Научная статья
УДК 519.214+519.177
DOI 10.25205/1560-750X-2024-27-2-131-143

ЗАКОН МАРЧЕНКО – ПАСТУРА ДЛЯ СПЕКТРА СЛУЧАЙНОГО ВЗВЕШЕННОГО ДВУДОЛЬНОГО ГРАФА

Анастасия Васильевна Надуткина[1], **Александр Николаевич Тихомиров**[2], **Дмитрий Анатольевич Тимушев**[3]

[1,2,3]Физико-математический институт, Коми НЦ УрО РАН, 167982, Сыктывкар, Россия
[2]Университет ВШЭ, 109028, Москва, Россия

[1]nadutkina.anastasiya@mail.ru,
[2]tikhomirov@ipm.komisc.ru,https://orcid.org/0000-0001-8228-4294
[3]timushev@ipm.komisc.ru

Аннотация

Рассматривается спектр случайного взвешенного двудольного графа. Установлено, что при определенных предположениях на вероятности ребер эмпирическая спектральная функция распределения матрицы смежности графа сходится к симметризованной функции распределения Марченко-Пастура.

Ключевые слова и фразы

Случайная матрица, закон Марченко-Пастура, случайный двудольный граф.

Источник финансирования

Работа выполнена в рамках проекта Коми научного центра УрО РАН “Математические проблемы теории стохастических и детерминированных систем, включая системы высокой размерности” (проект № 122040600066-5), а также в рамках Программы фундаментальных исследований Университета ВШЭ.

Для цитирования

Надуткина А. В., Тихомиров А. Н., Тимушев Д. А. Закон Марченко-Пастура для спектра случайного взвешенного двудольного графа // Математические труды, 2024, Т. 27, № 2, С. 131-143. DOI 10.25205/1560-750X-2024-27-2-131-143

ISSN 1560-750X
Математические труды, 2024, Том 27, № 2, С.131-143
Mat. Trudy, 2024, vol. 27, no. 2, pp. 131-143

MARCHENKO – PASTUR LAW FOR THE SPECTRUM OF A RANDOM WEIGHTED BIPARTITE GRAPH

Anastasiya V. Nadutkina[1], **Alexander N. Tikhomirov**[2], **Dmitrii A. Timushev**[3]

[1,2,3]Institute of Physics and Mathematics, Komi Scientific Center of the Ural Branch of the Russian Academy of Sciences, 660041, 167982, Syktyvkar, Russia
[2]Higher School of Economics University

[1]nadutkina.anastasiya@mail.ru, [2]tikhomirov@ipm.komisc.ru, [3]timushev@ipm.komisc.ru

Abstract

This study investigates the spectra of random weighted bipartite graphs. We establish that under specific assumptions on the edge probabilities, the empirical spectral distribution function of the graph's adjacency matrix converges to the symmetrized Marchenko-Pastur distribution function.

Keywords

Random matrices, Marchenko-Pastur law, random bipartite graphs.

Funding

The work was carried out within the framework of the project of the Komi Scientific Center of the Ural Branch of the Russian Academy of Sciences "Mathematical problems in the theory of stochastic and deterministic systems, including high-dimensional systems" (project No. 122040600066-5), as well as within the framework of the Fundamental Research Program of the Higher School of Economics University.

For citation

Nadutkina A.V., Tikhomirov A.N, Timushev D.A. Marchenko-pastur law for the spectrum of a random weighted bipartite graph // Mat. Trudy, 2024, V. 27, no. 2, pp. 131-143. DOI 10.25205/1560-750X-2024-27-2-131-143

§ 1. Введение

В традиционной теории графов как правило предполагается, что графы статичны и детерминированы. Однако в большинстве реальных сетей наблюдается некоторая степень случайности. Так, социальные сети могут иметь случайные связи, основанные на случайных встречах, а в интернете случайные связи возникают в соответствии с протоколами маршрутизации. Таким образом, рассмотрение одних только лишь детерминированных графов не сможет передать всю сложность этих систем. Модели случайных графов позволяют разрабатывать алгоритмы, которые более эффективны при работе с массивными наборами данных с миллионами или даже миллиардами вершин. В итоге, если детерминированные графы ценны для фундаментальных исследований, то случайные графы необходимы для отражения присущей сетям реального мира случайности и больших масштабов.

В последние три десятилетия появился большой объем литературы, посвященной исследованию случайных графов. Теория случайных графов претерпела быстрое и значительное развитие благодаря разнообразным приложениям в различных областях — от математики и физики до финансов, социальных наук, биологии, химии, информационных сетей (см., например, [1, 2]). Основы теории случайных графов были заложены в 1950-х годах Эрдешом и Реньи [4], объединив принципы теории графов и теории вероятностей в междисциплинарную область. Случайный граф Эрдеша-Реньи — это граф с n вершинами, ребра которого выбираются независимо с вероятностью p, лежащей в диапазоне $0 < p < 1$. Эта вероятность ребра может быть либо постоянной, либо функцией, зависящей от n. Спектральная теория графов является одной из ключевых областей исследования в теории графов, фокусируясь на свойствах графа, касающихся собственных значений и собственных векторов связанных с ним матриц, таких как матрица смежности, матрица Лапласа и нормализованная матрица Лапласа. Спектры этих матриц называются спектрами графа. Понимание спектров случайных графов очень важно для понимания свойств случайных графов. Например, спектр матриц смежности и Лапласиана графа коррелирует с его связностью и количеством вхождений определенных подграфов, а также с его хроматическим числом и числом независимости. Спектр нормализованной матрицы Лапласиана связан с диффузией на графах, случайными блужданиями на графах [2].

В данной работе исследуются случайные взвешенные двудольные графы. Эти графы состоят из двух отдельных множеств с n и m вершинами, каждое из которых не имеет связей внутри одного множества, что в сумме дает $n + m$ вершин в двудольном графе. Связи между вершинами возникают случайным образом, и матрицы смежности этих графов имеют блочную структуру. Такое разделение вершин на две части позволяет отразить особенности многих реальных сетей — от социальных взаимодействий, где пользователи общаются с другими пользователями, до рекомендательных систем, сопоставляющих товары с заинтересованными покупателями. Двудольные графы представляют собой мощный инструмент для моделирования подобных процессов [3, 5]. Добавление весов к двудольному графу значительно повышает его полезность, позволяя количественно оценить силу отношений, определить приоритетность конкретных взаимодействий и найти другие подходящие применения.

В работе [6] была доказана сходимость эмпирической спектральной функции распределения матрицы смежности обобщенного случайного графа к функции распределения полукругового закона Вигнера при ряде условий на вероятности ребер графа. В настоящей работе исследуются случайные взвешенные двудольные графы, расширяя установленный в [6] результат о сходимости симметризованной эмпирической спектральной функции распределения матрицы смежности двудольного графа к симметризованной функции распределения Марченко-Пастура.

Рассмотрим неориентированный взвешенный двудольный граф $G = (V_1, V_2, E, \omega)$, где V_1 и V_2 — конечные множества вершин с $|V_1| = n$, $|V_2| = m$, а n, m — величины одного порядка. Не умаляя общности будем считать, что $m \geq n$. Обозначим матрицу смежности графа через $\mathbf{A}$. Тогда $\mathbf{A}$ имеет следующую блочную структуру:

$$\mathbf{A} = \begin{bmatrix} \mathbf{O}_n & \mathbf{W} \\ \mathbf{W}^* & \mathbf{O}_m \end{bmatrix}.$$

Здесь $\mathbf{O}_n$ и $\mathbf{O}_m$ — нулевые матрицы размеров $n \times n$ и $m \times m$, соответственно. $\mathbf{W}$ — матрица размера $n \times m$, определенная, как $\mathbf{W} = \frac{1}{\sqrt{a_n}} \Xi \circ \mathbf{X}$, где $\mathbf{X} = (X_{jk})$ — матрица размера $n \times m$ с независимыми элементами X_{jk} с нулевым средним $\mathbb{E}X_{jk} = 0$, и единичной дисперсией $\mathbb{E}X_{jk}^2 = 1$, для всех $j \in \{1, \dots, n\}$ и $k \in \{1, \dots, m\}$, а $\Xi = (\xi_{jk})$ — $n \times m$-матрица с независимыми элементами Бернулли ξ_{jk}, то есть ξ_{jk} принимает значение 1 с вероятностью $\mathbb{E}\xi_{jk} = p_{jk}$ и значение 0 с вероятностью $1 - p_{jk}$. Положительную скалярную величину a_n мы определяем как

$$a_n = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{m} p_{jk},$$

и предполагаем, что $m = m(n)$, причем $\lim_{n \to \infty} n/m =: y$, где y — некоторая константа.

Пусть $s_1 \geq s_2 \geq \cdots \geq s_n$ обозначают сингулярные значения матрицы $\mathbf{W}$, упорядоченные в невозрастающем порядке. Рассмотрим соответствующую симметризованную функцию эмпирического спектрального распределения:

$$F_n(x) := \frac{1}{2n} \sum_{k=1}^{n} \Big(\mathbb{I}\{s_k \leq x\} + \mathbb{I}\{-s_k \leq x\} \Big).$$

Здесь $\mathbb{I}\{\cdot\}$ обозначает индикаторную функцию. Кроме того, мы определим $G_y(x)$ как симметризованную функцию распределения Марченко-Пастура с плотностью:

$$g_y(x) = \frac{1}{2\pi y|x|}\sqrt{(x^2-a^2)(b^2-x^2)}\mathbb{I}\{a^2 \le x^2 \le b^2\},$$

где $a = 1-\sqrt{y}$ и $b = 1+\sqrt{y}$.

Мы будем предполагать, что вероятность прореживания p_n и моменты элементов матрицы X_{jk} удовлетворяют ряду условий:

- условие $CP(0)$:
$$a_n \to \infty, \text{ при } n \to \infty;$$
- условие $CP(1)$:
$$\lim_{n\to\infty}\frac{1}{na_n}\sum_{j=1}^{n}\sum_{k=1}^{m}\left|p_{jk}-\frac{a_n}{m}\right| = 0;$$
- условие $CX(0)$: для любого $\tau > 0$
$$L_n(\tau) := \frac{1}{na_n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}\mathbb{E}X_{jk}^2\mathbb{I}\{|X_{jk}| > \tau\sqrt{a_n}\} \to 0, \text{ при } n \to \infty.$$

Основной результат статьи представляет следующая теорема.

Теорема 1. *Предположим, что условия $CP(0)$, $CP(1)$ и $CX(0)$ выполнены. Тогда*

$$\sup_x |F_n(x) - G_y(x)| \to 0 \textit{ по вероятности, при } n \to \infty.$$

§ 2. Доказательство теоремы

Пусть $\mathbf{R_A}(z) = (\mathbf{A} - z\mathbf{I}_{n+m})^{-1}$ — резольвента матрицы $\mathbf{A}$. Дополнение Шура дает

$$\mathbf{R_A}(z) = \begin{bmatrix} z(\mathbf{W}\mathbf{W}^* - z^2\mathbf{I}_n)^{-1} & \mathbf{W}(\mathbf{W}^*\mathbf{W} - z^2\mathbf{I}_m)^{-1} \\ (\mathbf{W}^*\mathbf{W} - z^2\mathbf{I}_m)^{-1}\mathbf{W}^* & z(\mathbf{W}^*\mathbf{W} - z^2\mathbf{I}_m)^{-1} \end{bmatrix}. \tag{1}$$

Обозначим через $s_y(z)$ преобразование Стильтьеса симметризованной функции распределения Марченко-Пастура $G_y(x)$, а через $s_n(z)$ — преобразование Стильтьеса функции распределения $F_n(x)$. Преобразование Стильтьеса для $G_y(x)$ имеет вид

$$s_y(z) = \frac{-z + \frac{1-y}{z} + \sqrt{\left(z - \frac{1-y}{z}\right)^2 - 4y}}{2y},$$

и преобразование Стильтьеса для $F_n(x)$

$$s_n(z) = \int_{-\infty}^{\infty} \frac{1}{x-z} dF_n(x).$$

Пусть $j \in \{1, \ldots, n\}$. Рассмотрим σ-алгебру $\mathfrak{M}^{(j)}$, порожденную элементами $\mathbf{A}$ за исключением строки и столбца с номером j. Символом $\mathbf{A}^{(j)}$ обозначим матрицу $\mathbf{A}$, у которой удалены строка и столбец с номером j. Аналогичным образом мы будем обозначать все объекты, определенные через $\mathbf{A}^{(j)}$, например резольвентную матрицу $\mathbf{R}^{(j)}$, ESD-преобразование Стильтьеса $s_n^{(j)}$ и так далее. Символом $\mathbb{E}_j$ обозначим условное математическое ожидание относительно σ-алгебры $\mathfrak{M}_j$.

Из уравнения (1) следует, что

$$s_n(z) = \frac{1}{n} \sum_{j=1}^{n} R_{jj} = \frac{1}{n} \sum_{l=1}^{m} R_{l+n,l+n} + \frac{m-n}{nz},$$

и

$$\begin{aligned} R_{jj}(z) &= \frac{1}{-z - \frac{1}{a_n} \sum_{l,k=1}^{m} X_{jk} X_{jl} \xi_{jk} \xi_{jl} R^{(j)}_{k+n,l+n}} \\ &= -\frac{1}{z + y s_n(z) - \frac{1-y}{z}} (1 - \varepsilon_j R_{jj}(z)), \end{aligned}$$

где

$$\begin{aligned} \varepsilon_j &= \varepsilon_{j1} + \varepsilon_{j2} + \varepsilon_{j3} + \varepsilon_{j4} + \varepsilon_{j5}, \\ \varepsilon_{j1} &= -\frac{1}{a_n} \sum_{k=1}^{m} X_{jk}^2 \big(\xi_{jk} - p_{jk}\big) R^{(j)}_{k+n,k+n}, \\ \varepsilon_{j2} &= -\frac{1}{a_n} \sum_{k=1}^{m} p_{jk} \big(X_{jk}^2 - 1\big) R^{(j)}_{k+n,k+n}, \\ \varepsilon_{j3} &= -\frac{1}{a_n} \sum_{k=1}^{m} \Big(p_{jk} - \frac{a_n}{m}\Big) R^{(j)}_{k+n,k+n}, \\ \varepsilon_{j4} &= \frac{1}{m} \sum_{k=1}^{m} R_{k+n,k+n} - \frac{1}{m} \sum_{k=1}^{m} R^{(j)}_{k+n,k+n}, \\ \varepsilon_{j5} &= -\frac{1}{a_n} \sum_{l \neq k}^{m} X_{jk} X_{jl} \xi_{jk} \xi_{jl} R^{(j)}_{k+n,l+n}. \end{aligned}$$

Заметим, что имеет место следующее очевидное неравенство (см. [8], Лемма 7.6):

$$\max_{k,l}\left\{\left|R_{k+n,l+n}(z)\right|,\left|R^{(j)}_{k+n,l+n}(z)\right|\right\}\le\frac{1}{v},\quad z=u+iv\quad v\ge 0. \tag{2}$$

Суммируя $R_{jj}(z)$, получаем:

$$s_n(z)=-\frac{1}{z+ys_n(z)-\frac{1-y}{z}}(1-T_n(z)),$$

где

$$T_n(z)=\frac{1}{n}\sum_{j=1}^{n}\varepsilon_jR_{jj}=\sum_{\nu=1}^{5}T_{n\nu}(z),\quad T_{n\nu}(z)=\frac{1}{n}\sum_{j=1}^{n}\varepsilon_{j\nu}R_{jj}.$$

Поскольку преобразование Стилтьеса $s_y(z)$ симметризованного распределения Марченко-Пастура удовлетворяет уравнению

$$s_y(z)=-\frac{1}{z+ys_y(z)-\frac{1-y}{z}},$$

для доказательства сходимости $s_n(z)\to s_y(z)$ достаточно показать, что $\mathbb{E}|T_n(z)|\to 0$ равномерно по $z=u+iv$, $|u|\le u_0$, $v\ge C$. Для этого нам понадобится ряд вспомогательных утверждений.

Лемма 1. *Пусть выполнены условия $CP(0)$, $CP(1)$ и $CX(0)$. Тогда*

$$\mathbb{E}|T_{n1}|\to 0,\ \textit{при } n\to\infty.$$

Доказательство. Имеем

$$
\begin{aligned}
\mathbb{E}|T_{n1}| &\le \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\varepsilon_{j1}R_{jj}| \le \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}X_{jk}^2\big(\xi_{jk}-p_{jk}\big)R^{(j)}_{k+n,k+n}\right| \\
&\le \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}X_{jk}^2\big(\xi_{jk}-p_{jk}\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|\le\tau\sqrt{a_n}\}\right| \\
&+ \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}X_{jk}^2\big(\xi_{jk}-p_{jk}\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\}\right| \\
&\le \frac{1}{va_n}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}X_{jk}^2\big(\xi_{jk}-p_{jk}\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|\le\tau\sqrt{a_n}\}\right|^2\right)^{\frac{1}{2}} \\
&+ \frac{1}{nva_n}\sum_{j=1}^{n}\sum_{k=1}^{m}\mathbb{E}\left|X_{jk}^2\big(\xi_{jk}-p_{jk}\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\}\right| \\
&\le \frac{1}{v^2a_n}\left(\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{m}\mathbb{E}\big(\xi_{jk}-p_{jk}\big)^2\mathbb{E}X_{jk}^4\mathbb{I}\{|X_{jk}|\le\tau\sqrt{a_n}\}\right)^{\frac{1}{2}} \\
&+ \frac{1}{nv^2a_n}\sum_{j=1}^{n}\sum_{k=1}^{m}\mathbb{E}\big|\xi_{jk}-p_{jk}\big|\mathbb{E}X_{jk}^2\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\} \\
&\le \frac{\tau\sqrt{a_n}}{v^2a_n}\left(\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}(1-p_{jk})\right)^{\frac{1}{2}} \\
&+ \frac{2}{nv^2a_n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}(1-p_{jk})\mathbb{E}X_{jk}^2\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\} \\
&\le \frac{\tau+2L_n(\tau)}{v^2}.
\end{aligned}
$$

Заметим, что мы можем заменить τ в условии $CX(0)$ на убывающую последовательность τ_n, стремящуюся к нулю, такую, что

$$L_n(\tau_n)\to 0$$

при $n\to\infty$. Это завершает доказательство леммы. □

Лемма 2. *Пусть выполнены условия $CP(0)$, $CP(1)$ и $CX(0)$. Тогда*

$$\mathbb{E}|T_{n2}|\to 0,\ \text{при } n\to\infty.$$

Доказательство. Определим случайные величины $\eta_{jk}=\big(X_{jk}^2-1\big)\cdot$ $\cdot\mathbb{I}\{|X_{jk}|\le\tau\sqrt{a_n}\}$. Легко видеть, что $\mathbb{E}\eta_{jk}=-\mathbb{E}\big(X_{jk}^2-1\big)\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\}$ и

$|\eta_{jk} - \mathbb{E}\eta_{jk}| \le 2\tau^2 a_n$. Из (2) и этих соотношений получаем

$$\begin{aligned}
\mathbb{E}|T_{n2}| &\le \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\varepsilon_{j2}R_{jj}| \le \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}p_{jk}\big(X_{jk}^2-1\big)R^{(j)}_{k+n,k+n}\right| \\
&\le \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}p_{jk}\big(X_{jk}^2-1\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|\le\tau\sqrt{a_n}\}\right| \\
&+ \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}\left|\sum_{k=1}^{m}p_{jk}\big(X_{jk}^2-1\big)R^{(j)}_{k+n,k+n}\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\}\right| \\
&\le \frac{1}{nva_n}\sum_{j=1}^{n}\mathbb{E}^{\frac12}\left|\sum_{k=1}^{m}p_{jk}R^{(j)}_{k+n,k+n}\big(\eta_{jk}-\mathbb{E}\eta_{jk}\big)\right|^2 \\
&+ \frac{2}{nva_n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}\mathbb{E}\big|X_{jk}^2-1\big|\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\}\mathbb{E}\big|R^{(j)}_{k+n,k+n}\big|.
\end{aligned}$$

Применяя неравенство Коши и (2), имеем

$$\begin{aligned}
\mathbb{E}|T_{n2}| &\le \frac{1}{va_n}\left(\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}^2\mathbb{E}\big|R^{(j)}_{k+n,k+n}\big|^2\mathbb{E}\big|\eta_{jk}-\mathbb{E}\eta_{jk}\big|^2\right)^{\frac12} \\
&+ \frac{2}{nv^2a_n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}\mathbb{E}X_{jk}^2\mathbb{I}\{|X_{jk}|>\tau\sqrt{a_n}\} \\
&+ \frac{2}{nv^2a_n}\sum_{j=1}^{n}\sum_{k=1}^{m}p_{jk}\mathbf{P}\{|X_{jk}|>\tau\sqrt{a_n}\} \\
&\le \frac{C}{v^2}\left(\tau+\frac{L_n(\tau)}{\tau^2a_n}+L_n(\tau)\right).
\end{aligned}$$

Заметим, что мы можем заменить τ в условии $CX(0)$ на убывающую последовательность τ_n, стремящуюся к нулю, такую, что

$$\tau_n^2 a_n \to \infty$$

и

$$L_n(\tau_n) \to 0$$

при $n \to \infty$. Это завершает доказательство леммы. □

Лемма 3. *Предполагая, что условия $CP(0)$, $CP(1)$ и $CX(0)$ выполнены, мы имеем следующую сходимость:*

$$\mathbb{E}|T_{n3}| \to 0, \ \textit{при } n \to \infty.$$

Доказательство. Из определения T_{n3} и неравенства (2) следует, что

$$\mathbb{E}|T_{n3}| \leq \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\varepsilon_{j3}R_{jj}| \leq \frac{1}{nva_n}\sum_{j=1}^{n}\sum_{k=1}^{m}\Big|p_{jk}-\frac{a_n}{m}\Big|\mathbb{E}\Big|R^{(j)}_{k+n,k+n}\Big|$$
$$\leq \frac{1}{nv^2a_n}\sum_{j=1}^{n}\sum_{k=1}^{m}\Big|p_{jk}-\frac{a_n}{m}\Big|.$$

Объединив это неравенство и условие $CP(1)$, получаем требуемый результат. □

Лемма 4. *Пусть выполнены условия $CP(0)$, $CP(1)$ и $CX(0)$. Тогда*

$$\mathbb{E}|T_{n4}| \to 0, \text{ при } n \to \infty.$$

Доказательство. Используя лемму 3.2 из статьи [7] для оценки разности следов резольвент, получаем

$$\mathbb{E}|T_{n4}| \leq \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\varepsilon_{j4}R_{jj}| \leq \frac{1}{nmv}\sum_{j=1}^{n}\mathbb{E}\Bigg|\sum_{k=1}^{m}R_{k+n,k+n}-\sum_{k=1}^{m}R^{(j)}_{k+n,k+n}\Bigg| \leq \frac{y}{nv^2}.$$

На этом доказательство леммы завершено. □

Лемма 5. *Пусть выполнены условия $CP(0)$, $CP(1)$ и $CX(0)$. Тогда*

$$\mathbb{E}|T_{n5}| \to 0, \text{ при } n \to \infty.$$

Доказательство. Применяя неравенство треугольника, неравенство Коши, неравенство (2) и вычисляя дисперсию квадратичной формы относительно переменных $\{X_{jk}\xi_{jk}\}_{k=1}^{m}$ при фиксированных значениях матриц $\mathbf{R}^{(j)}$, мы получим

$$\mathbb{E}|T_{n5}| \leq \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\varepsilon_{j5}R_{jj}| \leq \frac{1}{a_nv}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\Bigg|\sum_{l\neq k}^{m}X_{jk}X_{jl}\xi_{jk}\xi_{jl}R^{(j)}_{k+n,l+n}\Bigg|^2\right)^{\frac{1}{2}}$$
$$\leq \frac{1}{a_nv}\left(\frac{1}{n}\sum_{j=1}^{n}\sum_{l\neq k}^{m}p_{jk}p_{jl}\mathbb{E}|R^{(j)}_{k+n,l+n}|^2\right)^{\frac{1}{2}} \leq \frac{1}{a_n^{1/2}v^2}.$$

Здесь мы используем неравенство (см. [8], Лемма 7.6)

$$\sum_{l}^{m}p_{jl}|R^{(j)}_{k+n,l+n}|^2 \leq \frac{\operatorname{Im} R^{(j)}_{k+n,k+n}}{v} \leq \frac{1}{v^2}.$$

Тем самым, лемма доказана. □

Теперь мы готовы доказать основной результат.

Доказательство теоремы 1. Во-первых, заметим, что справедливы уравнения

$$ys_y^2(z) + \left(z - \frac{1-y}{z}\right)s_y(z) + 1 = 0,$$

$$ys_n^2(z) + \left(z - \frac{1-y}{z}\right)s_n(z) + 1 = T_n(z).$$

Отсюда получим

$$\left(\left|z - \frac{1-y}{z}\right| - \frac{2y}{v}\right)\left|s_n(z) - s_y(z)\right|$$
$$\leq \left|y\big(s_n^2(z) - s_y^2(z)\big) + \left(z - \frac{1-y}{z}\right)\big(s_n(z) - s_y(z)\big)\right| = |T_n(z)|.$$

Выбирая $v_0 \geq \sqrt{2(1-y)}$, имеем

$$\big|s_n(z) - s_y(z)\big| \leq C|T_n(z)|. \tag{3}$$

Так как $\mathbb{E}|T_n(z)| \leq \sum_{\nu=1}^{5}\mathbb{E}|T_{n\nu}(z)|$, то из лемм 1–5 получим $\mathbb{E}|T_n(z)| \to 0$. Отсюда следует, что

$$\mathbb{E}\big|s_n(z) - s_y(z)\big| \to 0 \text{ при } n \to \infty$$

равномерно по $z = u + iv$, $|u| \leq u_0$, $v \geq v_0$. Поскольку $s_n(z)$ — аналитическая функция и $|s_n(z)| \leq \frac{1}{v_0}$ для любого $z = u + iv$ и $v \geq v_0$, то из теоремы Монтеля следует, что семейство локально ограниченных аналитических функций $\{s_n(z)\}$ является предкомпактным подмножеством. Отсюда следует сходимость по вероятности

$$s_n(z) \to s_y(z) \text{ при } n \to \infty$$

для $z = u + iv$, $v \geq v_0$. Наконец, из сходимости преобразований Стильтьеса $s_n(z)$ следует сходимость симметризованных эмпирических функций распределения $F_n(x)$ к симметризованной функции распределения Марченко-Пастура по вероятности. Теорема 1 доказана. □

Список литературы

1. *Kolaczyk E. D. Statistical Analysis of Network Data.* New York: Springer, 2009.

2. *Newman M. Networks: An Introduction.* Oxford: OUP, 2010.

3. *Kunegis J., De Luca E. W. and Albayrak S.* The Link Prediction Problem in Bipartite Networks // *Proceedings of the Computational Intelligence for Knowledge-based Systems Design, and 13th International Conference on Information Processing and Management of Uncertainty.* 2010. P. 380-389.

4. *Erdős P. and Rényi A.* On Random Graphs I // *Publ. Math. Debrecen.* 1959. V. 6. P. 290-297.

5. *Pavlopoulos G. A., Kontou P. I., Pavlopoulou A., Bouyioukos C., Markou E., Bagos P. G.* Bipartite graphs in systems biology and medicine: a survey of methods and applications // *GigaScience.* 2018. V. 7, N 4. https://doi.org/10.1093/gigascience/giy014 .

6. *Tikhomirov A. N.* On the Wigner Law for Generalizided Random Graphs // *Sib. Adv. Math.* 2021. V. 31. P. 301-308.

7. *Götze F., Tikhomirov A. N.* The Rate of Convergence of Spectra of Sample Covariance Matrices // *Theory Probab. Appl.* 2010. V. 54, N 1. P. 129-140.

8. *Götze F., Tikhomirov A. N.* Optimal bounds for convergence of expected spectral distributions to the semi-circular law // *Probab. Theory Relat. Fields.* 2016. V. 165, N 1. P. 163-233.

Информация об авторах

Анастасия Васильевна Надуткина, аспирант

Александр Николаевич Тихомиров, доктор физико - математических наук, профессор
SPIN 3301-7012, AuthorID: 3467
Scopus Author ID 7103304377

Дмитрий Анатольевич Тимушев, кандидат физико - математических наук, старший научный сотрудник
SPIN 8132-1545, AuthorID: 124353
Scopus Author ID 16204078200

Author Information

Anastasiya V. Nadutkina, graduate student

Alexander N. Tikhomirov, Doctor of Physical and Mathematical Sciences, Professor
SPIN 3301-7012, AuthorID: 3467
Scopus Author ID 7103304377

ISSN 1560-750X
Математические труды, 2024, Том 27, № 2, С. 131-143
Mat. Trudy, 2024, vol. 27, no. 2, pp. 131-143

Dmitrii A. Timushev, Candidate of Physical and Mathematical Sciences, senior researcher
SPIN 8132-1545, AuthorID: 124353
Scopus Author ID 16204078200

Статья поступила в редакцию 04.06.2024;
одобрена после рецензирования 06.06.2024; принята к публикации 13.06.2024

The article was submitted 04.06.2024;
approved after reviewing 06.06.2024; accepted for publication 13.06.2024