НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ

Факультет математики

Специальный курс

СЛУЧАЙНЫЕ МАТРИЦЫ, СЛУЧАЙНЫЕ ПРОЦЕССЫ И ИНТЕГРИРУЕМЫЕ МОДЕЛИ

Лекции:
Поволоцкий Александр Маркович

Москва
2023

Оглавление

Введение

В данном курсе обсуждается круг вопросов, связанный с образованием предельных форм и наличием случайных отклонений, описываемых универсальными вероятностными законами. Ситуации, в которых сложные многокомпонентные системы при огрубленном рассмотрении ведут себя простым универсальным образом, не зависящим от деталей исходной системы, хорошо знакомы как математикам, так и физикам. В теории вероятности они описываются утверждениями, известными как Закон Больших Чисел (ЗБЧ) и Центральная Предельная Теорема (ЦПТ). Однако стандартные примеры таких утверждений, содержащиеся в учебниках, относятся к суммам большого числа независимых случайных величин, и хотя класс подобных задач достаточно широк, требование независимости ограничивает область приложений получаемых результатов. Естественно, возникает вопрос, можно ли сформулировать столь же содержательные и универсальные утверждения о каких-либо случайных системах с большим числом степеней свободы, выходящие за рамки невзаимодействующего мира.

В последние годы было найдено множество замечательных связей между, на первый взгляд, совершенно различными задачами математики, а также математической и теоретической физики. С математической стороны это комбинаторные и вероятностные задачи о системах с большим числом степеней свободы. Среди них задача описания собственных значений матриц со случайными элементами, задачи о статистике случайных диаграмм Юнга, задачи о замощении различных областей плоскости доминошками или ромбиками, задачи о перечислении непересекающихся путей на решётках. С физической стороны это задачи статистической физики о распространении границ разделов между различными средами, потоках взаимодействующих частиц, полимерах в неупорядоченных средах и т.д.

Курс посвящён возникновению универсального поведения в различных моделях многокомпонентных случайных систем. С формальной точки зрения мы рассмотрим широкий класс многомерных случайных величин и случайных процессов со взаимодействием, заданным простыми локальными правилами. Оказывается, что в правильно выбранных, естественных единицах измерения такие системы обнаруживают удивительно схожее поведение. А именно, если число степеней свободы в рассматриваемых системах велико, то возникающие при соответствующем перемасштабировании наборы случайных величин демонстрируют «свойство концентрации» к некоторым неслучайным детерминистическим предельным формам. Случайные отклонения от этих предельных форм, «флуктуации», измеренные в характерном «флуктуационном» масштабе, описываются небольшим набором универсальных вероятностных распределений, функциональный вид которых зависит только от глобальных симметрий системы, и не зависит от «микроподробностей» исходной задачи.

Вывод точного, явного вида универсальных распределений — это задача, которую можно решить аналитически лишь для систем, обладающих специальной математической структурой, которая обеспечивает наличие большого количества симметрий и которая оказывается тесно связанной с понятием интегрируемости. Эта структура стоит за множеством красивых точных математических результатов, которые в скейлинговом пределе приводят утверждениям, аналогичным ЗБЧ и ЦПТ. Хотя курс не следует какому либо учебному пособию, большую часть материала можно найти в книгах и курсах [1, 2, 3, 4, 5, 6, 7], приведенных в списке литературы ниже.

Благодарности

Автор благодарит Хайдара Нурлигареева за составление конспектов лекций и подбор упражнений, которые легли в основу этого текста.

Литература

[1] Anderson G. W., Guionnet A., Zeitouni O. An introduction to random matrices. No. 118. — Cambridge university press, 2010.

[2] Bai Z., Silverstein J. W. Spectral analysis of large dimensional random matrices. — Springer, 2010. — Vol. 20.

[3] Mingo J. A., Speicher R. Free probability and random matrices. — Springer, 2017. — Vol. 35.

[4] Speicher R. Lecture Notes on"Free Probability Theory-// arXiv preprint arXiv:1908.08125. — 2019.

[5] Мета М. Л. Случайные матрицы // М.: МЦНМО. — 2012.

[6] Forrester P. J. Log-gases and random matrices (LMS-34). — Princeton University Press, 2010.

[7] Borodin A., Gorin V. Lectures on integrable probability // Probability and statistical physics in St. Petersburg. — 2016. — Vol. 91. — P. 155–214.

Глава 1

Необходимые сведения из теории вероятностей

В этом разделе мы напоминаем основные понятия теории вероятности, которые будут служить языком нашего изложения, а также приведем доказательства нескольких утверждений, которые понадобятся нам в дальнейшем.

1.1 Вероятностное пространство, случайные величины и их характеристики

Любой математический текст, посвященный обсуждению вероятностных задач, начинается с постулирования вероятностного пространства. Вероятностное пространство есть тройка $(\Omega, \mathcal{F}, \mathbb{P})$, включающая

- пространство элементарных исходов Ω,
- сигма-алгебру его подмножеств $\mathcal{F} \subset 2^{\Omega}$
- и вероятностную меру $\mathbb{P} : \mathcal{F} \to [0, 1]$, ставящую в соответствие каждому множеству из $\mathcal{F}$ число от нуля до единицы.

На практике обычно изучается не само вероятностное пространство, а случайные величины, т.е. измеримые функции на нем

$$\xi : \Omega \to \mathfrak{X},$$

принимающие значения в некотором пространстве $\mathfrak{X}$, наделенном топологией и, соответственно, сигма-алгеброй борелевских множеств. В общем случае мы обычно будем понимать под $\mathfrak{X}$ польское (т.е. полное сепарабельное метрическое) пространство, в простейшем варианте сводящееся к $\mathbb{R}$ или подобным.

Рассмотрим простейший случай $\mathfrak{X} = \mathbb{R}$. Для случайной величины $\xi : \Omega \to \mathbb{R}$ можно определить матожидание

$$\mathbb{E}\xi := \int_{\Omega} \xi(\omega) d\mathbb{P} \tag{1.1}$$

Полезное понятие, обобщающее матожидание, – условное матожидание. Пусть $\mathcal{F}' \subset \mathcal{F}$ – σ-подалгебра $\mathcal{F}$. Условным матожиданием случайной величины ξ относительно $\mathcal{F}'$

называется случайная величина $\mathbb{E}(\xi|\mathcal{F}')$, измеримая относительно $\mathcal{F}'$, такая что для любой функции $\eta : \Omega \to \mathbb{R}$, также измеримой относительно $\mathcal{F}'$, имеем

$$\mathbb{E}[\mathbb{E}(\xi|\mathcal{F}')\eta] = \mathbb{E}\xi\eta. \tag{1.2}$$

Как можно охарактеризовать случайную величину ξ на $\mathbb{R}$, так чтобы, не вникая в детали отображения $\xi(\omega)$, задать меру множеств, отображаемых в те или иные подмножества $\mathbb{R}$? Для этого служит функция распределения вероятности

$$F_\xi(x) := \mathbb{P}\left(\{\omega : \xi(\omega) \leq x\}\right), \tag{1.3}$$

а если мера на $\mathbb{R}$, индуцируемая отображением ξ, абсолютно непрерывна относительно меры Лебега, то и плотность вероятности

$$f_\xi(x) = F'_\xi(x). \tag{1.4}$$

Аналогично можно рассматривать несколько случайных величин на одном и том же вероятностном пространстве. Они характеризуются совместными функциями распределения

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \mathbb{P}\left(\{\omega : \xi_1(\omega) \leq x_1, \ldots, \xi_n(\omega) \leq, x_n\}\right) \tag{1.5}$$

При этом мы говорим, что случайные величины независимы, если их совместная функция распределения распадается в произведение индивидуальных функций распределения,

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \prod_{k=1}^{n} F_{\xi_k}(x_k), \tag{1.6}$$

что немедленно влечет подобную факторизацию матожиданий произведений, как самих случайных величин, так и их функций.

Альтернативное описание распределения случайной величины ξ дает характеристическая функция

$$\varphi_\xi(x) = \mathbb{E}(e^{ix\xi}), \tag{1.7}$$

которая, будучи преобразованием Фурье меры, определяет распределение единственным образом.

Близко связанное с характеристической функцией понятие производящей функции моментов

$$M_\xi(x) = \mathbb{E}(e^{x\xi}) = \sum_{n=0}^{\infty} \frac{m_n x^n}{n!}. \tag{1.8}$$

определено, когда все моменты

$$m_n = \mathbb{E}(\xi^n) \tag{1.9}$$

распределения случайной величины ξ конечны. Ее можно понимать либо как формальный ряд, либо как аналитическую в некоторой окрестности нуля функцию. В первом случае экспонента под знаком матожидания в (1.8) также понимается как формальный ряд. Во втором предполагается, что ряд в правой части (1.8) сходится в некоторой окрестности нуля, что накладывает дополнительные ограничения на скорость убывания вероятности на бесконечности. Получившуюся аналитическую функцию можно аналитически продолжить из области сходимости как минимум на всю мнимую ось. Результат естественно совпадает с характеристической функцией.

Поскольку знание характеристической функции эквивалентно знанию распределения, мы подходим к ответу на естественный вопрос о том, можно ли, зная все моменты (кумулянты) случайной величины, однозначно восстановить распределение. Эта задача известна как *проблема моментов.* Одно из её решений состоит как раз в том, чтобы производящая функция моментов $M_\xi(x)$ была аналитической в некоторой окрестности нуля, то есть чтобы моменты убывали достаточно быстро. Это требование эквивалентно *критерию Рисса*:

$$\lim_{n\to\infty}\frac{m_n^{\frac{1}{n}}}{n}<\infty. \tag{1.10}$$

В частности, если носитель распределения конечен, то есть $\mathrm{Supp}(\xi)\in[-M,M]$ для некоторой константы M, то при всех $n\in\mathbb{N}$ выполнено неравенство $m_n\leqslant M^n$ и проблема моментов имеет единственное решение.

Замечание 1.1. *Наиболее полное и чуть менее ограничительное решение той же проблемы дается критерием Карлемана, который требует сходимости ряда обратных четных моментов.*

$$\sum_{n=0}^{\infty}\frac{1}{m_{2n}}<\infty. \tag{1.11}$$

Моменты не всегда являются удобной характеристикой распределения. Зачастую удобнее пользоваться *кумулянтами*, являющимися полиномиальными комбинациями моментов . Кумулянты определяются как коэффициенты c_n ряда Тейлора логарифма производящей функции моментов, который будем называть *производящей функцией кумулянтов*:

$$C_\xi(x)=\ln M_\xi(x)=\ln\mathbb{E}(e^{x\xi})=\sum_{n=0}^{\infty}\frac{c_nx^n}{n!}. \tag{1.12}$$

Кумулянты исторически называют также *полуинвариантами* или *семиинвариантами.*

Упражнение 1.2. *Пусть случайные величины $\xi_1,\dots,\xi_n$ являются независимыми, а числа $a_1,\dots,a_n$ представляют собой некоторый набор констант. Докажите, что (при условии, что указанные величины существуют):*
а) $\varphi_{a_1\xi_1+\ldots+a_n\xi_n}(x)=\varphi_{\xi_1}(a_1x)\cdot\ldots\cdot\varphi_{\xi_n}(a_nx)$,
б) $M_{a_1\xi_1+\ldots+a_n\xi_n}(x)=M_{\xi_1}(a_1x)\cdot\ldots\cdot M_{\xi_n}(a_nx)$,
в) $C_{a_1\xi_1+\ldots+a_n\xi_n}(x)=C_{\xi_1}(a_1x)+\ldots+C_{\xi_n}(a_nx)$.

Таким образом, как видно из упражнения 1.2, важной особенностью кумлянтов является их аддитивность: кумулянты суммы независимых случайных величин — это сумма кумулянтов этих величин.

Кумулянты и моменты содержат одинаковую информацию и могут быть однозначно выражены друг через друга формальным переразложением в ряд логарифмированного (соответственно, экспоненциированного) ряда для производящей функции моментов (кумулянтов).

Упражнение 1.3. *Докажите, что между первыми тремя моментами и кумулянтами выполняются следующие соотношения*

$$m_1=c_1,\qquad m_2 = c_2+c_1^2,\qquad m_3=c_3+3c_1c_2+c_1^3 \tag{1.13}$$
$$c_1=m_1,\qquad c_2 = m_2-m_1^2,\qquad c_3=m_3-3m_1m_2+2m_1^3 \tag{1.14}$$

В частности, $c_1=\mathbb{E}(\xi)$ и $c_2=\mathbb{D}(\xi)$.

Общая формула, выражающая моменты через кумулянты, имеет вид

$$m_n = \sum_{\pi \in \mathcal{P}(n)} \prod_{I \in \pi} c_{|I|}. \tag{1.15}$$

Здесь суммирование ведётся по множеству

$$\mathcal{P}(n) = \{\pi = \{I_1, \ldots, I_k\} : [n] = \bigcup_{i=1}^{k} I_i; I_i \bigcap I_j = \emptyset, i \neq j; k \in [n]\}. \tag{1.16}$$

разбиений множества

$$[n] := \{1, \ldots, n\}$$

на все возможные подмножества.

Приведенные формулы моментов и кумулянтов одной случайной величины – частный случай более общих формул для характеристик набора случайных величин. Пусть $\xi_1, \ldots, \xi_n$ – набор из n случайных величин, не обязательно различных, $n \in \mathbb{N}$. Определим совместные моменты как полилинейные функционалы степени n

$$m_n(\xi_1, \ldots, \xi_n) := \mathbb{E}(\xi_1 \cdots \xi_n). \tag{1.17}$$

Тогда можно определить и совместные кумулянты, как полилинейные функционалы, $c_k(x_1, \ldots, x_k)$, связанные с моментами соотношением, частным случаем, которого является соотношение (1.15),

$$m_n(\xi_1, \ldots, \xi_n) = \sum_{\pi \in \mathcal{P}(n)} \prod_{I \in \pi} c_{|I|}(\xi_I), \tag{1.18}$$

где $I = (i_1, \ldots, i_{|I|})$ – блоки разбиения π и мы ввели обозначение $\xi_I = (\xi_{i_1}, \ldots, \xi_{i_{|I|}})$. К этому соотношению можно также прийти, введя производящие функции совместных моментов и совместных кумулянтов для наборов случайных величин, и потребовав, чтобы, так же как и в случае одной случайной величины они были связаны между собой соотношением (1.12). В частности из этой связи следует следующий важный факт: кумулянты вычисленные на наборе случайных величин, среди которых имеется хотя бы одна, независимая от остальных, равны нулю.[1] Позже при обсуждении свободной вероятности мы познакомимся с некоммутативными аналогами этого факта и соотношения (1.18).

Охарактеризуем на языке моментов, кумулянтов и их производящих функций некоторые важнейшие примеры распределений.

Пример 1.4. *Пусть случайная величина $\xi = a$ задаётся* дельта-мерой Дирака

$$\delta_a(A) = \begin{cases} 1, & \text{если } a \in A \\ 0, & \text{если } a \notin A, \end{cases} \tag{1.19}$$

т.е. $\mathbb{P}(\xi = a) = 1$*, а её функция распределения представляет собой* функцию Хэвисайда*:*

$$F_\xi(x) = \theta(x - a) = \begin{cases} 1, & \text{если } x > a \\ 0, & \text{если } x \leqslant a. \end{cases} \tag{1.20}$$

[1]Эти факты также хорошо известны специалистам по квантовой теории поля и статистической физике, где совместные моменты и кумулянты известны под названием несвязные и связные корреляционные функции соответственно.

Тогда производящие функции моментов и кумулянтов имеют вид

$$M_\xi(x) = e^{xa}, \qquad C_\xi(x) = xa.$$

То есть единственный ненулевой кумулянт дельта-распределения — первый: $c_1 = a$.

Пример 1.5. *Второй и последний пример распределения, у которого почти все кумулянты равны нулю, а значит,* $C_\xi(x)$ *— многочлен, это распределение Гаусса. Для нормально распределённой случайной величины* $\xi \sim \mathcal{N}(\mu, \sigma^2)$*, задаваемой плотностью*

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{1.21}$$

имеем

$$M_\xi(x) = \exp\left(\mu x + \frac{\sigma^2 x^2}{2}\right), \qquad C_\xi(x) = \mu x + \frac{\sigma^2 x^2}{2}. \tag{1.22}$$

В частности, $c_1 = \mu$, $c_2 = \sigma^2$ *и* $c_k = 0$ *при* $k > 2$.

Пример 1.6. *Ещё одно важное дискретное распределение, распределение Пуассона* $\mathrm{Poi}(\lambda)$*, выделено тем, что все его кумулянты одинаковы:*

$$\mathbb{P}(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \tag{1.23}$$

$$M_\xi(x) = e^{e^{\lambda x}}, \qquad C_\xi(x) = e^{\lambda x} \tag{1.24}$$

и $c_k = \lambda$ *при всех* $k \in \mathbb{N}$.

Упражнение 1.7. *Убедитесь в справедливости формул* (1.22) *и* (1.24)*).*

Упражнение 1.8. *Докажите, что чётные моменты* полукругового распределения Вигнера *с плотностью* $f_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\cdot\mathbb{I}_{\{|x|\leqslant 2\}}$ *даются* числами Каталана*, а нечётные равны нулю:*

$$m_{2k} = C_k = \frac{C_{2k}^k}{k+1}, \qquad m_{2k+1} = 0. \tag{1.25}$$

.

1.2 Сходимость последовательностей случайных величин и предельные теоремы.

Определение 1.9. *Пусть* $\{\xi_n\}_{n\in\mathbb{N}}$ *— последовательность случайных величин определённых на одном вероятностном пространстве. Говорят, что эта последовательность сходится к случайной величине* ξ

- почти наверное *(или* почти всюду*), и пишут*

  $$\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi,$$

  если

  $$\mathbb{P}\left(\xi_n \xrightarrow[n\to\infty]{} \xi\right) = \mathbb{P}\left(\omega\colon \xi_n(\omega) \xrightarrow[n\to\infty]{} \xi(\omega)\right) = 1, \tag{1.26}$$

  т.е. имеет место поточечная сходимость везде, кроме, быть может, множества меры нуль;

- по вероятности, *и пишут*
$$\xi_n \xrightarrow[n\to\infty]{\mathbb{P}} \xi,$$
если для любого $\varepsilon > 0$*:*
$$\mathbb{P}\left(|\xi_n - \xi| > \varepsilon\right) = \mathbb{P}\left(\omega\colon |\xi_n(\omega) - \xi(\omega)| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0; \tag{1.27}$$

- по распределению или слабо, *и пишут*
$$\xi_n \xrightarrow[n\to\infty]{D} \xi,$$
если для любой непрерывной ограниченной функции $g \in C_b$*:*
$$\mathbb{E}(g(\xi_n)) \xrightarrow[n\to\infty]{} \mathbb{E}(g(\xi)), \tag{1.28}$$
или же, что эквивалентно сказанному выше, если последовательность функций распределения сходится к предельной функции распределения во всех её точках непрерывности:
$$\lim_{n\to\infty} F_{\xi_n}(t) \to F_\xi(t) \quad \forall\, t : F_\xi \in C(t); \tag{1.29}$$

- в L^p, *и пишут*
$$\xi_n \xrightarrow[n\to\infty]{\mathrm{L^p}} \xi,$$
если
$$\lim_{n\to\infty} \mathbb{E}(|\xi_n - \xi|^p) = 0. \tag{1.30}$$

Упражнение 1.10. *Докажите эквивалентность определений сходимости по распределению, заданных формулами* (1.28) *and* (1.29).

Упражнение 1.11. *Докажите следующие свойства сходимостей:*
а) если $\xi_n \xrightarrow[n\to\infty]{\text{п.н.}} \xi$*, то* $\xi_n \xrightarrow[n\to\infty]{\mathbb{P}} \xi$*, но обратное, вообще говоря, не верно;*
б) если $\xi_n \xrightarrow[n\to\infty]{\mathrm{L^p}} \xi$ *и* $p \geqslant 2$*, то* $\xi_n \xrightarrow[n\to\infty]{\mathbb{P}} \xi$*, но обратное, вообще говоря, не верно;*
в) если $\xi_n \xrightarrow[n\to\infty]{\mathbb{P}} \xi$*, то* $\xi_n \xrightarrow[n\to\infty]{D} \xi$*, но обратное, вообще говоря, не верно;*
г) если $\xi_n \xrightarrow[n\to\infty]{D} \xi$ *и* ξ *вырожденная случайная величина, т.е.* $\mathbb{P}(\xi = a) = 1$*, то* $\xi_n \xrightarrow[n\to\infty]{\mathbb{P}} \xi$*.*

Сходимость по распределению — это частный случай *слабой сходимости* мер; отчасти упражнение 1.11 объясняет последнее называние. Заметим, что здесь не обязательно требовать, чтобы члены последовательности и предельная случайная величина были определены на одном вероятностном пространстве. Отметим также, что хотя из сходимости по вероятности сходимость почти наверное не следует, *теорема Рисса* гарантирует нам, что из исходной последовательности можно извлечь подпоследовательность, которая будет сходиться к пределу почти наверное.

Опишем основные инструменты, используемые для доказательства сходимости.

Лемма 1.12. *[Неравенство Маркова] Пусть* ξ *— неотрицательная случайная величина с конечным математическим ожиданием:* $\xi \geqslant 0$ *и* $\mathbb{E}\xi \leqslant \infty$*. Тогда для любого* $\varepsilon > 0$ *справедливо следующее неравенство:*
$$\mathbb{P}\big(\xi > \varepsilon\big) \leqslant \frac{\mathbb{E}\xi}{\varepsilon}. \tag{1.31}$$

В частности, если при некотором $k \in \mathbb{N}$ для случайной величины ξ выполнено неравенство $\mathbb{E}|\xi - \mathbb{E}\xi|^k < \infty$, то

$$\mathbb{P}\big(|\xi - \mathbb{E}\xi| > \varepsilon\big) \leqslant \frac{\mathbb{E}|\xi - \mathbb{E}\xi|^k}{\varepsilon^k} \tag{1.32}$$

При $k = 2$ неравенство (1.32) превращается в *неравенство Чебышева* и используется для доказательства сходимости по вероятности: последняя имеет место, когда $\mathbb{D}\xi_n \xrightarrow[n\to\infty]{} 0$.

Наиболее сложный случай сходимости — сходимость почти наверное, поскольку для его обоснования требуется привлекать утверждения о событиях, включающих бесконечное число членов последовательности. Однако если типичная последовательность сходится достаточно быстро, для доказательства также можно воспользоваться неравенством Маркова в комбинации с леммой Бореля-Контелли.

Лемма 1.13. *[Лемма Бореля-Кантелли] Пусть $\{A_n\}_{n\in\mathbb{N}}$ — последовательность случайных событий. Обозначим*

$$A = \limsup_{n\to\infty} A_n \equiv \bigcap_{n=1}^{\infty}\left(\bigcup_{m=n}^{\infty} A_m\right) = \left\{\,\omega\colon\ \forall n \in \mathbb{N}\ \exists m \in \mathbb{N}, m > n\colon \omega \in A_m\,\right\},$$

т.е. A — это исходы, которые происходят бесконечно часто. Тогда:
I) если числовой ряд $\sum\limits_{n=1}^{\infty} \mathbb{P}(A_n)$ сходится, то $\mathbb{P}(A) = 0$;
II) если все события A_n независимы в совокупности и числовой ряд $\sum\limits_{n=1}^{\infty} \mathbb{P}(A_n)$ расходится, то $\mathbb{P}(A) = 1$.

Теперь для обоснования сходимости почти наверное можно действовать следующим образом. Рассмотрим событие $A_n = \{\omega : |\xi_n - \xi| > \varepsilon\}$. Если при некотором $k > 0$ ряд, составленный из $\mathbb{E}|\xi_n - \xi|^k$, сходится, то первая часть леммы Бореля-Контелли совместно с неравенством Маркова гарантируют сходимость почти наверное.

Вернемся к сходимости по распределению. Незаменимый инструмент для доказательства такой сходимости дает теорема Леви о непрерывности.

Теорема 1.14. *[Теорема Леви о непрерывности] Пусть $\big\{\xi_n\big\}_{n\in\mathbb{N}}$ — последовательность случайных величин. Тогда если $\xi_n \xrightarrow[n\to\infty]{D} \xi$, то для любого $t \in \mathbb{R}$:*

$$\varphi_{\xi_n}(t) \xrightarrow[n\to\infty]{} \varphi_\xi(t).$$

Обратно, если $\varphi(t) \in C(0)$ — функция действительного аргумента, непрерывная в нуле, и для каждого $t \in \mathbb{R}$ имеет место сходимость $\varphi_{\xi_n}(t) \xrightarrow[n\to\infty]{} \varphi(t)$, то $\varphi(t)$ является характеристической функцией случайной величины ξ, для которой $\xi_n \xrightarrow[n\to\infty]{D} \xi$.

Часто вместо анализа сходимости характеристической функции удобнее исследовать сходимость моментов (кумулянтов) распределения.

Определение 1.15. *Пусть для всех $k, n \in \mathbb{N}$ моменты*

$$m_n^{(k)} = \mathbb{E}(\xi_k^n) \quad \text{и} \quad m_n = \mathbb{E}(\xi^n)$$

случайных величин ξ_k и ξ соответственно конечны для любого порядка n. Будем говорить, что ξ_n сходится к ξ в смысле моментов*, если для каждого $n \in \mathbb{N}$:*

$$\lim_{k\to\infty} m_n^{(k)} = m_n. \tag{1.33}$$

Вообще говоря, сходимости в смысле моментов и по распределению не эквивалентны и ни одна из них не влечет другую.

Упражнение 1.16. *На примере меры*

$$\left(1 - \frac{1}{n}\right)\delta_0 + \frac{1}{n}\delta_n$$

убедитесь в том, что из сходимости по распределению (и даже по вероятности), сходимость в смысле моментов не следует.

Тем не менее, в некоторых случаях можно утверждать, что из сходимости в смысле моментов следует сходимость по распределению. Так же, как в проблеме моментов, для этого достаточно, чтобы моменты росли не слишком быстро. В частности, если и ξ, и все ξ_n имеют компактный носитель, то сходимость следует из *теоремы Вейерштрасса о полиномиальной аппроксимации.*

Теорема 1.17. *[Теорема Вейерштрасса о полиномиальной аппроксимации] Пусть f — функция, определённая на отрезке $[a, b]$ и непрерывная на этом отрезке. Тогда для любого $\varepsilon > 0$ существует такой многочлен $P \in \mathbb{R}[x]$, что для всех $x \in [a, b]$ выполняется $|f(x) - p(x)| < \varepsilon$.*

В свою очередь, сходимость математических ожиданий многочленов от случайной величины немедленно следует из сходимости моментов.

От требования ограниченности носителя можно избавиться, наложив определённые условия на предельную случайную величину. А именно, достаточно, чтобы моменты предельной случайной величины гарантировали единственность решения проблемы моментов, например, удовлетворяя критерию Карлемана. Мы докажем более слабое утверждение, которое пригодится нам в дальнейшем, потребовав компактности носителя, но лишь для предельной величины.

Лемма 1.18. *Пусть последовательность ξ_n сходится в смысле моментов к случайной величине ξ с ограниченным носителем, т.е.* $\operatorname{Supp}(\xi) \subset [-a, a]$ *для некоторого $a > 0$. Тогда имеет место сходимость по распределению:*

$$\xi_n \xrightarrow[n\to\infty]{D} \xi.$$

Доказательство. Рассмотрим произвольную ограниченную непрерывную функцию $g(x)$, такую, что $|g(x)| < b$ для некоторого $b > 0$. Докажем, что

$$\left| \int_{\mathbb{R}} g\, dF_\xi - \int_{\mathbb{R}} g\, dF_{\xi_n} \right| \xrightarrow[n\to\infty]{} 0.$$

Для этого разобьём область интегрирования на отрезок $A = [-2a, 2a]$ и его дополнение. По теореме Вейерштрасса о полиномиальной аппроксимации (теорема 1.17) для любого $\varepsilon > 0$ можно найти многочлен $P(x)$, для которого верно

$$\big|g(x) - P(x)\big| < \frac{\varepsilon}{8} \quad \forall\, x \in A.$$

Кроме того, имеет место неравенство

$$\begin{aligned}\left|\int_{\mathbb{R}} g\,dF_\xi - \int_{\mathbb{R}} g\,dF_{\xi_n}\right| \;\leqslant\; & \int_A |g-P|\,dF_\xi + \int_A |g-P|\,dF_{\xi_n} \\ & + \left|\int_{\mathbb{R}} P\,dF_\xi - \int_{\mathbb{R}} P\,dF_{\xi_n}\right| + \int_{\mathbb{R}\setminus A} (|g|+|P|)\;dF_{\xi_n},\end{aligned}$$

где в правой части уже опущены два интеграла по dF_ξ по области $\mathbb{R}\setminus A$, поскольку там dF_ξ равен нулю тождественно. Ясно, что первые два слагаемых правой части в сумме меньше $\varepsilon/4$ в силу выбора многочлена P и вероятностности меры. Выбрав достаточно большое n, можно добиться, чтобы третье слагаемое также не превышало $\varepsilon/4$; здесь мы пользуемся сходимостью моментов и тем, что P — многочлен. Наконец, четвертое слагаемое оценивается с помощью аргумента, использованного при доказательстве неравенства Маркова (лемма 1.12). В самом деле, пусть $2q$ — чётное число, большее или равное степени многочлена P и $p \in \mathbb{N}$ — некоторое натуральное число. Тогда найдётся такая константа c, что $|P(x)| \leqslant c(1+|x|^{2q})$, откуда

$$\int_{\mathbb{R}\setminus A} (|g|+|P|)\;dF_{\xi_n} \leqslant \int_{\mathbb{R}\setminus A} \big(b + c(1+|x|^{2q})\big)\,dF_{\xi_n} \leqslant \left(\frac{b+c}{(2a)^{2(q+p)}} + \frac{c}{(2a)^{2p}}\right)\int_{\mathbb{R}} x^{2(q+p)}\,dF_{\xi_n}.$$

Интеграл в правой части сходится к $m_{2(q+p)} \leqslant a^{2(p+q)}$. Поэтому, увеличивая p, можно добиться того, чтобы для достаточно больших n правая часть не превышала $\varepsilon/4$. □

Наконец, приведём формулировки обещанных выше ЗБЧ и ЦПТ, и докажем их, воспользовавшись теоремой Леви о непрерывности.

Теорема 1.19. *[ЗБЧ] Пусть (ξ_n) — последовательность независимых одинаково распределённых случайных величин, обладающих конечным математическим ожиданием $\mathbb{E}(\xi_n) = \mu$. Тогда*

$$\overline{\xi}_n = \frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} \xi_k \xrightarrow[n\to\infty]{\mathbb{P}} \mu. \tag{1.34}$$

Теорема 1.20. *[ЦПТ] Если же, помимо конечных математических ожиданий, (ξ_n) имеют ещё и конечную дисперсию $\mathbb{D}(\xi_n) = \sigma^2$, то*

$$\zeta_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{D} \eta, \tag{1.35}$$

где случайная величина η имеет стандартное нормальное распределение $\mathcal{N}(0,1)$.

Доказательство. Прежде всего, заметим, что если все моменты ξ_n существуют, то из независимости и аддитивности кумулянтов мы немедленно устанавливаем, что все кумулянты случайных величин $\overline{\xi}_n$ и ζ_n стремятся к нулю при $n \to \infty$, за исключением $\mathbb{E}(\overline{\xi}_n) \to \mu$ в первом случае, и $\mathbb{E}(\zeta_n^2) \to 1$ во втором. Однако нет надобности ограничивать класс распределений; общий случай следует из предельных переходов для характеристической функции и теоремы Леви (теорема 1.14). Действительно, существование математического ожидания гарантирует существование первой прозводной $\varphi_{\xi_1}(x)$ в $x = 0$, откуда вытекает, что

$$\varphi_{\xi_1}(x) = 1 + \mathrm{i}\mu x + o(x), \qquad x \to 0.$$

В силу линейности (упражнение 1.2) это означает, что

$$\varphi_{\overline{\xi}_n}(x) = \left(\varphi_{\xi_1}\left(\frac{x}{n}\right)\right)^n = \left(1 + \mathrm{i}\,\frac{\mu x}{n} + o\left(\frac{x}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^{\mathrm{i}\mu x}.$$

Следовательно, $\overline{\xi}_n \xrightarrow[n\to\infty]{D} \mu$, а значит, в силу упражнения 1.11 мы имеем $\overline{\xi}_n \xrightarrow[n\to\infty]{\mathbb{P}} \mu$.

Аналогично, поскольку характеристическая функция случайной величины

$$\eta_k = \frac{\xi_k - \mu}{\sigma}$$

имеет вид

$$\varphi_{\eta_1}(x) = 1 - \frac{x^2}{2} + o(x^2), \qquad x \to 0,$$

для $\zeta_n = \sum\limits_{k=1}^{n} \dfrac{\eta_k}{\sqrt{n}}$ мы имеем

$$\varphi_{\zeta_n}(x) = \left(\varphi_{\eta_1}\left(\frac{x}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{x^2}{2n} + o\left(\frac{x^2}{n}\right)\right)^n \xrightarrow[n\to\infty]{} \exp\left(-\frac{x^2}{2}\right).$$

Следовательно, согласно теореме Леви, $\zeta_n \xrightarrow[n\to\infty]{D} \eta$, к чему мы и стремились. □

Замечание 1.21. *Существует также усиленный закон больших чисел, гарантирующий сходимость почти наверное при условии совокупной независимости всех членов последовательности и конечности матожидания.*

На практике ЗБЧ и ЦПТ означают, что при больших n величина S_n детерминистически растёт линейно по n. Случайные отклонения от детерминистического линейного роста, которые могут быть обнаружены с конечной вероятностью, имеют порядок величины $O(\sqrt{n})$ и описываются нормальным распределением, независимо от деталей исходных случайных величин:

$$S_n \sim \mu n + \sigma\sqrt{n} \cdot \mathcal{N}(0, 1). \tag{1.36}$$

1.3 Мартингалы и концентрация меры

ЗБЧ это пример концентрации меры, вокруг неслучайной величины. Этот результат, будучи в большой степени универсальным, тем не менее, ограничен суммами последовательностей независимых случайных величин. Еще одним классом процессов, в которых наблюдается подобное явление, не основанных на независимости случайных величин, являются мартингалы.

Определение 1.22. *Мартингалом (в дискретном времени) относительно последовательности сигма алгебр $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \ldots$ называют последовательность случайных величин $(\chi_n)_{n\in\mathbb{N}}$ такую, что для всех $n \in \mathbb{N}$ величина χ_n измерима относительно $\mathcal{F}_n$, $\mathbb{E}|\chi_n| < \infty$ и*

$$\mathbb{E}(\chi_n|\mathcal{F}_{n-1}) = \chi_{n-1} \tag{1.37}$$

Видно, что мартингал это процесс, который в среднем никуда не смещается. Поэтому не удивительно, что приятной особенностью мартингалов являются различные утверждения о их сходимости. Один из инструментов, который позволяет их делать — мартингальные неравенства. Докажем одно такое неравенство, котрое будет нам полезно для доказательства свойства концентрации меры.

Лемма 1.23 (Неравенство Азумы)**.** *Пусть* $(\chi_n, \mathcal{F}_n)_{n\in\mathbb{Z}_{\geq 0}}$ *- мартингал с ограниченными приращениями, т.е. для некоторых* $a_n > 0$, $n = 1, 2, \dots$

$$|\chi_{n+1} - \chi_n| \leq a_{n+1}, \quad \forall n \in \mathbb{Z}_{\geq 0}. \tag{1.38}$$

тогда справедлива следующая оценка

$$\mathbb{P}(|\chi_0 - \chi_n| > t) \leq 2e^{-\dfrac{t^2}{2\sum_{k=1}^n a_k^2}} \tag{1.39}$$

Доказательство. Запишем экспоненциальное неравенство Маркова для величины $(\chi_0 - \chi_n)$

$$\mathbb{P}(\chi_n - \chi_0 > t) = \mathbb{P}(e^{\lambda(\chi_n - \chi_0)} > e^{\lambda t}) \leq \frac{\mathbb{E}e^{\lambda(\chi_n-\chi_0)}}{e^{\lambda t}}, \tag{1.40}$$

где предпоалагается, что $\lambda > 0$. Оценим знаменатель в правой части. Для этого введем линейную функцию

$$h_a(x) = \operatorname{ch} a\lambda + \frac{x}{a}\operatorname{sh} a\lambda, \qquad a > 0 \tag{1.41}$$

и заметим, что в силу выпуклости экспоненты неравенство

$$e^{\lambda x} \leq h_a(x) \tag{1.42}$$

выполняется при $-a \leq x \leq a$. Тогда

$$\begin{aligned}
\mathbb{E}e^{\lambda(\chi_n-\chi_0)} &= \mathbb{E}\big(\mathbb{E}(e^{\lambda(\chi_n-\chi_{n-1})}|\mathcal{F}_{n-1})e^{\lambda(\chi_{n-1}-\chi_0)}\big) && (1.43)\\
&\leq \mathbb{E}\big(\mathbb{E}(h_{a_n}(\chi_n - \chi_{n-1})|\mathcal{F}_{n-1})e^{\lambda(\chi_{n-1}-\chi_0)}\big) = h_{a_n}(0)\mathbb{E}e^{\lambda(\chi_{n-1}-\chi_0)} && (1.44)\\
&= \operatorname{ch}(\lambda a_n)\mathbb{E}e^{\lambda(\chi_{n-1}-\chi_0)} \leq e^{\frac{\lambda^2 a_n^2}{2}}\mathbb{E}e^{\lambda(\chi_{n-1}-\chi_0)} \leq e^{\frac{\lambda^2\sum_{k=1}^n a_k^2}{2}} && (1.45)
\end{aligned}$$

где в первом переходе мы воспользовались измеримостью разности $(\chi_{n-1} - \chi_0)$ относительно $\mathcal{F}_n$, во втором неравенством (1.42), в третьем тем, что χ_n – мартингал, и наконец тем, что гиперболический косинус мажорируется экспонентой от квадратичной функции. Тогда из (1.40) имеем

$$\mathbb{P}(\chi_n - \chi_0 > t) \leq e^{\frac{\lambda^2}{2}\sum_{k=1}^n a_k^2 - \lambda t} \leq e^{-\frac{t^2}{2\sum_{k=1}^n a_k^2}}, \tag{1.46}$$

где во втором неравенстве мы минимизировали экспоненту по λ. Для $\mathbb{P}(\chi_n - \chi_0 < -t)$ верна такая же оценка, и наконец

$$\mathbb{P}(|\chi_n - \chi_0| > t) \leq \mathbb{P}(\chi_n - \chi_0 < -t) + \mathbb{P}(\chi_n - \chi_0 > t) \leq 2e^{-\frac{t^2}{2\sum_{k=1}^n a_k^2}}. \tag{1.47}$$

□

Важный пример мартингала — мартингал Дуба.

Определение 1.24. *[мартингал Дуба] Пусть $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \dots$ - последовательность сигма-алгебр, а χ — произвольная случайная величина, такая что $\mathbb{E}|\chi| < \infty$. Тогда последовательность случайных величин*

$$\chi_n := \mathbb{E}(\chi|\mathcal{F}_n), \quad n \in \mathbb{N} \tag{1.48}$$

образует мартингал по отношению $(\mathcal{F}_n)_{n\in\mathbb{N}}$

Действительно, имеет место цепочка равенств

$$\mathbb{E}(\chi_n|\mathcal{F}_{n-1}) = \mathbb{E}\big(\mathbb{E}(\chi|\mathcal{F}_n)|\mathcal{F}_{n-1}\big) = \mathbb{E}\big(\chi|\mathcal{F}_{n-1}\big) = \chi_{n-1}$$

где во втором равенстве использовано телескопическое свойство условного матожидания по отношению к вложенным сигма-алгебрам и кроме того конечность матожидания модуля χ_n следует из аналогичного свойства величины χ,

$$\mathbb{E}|\chi_n| = \mathbb{E}|\mathbb{E}(\chi|\mathcal{F}_n)| \leq \mathbb{E}\mathbb{E}(|\chi||\mathcal{F}_n) = \mathbb{E}|\chi| < \infty.$$

Применение неравенства Азумы к мартингалу Дуба позволяет доказывать свойство концентрации меры.

Пусть $\chi = (\chi_n)_{1\leq n\leq m}$ — набор независимых случайных величин, а $f(x_1, \dots, x_m)$ — c-липшицева вещественнозначная функция m переменных, что означает, что изменение функции при варьировании одной переменной конечно, т.е.

$$|f(x_1, \dots x_{k-1}, x_k, x_{k+1}, \dots, x_m) - f(x_1, \dots x_{k-1}, x'_k, x_{k+1}, \dots, x_m)| \leq a_k. \tag{1.49}$$

для некоторого набора $a_k > 0, k = 1, \dots, m$. Из этого следует, что $\mathbb{E}f(\chi_1, \dots, \chi_m) < \infty$. Покажем также, что выполнено неравенство

$$\Big|\mathbb{E}(f(\chi_1, \dots, \chi_m)|\chi_k, \dots, \chi_m) - \mathbb{E}(f(\chi_1, \dots, \chi_m)|\chi_{k+1}, \dots, \chi_m)\Big| \leq a_k. \tag{1.50}$$

Для этого введем обозначение $\chi^{(i,j)} = \{\chi_i, \dots, \chi_j\}$, где $1 \leq i < j \leq m$ для подмножеств χ состоящих из элементов с номерами от i до j. Тогда

$$\Big|\mathbb{E}(f(\chi)|\chi^{(k,m)}) - \mathbb{E}(f(\chi)|\chi^{(k+1,m)})\Big| = \Big|\mathbb{E}[\mathbb{E}[f(\chi)|\chi\backslash\chi_k] - f(\chi)|\chi^{(k+1,m)}]\Big|. \tag{1.51}$$

Заметим, что в виду независимости $\chi_1, \dots, \chi_k$ имеет место равенство

$$\mathbb{E}[f(\chi)|\chi\backslash\chi_k] = \int_{\mathbb{R}} f(\chi^{(1,k-1)}, x, \chi^{(k+1,m)}) dF_{\chi_k}(x), \tag{1.52}$$

где $F_{\chi_k}(x_k)$ – функция распределения χ_k. Используя неравенство Гёльдера, мы приходим к оценке

$$|\mathbb{E}[f(\chi)|\chi\backslash\chi_k] - f(\chi)| = \left|\int_{\mathbb{R}} \big(f(\chi^{(1,k-1)}, x, \chi^{(k+1,m)}) - f(\chi)\big)\, dF_{\chi_k}(x)\right| \leq a_k, \tag{1.53}$$

которая вкупе с (1.51) даёт (1.50). Отсюда, применяя неравенство Азумы, получим неравенство Мак-Диармида

$$\mathbb{P}\big(|f(\chi_1, \dots, \chi_m) - \mathbb{E}f(\chi_1, \dots, \chi_m)| > t\big) \leq 2e^{-\frac{t^2}{2\sum_{k=1}^m a_k^2}}. \tag{1.54}$$

Приведем пример его применения.

Пример 1.25. *Пусть $f_m(x_1, \dots, x_m)$, $m \in \mathbb{N}$, — последовательность измеримых c-липшицевых функций, для каждой из которых неравенство (1.50) выполнено с $a_1 = \dots = a_m = a(m) = O(1/\sqrt{m(\ln m)^{(1+\epsilon)}})$ с $\epsilon > 0$ при $m \to \infty$.. Тогда из неравенства Мак-Диармида и леммы Бореля-Контелли следует следующее свойство концентрации меры*

$$f_m(\chi_1, \dots, \chi_m) - \mathbb{E}f_m(\chi_1, \dots, \chi_m) \xrightarrow[m\to\infty]{\text{п.н.}} 0. \tag{1.55}$$

1.4 Слабая сходимость случайных мер

До сих пор мы обсуждали сходимость последовательностей случайных величин, принимающих значения в $\mathbb{R}$. При этом выяснилось, что случайные величины, построенные как функции большого числа других случайных величин, зачастую демонстрируют универсальное типичное поведение, которое проявляется в их сходимостях к неслучайным или универсальным случайным пределам, определяемым утверждениями типа ЗБЧ и ЦПТ. В более общем случае при рассмотрении многокомпонентных случайных систем, таких как случайные матрицы, уместно ставить вопросы о более детальном описании их типичного поведения в пределе когда число компонент неограниченно растет. Для этого хорошо приспособлен язык случайных мер и сходимости их последовательностей.

Пусть $\mathfrak{X}$ – польское пространство, $\mathcal{B}(\mathfrak{X})$ — борелевская сигма-алгебра его подмножеств, а $\mathcal{P}(\mathfrak{X})$ — пространство вероятностных борелевских мер на $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$. В секции 1 было дано понятие слабой сходимости мер, которое можно определять через сходимость интегралов от ограниченных непрерывных функций. Отметим, что слабая сходимость может быть метризуема, например, при помощью метрики Леви-Прохорова, в случае $\mathfrak{X} = \mathbb{R}$ имеющую вид

$$\ell(\mu, \nu) = \inf\big\{\varepsilon > 0\colon\ F_\mu(x-\varepsilon) - \varepsilon \leqslant F_\nu(x) \leqslant F_\mu(x+\varepsilon) + \varepsilon, \quad \forall x \in \mathbb{R}\big\}, \tag{1.56}$$

с которой слабая сходимость превращается в привычную сходимость в метрическом пространстве. С такой метрикой пространство $\mathcal{P}(\mathfrak{X}))$ само является польским пространством.

Приведем также определение метрики Леви-Прохорова для случая общих метрических пространств $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$. Пусть $A \in \mathcal{B}(\mathfrak{X})$ – борелевское множество. Для $\epsilon > 0$ определим множество

$$A^\epsilon = \bigcup_{p \in A} B_\epsilon(p), \tag{1.57}$$

как объединение шаров где $B_\epsilon(p)$ радиуса ϵ с центрами в каждой точке множества A. Тогда для мер $\mu, \nu \in \mathcal{P}(\mathfrak{X})$

$$\ell(\mu, \nu) = \inf\big\{\varepsilon > 0\colon\ \mu(A) \le \nu(A^\varepsilon) + \varepsilon \ \&\ \nu(A) \le \mu(A^\varepsilon) + \varepsilon, \quad \forall A \in \mathcal{B}(\mathfrak{X})\big\}, \tag{1.58}$$

Слабая сходимость[2] позволяет задать топологию в $\mathcal{P}(\mathbb{R})$, а следовательно и борелевскую сигма-алгебру $\mathcal{F}_\mathcal{P}$. Если $(\Omega, \mathcal{F}, \mathbb{P})$ — некоторое вероятностное пространство, то случайная вероятностная мера — это случайный элемент на этом пространстве, который принимает значения в $\mathcal{P}(\mathbb{R})$. Иными словами, случайная вероятностная мера — это измеримое отображение $(\Omega, \mathcal{F}) \to (\mathcal{P}(\mathbb{R}), \mathcal{F}_\mathcal{P})$, определённое правилом $\omega \mapsto \mu(\omega)$. К таким мерам применимы обычные понятия сходимости, принятые в теории вероятности, с которыми мы имели дело в разделе 1.

Определение 1.26. *Пусть $\mu, \mu_1, \mu_2 \dots$ — случайные вероятностные меры, определенные на одном вероятностном пространстве. Будем говорить, что последовательность* (m_n) слабо сходится *к* μ почти наверное (по вероятности, в среднем), *и писать*

$$\mu_n \xrightarrow[n\to\infty]{\text{п.н.}(\mathbb{P},\mathbb{E})} \mu, \tag{1.59}$$

[2]С точки зрения функционального анализа слабая сходимость мер — это *-слабая сходимость линейного функционала в двойственном пространстве.

тогда и только тогда, когда

$$\int_{\mathbb{R}} g(x)\, d\mu_n \xrightarrow[n\to\infty]{\text{п.н.}(\mathbb{P},\mathbb{E})} \int_{\mathbb{R}} g(x)\, d\mu, \quad \forall g(x) \in C_b(\mathbb{R}).$$

Пример 1.27. *Пусть $\lambda_1, \ldots, \lambda_n$ — выборка из некоторого распределения $F_\lambda(x)$ (то есть случайные величины $\lambda_1, \ldots, \lambda_n$ — независимы и одинаково распределены как некоторая случайная величина λ распределенная с мерой m_λ). Рассмотрим* эмпирическую выборочную меру

$$L_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k}. \tag{1.60}$$

Тогда

$$L_n \xrightarrow[n\to\infty]{\text{п.н.}} m_\lambda \tag{1.61}$$

Действительно, поскольку $\mathbb{I}_{\lambda<x}$ распределена как бернуллиевская случайная величина, то Усиленный Закон Больших Чисел даёт нам

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}_{\{\lambda_k<x\}} \xrightarrow[n\to\infty]{\text{п.н.}} F_\lambda(x). \tag{1.62}$$

Из поточечной почти наверное сходимости последовательности функций распределения следует слабая сходимость мер почти наверное. Предел (1.61) *можно интерпретировать как пример Закона Больших Чисел для эмпирической спектральной меры диагональной матрицы, элементы которой являются одинаково распределенными случайными величинами.*

Глава 2

Метод моментов и закон Вигнера

Обсуждая ЗБЧ и ЦПТ, мы имели дело с независимыми случайными величинами. В дальнейшем мы будем пытаться высказывать подобного рода утверждения, скорее, для наборов зависимых величин. Оказывается, существуют широкие классы систем, демонстрирующих на больших масштабах характерное поведение, которое не зависит от микроскопических подробностей, и первым таким классом, который мы рассмотрим, станут случайные матрицы.

Впервые случайные матрицы возникли в физике при изучении спектров ядер больших атомов. Несмотря на то, что задача диагонализации гамильтонианов больших взаимодействующих систем чрезвычайно сложна и не решается точно, оказалось, что статистически уровни энергии ведут себя как собственные значения некоторого класса случайных матриц. Позже выяснилось, что подобные утверждения справедливы и для других сложных систем, например, таких, как квантовые биллиарды, квантовые точки и др. С другой стороны, случайные матрицы представляют интерес и для математиков, например, потому что поведение, подобное собственным значениям случайных матриц, обнаруживается в структуре комплексных нулей ζ-функции Римана, которая, в свою очередь, тесно связана с распределением простых чисел.

Мы начнём с круга вопросов, касающихся изучения спектра случайных матриц. Как уже было упомянуто, изучение случайных матриц было мотивировано вопросом о типичном виде ядерных спектров, которые определяются наборами собственных чисел многомерных самосопряженных операторов. Основная идея состоит в том, что в достаточно общей ситуации типичный вид этого спектра не зависит от деталей отдельных матричных элементов, а определяется лишь симметриями оператора. Поэтому мы будем считать матричные элеметы случайными величинами, а цель будет заключаться в том, чтобы попытаться ответить на вопрос, как выглядит типичной спектр очень большой эрмитовой матрицы со случайными матричными элементами. Этот вопрос был впервые поставлен и решен математиком и физиком Юджном Вигнером в 1958 году сначала для матриц со случайными элементами, задаваемыми распределением Бернулли, а потом и для произвольных распределений.

Для простоты мы ограничимся случаем вещественных симметричных матриц, хотя те же утверждения относятся и к комплекснозначным эрмитовым матрицам, а также к вещественно-кватернионным эрмитовым матрицам.

Определение 2.1. *Матрицей Вигнера будем называть вещественную симметричную матрицу $W = (w_{ij}) \in \mathbb{R}^{n\times n}$, элементы $w_{ij} = w_{ji}$ которой на главной диагонали и выше являются независимыми случайными величинами, одинаково распредёленными выше* *главной диагонали, а также одинаково (но, возможно, по-другому), распределёнными на главной диагонали:*

$$w_{ij} = w_{ji} \sim \xi, \quad 1 \leqslant i < j \leqslant n; \qquad \mathbb{E}(\xi) = 0, \quad \mathbb{D}(\xi) = \sigma^2 < \infty; \tag{2.1}$$

$$w_{ii} \sim \eta, \qquad 1 \leqslant i \leqslant n; \qquad \mathbb{E}(\eta) = 0, \qquad \mathbb{D}(\eta) < \infty. \tag{2.2}$$

Для упрощения технических деталей на протяжении всего этого раздела мы будем также предполагать, что все моменты распределений случайных величин ξ и η конечны.

Вопрос, на который мы хотим ответить в этом разделе, звучит следующим образом: «Как выглядит типичный спектр большой Вигнеровской матрицы?» Ответ оказывается удобно сформулировать в виде Закона Больших Чисел для эмпирической спектральной меры вигнеровской матрицы, отнормированной так, чтобы спектр находился в ограниченной области.

Определение 2.2. *Пусть $M = M^\dagger \in \mathbb{C}^{n\times n}$ — эрмитова матрица*[1]*, обладающая собственными значениями $\{\lambda_1, \dots, \lambda_n\}$. Всякой такой матрице можно сопоставить* эмпирическую спектральную меру *(ЭСМ)*

$$L_M = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k} \tag{2.3}$$

Поскольку матрицы Вигнера — случайные, их ЭСМ — также случайная мера. Однако если выбрать правильный масштаб, то мы увидим, что чем больше размер матрицы, тем менее вид спектра, а следовательно и ЭСМ, зависит от исходного распределения её матричных элементов, и тем ближе они становятся к единому универсальному закону. В пределе для ЭСМ вигнеровских матриц, перемасштабированных так, чтобы при увеличении размера спектр оставался в ограниченной области, справедлив аналог ЗБЧ, а именно, ЭСМ стремится к полукруговому закону Вигнера.

Теорема 2.3. *[Теорема Вигнера]*
Пусть $(w_{ij})_{1\leqslant i<j}$ и $(w_{ii})_{i\geqslant 1}$ — последовательности независимых одинаково распределенных случайных величин, определенных на одном вероятностном пространстве, причём

$$\mathbb{E}w_{12} = \mathbb{E}w_{11} = 0, \tag{2.4}$$

$$\mathbb{D}w_{12} = \sigma^2, \quad \mathbb{D}w_{11} < \infty, \tag{2.5}$$

$$\mathbb{E}w_{12}^k < \infty, \quad \mathbb{E}w_{ii}^k < \infty, \quad k \geqslant 2, \tag{2.6}$$

а $W_n \in \mathbb{R}^{n\times n}$ — матрица Вигнера с матричными элементами $w_{ij} = w_{ji}$, $1 \leqslant i, j \leqslant n$. Введём перемасштабированную матрицу

$$M_n = \frac{1}{\sigma\sqrt{n}} \cdot W_n. \tag{2.7}$$

Тогда последовательность ЭСМ матриц M_n слабо сходится к полукруговому закону Вигнера почти наверное, то есть

$$L_{M_n} \xrightarrow[n\to\infty]{\text{п.н.}} \mu_{sc}, \tag{2.8}$$

[1]Знак † означает транспонирование и комплексное сопряжение.

где мера μ_{sc} задаётся плотностью

$$f_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2} \cdot \mathbb{I}_{\{|x|\leqslant 2\}}.$$

Доказательство. Для доказательства теоремы Вигнера мы будем использовать *метод моментов.* Суть его заключается в том, чтобы сначала доказать сходимость последовательности моментов ЭСМ к моментам полукругового закона, подобно тому, как мы это делали при доказательстве ЗБЧ в разделе 1. Если имеет место сходимость моментов почти наверное, то в силу ограниченности носителя полукругового распределения и леммы 1.18 мы будем иметь и слабую сходимость ЭСМ с вероятностью единица.

Для произвольной меры μ на $\mathbb{R}$ введём следующее обозначение

$$\langle \mu, f \rangle := \int_{\mathbb{R}} f(x)\, \mu(dx). \tag{2.9}$$

В частности $= m_k(\mu) = \langle \mu, x^k \rangle$ — это k-ый момент случайной величины, заданной мерой μ. Определим усредненную меру $\bar{L}_{M_n}$, которая характеризуется тем свойством, что для любой измеримой функции $f(x)$ выполнено соотношение

$$\langle \bar{L}_{M_n}, f(x) \rangle = \mathbb{E}_W \langle L_{M_n}, f(x) \rangle, \tag{2.10}$$

где нижний индекс W подчеркивает, что матожидание берется по отношению к мере на матрицах Вигнера.

Стратегия доказательства следующая. Сначала доказывается сходимость моментов в среднем к моментам полукругового распределения. Следом — сходимость моментов почти наверное к своим средним значениям. Таким образом, доказательство сводится к следующим трем утверждениям

(a) $\langle \mu_{sc}, x^{2k} \rangle = C_k$, где $C_k = \frac{1}{k+1}\binom{2k}{k}$ — числа Каталана,

(b) $\langle \bar{L}_{M_n}, x^k \rangle \xrightarrow[n\to\infty]{} \langle \mu_{sc}, x^k \rangle$,

(c) $\langle L_{M_n}, x^k \rangle \xrightarrow[n\to\infty]{\text{п.н.}} \langle \bar{L}_{M_n}, x^k \rangle$.

Отметим, что доказательство утверждения (a) было сформулировано в разделе 1 в виде упражнения 1.8.

Итак, покажем, что $\langle \bar{L}_{M_n}, x^k \rangle \xrightarrow[n\to\infty]{} \langle \mu_{sc}, x^k \rangle$. Прежде всего, заметим, что моменты ЭСМ — это следы степеней матрицы:

$$\langle L_{M_n}, x^k \rangle = \frac{1}{n}\sum_{i=1}^{n} \lambda_i^k = \frac{1}{n} \cdot \operatorname{Tr} M_n^k = \frac{1}{n(\sigma\sqrt{n})^k} \cdot \operatorname{Tr} W_n^k. \tag{2.11}$$

То же самое справедливо и для моментов усредненной меры

$$\langle \bar{L}_{M_n}, x^k \rangle = \frac{1}{n^{1+k/2}\sigma^k} \cdot \mathbb{E}_W\big(\operatorname{Tr} W_n^k\big), \tag{2.12}$$

с той лишь разницей, что справа стоит матожидание следа, или, в силу линейности следа, след мотожидания степеней матрицы W_n.

Выразим следы через матричные элементы и попробуем вычислить матожидания от полученных комбинаций матричных элементов. Чтобы понять, как ведёт себя величина $\langle \bar{L}_{M_n}, x^k \rangle$ в пределе $n \to \infty$, сначала вычислим следы $\operatorname{Tr} W_n^k$ при малых значениях k.

При $k = 1$, очевидно, получаем $\mathbb{E}_W\big(\operatorname{Tr} W_n\big) = \mathbb{E}_W(w_{11}) = 0$ и

$$\langle \bar{L}_{M_n}, x \rangle = 0. \tag{2.13}$$

При $k = 2$ справедлива следующая цепочка равенств:

$$\begin{aligned}\mathbb{E}_W\big(\operatorname{Tr} W_n^2\big) &= \mathbb{E}_W\left(\sum_{i=1}^{n}\sum_{k=1}^{n} w_{ik}w_{ki}\right) = \qquad (2.14)\\ &= \sum_{i\neq k}\mathbb{E}_W(w_{ik}^2) + \sum_{i=1}^{n}\mathbb{E}_W(w_{ii}^2) = n(n-1)\sigma^2 + n\,\mathbb{D}_W(w_{11}).\end{aligned}$$

Таким образом, при $n \to \infty$ мы получим

$$\langle \bar{L}_{M_n}, x^2 \rangle = \frac{1}{n}\,\mathbb{E}_W\big(\operatorname{Tr} M_n^2\big) = \frac{n(n-1)\sigma^2 + n\,\mathbb{D}_W(Y)}{n^2\sigma^2} \xrightarrow[n\to\infty]{} 1. \tag{2.15}$$

При $k = 3$ из независимости случайных величин w_{ik} для различных пар (i, k) имеем

$$\langle \bar{L}_{M_n}, x^3 \rangle = \frac{1}{n^{5/2}} \cdot \sum_{i=1}^{n}\mathbb{E}_W(w_{ii}^3) = \frac{n}{n^{5/2}} \cdot \mathbb{E}_W(w_{11}^3) \xrightarrow[n\to\infty]{} 0. \tag{2.16}$$

Что получится в общем случае? Как видно из разобранных примеров, для вычисления пределов моментов произвольного порядка нужно просуммировать вклады величин

$$\mathcal{W}(\pi) = \mathbb{E}\big(w_{i_1 i_2} w_{i_2 i_3} \dots w_{i_k i_1}\big) \tag{2.17}$$

с соответствующими весами, где $\pi = (i_1, \dots, i_k)$ — набор из k натуральных чисел, принимающих произвольные значения из множества $\{1, \dots, n\}$. Для решения этой задачи удобно использовать язык теории графов. Пусть $G = (V, E)$ — полный граф на n вершинах с петлей в каждой вершине.[2] Будем считать, что вершины пронумерованы числами от единицы до n, то есть $V = \{1, \dots, n\}$. Тогда каждому выражению вида (2.17) можно сопоставить замкнутый путь

$$\pi = \big(i_1 \to i_2 \to i_3 \to \dots \to i_k \to i_1\big) \tag{2.18}$$

с весом $\mathcal{W}(\pi)$ определённым формулой (2.17). Таким образом, выражение для моментов сводится к сумме

$$\langle \bar{L}_{M_n}, x^k \rangle = \frac{1}{n^{k/2+1}\sigma^k} \sum_{\pi(k)} \mathcal{W}(\pi(k)) \tag{2.19}$$

по путям $\pi(k)$ из k шагов на графе G.

Далее, заметим, что различные пути, которые переводятся друг в друга перенумерацией вершин, имеют один и тот же вес. Будем говорить, что пути π и π' принадлежат одному классу эквивалентности, $\pi \sim \pi'$, если $\pi' = \tau(\pi)$, где $\tau \in S_n$. Здесь мы имеем в виду,

[2] Полным графом называется граф, у которого каждая вершина соединена ребрами со всеми остальными. В нашем случае мы также добавляем в каждую вершину петлю — ребро, соединяющее вершину с самой собой.

что $\pi = (i_1 \to i_2 \to i_3 \to \ldots \to i_k \to i_1)$, $\pi' = (\tau(i_1) \to \tau(i_2) \to \tau(i_3) \to \ldots \tau(i_k) \to \tau(i_1))$, и, как следствие, $\mathcal{W}(\pi') = \mathcal{W}(\pi)$. Таким образом, суммирование по путям, сводится к суммированию по множеству классов эквивалентности $\mathcal{P} = \{\pi\}/\sim$:

$$\langle \bar{L}_{M_n}, x^k \rangle = \frac{1}{n^{k/2+1}\sigma^k} \sum_{\pi(k)\in\mathcal{P}} \|\pi(k)\| \cdot \mathcal{W}(\pi(k)), \tag{2.20}$$

где $\|\pi(k)\|$ — число путей в классе эквивалентности. Правда, пока остаётся вопрос, как пересчитать пути каждом таком классе. Оказывается, это легко сделать, если явно выделить классы эквивалентности путей, проходящих через фиксированное число вершин.

Обозначим через $\pi(k,m)$ путь из k шагов, проходящих в точности через m вершин, $\|\pi(k,m)\|$ — число элементов в его классе эквивалентности. Очевидно, что выполняется равенство $\|\pi(k,m)\| = n(n-1)\ldots(n-m+1) = A_n^m$. Поэтому имеет место соотношение

$$\langle \bar{L}_{M_n}, x^k \rangle = \frac{1}{n^{k/2+1}\sigma^k} \sum_{m=1}^{k} \sum_{\pi(k,m)\in\mathcal{P}} A_n^m \cdot \mathcal{W}(\pi(k,m)). \tag{2.21}$$

Замечательной особенностью выражения, стоящего под знаком суммы, является тот факт, что общее число вершин в графе G влияет лишь на известный коэффициент A_n^m, а всё остальное зависит только от k и m, то есть остается конечным в пределе $n \to \infty$.

Какие пути вносят вклад в эту сумму? Прежде всего, заметим, что в силу условия (2.4) ненулевой вклад дают только те пути, в которых каждое пройденное ребро встречается не менее двух раз. Следовательно, учитывать имеет смысл лишь те пути, в которых число использованных рёбер не превышает $[k/2]$. Пусть $\tilde{G} = (\tilde{V}, \tilde{E})$ — подграф графа G, для которого $\|\tilde{E}\| \leqslant [k/2]$ и $\tilde{V} = m$, а через рёбра и вершины которого проходит путь $\pi(k,m)$. Сколько в таком графе может быть вершин, то есть каково число m? Чтобы ответить на этот вопрос, сформулируем следующее известное из теории графов утверждение.

Лемма 2.4. *Пусть $G = (V,E)$ конечный связный граф. Тогда $\|V\| \leqslant \|E\| + 1$, причём равенство имеет место тогда и только тогда, когда G — дерево.*

Следовательно, число рёбер удовлетворяет неравенству $m \leqslant [k/2] + 1$, откуда

$$\langle \bar{L}_{M_n}, x^k \rangle = \frac{1}{n^{k/2+1}\sigma^k} \sum_{m=1}^{[k/2]+1} \sum_{\pi(k,m)\in\mathcal{P}} A_n^m \cdot \mathcal{W}(\pi(k,m)). \tag{2.22}$$

При $n \to \infty$ величина A_n^m растет как $A_n^m \asymp n^m$. В случае нечётных k это означает, что моменты стремятся к нулю. Если же k — чётное число, то в пределе выживают только слагаемые, соотвествующие $m = k/2 + 1$. Как следует из леммы 2.4, подграф $\tilde{G}$ в этом случае является деревом, причём пути $\pi(k, k/2+1)$ проходят по каждому ребру ровно два раза. Вес таких путей не зависит от класса эквивалентности и равен

$$\mathcal{W}(\pi(k, k/2+1)) = \sigma^k. \tag{2.23}$$

Таким образом, предельное выражение для чётных моментов равно числу классов эквивалентности путей $\pi(k, k/2+1)$, то есть

$$\langle \bar{L}_{M_n}, x^k \rangle = \begin{cases} \|\{\pi(k, k/2+1) \in \mathcal{P}\}\|, & k \in 2\mathbb{N} \\ 0, & k \in 2\mathbb{N}+1 \end{cases} \tag{2.24}$$

Остается предъявить процедуру построения неэквивалентных путей из k шагов, обходящих деревья с $(k/2+1)$ вершинами так, что каждое ребро проходится два раза. Будем считать, что мы обходим вершины $\tilde{V} = (1, \dots, k/2+1)$, а итоговый результат представляет собой последовательность вершин $(v(0), \dots, v(k))$. Без ограничения общности можно положить $v(0) = 1$ и $v(1) = 2$, то есть первый шаг делается вдоль ребра $1 \to 2$. В дальнейшем, поскольку пройденные ребра не должны образовывать циклов, на каждом шаге мы можем либо перейти в ещё не посещённую вершину, либо вернуться в вершину, в которой были на предыдущем шаге. То есть имеется две возможности:

(i) $v(i) \to v(i+1) = \max(v(0), \dots, v(i)) + 1,$

(ii) $v(i) \to v(i-1).$

Путь заканчивается возвращением на шаге k в исходную вершину: $v(k) = v(0) = 1$.

Далее, сопоставим последовательности $(v(0), \dots, v(k))$ последовательность из плюс-минус единиц $u = (u_1, \dots, u_k)$, заданную следующим правилом: для всех $j = 1, \dots, k$ шагу $v(j) \to v(j+1)$ типа (i) — переход в новую вершину — ставим в соответствие $u_{j+1} = 1$, а с шагом типа (ii) — возвращение — соотносим $u_{j+1} = -1$,. Ясно, что сумма членов такой последовательности на каждом шаге неотрицательна, то есть

$$\sum_{i=1}^{l} u_i \geqslant 0, \qquad 1 \leqslant l \leqslant k,$$

а полная сумма равна нулю: $u_1 + \ldots + u_k = 0$. Поэтому при замене единиц на левые скобки, а минус единиц — на правые получается *правильная скобочная последовательность*, в которой каждой левой скобке соответствует правая, а любой начальный кусок содержит не меньше левых скобок, чем правых. Также, по последовательности u можно построить *путь Дика* — ломаную линию на плоскости, представляющую собой путь из точки $(0,0)$ в точку $(2k, 0)$, составленный из векторов $(1, u_i)$, а потому не опускающийся ниже оси абсцисс Ox (рис. 2.1).

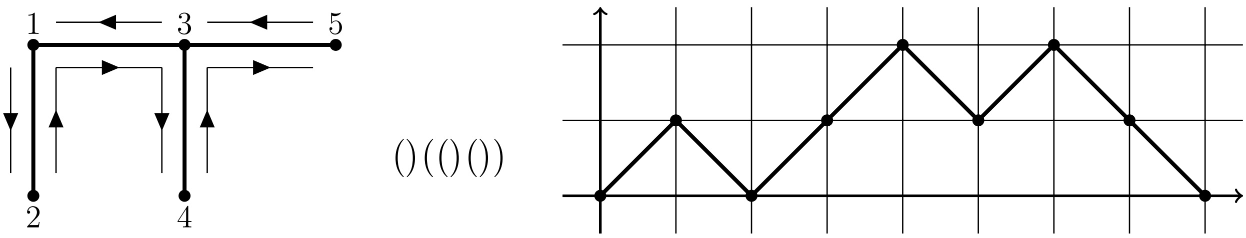

Рис. 2.1:

Итак, каждому обходу дерева можно сопоставить правильную скобочную последовательность или путь Дика. Очевидно и обратное: по каждому пути Дика можно построить представитель класса эквивалентности замкнутых путей из k шагов, проходящих ровно два раза по каждому ребру некоторого дерева $\tilde{G}$ с $\|\tilde{V}\| = k/2+1$ вершинами. Хорошо известно, что количество путей Дика, а следовательно, и выражение для моментов чётного порядка k даётся числом Каталана $C_{k/2}$. Таким образом, сходимость моментов в среднем, то есть утверждение (b), доказана.

Доказательство утверждения (c) мы оставляем читателю в виде упражнения. Обозначим лишь общий план действий. Чтобы убедиться, что имеет место сходимость почти наверное, достаточно исследовать поведение дисперсии $\mathbb{D}\langle L_{M_n}, x^k\rangle$. Доказательство снова проводится методом моментов, но на этот раз необходимо провести суммирование по парам путей из k шагов. Сначала имеет смысл удостовериться в том, что пары путей, не имеющих общих рёбер, не дают вклада в дисперсию. Этот факт помогает установить, что дисперсия стремится к нулю при $n \to \infty$, откуда с помощью неравенства Чебышёва следует сходимость по вероятности. Если же явно выделить пары путей, ответственных за ведущий ненулевой вклад, то можно показать, что дисперсия убывает достаточно быстро, а потому можно воспользоваться неравенством Маркова и леммой Бореля-Кантелли для доказательства сходимости почти наверное. $\square$

Упражнение 2.5. *Докажите, что число правильных скобочных последовательностей из $2k$ скобок совпадает с количеством путей Дика длины $2k$ и равно k-ому числу Каталана C_k.*

Упражнение 2.6. *Проведите соответствующие вычисления и убедитесь, что в условиях теоремы Вигнера имеет место сходимость $\langle L_{M_n}, x^k\rangle \xrightarrow[n\to\infty]{\text{п.н.}} \langle \bar{L}_{M_n}, x^k\rangle$.*

В приведённом доказательстве мы построили биекцию между классами эквивалентности обходов деревьев и путями Дика или правильными скобочными последовательностями. Дадим ещё пару комбинаторных интерпретаций этого результата.

Определение 2.7. *Дерево с помеченной вершиной* — корнем — *называется* корневым деревом *или* ориентированным деревом. Потомками *данной вершины дерева, называются те вершины, для которых любой путь, соединяющий их с корнем, проходит через данную вершину.* Упорядоченным *называется дерево, у которого множество* сыновей *(ближайших потомков) каждой вершины упорядочено. Эквивалентно можно говорить о планарном дереве: ориентация дерева на плоскости задает естественный порядок его вершин.*

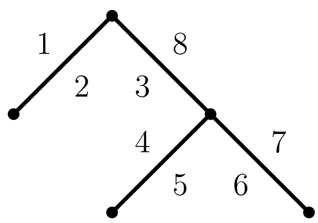

Рис. 2.2:

Заметим, что любое дерево в совокупности с порядком обхода его вершин задаёт упорядоченное корневое дерево или, эквивалентно, планарное дерево (рис. 2.2). Планарные деревья, таким образом, состоят в биекции с путями Дика (ср. с рис. 2.1).

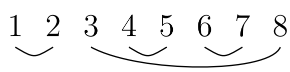

Рис. 2.3:

;

Ещё один важный комбинаторный объект, для которого можно установить биективное соответствие с путями Дика, правильными скобочными последовательностями

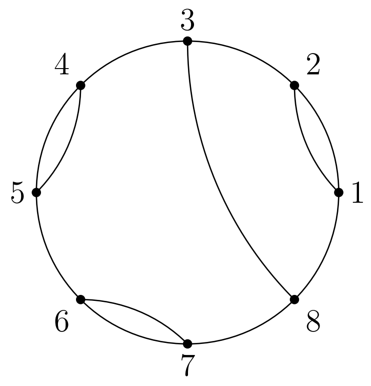

Рис. 2.4:

и планарными корневыми деревьями — это так называемое неперекрёстное разбиение. Оно может быть построено следующим образом. Перенумеруем шаги обходящего дерево пути числами от 1 до k. Выпишем эти числа подряд и соединим снизу дугами пары чисел, которые соответствуют шагам вдоль одного и того же ребра, сделанным с разных сторон (рис. 2.3). Ясно, что так как эта конструкция построена по обходу планарного дерева, дуги можно провести так, чтобы они не пересекались.

Определение 2.8. *Разбиение упорядоченного множества* $\{1,\ldots,k\}$ *на подмножества называется* неперекрёстным разбиением*, если для любых четырёх его элементов* $1 \leqslant a < b < c < d \leqslant k$*, а также для любых двух подмножеств разбиения* A *и* B *из условия* $a, c \in A$ *и* $b, d \in B$ *следует, что* $A = B$*. Непересрёстное разбиение на множества из двух элементов будем называть* попарным неперекрёстным разбиением или неперекрестным парасочитанием.

Очевидно, разбиение будет неперекрёстным, если его можно изобразить, соединив между собой элементы каждого подмножества непересекающимися дугами, нарисованными под строчной записью множества $\{1,\ldots,2k\}$, как это было сделано на рис. 2.3. Эквивалентно, можно расположить элементы множества на окружности; в этом случае неперекрёстное разбиение изображается непересекающимися хордами, соединяющими элементы подмножеств (рис. 2.4). Как видно из рассуждения выше, моменты полукругового распределения определяются числом попарных неперекрёстных разбиений множества из $2k$ элементов.

Напомним, что суммирование по разбиениям уже встречалось нам в разделе 1 при обсуждении связи между моментами и кумулянтами в обычной теории вероятности. В частности, в силу того, что у стандартного стандартного нормального распределения $\mathcal{N}(0,1)$ лишь один кумулянт $c_2 = 1$ не равен нулю, момент порядка $2k$ равен числу парных разбиений (парасочитаний) в множестве из $2n$ элементов. Поэтому в некотором смысле закон Вигнера аналог распределения Гаусса. В дальнейшем мы увидим, что теории свободной вероятности суммирование по неперекрёстным разбиениям играет ту же роль, что суммирование по любым разбиениям, которое возникало выше в связи с переходом от моментов к кумулянтам.

Глава 3

Метод преобразования Стилтьеса и закон Марченко-Пастура

В этом разделе мы изучим спектр случайных матриц другого типа, выборочных ковариационных матриц, построенных по случайной выборке из некоторого многомерного распределения. Такие матрицы возникают приложениях к задачам математической статистики, и вопрос о виде их спектра был впервые поставлен задолго до того, как Ю. Вигнером были введены в рассмотрение случайные эрмитовы матрицы. Впервые задача о точном распределении спектра выборочной коварицонной матрицы, построенной по выборке из многомерного нормального распределения, была решена Джоном Вишертом в 1928 году. В этом разделе мы обсудим более поздний результат, о асимптотическом виде спектра больших выборочных коварицонных матриц, построенных по выборке из многомерных распределений широкого класса. Эта задача была решена советскими математиками А. В. Марченко и Л.А. Пастуром в 1967-ом году. Мы увидим, что как и в случае закона Вигнера, асимптотический вид спектра не зависит от деталей распределения матричных элементов.

3.1 Статистики многокомпонентных выборок и сингулярные числа матриц.

Пусть $\mathbf{x} = (x_1, \ldots, x_n)^T$ — n-компонентная случайная величина, заданная некоторым n-мерным распределением, и пусть имеется случайная выборка

$$X_m = (\mathbf{x}_1, \ldots, \mathbf{x}_m) = \begin{pmatrix} x_{1,1} & \ldots & x_{1,m} \\ \cdots & \cdots & \cdots \\ x_{n,1} & \ldots & x_{n,m} \end{pmatrix}$$

размера m из такого распределения. Основными статистиками, традиционно характеризующими выборку, являются вектор выборочного среднего

$$\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} = \frac{1}{m}(\mathbf{x}_1 + \cdots + \mathbf{x}_m),$$

и выборочная коварицонная матрица с матричными элементами

$$q_{ij} = \frac{1}{m-1}\sum_{i=1}^{m}(x_{i,k} - \overline{x}_i)(x_{j,k} - \overline{x}_j).$$

Эти статистики можно использовать для оценки параметров исходного распределения, в частности анализируя их собственные значения. Существуют также много других различных статистик. В дальнейшем мы будем интересоваться задачей о спектре похожей матрицы

$$R_m = \frac{1}{m} X_m X_m^{\dagger}, , \tag{3.1}$$

которую мы также будем называть коварицонной выборочной матрицей, построенной по n-мерному распределению, в котором все n компонент независимы и одинаково распределены. Её матричные элементы

$$r_{ij} = \frac{1}{m}\sum_{i=1}^{m} x_{i,k} x_{j,k}^{*},$$

которые для простоты определены без сдвига элементов выборки на выборочное среднее. Как, будет замечено ниже, такой сдвиг не меняет асимптотического вида спектра. Звездочка над матричным элементом означает комплексное сопряжение, а знак † над матрицей — эрмитово сопряжение, полученное транспонированием и комплексным сопряжением матрицы. Здесь мы имеем в виду, что мы будем рассматривать не только вещественные, но и комплекснозначные случайные матрицы X. Тогда матрица определенная в (3.1) будет эрмитовой, и следовательно имеющей чисто вещественный спектр. Нетрудно проверить, что она также будет неотрицательно определена, т.е.

$$(\mathbf{v}, XX^{\dagger}, \mathbf{v}) \geq 0 \tag{3.2}$$

для любых векторов $\mathbf{v}$, где под скалярным произведением понимается эрмитово скалярное произведение

$$(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} u_i^{*} v_i. \tag{3.3}$$

Все собственные значения такой матрицы неотрицательны.

В этом разделе мы попробуем ответить на вопрос, каков типичный вид спектра такой матрицы, когда она очень большая. Также, об этой задаче можно думать, как о задаче о распределении квадратов сингулярных чисел прямоугольной матрицы с независимыми одинаково распределенными матричными элементами. Пусть X прямоугольная матрица. Для любой прямоугольной матрицы можно построить полярное разложение

$$X = |X|U,$$

где U унитарная матрица, $UU^{\dagger} = I$, а матрица $|X| = \sqrt{XX^{\dagger}}$ — неотрицательная эрмитова матрица (эрмитова матрица с неотрицательными собственными значениями). Собственные числа матрицы $|X|$ называются сингулярными числами матрицы X. Следовательно собственные числа матрицы $R = XX^{\dagger}$ ни что иное, как квадраты сингулярных чисел.

Упражнение 3.1. *Пользуясь результатами предыдущего раздела найдите асимптотическую спектральную плотность квадратов сингулярных чисел вигнеровских матриц. Хотя хотя матричные элементы матрицы Вигнера сверху и снизу от главной диагонали не независимы, как мы увидим ниже, асимптотическое распределение ее сингулярных чисел точно такое же как для квадратной матрицы с независимыми одинаково распределенными элементами.*

Прежде чем обратиться к решению намеченной задачи, нам потребуется ввести несколько новых инструментов.

3.2 Преобразование Стилтьеса

Мощный инструмент для работы с вероятностными мерами, особенно эффективный для рабыты с ЭСМ случайных матриц, — преобразование Стилтьеса.

Определение 3.2. *Пусть ξ — случайная величина, обладающая функцией распределения $F_\xi(x)$. Тогда на комплексной области $\mathbb{C} \setminus \mathrm{Supp}(dF_\xi(x))$ мы можем определить функцию*

$$s_\xi(z) = \int\limits_{-\infty}^{+\infty} \frac{1}{\lambda - z}\, dF_\xi(\lambda), \tag{3.4}$$

которая называется преобразованием Стилтьеса *меры $dF_\xi(x)$.*

Из определения преобразования Стилтьеса следуют его следующие простые свойства

1. Функция $s_\xi(z)$ аналитична в верхней полуплоскости $\mathbb{C}^+$.
2. Поведение на бесконечности: $\lim\limits_{z\to\infty} e^{\mathrm{i}\epsilon} z s_\xi(e^{\mathrm{i}\epsilon} z) = -1, \quad \epsilon \neq 0.$
3. Знакопостоянная мнимая часть: $\Im s_\xi(z) > 0$ при $z \in \mathbb{C}^+$,
4. Ограниченность: $|s_\xi(z)| \leq 1/\Im z$, $z \in \mathbb{C}^+$
5. Симметрия относительно отражения: $s_\xi(\overline{z}) = \overline{s_\xi(z)}$.

Справедливо и обратное утверждение.

Утверждение 3.3. *Если функция $s(z)$ имеет свойства 1-5, то существует случайная величина ξ, такая что $s(z) = s_\xi(z)$ — преобразование Стилтьеса ее распределения.*

Чтобы восстановить распределение по преобразованию Стилтьеса, можно воспользоваться следующей формулой обращения.

Утверждение 3.4. *Пусть x — точка непрерывности функции распределения $F_\xi(x)$. Тогда имеет место обратное преобразование Стилтьеса:*

$$F_\xi(x) = \pm \lim_{y\to 0^+} \frac{1}{\pi} \int\limits_{-\infty}^{x} \mathrm{Im}\big(s_\xi(z \pm \mathrm{i}y)\big)\, dz. \tag{3.5}$$

Доказательство. Чтобы убедиться в справедливости формулы (3.5), заметим, что

$$\frac{1}{\pi}\int\limits_a^b \mathrm{Im}\big(s_\xi(z+\mathrm{i}y)\big)\,dz = \frac{1}{\pi}\int\limits_a^b \frac{dz}{2\mathrm{i}}\int\limits_{-\infty}^{+\infty}\left(\frac{1}{\lambda-(z+iy)}-\frac{1}{\lambda-(z-iy)}\right)dF_\xi(\lambda) =$$

$$= \frac{1}{\pi}\int\limits_a^b dz\int\limits_{-\infty}^{+\infty}\frac{y}{(\lambda-z)^2+y^2}\,dF_\xi(\lambda) = \frac{1}{\pi}\int\limits_{-\infty}^{+\infty}dF_\xi(\lambda)\int\limits_a^b\frac{y\,dz}{(\lambda-z)^2+y^2} =$$

$$= \frac{1}{\pi}\int\limits_{-\infty}^{+\infty}\left(\mathrm{arctg}\left(\frac{b-\lambda}{y}\right)-\mathrm{arctg}\left(\frac{a-\lambda}{y}\right)\right)dF_\xi(\lambda).$$

Воспользовавшись свойством $\lim\limits_{y\to 0^+}\mathrm{arctg}\left(\dfrac{x}{y}\right)=\dfrac{\pi}{2}\cdot\mathrm{sign}(x)$, получаем

$$\frac{1}{\pi}\int\limits_a^b \mathrm{Im}\big(s_\xi(z+\mathrm{i}y)\big)\,dz \xrightarrow[y\to 0^+]{} \frac{1}{\pi}\int\limits_{-\infty}^{+\infty}\frac{\pi}{2}\big(\mathrm{sign}(b-\lambda)-\mathrm{sign}(a-\lambda)\big)\,dF_\xi(\lambda) = F(b)-F(a).$$

Остаётся устремить a к $-\infty$. Аналогично доказывается второе равенство. □

Замечание 3.5. *Из доказательства также не трудно видеть, что если x– точка разрыва,*

$$F_\xi(x)-F_\xi(x^-)=\mu_\xi(x)\neq 0,$$

то формула обратного преобразования Стилтьеса принимает вид

$$F_\xi(x)-\frac{1}{2}\mu_\xi(x) = \pm\lim_{y\to 0^+}\frac{1}{\pi}\int\limits_{-\infty}^{x}\mathrm{Im}\big(s_\xi(z\pm\mathrm{i}y)\big)\,dz \tag{3.6}$$

Следствие 3.6. *Из этого утверждения в частности следует, что если случайная величина ξ обладает плотностью $f_\xi(x)$, формула обращения сводится к вычислению мнимой части преобразования Стилтьеса.*

$$f_\xi(x) = \frac{1}{\pi}\lim_{y\to 0^+}\Im s_\xi(x+\mathrm{i}y) \tag{3.7}$$

Если предел в правой части существует, то функция распределения дифференцируема в точке x и дается пределом мнимой части преобразования Стилтьеса из верхней полуплоскости. Эта формула есть частный случай одной из формул Племеля-Сохоцкого (см. приложение 4.1).

Следствие 3.7.

Подобно тому, как преобразование Лапласа распределения вероятности играет роль экспоненциальной производящей функции моментов, преобразование Стилтьеса — обычная производящая функция моментов.

Утверждение 3.8. *Пусть случайная величина ξ имеет компактный носитель,* $\mathrm{Supp}(dF_\xi(x))\subset$ *$[a;b]$, где $-\infty<a<b<\infty$. Тогда для $z\in\mathbb{C}$, удовлетворяющих условию $|z|>max(|a|,|b|)$, справедливо соотношение*

$$s_\xi(z) = -\frac{1}{z}\sum_{n=0}^{\infty}\frac{m_n(\xi)}{z^n}. \tag{3.8}$$

Так же как преобразование Фурье, оно же характеристическая функция, преобразование Стилтьеса представляет собой удобный инструмент для доказательства слабой сходимости вероятностных мер.

Теорема 3.9. *[Теорема о непрерывности преобразования Стилтьеса]*
Пусть $\mu, \mu_1, \mu_2, \ldots$ — вероятностные меры на $\mathbb{R}$. Для того, чтобы слабая сходимость $\mu_n \underset{n\to\infty}{\Longrightarrow} \mu$ имела место необходимо и достаточно, чтобы имела место сходимость $s_{\mu_n}(z) \underset{n\to\infty}{\Longrightarrow} s_\mu(z)$.[1]

Это утверждение можно распространить и на слабую сходимость случайных мер почти наверное. Однако имеется небольшая тонкость, которая требует некоторых дополнительных рассуждений. Дело в том, что даже если доказана поточечная сходимость преобразований Стилтьеса почти наверное, т.е.

$$\mathbb{P}\{\omega \in \Omega : s_{\mu_n}(z,\omega) \underset{n\to\infty}{\longrightarrow} s_\mu(z,\omega)\} = 1, \quad \forall z \in \mathbb{C}^+, \tag{3.9}$$

для каждой точки z все еще имеется множество меры ноль, где сходимость не гарантирована. Таких точек континуум, а мера множества, полученного объединением континуума множеств меры ноль, вообще говоря не обязана быть нулем. Для разрешения этой проблемы достаточно воспользоваться свойством аналитичности преобразования Стилтьеса и теоремой Витали о сходимости.

Теорема 3.10 (Витали)**.** *Пусть $f_1, f_2, \ldots$ — последовательность функций, равномерно ограниченных и аналитических в некотором связном открытом подмножестве D комплексной плоскости $\mathbb{C}$, и последовательности $(f_n(z))$ сходятся поточечно в некотором подмножестве множества D, содержащем хотя бы одну предельную точку в D. Тогда существует функция $f(z)$, аналитическая в D, такая что*

$$f_n(z) \underset{n\to\infty}{\longrightarrow} f(z), \quad \forall z \in D.$$

Предположим, что (3.9) доказано. Возьмем в качестве множества D подмножество верхней полуплоскости $\{z \in \mathbb{C}^+ : \Im z > 1/a\}$, где $|s_{\mu_n}(z,\omega)| \le a$. Выберем в нем последовательность $(z_m)_{m\in\mathbb{N}}$, сходящуюся в некоторой к некоторой точке в D. Каждая из последовательностей $(s_{\mu_n}(z_m,\omega))_{n\in\mathbb{N}}$ сходится для всех ω кроме быть может множеств меры ноль. Множество событий где расходится хотя бы одна из последовательностей при при каком-то $m = 1, 2, \ldots$, будучи счетным объединением множеств нулевой меры, также имеет нулевую меру. Значит последовательность преобразований Стилтьеса сходится с вероятностью единица для всех z_m одновременно, что в силу теоремы Витали гарантирует одновременную сходимость почти наверное во всех точках верхней полуплоскости, откуда имеем

[1] Так же как и в случае теоремы Леви, в утверждении о достаточности важен тот факт, что предельная функция есть преобразование Стилтьеса именно вероятностной меры. В общем случае сходимость преобразований Стилтьеса влечет за собой не слабую, а так называемую грубую сходимость (vaugue convergence), которая также как в (1.28) определяется через сходимость интегралов, в которых, однако, в качестве пробных функций вместо непрерывных ограниченных функций используются непрерывные функции с компатным носителем. При грубой сходимости предел последовательности вероятностных мер не обязан быть вероятностной мерой, так как часть массы может уйти на бесконечность. Однако если известно, что предел — это также вероятностная мера, слабая равносильна грубой.

Теорема 3.11. *Пусть* $(\Omega, \mathcal{F}, \mathbb{P})$ *– вероятностное пространство, а* $\mu(\omega), \mu_1(\omega), \mu_2(\omega), \ldots$ *— случайные вероятностные меры на* $\mathbb{R}$, *т.е. измеримые отображения из* Ω *в пространство борелевских вероятностных мер на* $\mathbb{R}$. *Для слабой сходимости почти наверное*

$$\mu_n(\omega) \xrightarrow[n\to\infty]{\text{п.н.}} \mu(\omega)$$

необходима и достаточна сходимость

$$s_{\mu_n}(z,\omega) \xrightarrow[n\to\infty]{\text{п.н.}} s_\mu(z,\omega).$$

для всех $z \in \mathbb{C}^+$.

3.3 Преобразование Стилтьеса эмпирической спектральной меры.

Перейдём теперь к вопросу о том, как преобразование Стилтьеса может быть использовано применительно к матрицам. Рассмотрим произвольную *эрмитову матрицу* $R \in \mathbb{C}^{n\times n}$, то есть матрицу, удовлетворяющую соотношению $R = R^\dagger$. Такую матрицу можно диагонализовать с помощью некоторой унитарной матрицы U,

$$R \to URU^\dagger = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

где $\lambda_1, \ldots, \lambda_n$ — собственные значения матрицы R. Введём соответствующие им эмпирическую спектральную меру и эмпирическую функцию распределения:

$$L_R = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k}, \qquad F_R(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}_{x\geqslant\lambda_k}. \tag{3.10}$$

Тогда соответствующее мере L_R преобразование Стилтьеса имеет вид

$$s_R(z) = \int\limits_{-\infty}^{+\infty} \frac{1}{\lambda - z}\, dF_R(\lambda) = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\lambda_k - z} = \frac{1}{n}\,\mathrm{Tr}\big(R - zI_n\big)^{-1}, \tag{3.11}$$

где $I_n \in \mathbb{C}^{n\times n}$ — единичная матрица размера n.

Рассмотрим теперь квадратные матрицы, полученные как произведение двух прямоугольных матриц. Очевидно, что размер такой квадратной матрицы зависит от того, в каком порядке стоят сомножители. В то же время при перестановке сомножителей ранг произведения не меняется. Что можно сказать о собственных значениях в этом случае?

Лемма 3.12. *Пусть* $m \leqslant n$, *а матрицы* $A \in \mathbb{C}^{n\times m}$ *и* $B \in \mathbb{C}^{m\times n}$ *таковы, что их произведения являются эрмитовыми матрицами:* $(AB)^\dagger = AB$, $(BA)^\dagger = BA$. *Тогда* n *собственных значений матрицы* AB *содержат* m *собственных значений матрицы* BA *и* $n - m$ *нулевых собственных значений.*

Доказательство. Заметим, сначала, что если $m = n$, то есть матрицы A и B квадратные и A — обратимая матрица, то выполняется равенство $BA = A^{-1}(AB)A$, то есть одно произведение получается из другого сопряжением, и следовательно их собственные значения совпадают.

Если обе матрицы вырождены, можно немного изменить некоторые их матричные элементы на малую величину ϵ, чтобы сделать их невырожденными, и использовать предыдущий аргумент. Поскольку собственные значения, это корни характеристического многочлена, равенство двух многочленов сохранится и в пределе $\epsilon \to 0$, некоторые собственные значения обратятся в ноль. Таким образом утверждение доказано для квадратных матриц.

Пусть теперь $m < n$. Можно достроить матрицы A и B до квадратных матриц $\tilde{A}$ и $\tilde{B}$ размера $n \times n$, добавив в них нули, и воспользоваться утверждением, доказанным для квадратных матриц. Действительно, матрица $\tilde{A}\tilde{B}$ совпадает с матрицей AB, а матрица $\tilde{A}\tilde{B}$ имеет блочный вид

$$\tilde{B}\tilde{A} = \begin{pmatrix} BA & 0 \\ 0 & 0 \end{pmatrix}.$$

Поскольку собственные значения квадратных матриц $\tilde{A}\tilde{B}$ и $\tilde{B}\tilde{A}$ совпадают, получим требуемое утверждение. □

Два важных следствия будут использоваться далее.

Следствие 3.13. *Имеет место соотношение для определителей*

$$\det(\lambda I_n - AB) = \lambda^{n-m} \det(\lambda I_m - BA) \tag{3.12}$$

Его важный частный случай при $\lambda = 1$ справедлив не только для матриц но и для бесконечномерных операторов в гильбертовых пространствах.

Следствие 3.14. *Для всех $z \in \mathbb{C} \setminus \mathbb{R}$ справедливо соотношение*

$$\frac{m}{n} \cdot s_{BA}(z) = s_{AB}(z) + \frac{n-m}{n} \cdot \frac{1}{z}. \tag{3.13}$$

Это соотношение справедливо независимо от того, какое из чисел n, m больше.

Доказательство. Поскольку собственные значения для AB — это собственные значения для BA, а также ещё $n - m$ нулевых собственных значений, то добавляя соответствующее им слагаемое $(n-m)\delta_0$ к мере L_{BA}, получаем требуемое утверждение. Равенство не изменится, если одновременно поменять местами n с m и AB с BA. □

Теперь мы готовы сформулировать теорему Марченко-Пастура.

Теорема 3.15. *[Закон Марченко-Пастура] Пусть $(x_{ij})_{(i,j)\in\mathbb{N}^2}$ – двойная последовательность независимых одинаково распределенных комплексных случайных величин, определенных на одном вероятностном пространстве, с независимыми мнимыми и вещественными частями, таких что*

$$\mathbb{E}x_{11} = 0, \quad \mathbb{E}|x_{11}|^2 = \mathbb{E}(\Im x_{11})^2 + \mathbb{E}(\Re x_{11})^2 = 1, \quad \mathbb{E}\,|x_{11}|^s < \infty, \quad s \le 8,$$

а $X^{(n,m)} \in \mathbb{C}^{n\times m}$ составленные из $n \times m$ первых из них прямоугольные матрицы

$$X^{(n,m)} = \frac{1}{\sqrt{m}}\{x_{ij} : 1 \le i \le n, 1 \le j \le m\},$$

нормированные таким образом, чтобы дисперсия матричных элементов была $\mathbb{D}(X_{ij}^{n\times m}) = 1/m$.

Пусть $(m_k)_{k\in\mathbb{N}}$ *и* $(n_k)_{k\in\mathbb{N}}$ *– последовательности натуральных чисел, такие что* $n_k, m_k \xrightarrow[k\to\infty]{} \infty$*, так что* $n_k/m_k \xrightarrow[k\to\infty]{} c \in (0;\infty)$*. Тогда, последовательность ЭСМ* $(L_{R_k})_{k\in\mathbb{N}}$ *случайных эмпирических спектральных мер матриц* $R_k = X^{(n_k,m_k)}\cdot\left(X^{(n_k,m_k)}\right)^\dagger$ *сходится слабо почти наверное,*

$$L_{R_k} \xRightarrow[k\to\infty]{\text{п.н.}} \mu_{MP}, \tag{3.14}$$

к закону Марченко-Пастура

$$\mu_{MP} = \mu_{MP}^{\mathrm{ac}} + (1 - c^{-1})\delta_0(x)\mathbb{I}_{\{c>1\}}, \tag{3.15}$$

где μ_{MP}^{ac} *– абсолютно непрерывная относительно меры Лебега часть, задаваемая плотностью*

$$f_{MP}(x) = \frac{\sqrt{(a_+ - x)(x - a_-)}}{2\pi xc}\mathbb{I}_{\{a_-<x<a_+\}}, \quad a_\pm = (1 \pm \sqrt{c})^2. \tag{3.16}$$

Упражнение 3.16. *Используя утверждение теоремы 3.15, покажите, что последовательность эмпирических случайных мер, построенных по сингулярным числам случайных комплексных квадратных матриц с независимыми одинаково распределенными матричными элементами, при увеличении размера матриц почти наверное слабо сходится к половине полукругового закон Вигнера .*

Согласно теореме 3.11 для доказательства слабой сходимости ЭСМ достаточно доказать поточечную сходимость её преобразования Стилтьеса. Это доказательство состоит из двух шагов.

1. Для любого фиксированного $z \in \mathbb{C}^+$, $s_{R_k(z)} - \mathbb{E}s_{R_k(z)} \xrightarrow[k\to\infty]{\text{п.н.}} 0$.

2. Для любого фиксированного $z \in \mathbb{C}^+$, $\mathbb{E}s_{R_k(z)} \xrightarrow[k\to\infty]{} s_{MP}(z)$, где

$$s_{MP}(z) = \frac{1 - c - z + \sqrt{(z - a_+)(z - a_-)}}{2cz}, \quad a_\pm = (1 \pm \sqrt{c})^2, \tag{3.17}$$

   преобразование Стилтьеса распределения Марченко-Пастура.

Упражнение 3.17. *Докажите, что преобразование Стильтьеса распределения Марчеко-Пастура (3.15,3.16) дается формулой (3.17).*

Первый шаг, доказательство сходимости $s_{R_k}(z)$ к своему среднему, можно можно провести так же, как с использованием неравенства Мак Диармида доказывалась сходимость в примере 1.25. А именно, сходимость преобразования Стилтьеса ЭСМ матриц R_k следствие того факта, что при изменении ЭСМ вызванное изменением любого из столбцов матрицы $X^{(n_k,m_k)} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ мало. Это общее важное свойство устойчивости собственных значений матрицы по отношению к возмущениям малого ранга. Обсудим его детальнее перед тем, как перейти к техническим деталям доказательства сходимости.

3.4 Теорема о перемежаемости и возмущение спектра

Теорема 3.18 (Теорема Коши о перемежаемости.). *Пусть $A = A^\dagger \in \mathbb{C}^{n\times n}$ — эрмитова матрица, а матрица $A^{(1)}$ получена из матрицы A вычёркиванием первой строки и первого столбца. Пусть, далее, $\lambda_1 \le \ldots \le \lambda_n$ — собственные значения матрицы A, а $\lambda_1^{(1)} \le \ldots \le \lambda_{n-1}^{(1)}$ — собственные значения матрицы $A^{(1)}$. Тогда справедливы неравенства*

$$\lambda_1 \le \lambda_1^{(1)} \le \lambda_2 \le \ldots \le \lambda_{n-1} \le \lambda_{n-1}^{(1)} \le \lambda_n. \tag{3.18}$$

Доказательство. Проведем доказательство для случая общего положения, когда собственные значения матрицы $A^{(1)}$ строго упорядочены, $\lambda_1^{(1)} < \ldots < \lambda_{n-1}^{(1)}$. Считатя что они известны, решим задачу на собственные значения матрицы A. Пусть $\bar v$ — собственный вектор матрицы A, отвечающий собственному значению λ, $\bar v^{(1)}$ — он же, но без первого элемента, то есть $\bar v^\dagger = (v_1, \ldots, v_n)$, $(\bar v^{(1)})^\dagger = (v_2, \ldots, v_n)$. Представив матрицу A как

$$A = \begin{pmatrix} \alpha & \bar a^\dagger \\ \bar a & A^{(1)} \end{pmatrix},$$

мы можем переписать условие $A\bar v = \lambda \bar v$ в виде системы уравнений

$$\left\{ \begin{array}{rcl} \alpha v_1 + \bar a^\dagger \bar v^{(1)} & = & \lambda v_1 \\ \bar a v_1 + A^{(1)} \bar v^{(1)} & = & \lambda \bar v^{(1)} \end{array} \right. \quad \Leftrightarrow \quad \left\{ \begin{array}{rcl} (\lambda - \alpha) v_1 & = & \bar a^\dagger \bar v^{(1)} \\ \bar a v_1 & = & \big(\lambda I_{n-1} - A^{(1)}\big) \bar v^{(1)} \end{array} \right. .$$

Таким образом, $\bar v^{(1)} = \big(\lambda I_{n-1} - A^{(1)}\big)^{-1} \bar a v_1$. Подставляя это выражение в первое уравение и сокращая на v_1 (в предположении $v_1 \ne 0$ в ситауции общего положения), получим

$$\lambda - \alpha = \bar a^\dagger \big(\lambda I_{n-1} - A^{(1)}\big)^{-1} \bar a. \tag{3.19}$$

Пусть теперь $\bar v_k^{(1)}$ — собственный вектор матрицы $A^{(1)}$, отвечающий собственному значению $\lambda_k^{(1)}$, то есть $A^{(1)} \bar v_k^{(1)} = \lambda_k^{(1)} \bar v_k^{(1)}$. Разложим вектор $\bar a$ по базису $\bar v_1^{(1)}, \ldots, \bar v_{n-1}^{(1)}$,

$$\bar a = \sum_{k=1}^{n-1} \big(\bar a, \bar v_k^{(1)}\big) \cdot \bar v_k^{(1)},$$

после чего подставим это разложение в формулу (3.19). Получим

$$\lambda - \alpha = \sum_{k=1}^{n-1} \sum_{l=1}^{n-1} \big(\bar a, \bar v_k^{(1)}\big) \cdot \big(\bar v_k^{(1)}\big)^\dagger \big(\lambda I_{n-1} - A^{(1)}\big)^{-1} \bar v_l^{(1)} \cdot \big(\bar a, \bar v_l^{(1)}\big)$$

или

$$\lambda - \alpha = \sum_{k=1}^{n-1} \frac{\Big|\Big|\big(\bar a, \bar v_k^{(1)}\big)\Big|\Big|^2}{\lambda - \lambda_k^{(1)}}. \tag{3.20}$$

На формулу (3.20) можно смотреть как на уравнение относительно переменной λ. Левая часть этого уравнения монотонно возрастает на всей числовой прямой. Правая, напротив, на каждом промежутке вида $\big(\lambda_{k-1}^{(1)}; \lambda_k^{(1)}\big)$ является убывающей функцией, которая при этом принимает все действительные значения, а на лучах $\big(-\infty; \lambda_1^{(1)}\big)$ и $\big(\lambda_{n-1}^{(1)}; +\infty\big)$ она монотонно убывает, принимая значения $(-\infty; 0)$ и $(0; +\infty)$ соответственно (рис. 3.1).

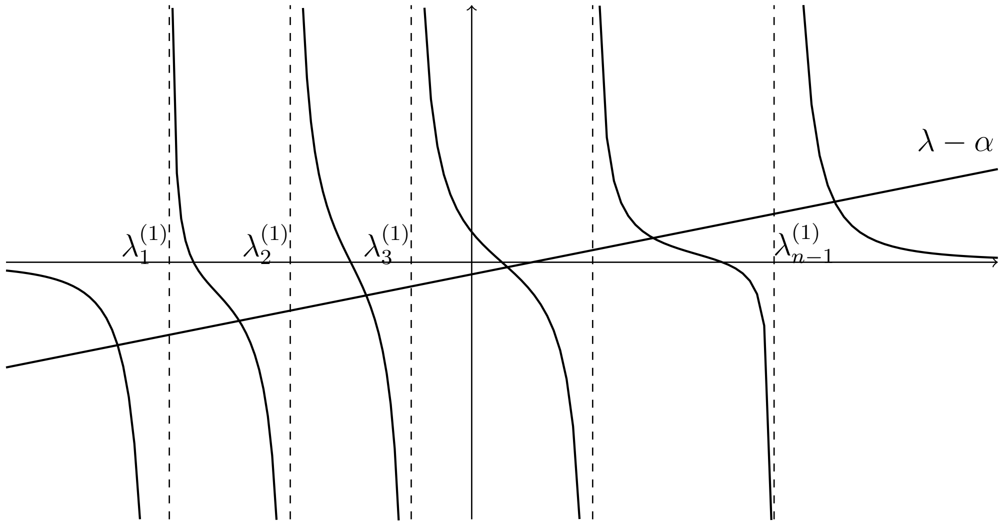

Рис. 3.1:

Следовательно, уравнение (3.20) имеет ровно n решений, которые удовлетворяют неравенству $\lambda_1 < \lambda_1^{(1)} < \lambda_2 < \ldots < \lambda_{n-1} < \lambda_{n-1}^{(1)} < \lambda_n$.

В случае матрицы не общего положения можно изменить матричные элементы на малые добавки, чтобы вернуться в ситуацию общего положения. Поскольку собственные значения являются корнями соответствующих характеристических многочленов, при стремлении малых добавок к нулю свойство их перемежаемости сохранится, быть может с заменой строгих неравенств на нестрогие. □

Перемежаемость собственных значений приводит к нескольким важным следствиям, в частности к тому, что преобразования Стилтьеса ЭСМ матриц A и $A^{(1)}$ отличаются не слишком сильно.

Следствие 3.19. *В условиях леммы 3.18 для любого $z \in \mathbb{C}^*$ имеет неравенство*

$$|s_A(z) - s_{A^{(1)}}(z)| \leq \frac{\pi}{n \cdot \operatorname{Im} z}. \tag{3.21}$$

Доказательство. Заметим, что согласно лемме 3.18 эмпирические функции распределения $F_A(x)$ и $F_{A^{(1)}}(x)$ отличаются друг от друга не слишком сильно:

$$|F_A(x) - F_{A^{(1)}}(x)| \leqslant \frac{1}{n}. \tag{3.22}$$

Поэтому, используя это неравенство и интегрируя по частям, получаем

$$|s_A(z) - s_{A^{(1)}}(z)| = \left| \int\limits_{-\infty}^{+\infty} \frac{dF_A(\lambda)}{\lambda - z} - \int\limits_{-\infty}^{+\infty} \frac{dF_{A^{(1)}}(\lambda)}{\lambda - z} \right| =$$

$$= \left| \frac{F_A(\lambda) - F_{A^{(1)}}(\lambda)}{\lambda - z} \right|_{-\infty}^{+\infty} + \int\limits_{-\infty}^{+\infty} \frac{F_A(\lambda) - F_{A^{(1)}}(\lambda)}{(\lambda - z)^2}\, d\lambda \Bigg| \leqslant \frac{1}{n} \int\limits_{-\infty}^{+\infty} \frac{d\lambda}{(\lambda - z)^2} =$$

$$= \frac{1}{n} \int\limits_{-\infty}^{+\infty} \frac{d\lambda}{(\lambda - \operatorname{Re} z)^2 + (\operatorname{Im} z)^2} = \frac{1}{n} \cdot \frac{1}{\operatorname{Im} z} \cdot \operatorname{arctg} \left(\frac{\lambda - \operatorname{Re} z}{\operatorname{Im} z} \right) \Bigg|_{-\infty}^{+\infty} = \frac{\pi}{n \cdot \operatorname{Im} z}.$$

□

Применяя теорему о перемежаемости несколько раз можно обобщить её на матрицы $A^{(k)}$, полученные из A стиранием нескольких столбцов и строк с одинаковыми номерами. При этом условие перемежаемости заменится на (3.18) неравенства

$$\lambda_i \leq \lambda_i^{(k)} \leq \lambda_{i+k}, \quad i =, 1, \dots, n-k, \tag{3.23}$$

что в свою очередь приведет к появлению множителя k в числителе а в правой части аналогов формул (3.21,3.22). В свою очередь все эти неравенства есть частный случай более общего неравенства для любых эрмитовых матриц.

Следствие 3.20. *Пусть $A, B \in \mathbb{C}^{n \times n}$ – эрмитовы матрицы. Для любого $z \in \mathbb{C}^+$ имеет неравенство*

$$|s_A(z) - s_B(z)| \leq \frac{\pi \cdot \operatorname{rank}(A-B)}{n \cdot \operatorname{Im} z}. \tag{3.24}$$

Доказательство. Пусть $k = \operatorname{rank}(A-B) > 0$. Заметим, что обе стороны неравенства 3.24 инвариантны относительно унитарных преобразований. Найдется унитарное преобразование U, приводящее матрицу $C = A - B$ к следующему блочному виду

$$UCU^\dagger = \begin{pmatrix} \tilde{C} & 0 \\ 0 & 0 \end{pmatrix}, \quad UAU^\dagger = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad UBU^\dagger = \begin{pmatrix} B_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{3.25}$$

где $\tilde{C} = A_{11} - B_{11} \in \mathbb{C}^{k \times k}$ – эрмитова матрица полного ранга k, Согласно теореме о перемежаемости и, в частности, формуле (3.23), собственные значения матриц A, B и A_{22} связаны неравенствами

$$\max\{\lambda_j(A), \lambda_j(B)\} \leq \lambda_j(A_{22}) \leq \min\{\lambda_{j+k}(A), \lambda_{j+k}(B)\}, \quad j = 1, \dots, n-k, \tag{3.26}$$

откуда следует, что для $x \in [\lambda_{j-1}(A_{22}), \lambda_j(A_{22}))$ мы имеем

$$\max\{\lambda_{j-1}(A), \lambda_{j-1}(B)\} \leq x < \min\{\lambda_{j+k}(A), \lambda_{j+k}(B)\}, \tag{3.27}$$

т.е.

$$\frac{j-1}{n} \leq F_A(x), F_B(x) < \frac{j+k}{n}, \tag{3.28}$$

и мы приходим к (3.24). □

Можно сделать похожее утверждение о том, что разность преобразований Стилтьеса ЭСМ коварицонных выборочных матриц ограничена рангом разности исходных прямоугольных матриц.

Следствие 3.21. *Пусть $A, B \in \mathbb{C}^{n \times m}$ – прямоугольные комплексные матрицы. Для любого $z \in \mathbb{C}^+$ имеет место неравенство*

$$|s_{AA^\dagger}(z) - s_{BB\dagger}(z)| \leq \frac{\pi \cdot \operatorname{rank}(A-B)}{\min\{n, m\}(\operatorname{Im}\sqrt{z})^2}. \tag{3.29}$$

Доказательство. Сначала предположим, что $n < m$. Заметим, что собственные значения положительно определенной матрицы $AA^\dagger$, они же квадраты сингулярных чисел матрицы A, так же дают квадраты собственных значений эрмитовой матрицы $(n+m)\times(n+m)$ вида

$$\tilde{A} = \begin{pmatrix} 0 & A^\dagger \\ A & 0, \end{pmatrix}$$

имеющей симметричный относительно нуля спектр. Кроме того у матрицы $\tilde{A}$ есть еще $(m-n)$ нулевых собственных значений. Следовательно, преобразование Стилтьеса спектральной меры матрицы $\tilde{A}$ следующим образом выражается через преобразование Стилтьеса спектральной меры матрицы $AA^\dagger$

$$s_{\tilde{A}}(z) = \frac{1}{n+m}\left(-\frac{m-n}{z} + \sum_{k=1}^{n}\left(\frac{1}{\lambda_k - z} - \frac{1}{\lambda_k + z}\right)\right) = \frac{2nz}{n+m}s_{AA^\dagger}(z^2) - \frac{1}{z}\frac{m-n}{m+n} \quad (3.30)$$

Применим теорему о перемежаемости к квадратной матрице $\tilde{A}$, и, заметив, что

$$\mathrm{Rank}(\tilde{A}) = 2\mathrm{Rank}(AA^\dagger), \quad (3.31)$$

придем к сделанному утверждению. Случай $m < n$ рассматривается аналогично. $\square$

3.5 Доказательство теоремы Марченко-Пастура

Зафиксируем $z \in \mathbb{C}^+$.

Шаг 1.

Рассмотрим преобразования Стилтьеса ЭСМ двух матриц $XX^\dagger$ и $X'X'^\dagger$, построенных из матриц $X, X' \in \mathbb{C}^{n\times n}$, отличающихся одним столбцом,

$$X = (\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_m), \quad X' = (\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{x}'_k, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_m).$$

Согласно следствию 3.21 для любого $z \in \mathbb{C}^+$ разность их преобразований Стилтьеса ограничена.

$$|s_{XX^\dagger}(z) - s_{X'X'^\dagger}(z)| \leq \frac{\pi}{n\cdot \operatorname{Im} z}. \quad (3.32)$$

Поэтому, также как в примере 1.25, воспользовавшись неравенством Мак Диармида, неравенством Маркова и леммой Бореля-Контелли, мы приходим к следующему утверждению

Лемма 3.22. *В условиях теоремы 3.15 для любого фиксированного $z \in \mathbb{C}^+$ имеет место сходимость*

$$s_{R_k(z)} - \mathbb{E}s_{R_k(z)} \xrightarrow[k\to\infty]{п.н.} 0. \quad (3.33)$$

Шаг 2.

Идея второго шага доказательства – построить рекуррентную процедуру вычисления преобразования Стилтьеса искомой ЭСМ, связав выражения для матриц разных размеров.

Лемма 3.23. *Пусть $X \in \mathbb{C}^{n\times m}$ – комплексая матрица. Запишем ее в виде*

$$X = (\mathbf{y}_1, \dots, \mathbf{y}_n)^\dagger, \tag{3.34}$$

где $\mathbf{y}_i \in \mathbb{C}^m$ соответствующие столбцы (а после транспонирования $\mathbf{y}_i^\dagger$ – строки), и пусть

$$Y_i = (\mathbf{y}_1, \dots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \dots, \mathbf{y}_n)^\dagger, \quad i = 1, \dots, n \tag{3.35}$$

подматрицы полученные из X вычеркиванием строки с номером $i = 1, \dots, n$. Тогда

$$\left[\left(XX^\dagger - zI_n\right)^{-1}\right]_{ii} = -\frac{1}{z} \cdot \frac{1}{1 + \mathbf{y}_i{}^\dagger\left(Y_i^\dagger Y_i - zI_m\right)^{-1}\mathbf{y}_i}. \tag{3.36}$$

Доказательство. Докажем утверждение для $i = 1$. Мы воспользуемся леммой 4.5 о дополнении Шура, доказанной в приложении 4.2. Эта лемма дает формулу обращения блочной матрицы вида

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

также в виде блочной матрицы. В частности, при размере $k\times k$ верхнего диагонального блока A матричные элементы такого же верхнего диагонального блока имеют вид

$$\left[M^{-1}\right]_{ij} = \left[\left(A - BD^{-1}C\right)^{-1}\right]_{ij}, \quad 1 \le i, j \le k. \tag{3.37}$$

Пусть $X \in \mathbb{C}^{n\times m}$. Тогда можем записать

$$X = \begin{pmatrix} \mathbf{y}^\dagger \\ Y \end{pmatrix}, \qquad X^\dagger = \begin{pmatrix} \mathbf{y} & Y^\dagger \end{pmatrix}, \qquad XX^\dagger = \begin{pmatrix} \mathbf{y}^\dagger\mathbf{y} & \mathbf{y}^\dagger Y^\dagger \\ Y\mathbf{y} & YY^\dagger \end{pmatrix}. \tag{3.38}$$

Формула (3.37) с $k = 1$, примененная к матрице

$$M = XX^\dagger - zI_n = \begin{pmatrix} \mathbf{y}^\dagger\mathbf{y} - z & \mathbf{y}^\dagger Y^\dagger \\ Y\mathbf{y} & YY^\dagger - zI_{n-1} \end{pmatrix}. \tag{3.39}$$

даёт

$$\left[\left(XX^\dagger - zI_n\right)^{-1}\right]_{11} = \frac{1}{(\mathbf{y}\mathbf{y}^\dagger - z) - \mathbf{y}^\dagger Y^\dagger (YY^\dagger - zI_{n-1})^{-1} Y\mathbf{y}} =$$
$$= \frac{1}{-z + \mathbf{y}^\dagger\left(I_m - Y^\dagger(YY^\dagger - zI_{n-1})^{-1}Y\right)\mathbf{y}} = \frac{1}{-z - z\cdot \mathbf{y}^\dagger\left(Y^\dagger Y - zI_m\right)^{-1}\mathbf{y}},$$

что соответствует формуле (3.37). □

Упражнение 3.24. *Проверьте, что в указанных выше условиях выполняется равенство:*

$$I_m - Y^\dagger(YY^\dagger - zI_{n-1})^{-1}Y = z\left(zI_m - Y^\dagger Y\right)^{-1}. \tag{3.40}$$

Просуммировав обе части (3.37) и разделив на n, получим преобразование Стильтьеса матрицы $XX^\dagger$ в левой части

$$s_{XX^\dagger}(z) = \frac{1}{n}\mathrm{Tr}\left[\left(XX^\dagger - zI_n\right)^{-1}\right] = -\frac{1}{nz}\sum_{i=1}^{n}\frac{1}{1 + \mathbf{y}_i{}^\dagger\left(Y_i^\dagger Y_i - zI_m\right)^{-1}\mathbf{y}_i}. \tag{3.41}$$

Покажем что, если в качестве матрицы X брать матрицы X_k из условия, то правую часть можно с любой точностью выразить через него же. Во первых убедимся, что присутствующее в знаменателе произведение $\mathbf{y}_i{}^\dagger\left(Y_i^\dagger Y_i - zI_m\right)^{-1}\mathbf{y}_i$ близко к преобразованию Стилтьеса ЭСМ матрицы $Y_i^\dagger Y_i$.

Лемма 3.25. *Пусть $(x_i)_{i\in\mathbb{N}}$ – последовательность независимых одинаково распределенных случайных величин, определенных на одном вероятностном пространстве, таких что $\mathbb{E}x_1 = 0$, $\mathbb{D}|x_1| = 1$ и $\mathbb{E}|x_1|^8 < \infty$, а*

$$\mathbf{y}_m = \frac{1}{\sqrt{m}}(x_1, \ldots, x_m)^T \in \mathbb{C}^m, \quad m \in \mathbb{N}$$

столбцы из m таких переменных, нормированные так, чтобы дисперсия компонент была равна $1/m$. Пусть, кроме того, $\{A_n \in \mathbb{C}^{n\times n}\}_{n\in\mathbb{N}}$ – последовательность независимых от $(x_i)_{i\in\mathbb{N}}$ комплексных квадратных матриц с ограниченной спектральной нормой. Тогда

$$\mathbf{y}_m^\dagger A_m \mathbf{y}_m - \frac{1}{m}\,\mathrm{Tr} A_m \xrightarrow[m\to\infty]{\text{п.н.}} \tag{3.42}$$

Доказательство. Сходимость в среднем (а точнее тождественное зануление матожидания всех членов последовательности) обеспечивается независимостью компонент векторов $\mathbf{y}_m$, из которой следует, что

$$\mathbb{E}\big(y_{m,i}^\dagger y_{m,j}\big) = \frac{\delta_{ij}}{m},$$

откуда имеем

$$\mathbb{E}\big(\mathbf{y}_m^\dagger A_m \mathbf{y}_m\big) = \sum_{i=1}^m \sum_{i=1}^m \mathbb{E}\big((A_m)_{ij} y_i^\dagger y_j\big) = \frac{1}{m}\,\mathbb{E}\mathrm{Tr} A_m. \tag{3.43}$$

Для доказательства сходимости почти наверное достаточно убедиться в том, что для некоторого $k \in \mathbb{N}$, $C > 0$ и $\varepsilon > 0$ справедлива оценка

$$\mathbb{E}\left(\left|y^\dagger A_m y - \frac{1}{m}\,\mathrm{Tr} A_m\right|^k\right) \leqslant \frac{C}{m^{1+\varepsilon}}, \tag{3.44}$$

после чего достаточно будет воспользоваться леммой 1.13 Бореля-Кантелли.

Проверку неравенства (3.44) для $k = 4$ оформим в виде нижеследующей леммы.

Лемма 3.26. *В условиях леммы 3.25 справедлива оценка*

$$\mathbb{E}\left(\left|y^\dagger A_m y - \frac{1}{m}\,\mathrm{Tr} A_m\right|^4\right) = O\left(\frac{1}{m^2}\right), m \to \infty.$$

Доказательство этой леммы мы оставляем читателю, заметив лишь, что её методика воспроизводит аргументы, использованные при доказательстве теоремы Вигнера методом моментов, с той разницей, что здесь требуется анализ моментов лишь до четвертого порядка включительно. □

Упражнение 3.27. *Докажите лемму 3.26.*

Обозначим через $Y_{k,i}$ матрицу, полученную из матрицы $X_k = X^{(n_k, m_k)}$ вычеркиванием строки с номером i, а соответствующую строку $\mathbf{y}_{k,i}^\dagger$. Поскольку для любой эрмитовой матрицы $M \in \mathbb{C}^{m\times m}$ спектральная норма матрицы $\big(M - zI_m\big)^{-1}$ ограничена величиной $1/|\Im z|$, из доказанной леммы следует, что

$$\begin{aligned} \mathbf{y}_{k,i} i^\dagger \big(Y_{k,i}^\dagger Y_{k,i} - zI_m\big)^{-1} \mathbf{y}_{k,i} &= s_{Y_{k,i}^\dagger Y_{k,i}}(z) + \delta_{k,i} \\ &= c_k s_{Y_{k,i} Y_{k,i}^\dagger}(z) + (c_k - 1)/z + \delta_{k,i} \\ &= c_k\,\mathbb{E} s_{X_k X_k^\dagger}(z) + (c_k - 1)/z + \delta'_{k,i}, \end{aligned} \tag{3.45}$$

где $c_k = n_k/m_k$, а $\delta_{k,i}, \delta'_{k,i}$ – последовательности случайных величин сходящихся к нулю почти наверное при $k \to \infty$ для каждого i. Здесь в первом равенстве мы воспользовались утверждением леммы 3.25, во втором формулой (3.13), связывающей преобразования Стилтьеса ЭСМ произведения матриц и переставленных матриц, а в третьем тем, что $s_{Y_{k,i}Y^\dagger_{k,i}}(z)$ и $s_{X_kX^\dagger_k}(z)$ отличаются на $O(1/n_k)$ согласно следствию 3.21, тем что $c_k \to_{k\to\infty} c$, и леммой 3.22 о сходимости преобразования Стилтьеса к своему матожиданию.

Подставляя X_k вместо X в формулу (3.41), и вычисляя матожидание от обеих частей мы получаем

$$\bar{s}_{R_k}(z) = \frac{1}{1 - c_k - z - c_k z \bar{s}_{R_k}(z)} + \epsilon_k(z), \tag{3.46}$$

где мы ввели обозначения $\mathbb{E}s_{R_k}(z) = \bar{s}_{R_k}(z)$ и

$$\epsilon_k(z) = \frac{1}{n_k}\sum_{i=1}^{n_k} \mathbb{E}\left(\frac{1}{1 - c_k - z - c_k z \bar{s}_{R_k}(z)} - \frac{1}{1 - c_k - z - c_k z \bar{s}_{R_k}(z) - z\delta'_{k,i}}\right). \tag{3.47}$$

Несложно проверить, что $\epsilon_k(z) \to 0$ при $k \to \infty$. Это следует из того, что $\delta'_{k,i} \to 0$, т.е. разность знаменателей дробей под знаком суммы стремится к нулю почти наверное и, следовательно, в среднем, тогда как сами знаменатели отделены от нуля при $z \in \mathbb{C}^+$. Действительно, можно сделать следующие оценки

$$|1 - c_k - z - c_k z \bar{s}_{R_k}(z)| \geq |\Im(1 - c_k - z - c_k z \bar{s}_{R_k}(z))| \tag{3.48}$$

и

$$\begin{aligned} \Im(1 - c_k - z - c_k z \bar{s}_{R_k}(z)) &= -\Im(z + c_k z \bar{s}_{R_k}(z)) \\ &= -\Im z\left(1 + c_k\,\mathbb{E}\int_{\mathbb{R}} \frac{x dF_{L_{R_k}}(x)}{|x-z|^2}\right) \leq -\Im z, \end{aligned} \tag{3.49}$$

где последнее неравенство следует из того факта, что $\mathrm{Supp}(dF_{R_k}) \subset \mathbb{R}_{\geq 0}$.

Уравнение (3.46) имеет два решения

$$s_k^{\pm}(z) = \frac{1 - c_k - z + c_k z\epsilon_k(z) \pm \sqrt{(1 - c_k - z - c_k z \epsilon_k(z))^2 - 4c_k z}}{2c_k z}. \tag{3.50}$$

Нам осталось доказать, что истинным является решение $s_k^+(z)$, которое очевидно сходится к (3.17) в пределе $k \to \infty$. Для этого заметим, что из свойств преобразования Стилтьеса следует, что $\bar{s}_{R_k}(z) \to 0$ в пределе $\Im z \to \infty$ при фиксированном $\Re z$. Из уравнения (3.46 видно, что что при этом также $\epsilon_k(z) \to 0$. Поскольку этому условию удовлетворяет только функция $s^+(z)$, существует некоторая область в $\mathbb{C}^+$, в которой $\Im z$ достаточно велико, где равенство $\bar{s}_{R_k}(z) = s_k^+(z)$ точно имеет место. Предположим, что это верно не всюду в $\mathbb{C}^+$. Тогда, в силу непрерывности преобразования Стилтьеса существует точка $z_0 \in C^+$, в которой выполнено равенство $s_k^+(z_0) = s_k^-(z_0)$. Из этого равенства следует равенство нулю подкоренного выражения в (3.50), которое имеет место если

$$\epsilon_k(z_0) = \frac{1 - c_k - z_0 \pm 2\sqrt{c_k z_0}}{c_k z_0}. \tag{3.51}$$

Если подставить это выражение назад в (3.50), мы увидим, что из неравенства (3.49 следует, что в (3.51) должен быть выбран знак плюс, откуда имеем

$$\bar{s}_{R_k}(z_0) = \frac{1 - c_k - z_0 + \sqrt{c_k z_0}}{c_k z_0} \tag{3.52}$$

Наконец заметим, что согласно следствию 3.14 преобразование Стилтьеса $s_{R'_k}$ ЭСМ матрицы $R'_k = X_k^\dagger X_k$, полученной из R_k перестановкой сомножителей, связана $s_{R_k}(z)$ соотношением 3.13. То же самое справедливо для их матожиданий, т.е.

$$\bar{s}_{R'_k}(z_0) = c_k \bar{s}_{R_k}(z_0) + \frac{c_k - 1}{z_0}. \tag{3.53}$$

Подставляя сюда (3.52), приходим к соотношению

$$1 + \bar{s}_{R'_k}(z_0) = \frac{1}{\sqrt{c_k z_0}}, \tag{3.54}$$

которое очевидно не может быть верным, поскольку мнимая часть левой части положительна, а правой отрицательна. Таким образом мы показали, что $\bar{s}_{R_k}(z) = s_k^+(z)$ для всех $z \in \mathbb{C}^+$, а следовательно.

$$\lim_{k\to\infty} \bar{s}_{R_k}(z) = \frac{1 - c - z + \sqrt{(z - a_+)(z - a_-)}}{2cz}, \qquad \forall z \in \mathbb{C}^+, \tag{3.55}$$

где, $a_\pm = (1 \pm \sqrt{c})^2$. Таким образом мы доказали сходимость матожидания преобразования Стилтьеса ЭСМ матрицы R_k к правой части (3.55), а с учетом леммы 3.22 и сходимость почти наверное самого преобразования Стилтьеса.

В правой части (3.55) мы видим формулу (3.17)) преобразования Стилтьеса $s_{MP}(z)$ распределения Марченко-Пастура. Действительно, при $0 < c < 1$ функция аналитична на комплексной плоскости с разрезом на отрезке $[a_-, a_+]$ соединяющим две точки ветвления квадратного корня. В этом случае плотность распределения может быть найдена с помощью формул Сохоцкого-Племеля

$$f_{MP}(z) = \lim_{y\to 0^+} \frac{1}{\pi} \operatorname{Im} s_{MP}(z + \mathrm{i}y) = \frac{1}{\pi} \cdot \frac{\sqrt{(a_+ - x)(x - a_-)}}{2cx}$$

При $c > 1$ функция $s_{MP}(z)$ nfr;t имеет полюс $z = 0$,

$$s_{MP}(z) \sim -\frac{1 - c^{-1}}{z},$$

который говорит о том, что в точке $x = 0$ имеется атом, с вероятностью $1 - c^{-1}$.

Глава 4

Вспомогательные сведения

4.1 Формулы Сохоцкого-Племеля

Теорема 4.1. *[Формула Сохоцкого-Племеля] Пусть L — направленный гладкий контур в комплексной плоскости, а $f(z)$ — абсолютно интегрируемая на L функция. Введём следующие обозначения:*

$$g(x) = \int\limits_L \frac{f(z)\,dz}{z-x}, \quad \forall x \in \mathbb{C} \tag{4.1}$$

и

$$g(x^{\pm}) = \lim_{y \to x^{\pm}} g(y) \quad \forall x \in L, \tag{4.2}$$

где знак плюс (минус) соответствует пределу слева (справа) от контура, соответственно. Тогда, если в некоторой окрестности точки $x_0 \in L$ функция g(x) удовлетворяет условию Гёльдера, т.е. $|g(x') - g(x'')| \le a|x' - x''|^{\nu}$ для некоторых $a, \nu > 0$, справедливы соотношения

$$g(x_0^+) - g(x_0^-) = 2\pi\mathrm{i} \cdot f(x_0), \tag{4.3}$$

$$g(x_0^+) + g(x_0^-) = 2 \cdot V.P. \int\limits_L \frac{f(z)\,dz}{z-x_0}, \tag{4.4}$$

где интеграл теперь понимается в смысле главного значения.

4.2 Дополнение Шура

Лемма 4.2. *[Лемма о дополнении Шура] Пусть матрица $M \in \mathbb{C}^{n\times n}$ имеет блочный вид*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{4.5}$$

где $A \in \mathbb{C}^{k\times k}$, $B \in \mathbb{C}^{k\times(n-k)}$, $C \in \mathbb{C}^{(n-k)\times k}$, $D \in \mathbb{C}^{(n-k)\times(n-k)}$. Тогда обратная матрица имеет вид

$$M^{-1} = \begin{pmatrix} \left(A - BD^{-1}C\right)^{-1} & -\left(A - BD^{-1}C\right)^{-1} BD^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D - CA^{-1}B\right)^{-1} \end{pmatrix} \tag{4.6}$$

при условии, что указанные матрицы обратимы. В частности при всех $1 \leqslant i, j \leqslant k$ *имеет место следующая формула:*

$$\left[M^{-1}\right]_{ij} = \left[\left(A - BD^{-1}C\right)^{-1}\right]_{ij}, \tag{4.7}$$

Доказательство. Главная идея заключается в том, чтобы представить матрицу M в виде блочного *разложения Гаусса*:

$$M = \begin{pmatrix} I_k & * \\ 0 & I_{n-k} \end{pmatrix} \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \begin{pmatrix} I_k & 0 \\ * & I_{n-k} \end{pmatrix}. \tag{4.8}$$

Непосредственным вычислением можно проверить, что такое разложение имеет вид

$$M = \begin{pmatrix} I_k & BD^{-1} \\ 0 & I_{n-k} \end{pmatrix} \begin{pmatrix} H & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_k & 0 \\ D^{-1}C & I_{n-k} \end{pmatrix}, \tag{4.9}$$

где блок H дается выражением

$$H = A - BD^{-1}C \tag{4.10}$$

Теперь обращение матрицы M сводится к обращению её сомножителей и перемножению их в обратном порядке. При это мы используем, что блочно-диагональная матрица обращается поблочно, а обращение (блочной) верхне- или нижнетреугольной матрицы с единицами на диагонали сводится к замене знака ее внедиагонального блока на противоположный

$$\begin{pmatrix} I_k & X \\ 0 & I_{n-k} \end{pmatrix}^{-1} = \begin{pmatrix} I_k & -X \\ 0 & I_{n-k} \end{pmatrix}^{-1} \tag{4.11}$$

Следовательно,

$$M^{-1} = \begin{pmatrix} I_k & 0 \\ -D^{-1}C & I_{n-k} \end{pmatrix} \begin{pmatrix} H^{-1} & 0 \\ 0 & D^{-1} \end{pmatrix} \begin{pmatrix} I_k & -BD^{-1} \\ 0 & I_{n-k} \end{pmatrix} \tag{4.12}$$

$$= \begin{pmatrix} H^{-1} & -H^{-1}BD^{-1} \\ -D^{-1}CH^{-1} & D^{-1}CH^{-1}BD^{-1} + D^{-1} \end{pmatrix} \tag{4.13}$$

Можно еще упростить нижний правый блок (см. упражнение 4.3), что в конце концов формулу ур. (**??**). □

Упражнение 4.3. *Докажите, что в условиях леммы 4.2 справедливо* матричное тождество Вудбури *(см. формулы (4.12, 4.13)):*

$$(D - CA^{-1}B)^{-1} = D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1}. \tag{4.14}$$

Следствие 4.4. *В условиях леммы 4.2 справедливо соотношение*

$$\det M = \det D \cdot \det(A - BD^{-1}C). \tag{4.15}$$

Доказательство. Непосредственно вытекает из формулы (4.9). □