[7] MŁOTKOWSKI W,SAKUMA N,UEDA Y.Free self-decomposability and unimodality of the fuss-Catalan distributions[J].Journal of

Statistical Physics,2020,178(5):1055-1075.

[8] BEALS R,SZMIGIELSKI J.Meijer G-functions;a gentle introduction[J].Notices of the American Mathematical Society,2013,60(6):866.

[9] LIU D Z,WANG D,ZHANG L.Bulk and soft-edge universality for singular values of products of Ginibre random matrices[J].Annales de L’Institut Henri Poincaré,Probabilités et Statistiques,2016,52(4):1734-1762.

[10] DESHPANDE P,MENON K,SARKAR W.Refine-ments of the braid arrangement and two-parameter Fuss-Catalan numbers[J].Journal of Algebraic Combinatorics,2023,57(3):687-707.

[11] CHAFAÏ D.The dirichlet Markov ensemble[J].Journal of Multivariate Analysis,2010,101(3):555-567.

[12] BORDENAVE C,CAPUTO P,CHAFAÏ D.Circular law theorem for random Markov matrices[J].Probability Theory and Related Fields,2012,152(3):751-779.

[13] BAI Z D,YIN Y Q.Limit of the smallest eigenvalue of a large dimensional sample covariance matrix[J].The Annals of Probability,1993,21(3):1275-1294.

[14] TAO T,VU V.Random matrices;the circular law[J].Communications in Contemporary Mathematics,2008,10(2):261-307.

Limit of empirical spectral distribution for random Markov product matrices

Cheng Huihui, Song Minjie

(School of Mathematics and Statistics, North China University of Water Resources and Electric power, Zhengzhou 450046, China)

Abstract: Let $\{\mathbf{X}_{ij}^{(k)}, 1 \leq i, j \leq N; 1 \leq k \leq m\}$ be a sequence of independent and identically distributed non-negative ran-dom variables, with mean 1, finite variance $\sigma^2 > 0$. Then $\mathbf{M}^{(k)} = (\mathbf{X}_{ij}^{(k)} / \sum_{j=1}^N \mathbf{X}_{ij}^{(k)})_{N \times N} (1 \leq k \leq m)$ is a sequence of Markov random matrices. In this paper, we prove that as N tends to infinity, $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}$, the empirical spectral distribution of $\sigma^{-2} \Pi_m^T \Pi_m$, converges weakly to the Fuss-Catalan distribution almost surely, where $\Pi_m = N^{m/2} \mathbf{M}^{(m)} \cdots \mathbf{M}^{(2)} \mathbf{M}^{(1)}$ and $\lambda_1, \lambda_2, \cdots, \lambda_N$ are the eigenvalues of the real symmetric matrix $\Pi_m^T \Pi_m$. And the simulation results obtained through numerical simulation are consistent with the theoretical conclusion.

Keywords: Markov random matrices; empirical spectral distribution; product matrices; Fuss-Catalan distribution

[责任编辑 陈留院 杨浦]