М. А. ЛИФШИЦ

ЛЕКЦИИ ПО ГАУССОВСКИМ ПРОЦЕССАМ

Учебное пособие

ЛАНЬ®

САНКТ-ПЕТЕРБУРГ • МОСКВА • КРАСНОДАР
2016

ББК 22.171я73
Л 64

Лифшиц М. А.

Л 64 Лекции по гауссовским процессам: Учебное пособие. — СПб.: Издательство «Лань», 2016. — 192 с. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-2025-4

Цель этих лекций — представить быстрое и содержательное изложение ключевых аспектов теории гауссовских процессов, которые читателю необходимо понять и освоить для творческого овладения материалов. В первых главах рассматриваются основные понятия классической теории гауссовских процессов и мер. Ключевыми понятиями здесь являются ядро меры, интегральное представление процесса, изопериметрическое неравенство, принцип больших уклонений. Далее в лекциях отражён прогресс, достигнутый за последнее десятилетие и ещё недостаточно освещённый в литературе. Сюда можно отнести оценки вероятностей малых уклонений, разложения гауссовских векторов и задачи их бесконечномерного квантования.

Книга предназначена для студентов и аспирантов, обучающихся по направлениям «Математика» и «Прикладная математика», специализирующихся в области теории вероятностей, математической статистики, функционального анализа.

ББК 22.171я73

Рецензент

И. А. ИБРАГИМОВ — академик РАН, зав. лабораторией статистических методов ПОМИ РАН;

Я. Ю. НИКИТИН — доктор физико-математических наук, профессор, зав. кафедрой теории вероятностей и математической статистики математико-механического факультета СПбГУ.

Обложка
Е. А. ВЛАСОВА

Оглавление

Предисловие

Гауссовские процессы можно рассматривать как далеко идущее бесконечномерное обобщение классических нормально распределённых случайных величин. Соответствующая им теория является одной из наиболее продвинутых областей вероятностной науки и представляет собой мощный спектр инструментов вероятностного моделирования в различных теоретических и прикладных сферах, таких как математическая статистика, прогнозирование, финансы, передача информации, машинное обучение, и многих других.

Цель этих лекций - представить компактное и содержательное изложение ключевых аспектов теории гауссовских процессов, которые читателю необходимо понять и освоить для творческого овладения материалом. Книга предназначена прежде всего для студентов старших курсов, аспирантов и научных работников, занимающихся теоретической или прикладной математикой.

В первых главах рассматриваются основные понятия классической теории гауссовских процессов и мер. Ключевыми понятиями здесь являются гауссовская мера, её ядро, интегральное представление процесса, изопериметрическое неравенство, принцип больших уклонений. Все они пояснены и проиллюстрированы многочисленными тщательно подобранными примерами. Эта часть в основном следует монографии автора „Гауссовские случайные функции“ [19], но стиль изложения совсем иной. Поскольку краткость - приоритет для целей изучения и обучения, некоторые технические детали и доказательства опущены, так что принятый подход несколько менее формален, чем это принято в учебниках.

Разумеется, в книге отражён прогресс, достигнутый в области гауссовских процессов за последнее десятилетие. Сюда можно отнести неравенства, связанные с корреляционной гипотезой и другими экстремальными задачами, энтропийный подход к оценкам вероятностей малых уклонений, разложения гауссовских векторов, связи с теорией линейных операторов, задачи бесконечномерного квантования (дискретизации) случайных процессов.

Столь краткий текст никоим образом не претендует на то, чтобы дать полный обзор огромного поля научных исследований теоретического и прикладного характера, каким является современная теория гауссовских случайных процессов. Некоторые указания для дальнейшего ознакомления даны в главе „Что читать дальше“.

В университетской среде эти лекции можно положить в основу семестрового курса для аспирантов и старшекурсников. За последние годы такие курсы автору довелось читать в России (Санкт-Петербургский государственный университет), Германии (Технический университет Дармштадта), США (Технологический институт Джорджии), Финляндии (Технологический университет Хельсинки), Франции (Университет Лилль-I) и Швеции (Университет Линчёпинга). Автор признателен перечисленным учебным заведениям за возможность преподавать любимый предмет в их стенах.

Глава 1

Гауссовские векторы и распределения

Гауссовские векторы и гауссовские распределения в бесконечномерных пространствах появились в связи с необходимостью интерпретации нормального распределения в рамках теории случайных процессов. По своей простоте, важности и продвинутости результатов теория гауссовских процессов занимает едва ли не ведущее место во всей теории случайных процессов.

1.1 Одномерные объекты

Вещественная случайная величина X *нормально распределена* со средним $a \in \mathbb{R}$ и дисперсией $\sigma^2 > 0$, если её плотность распределения по мере Лебега имеет вид

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ \frac{-(x-a)^2}{2\sigma^2} \right\}.$$

Это распределение мы будем обозначать $N(a, \sigma^2)$ и писать $X \sim N(a, \sigma^2)$. Нормальное распределение с нулевой дисперсией $N(a, 0)$ – это распределение, сосредоточенное в точке a.

Если $X \sim N(a, \sigma^2)$, то характеристическая функция и преобразование Лапласа величины X записываются в виде

$$\begin{aligned} \mathbb{E}e^{itX} &= \exp\left\{iat - \frac{\sigma^2 t^2}{2}\right\}, \\ \mathbb{E}e^{tX} &= \exp\left\{at + \frac{\sigma^2 t^2}{2}\right\}. \end{aligned}$$

Из формулы для характеристической функции следует свойство *устойчивости*: если случайные величины $X_1 \sim N(a_1, \sigma_1^2)$ и $X_2 \sim N(a_2, \sigma_2^2)$ независимы, то $X_1+X_2 \sim N(a_1+a_2, \sigma_1^2+\sigma_2^2)$.

Семейство нормальных величин и распределений также инвариантно относительно линейных преобразований: если $X \sim N(a, \sigma^2)$, то

$$cX + d \sim N(d + ca, c^2\sigma^2).$$

Математическое ожидание и дисперсия нормальной случайной величины совпадают с параметрами её распределения:

$$\mathbb{E}X = a, \qquad \mathbb{D}X = \sigma^2.$$

Среди нормальных распределений выделяют *стандартное* нормальное распределение $N(0,1)$, а его функцию распределения обозначают $\Phi(r)$. Иными словами,

$$\Phi(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} \exp\left\{\frac{-x^2}{2}\right\} dx.$$

Отметим быстрое убывание хвостов нормального распределения на бесконечности:

$$\Phi(-r) = 1 - \Phi(r) \sim \frac{1}{\sqrt{2\pi} r} \exp\left\{\frac{-r^2}{2}\right\} \qquad \text{при } r \to \infty.$$

Из величины X со стандартным нормальным распределением преобразованием $X \mapsto Y = \sigma X + a$ можно получить величину с произвольным нормальным распределением $N(a, \sigma^2)$.

1.2 Многомерные объекты

Случайный вектор $X = (X_j)_{j=1}^n \in \mathbb{R}^n$ называется *стандартным гауссовским*, если его компоненты независимы и имеют стандартное нормальное распределение. Распределение X имеет плотность

$$p(x) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ \frac{-(x,x)}{2} \right\}, \qquad x \in \mathbb{R}^n.$$

Мы можем дать два эквивалентных определения гауссовского вектора общего вида в $\mathbb{R}^n$.

Определение 1.1 *Случайный вектор $Y \in \mathbb{R}^n$ называется гауссовским, если он может быть записан как $Y = a + LX$, где X – стандартный гауссовский случайный вектор, $a \in \mathbb{R}^n$, и $L : \mathbb{R}^n \mapsto \mathbb{R}^n$ – линейное отображение.*

Определение 1.2 *Случайный вектор $Y \in \mathbb{R}^n$ называется гауссовским, если (v, Y) является нормальной случайной величиной при всех $v \in \mathbb{R}^n$.*

Из определения 1.1 легко следует определение 1.2. Действительно,

$$(a + LX, y) = (a, y) + (X, L^*y) = (a, y) + \sum_{j=1}^{n} (L^*y)_j X_j$$

имеет нормальное распределение в силу устойчивости.

В многомерном пространстве не имеет смысла давать определение гауссовского распределения через плотность, так как во многих случаях (тех, где оператор L вырожден, т.е. его образ не совпадает со всем пространством $\mathbb{R}^n$) её просто не существует.

Мы будем придерживаться определения 1.2 как более удобного для дальнейших обобщений: в большинстве интересных пространств стандартного гауссовского вектора X, необходимого в определении 1.1, просто не существует.

Подобно записи $N(a, \sigma^2)$ в одномерном случае, семейство n-мерных гауссовских распределений также допускает разумную параметризацию. Напомним, что для любого случайного вектора $Z = (Z_j) \in \mathbb{R}^n$ *математическое ожидание* может пониматься покоординатно, т.е. $\mathbb{E}Z = (\mathbb{E}Z_j)$, а *ковариационный оператор* $K_Z : \mathbb{R}^n \mapsto \mathbb{R}^n$ определяется соотношением

$$cov((v_1, Z), (v_2, Z)) = (v_1, K_Z v_2).$$

Математическое ожидание $\mathbb{E}Z$ и ковариационный оператор K_Z заведомо существуют, если все координаты вектора Z имеют конечные вторые моменты. На значение математического ожидания не накладывается никаких ограничений, в то время как ковариационный оператор обязательно будет неотрицательно определённым и симметричным. Иными словами, в некотором ортонормальном базисе (e_j) он имеет диагональную форму $K_Z e_j = \lambda_j e_j$, причём $\lambda_j \geq 0$.

Мы пишем $Y \sim N(a, K)$, если Y – гауссовский вектор с математическим ожиданием a и ковариационным оператором K. В частности, для стандартного гауссовского вектора верно $X \sim N(0, E_n)$, где $E_n : \mathbb{R}^n \mapsto \mathbb{R}^n$ – тождественный оператор.

Предложенная запись порождает естественные вопросы:

- Существует ли $N(a, K)$ для всех $a \in R^n$ и всех симметричных неотрицательно определённых операторов K ?

- Единственно ли распределение $N(a, K)$?

- Все ли распределения гауссовских векторов имеют форму $N(a, K)$?

Дадим положительные ответы на все эти вопросы. Действительно, пусть $a \in R^n$, а K – симметричный неотрицательно определённый оператор. Приведём K к диагональному виду (см. выше) и определим оператор $L = K^{1/2}$ соотношениями $Le_j = \lambda_j^{1/2} e_j$. Рассмотрим вектор $Y = a + LX$. Мы уже видели, что он гауссовский. Легко также проверяется, что он имеет математическое ожидание a и ковариационный оператор K.

Единственность $N(a, K)$ следует из того, что пара (a, K) определяет распределение (v, Y) как $N((v, a), (v, Kv))$, откуда по теореме Крамера–Вольда однозначно восстанавливается и всё распределение вектора. Заодно определяется и характеристическая функция

$$\mathbb{E}e^{i(v,Y)} = \exp\left\{ i(v,a) - \frac{(v,Kv)}{2} \right\}.$$

Наконец, все компоненты гауссовского вектора будучи нормальными случайными величинами имеют вторые моменты. Поэтому каждый гауссовский вектор имеет математическое ожидание и ковариационный оператор, т.е. любое гауссовское распределение может быть записано в виде $N(a, K)$.

Упражнение 1.1 Пусть случайный вектор Y удовлетворяет определению 1.2. Докажите, что он удовлетворяет и определению 1.1.

Упражнение 1.2 Пусть все координаты случайного вектора Y – нормальные случайные величины. Следует ли отсюда, что вектор Y – гауссовский?

Как и в одномерном случае, гауссовость сохраняется при суммировании независимых векторов (свойство устойчивости) и при линейном преобразовании. Если $X_1 \sim N(a_1, K_1)$ и $X_2 \sim N(a_2, K_2)$ независимы, то $X_1+X_2 \sim N(a_1+a_2, K_1+K_2)$.

Если $L : \mathbb{R}^n \mapsto \mathbb{R}^n$ – линейный оператор, вектор $h \in \mathbb{R}^n$, и $X \sim N(a, K)$, то

$$LX + h \sim N(h + La, LKL^*).$$

Распределение нормы гауссовского вектора

Пусть $X = (X_j)_{j=1}^n \in \mathbb{R}^n$ стандартный гауссовский вектор. Из формулы для его плотности распределения следует, что

$$\mathbb{P}\{||X|| \le R\} = c \int_0^R r^{n-1} \exp\{-r^2/2\} \mathrm{d}r,$$

с нормировкой $c = 2^{1-n/2}\Gamma(n/2)^{-1}$. Мы также знаем, что

$$\mathbb{E}||X||^2 = \sum_{j=1}^n \mathbb{E}X_j^2 = n.$$

Более того, основная масса стандартного гауссовского распределения при *больших* n лежит в полосе ширины порядка константы вокруг значения $\sqrt{n}$. Действительно, к сумме

$$||X||^2 = \sum_{j=1}^{n} X_j^2$$

можно применять закон больших чисел и центральную предельную теорему, так что

$$\frac{1}{n}\sum_{j=1}^{n} X_j^2 \Rightarrow 1, \qquad \text{откуда } \frac{||X||}{\sqrt{n}} \Rightarrow 1.$$

Более того, для полоски ширины r

$$\begin{aligned} & \mathbb{P}\left\{\big|\, ||X|| - \sqrt{n}\big| \le r\right\} \\ = & \ \mathbb{P}\left\{\frac{(\sqrt{n}-r)^2 - n}{\sqrt{n}} \le \frac{\sum_{j=1}^{n} X_j^2 - n}{\sqrt{n}} \le \frac{(\sqrt{n}+r)^2 - n}{\sqrt{n}}\right\} \\ \to & \ 2\Phi\left(\frac{2r}{\sqrt{\mathbb{D}(X_j^2)}}\right) - 1 = 2\Phi\left(\sqrt{2}\, r\right) - 1. \end{aligned}$$

Проделанные выкладки показывают, что в пространствах большой размерности *стандартное гауссовское распределение напоминает равномерное распределение на сфере* соответствующего радиуса.

1.3 Гауссовские объекты в линейных пространствах

Пусть $\mathcal{X}$ – линейное топологическое пространство (дополнительные свойства которого будут уточнены ниже). Обозначим $\mathcal{X}^*$ пространство линейных непрерывных функционалов на $\mathcal{X}$. Двойственность между пространствами $\mathcal{X}$ и $\mathcal{X}^*$, т.е. значение функционала $f \in \mathcal{X}^*$ на элементе $x \in \mathcal{X}$, мы будем обозначать (f, x), вместо более традиционного $f(x)$. *Случайный вектор* X со значениями в $\mathcal{X}$ это измеримое отображение $X : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto \mathcal{X}$. При этом σ-алгебра, которая рассматривается в $\mathcal{X}$, должна быть достаточно большой, чтобы относительно неё были измеримы все линейные непрерывные функционалы.

Вполне аналогично конечномерному случаю определяются гауссовские векторы, их математические ожидания и ковариационные операторы.

Случайный вектор $X \in \mathcal{X}$ называется *гауссовским*, если (f, X) является нормальной случайной величиной при всех $f \in \mathcal{X}^*$.

Вектор $a \in \mathcal{X}$ называется *математическим ожиданием* случайного вектора $X \in \mathcal{X}$, если $\mathbb{E}(f, X) = (f, a)$ при всех $f \in \mathcal{X}^*$. Мы пишем в этом случае $a = \mathbb{E}X$. Линейный оператор $K : \mathcal{X}^* \mapsto \mathcal{X}$ называется *ковариационным оператором* вектора $X \in \mathcal{X}$, если $cov((f_1, X), (f_2, X)) = (f_1, K f_2)$ при всех $f_1, f_2 \in \mathcal{X}^*$. Мы пишем в этом случае $K = cov(X)$. Ковариационный оператор необходимо будет симметричным, т.е.

$$(f, Kg) = (g, Kf), \qquad \forall f, g \in \mathcal{X}^*,$$

и неотрицательно определённым, т.е.

$$(f, Kf) \geq 0, \qquad \forall f \in \mathcal{X}^*.$$

Как ясно из определения гауссовского вектора, оно имеет смысл только в том случае, когда запас непрерывных линейных функционалов на $\mathcal{X}$ достаточно богат. Если, например, $\mathcal{X}^* = \{0\}$, то такому определению удовлетворяет любой случайный вектор, что делает понятие гауссовости бессмысленным. Обычно рассматривается какая-то из трех нижеследующих ситуаций возрастающей степени общности:

(1) $\mathcal{X}$ – сепарабельное банахово пространство, например, $\mathbb{C}[0, 1]$, $L_p[0, 1]$ и т.п.

(2) $\mathcal{X}$ – полное сепарабельное локально выпуклое метризуемое линейное топологическое пространство, например, $\mathbb{C}[0, \infty)$, $\mathbb{R}^\infty$, и т.п.

(3) $\mathcal{X}$ – локально выпуклое линейное топологическое пространство, а вектор X таков, что его распределение является радоновой мерой.

В первых двух случаях любая мера является радоновой, так что случай (3) является наиболее общим из перечисленных.

В дальнейшем на протяжении всего курса мы молчаливо предполагаем, что выполнено одно из этих условий (т.е. как минимум – условие (3)), и называем их *обычными условиями*.

Как и в конечномерном случае, мы пишем, что вектор X имеет распределение $N(a, K)$, если X – гауссовский вектор с математическим ожиданием a и ковариационным оператором K.

Снова возникают всё те же три вопроса:

- Существует ли $N(a, K)$ для всех $a \in \mathcal{X}$ и всех симметричных неотрицательно определённых операторов $K : \mathcal{X}^* \mapsto X$?

- Единственно ли распределение $N(a, K)$?

- Все ли распределения гауссовских векторов имеют форму $N(a, K)$?

Ответы на эти вопросы несколько отличны от приведённых в предыдущем пункте. Что касается первого вопроса, то существование $N(a, K)$ зависит только от K. Действительно, если случайный вектор X имеет распределение $N(a_1, K)$, то вектор $X + a_2 - a_1$ имеет распределение $N(a_2, K)$. С другой стороны, следующее упражнение показывает, что в общем случае далеко не всякому симметричному неотрицательно определённому оператору K соответствует распределение $N(0, K)$.

Упражнение 1.3 Пусть $\mathcal{X}$ – бесконечномерное сепарабельное гильбертово пространство. Тогда $\mathcal{X}^* = \mathcal{X}$ и тождественный оператор $E : \mathcal{X} \mapsto \mathcal{X}$ симметричен и неотрицательно определён. Докажите, что распределение $N(0, E)$ не существует. Подсказка: если бы имелся случайный вектор X с распределением $N(0, E)$, то было бы верно абсурдное равенство $\mathbb{P}(||X||^2 = \infty) = 1$.

Поиск критерия существования $N(0, K)$ является глубоко нетривиальной проблемой, а её решение зависит от пространства $\mathcal{X}$. Здесь этот вопрос умышленно не затрагивается (за исключением гильбертовых пространств), так как мы больше заинтересованы в исследовании свойств заведомо существующих объектов.

Зато на последние два вопроса *при обычных условиях* мы можем ответить положительно. А именно, математическое ожидание и ковариационный оператор существуют у любого гауссовского вектора (см. [19]). Поэтому его распределение принадлежит семейству $\{N(a, K)\}$. Далее, пара (a, K) определяет распределение величины (f, X) как $N((f, a), (f, Kf))$, откуда однозначно восстанавливается её характеристическая функция

$$\mathbb{E}e^{i(f,X)} = \exp\left\{i(f, a) - \frac{(f, Kf)}{2}\right\}.$$

Радоново распределение в $\mathcal{X}$ восстанавливается по своей характеристической функции однозначно, поэтому распределение $N(a, K)$ единственно.

Глава 2

Примеры гауссовских векторов

Пример 2.1 *(Стандартная гауссовская мера в $\mathbb{R}^\infty$).* Рассмотрим пространство всех последовательностей $\mathbb{R}^\infty$ с топологией покоординатной сходимости. Оно превращается в полное метрическое сепарабельное пространство введением подходящей метрики.

Например, в качестве такой метрики можно взять

$$\rho(x,y) = \sum_{j=1}^{\infty} 2^{-j} \min\{|x_j - y_j|, 1\}.$$

Таким образом наши „обычные условия“ выполнены. Напомним, что сопряжённое пространство $\mathcal{X}^* = \mathbf{c_0}$ - пространство всех *финитных* последовательностей, а двойственность имеет вид

$$(f,x) = \sum_{j=1}^{\infty} f_j x_j,$$

причём здесь сумма на самом деле конечна. В качестве вектора $X \in \mathcal{X}$ возьмём последовательность независимых $N(0,1)$-распределённых случайных величин. Тогда в силу устойчивости нормального распределения для любого $f \in \mathcal{X}^*$ случайная величина (f, X) имеет распределение $N(0, \sigma^2)$ с дисперсией $\sigma^2 = \sum_{j=1}^{\infty} f_j^2$. Таким образом, вектор X – гауссовский. Очевидно, что $\mathbb{E}X = 0$. Ковариационным оператором для X будет оператор вложения $K : \mathbf{c_0} \mapsto \mathbb{R}^\infty$. Действительно,

$$\begin{aligned} cov((f,X)(g,X)) &= \mathbb{E}(f,X)(g,X) \\ &= \mathbb{E}\left(\sum_{j=1}^{\infty} f_j X_j\right)\left(\sum_{j=1}^{\infty} g_j X_j\right) \\ &= \sum_{j_1=1}^{\infty}\sum_{j_2=1}^{\infty} f_{j_1} g_{j_2} \mathbb{E}\left(X_{j_1} X_{j_2}\right) \\ &= \sum_{j=1}^{\infty} f_j g_j = (f, Kg). \end{aligned}$$

Распределение вектора X мы будем называть *стандартной гауссовской мерой* в $\mathbb{R}^\infty$.

Пример 2.2 (Гауссовские векторы в гильбертовом пространстве.) Пусть $\mathcal{X}$ – сепарабельное гильбертово пространство со скалярным произведением $(\cdot,\cdot)$. Тогда $\mathcal{X}^*$ можно отождествить с $\mathcal{X}$. „Обычные условия“ очевидным образом выполнены. Для построения гауссовского вектора в $\mathcal{X}$ нам потребуются: ортонормальный базис (e_j) в $\mathcal{X}$, последовательность независимых $N(0,1)$-распределённых случайных величин (ξ_j) и последовательность неотрицательных чисел (σ_j), удовлетворяющая условию $\sum_{j=1}^{\infty} \sigma_j^2 < \infty$. Тогда X можно определить формулой

$$X = \sum_{j=1}^{\infty} \sigma_j \, \xi_j \, e_j \,, \tag{2.1}$$

причём ряд будет сходиться почти наверное по гильбертовой норме $\mathcal{X}$. Такое представление называется *разложением Кархунена –Лоэва*[1].

[1] K. Karhunen, M. Loève.

Для любого $f = \sum_j f_j e_j \in \mathcal{X}$ распределение случайной величины

$$(f, X) = \sum_{j=1}^{\infty} \sigma_j \, f_j \, \xi_j$$

будет нормальным с нулевым средним и дисперсией, равной $\sum_{j=1}^{\infty} \sigma_j^2 \, f_j^2$. Поэтому X – гауссовский случайный вектор и $\mathbb{E}X = 0$. Чтобы найти ковариационный оператор, вычислим

$$\begin{aligned} cov((f,X)(g,X)) &= \mathbb{E}(f,X)(g,X) \\ &= \mathbb{E}\left(\sum_{j=1}^{\infty} \sigma_j \, f_j \, \xi_j\right)\left(\sum_{j=1}^{\infty} \sigma_j \, g_j \, \xi_j\right) \\ &= \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} f_{j_1} \sigma_{j_1} g_{j_2} \sigma_{j_2} \mathbb{E}\left(\xi_{j_1} \xi_{j_2}\right) \\ &= \sum_{j=1}^{\infty} \sigma_j^2 f_j g_j. \end{aligned}$$

Таким образом,

$$(f, Kg) = \sum_{j=1}^{\infty} \sigma_j^2 f_j g_j.$$

Подставляя сюда элементы базиса, находим, что

$$K : g \mapsto \sum_{j=1}^{\infty} \sigma_j^2 g_j \, e_j.$$

Иначе говоря, K – диагональный оператор в базисе e_j с собственными числами σ_j^2.

Можно показать, что *любой* гауссовский вектор в гильбертовом пространстве допускает представление (2.1), см. [27]. Иными словами, гауссовское распределение с ковариационным оператором K существует в том и только в том случае, когда он имеет в некотором базисе диагональную форму с неотрицательными собственными числами, и сумма этих собственных чисел конечна.

Упражнение 2.1 Докажите сходимость ряда (2.1), определяющего вектор X.

Дальнейшие примеры гауссовских векторов и распределений связаны с понятием гауссовского случайного процесса. Напомним, что *случайный процесс* X с параметрическим множеством T – это семейство случайных величин $X(t,\omega), t \in T$, заданных на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Процесс X называется *гауссовским*, если для любых точек $t_1, \dots, t_n \in T$ распределение $(X(t_1), \dots, X(t_n))$ – гауссовское распределение в $\mathbb{R}^n$. Свойства гауссовского процесса полностью определяются математическим ожиданием $\mathbb{E}X(t), t \in T$, и ковариацией $cov(X(s), X(t)), s, t \in T$.

Если T – топологическое пространство, то говорят, что процесс X имеет непрерывные траектории, если при $\mathbb{P}$-почти всех $\omega \in \Omega$ функция $X(\cdot, \omega)$ непрерывна на T.

Пример 2.3 *(Винеровский процесс).* Пусть $\mathcal{X} = \mathbb{C}[0,1]$ – банахово пространство непрерывных функций на отрезке $[0,1]$ с супремум-нормой

$$||x|| = \max_{t \in [0,1]} |x(t)|$$

и соответствующей топологией равномерной сходимости. Сопряжённым здесь будет пространство $\mathcal{X}^* = \mathbb{M}[0,1]$ – пространство зарядов (знакопеременных мер) конечной вариации на отрезке $[0,1]$. Двойственность задаётся соотношением

$$(\mu, f) = \int_{[0,1]} f d\mu, \qquad \mu \in \mathbb{M}[0,1],\ f \in \mathbb{C}[0,1].$$

В качестве гауссовского вектора возьмём траектории *винеровского процесса* $W = W(t), 0 \le t \le 1$, т.е. процесса, удовлетворяющего условиям

$$\mathbb{E}W(t) = 0, \qquad \mathbb{E}W(s)W(t) = \min\{s, t\}.$$

Найдём математическое ожидание и ковариационный оператор вектора W. Поскольку

$$\mathbb{E}(\mu, W) = \mathbb{E}\int_{[0,1]} W d\mu = \int_{[0,1]} \mathbb{E}W(t)\mu(\mathrm{d}t) = 0,$$

то $\mathbb{E}W = 0$. Далее,

$$\begin{aligned} cov((\mu, W), (\nu, W)) &= \mathbb{E}(\mu, W)(\nu, W) \\ &= \mathbb{E} \int_{[0,1]} W d\mu \int_{[0,1]} W d\nu \\ &= \mathbb{E} \int_{[0,1]^2} W(s)W(t)\mu(\mathrm{d}s)\nu(\mathrm{d}t) \\ &= \int_{[0,1]^2} \mathbb{E}W(s)W(t)\ \mu(\mathrm{d}s)\nu(\mathrm{d}t) \\ &= \int_{[0,1]^2} \min\{s,t\}\ \mu(\mathrm{d}s)\nu(\mathrm{d}t). \end{aligned}$$

Таким образом,

$$(\mu, K\nu) = \int_{[0,1]^2} \min\{s,t\}\ \mu(\mathrm{d}s)\nu(\mathrm{d}t),$$

и мы находим, что

$$(K\nu)(s) = \int_{[0,1]} \min\{s,t\}\ \nu(\mathrm{d}t).$$

Напомним основные свойства винеровского процесса.

- Он *самоподобен* с показателем $1/2$, т.е. для любого $c > 0$ случайный процесс $Y(t) := \frac{W(ct)}{\sqrt{c}}$ также будет винеровским;
- Он имеет стационарные приращения;
- Он имеет независимые приращения;
- Он является марковским;
- Он допускает инверсию времени: процесс $Z(t) := t\,W(\frac{1}{t})$ также будет винеровским.

Наконец, важность винеровского процесса объясняется его фундаментальной ролью в стохастическом исчислении и предельных теоремах для случайных процессов (принцип инвариантности).

Внимательный читатель заметит, что никаких специальных свойств винеровского процесса, за исключением непрерывности траекторий, т.е. свойства $W \in \mathbb{C}[0,1]$, при вычислении его ковариации мы не использовали. Поэтому аналогично предыдущему мы имеем более общий пример.

Пример 2.4 *(Произвольный непрерывный гауссовский процесс).* Пусть T – произвольное компактное метрическое пространство, $\mathcal{X} = \mathbb{C}(T)$ – банахово пространство непрерывных функций на T с супремум-нормой $||x|| = \max_{t\in T} |x(t)|$ и соответствующей топологией равномерной сходимости. Сопряжённым здесь будет пространство $\mathcal{X}^* = \mathbb{M}(T)$ – пространство зарядов (знакопеременных мер) конечной вариации на T. Двойственность задаётся соотношением

$$(\mu, f) = \int_T f d\mu, \qquad \mu \in \mathbb{M}(T),\ f \in \mathbb{C}(T).$$

Пусть $X(t), t \in T$, – гауссовский случайный процесс с параметрическим множеством T, имеющий п.н. непрерывные траектории. Он полностью характеризуется функциями

$$a(t) := \mathbb{E}X(t), \qquad K(s,t) := cov(X(s), X(t)).$$

Тогда X можно рассматривать как случайный элемент пространства $\mathcal{X}$, причём его математическое ожидание и ковариационный оператор K находятся по формулам $\mathbb{E}X = a$ и

$$(K\nu)(s) = \int_T K(s,t)\, \nu(\mathrm{d}t). \tag{2.2}$$

Далее будут приведены несколько наиболее интересных гауссовских процессов, укладывающихся в эту схему.

Пример 2.5 *(Дробное броуновское движение[2]).* Пусть $\alpha \in (0,2]$. Гауссовский процесс $W^{(\alpha)}(t), t \in \mathbb{R}$, называется *дробным броуновским движением (ДБД)* с параметром α, если

$$\mathbb{E}W^{(\alpha)}(t) = 0, \qquad \mathbb{E}W^{(\alpha)}(s)W^{(\alpha)}(t) = \frac{1}{2}\left(|s|^\alpha + |t|^\alpha - |s-t|^\alpha\right).$$

[2]Fractional Brownian motion.

Выбор ковариационной функции может показаться странным, но есть более естественное эквивалентное определение

$$\mathbb{E}W^{(\alpha)}(t) = 0, W^{(\alpha)}(0) = 0,$$

$$\mathbb{E}\left|W^{(\alpha)}(s) - W^{(\alpha)}(t)\right|^2 = |s-t|^\alpha. \tag{2.3}$$

Заметим также, что при $\alpha = 1$ получается классический винеровский процесс. Если $\alpha = 2$, то мы получаем довольно вырожденный процесс, в котором траектория полностью определяется одним значением: $W^{(2)}(t) = tW^{(2)}(1)$.

Напомним основные свойства дробного броуновского движения, которые легко усмотреть из (2.3).

- Оно является H-самоподобным процессом при $H = \frac{\alpha}{2}$, то есть для любого $c > 0$ процесс $Y(t) := \frac{W^{(\alpha)}(ct)}{c^H}$ также будет дробным броуновским движением с параметром α;
- Оно имеет стационарные приращения;
- Его приращения зависимы (кроме винеровского случая $\alpha = 1$);
- Оно не является марковским (кроме винеровского случая $\alpha = 1$);
- При $\alpha \in (1,2]$ оно имеет свойства сильной зависимости[3]. Эти свойства мы не будем здесь обсуждать.

Как и винеровский процесс, ДБД играет важную роль в предельных теоремах для случайных процессов (особенно в случае сильной зависимости).

Упражнение 2.2 Проверьте эквивалентность двух определений дробного броуновского движения.

Упражнение 2.3 Найдите предел ковариационной функции дробного броуновского движения при $\alpha \to 0$. Как построить процесс с предельной ковариационной функцией?

[3]Long range dependence.

Пример 2.6 *(Гауссовские марковские процессы).* Известно, что ковариационная функция гауссовского марковского процесса имеет вид

$$K(s,t) = A(\min\{s,t\})\ B(\max\{s,t\}). \tag{2.4}$$

Процессы с такими ковариационными функциями можно легко получить преобразованием винеровского процесса. Положим $X(t) = f(t)W(g(t))$, где $g(\cdot)$ – возрастающая функция. Тогда

$$\mathbb{E}X(s)X(t) = f(s)f(t)\min\{g(s), g(t)\}$$

имеет вид (2.4) при $A = fg, B = f$. Обратно, при заданных A, B можно положить $f = B$, $g = A/B$. Примерами ковариационных функций вида (2.4) являются: $\min\{s,t\}$ (отвечает винеровскому процессу),

$$\min\{s,t\} - st = \min\{s,t\}(1 - \max\{s,t\})$$

(отвечает *броуновскому мосту*) и

$$e^{-|s-t|/2} = e^{\min\{s,t\}/2} e^{-\max\{s,t\}/2}$$

(отвечает *процессу Орнштейна–Уленбека*). Последний процесс является примером гауссовского стационарного процесса и имеет очень короткую память в том смысле, что он может быть записан в виде

$$X(t) = e^{-t/2}X(0) + V(t), \qquad t \geq 0,$$

где $V(t)$ не зависит от прошлого $\{X(s), s \leq 0\}$. Спектральное представление его ковариационной функции включает меру Коши и имеет вид

$$e^{-|t|/2} = \frac{2}{\pi}\int_{-\infty}^{\infty} e^{itv}\ \frac{\mathrm{d}v}{1+4v^2}.$$

Пример 2.7 *(Броуновский лист*[4] *или случайное поле Винера–Ченцова).* Гауссовский процесс $W(t), t \in \mathbb{R}^d_+$, называется *бро-*

[4] Brownian sheet.

уновским листом или *случайным полем Винера–Ченцова*, если

$$\mathbb{E}W(t) = 0, \qquad \mathbb{E}W(s)W(t) = \prod_{l=1}^{d} \min\{s_l, t_l\}. \tag{2.5}$$

При $d = 1$ получается классический винеровский процесс.

Ковариационная функция броуновского листа имеет красивую геометрическую интерпретацию. Свяжем с любой точкой $t \in \mathbb{R}^d_+$ параллелепипед

$$[0, t] := \{s \in \mathbb{R}^d : 0 \le s_l \le t_l, 1 \le l \le d\}.$$

Тогда

$$\prod_{l=1}^{d} \min\{s_l, t_l\} = \lambda^d \left([0, s] \cap [0, t]\right),$$

где λ^d– мера Лебега в $\mathbb{R}^d_+$.

Из (2.5) легко усмотреть, что $W(t)$ является H-самоподобным процессом при $H = \frac{d}{2}$. Это означает, что для любого $c > 0$ процесс $Y(t) := \frac{W(ct)}{c^H}$ также будет броуновским листом.

Броуновский лист обладает, подобно винеровскому процессу, определённой „независимостью приращений“ , которую мы не будем здесь определять.

Броуновский лист является частным случаем *тензорного произведения* случайных процессов, т.е. случайным полем с ковариацией

$$K(s, t) = \prod_{l=1}^{d} K_l(s_l, t_l),$$

где $K_l(\cdot, \cdot)$ – ковариационные функции одномерных процессов. Они могут не совпадать между собой. Например, известное *поле Кифера* является тензорным произведением винеровского процесса и броуновского моста, т.е. имеет ковариацию

$$K(s, t) = \min\{s_1, t_1\} \cdot (\min\{s_2, t_2\} - s_2 t_2)$$

при $s_1, t_1 \geq 0,\ 0 \leq s_2, t_2 \leq 1$. Поле Кифера может быть получено предельным переходом из варианта эмпирической функции распределения, учитывающего временну́ю составляющую, Напомним, что эта функция определена выборкой $\{X_i, 1 \leq i \leq n\}$ независимых случайных величин, равномерно распределённых на интервале $[0,1]$, по формуле

$$F_n(t,r) = \frac{\#\{i : X_i \leq r, i \leq tn\}}{n}.$$

При $n \to \infty$ случайные поля $\sqrt{n}\,(F_n(t,r) - tr)$ сходятся по распределению к полю Кифера.

Пример 2.8 *(Броуновская функция Леви*[5]*).* Гауссовский процесс $W^L(t), t \in \mathbb{R}^d$, называется *броуновской функцией Леви или полем Леви*, если $\mathbb{E}W^L(t) = 0$ и

$$\mathbb{E}W^L(s)W^L(t) = \frac{1}{2}\left(||s|| + ||t|| - ||s-t||\right). \tag{2.6}$$

Здесь $||\cdot||$ – обычная евклидова норма в $\mathbb{R}^d$.

При $d = 1$ броуновская функция Леви даёт пару независимых классических винеровских процессов (при $t \geq 0$ и при $t \leq 0$).

Как и в одномерном случае, эквивалентное определение имеет вид

$$\mathbb{E}W^L(t) = 0, W^L(0) = 0, \mathbb{E}\left|W^L(s) - W^L(t)\right|^2 = ||s-t||. \tag{2.7}$$

Из (2.6) легко усмотреть, что $W^L(t)$ является H-самоподобным процессом при $H = \frac{1}{2}$. Это означает, что для любого $c > 0$ процесс $Y(t) := \frac{W^L(ct)}{c^{1/2}}$ также будет броуновской функцией Леви.

Упражнение 2.4 Примеры 2.7 и 2.8 обобщают понятие винеровского процесса на случай d-параметрических случайных полей. Сформулируйте аналогичные обобщения для дробного броуновского движения с произвольным показателем $\alpha \in (0,2)$. Найдите соответствующие показатели самоподобия.

[5] Lévy's Brownian function, [16].

Пример 2.9 *((α, K)-дробное броуновское движение* [106]*).* Назовём гауссовский случайный процесс $W^{\alpha,K}(t), t \in \mathbb{R}_+$, (α, K)-*дробным броуновским движением*, если выполнены условия

$$\mathbb{E}W^{\alpha,K}(t) = 0$$

и

$$\mathbb{E}W^{\alpha,K}(s)W^{\alpha,K}(t) = \frac{1}{2^K}\left((t^\alpha + s^\alpha)^K - |t-s|^{\alpha K}\right), \quad t, s \geq 0.$$

При $K = 1$ получается обычное дробное броуновское движение W^α. Процесс $W^{\alpha,K}$ существует при значениях параметров $0 < \alpha \leq 2$, $0 < K \leq 2$, и $\alpha K \leq 2$. В следующей главе поясняется тесная связь этого процесса с обычным дробным броуновским движением $W^{\alpha K/2}$, см. упражнение 3.2.

Глава 3

Гауссовский белый шум

3.1 Определение белого шума и интеграла по нему

Многие гауссовские случайные функции удобно выразить или определить с помощью интеграла по белому шуму.

Пусть $(\mathcal{R}, \mathcal{A}, \nu)$ – измеримое пространство с мерой. Положим

$$\mathcal{A}_0 = \{A \in \mathcal{A} : \nu(A) < \infty\}.$$

Гауссовская случайная функция $\{\mathcal{W}(A), A \in \mathcal{A}_0\}$ называется *гауссовским белым шумом* с мерой контроля ν, если $\mathbb{E}\mathcal{W}(A) = 0$ и $\mathbb{E}\mathcal{W}(A)\mathcal{W}(B) = \nu(A \cap B)$. Основные свойства гауссовского белого шума:

- $\mathbb{D}\mathcal{W}(A) = \nu(A)$;

- Если множества $A_1, \ldots, A_n$ не пересекаются, то соответствующие случайные величины $\mathcal{W}(A_1), \ldots, \mathcal{W}(A_n)$ независимы;

- Если множества $A_1, \ldots, A_n$ не пересекаются, то

$$\sum_{j=1}^{n} \mathcal{W}(A_j) = \mathcal{W}\left(\bigcup_{j=1}^{n} A_j\right) \qquad \text{п.н.}$$

Упражнение 3.1 Выведите эти свойства из определения.

Интеграл по белому шуму $\int_{\mathcal{R}} f \mathrm{d}\mathcal{W}$ может быть определен для любой функции $f \in L_2(\mathcal{R}, \mathcal{A}, \nu)$. Сначала определяется интеграл от ступенчатой функции

$$\int_{\mathcal{R}} \left(\sum_j c_j \mathbf{1}_{A_j}\right) \mathrm{d}\mathcal{W} := \sum_j c_j \mathcal{W}(A_j), \qquad c_j \in \mathbb{R}, A_j \in \mathcal{A}_0,$$

и проверяется корректность этого определения, т.е.

$$\sum_j c_j \mathbf{1}_{A_j} = \sum_i b_i \mathbf{1}_{B_i} \curvearrowright \sum_j c_j \mathcal{W}(A_j) = \sum_i b_i \mathcal{W}(B_i).$$

Затем в классе ступенчатых функций устанавливаются линейность

$$\int_{\mathcal{R}} (cf) \mathrm{d}\mathcal{W} = c \int_{\mathcal{R}} f \mathrm{d}\mathcal{W}; \quad \int_{\mathcal{R}} (f+g) \mathrm{d}\mathcal{W} = \int_{\mathcal{R}} f \mathrm{d}\mathcal{W} + \int_{\mathcal{R}} g \mathrm{d}\mathcal{W}$$

и изометрическое свойство

$$\mathbb{E}\left(\int_{\mathcal{R}} f \mathrm{d}\mathcal{W} \cdot \int_{\mathcal{R}} g \mathrm{d}\mathcal{W}\right) = \int_{\mathcal{R}} f g \mathrm{d}\nu,$$

из которого, в частности, следует

$$\mathbb{D} \int_{\mathcal{R}} f \mathrm{d}\mathcal{W} = \int_{\mathcal{R}} |f|^2 \mathrm{d}\nu = ||f||_2^2.$$

Разумеется,

$$\mathbb{E} \int_{\mathcal{R}} f \mathrm{d}\mathcal{W} = 0.$$

Далее воспользуемся плотностью класса ступенчатых функций в $L_2(\mathcal{R}, \mathcal{A}, \nu)$. Для любой функции $f \in L_2(\mathcal{R}, \mathcal{A}, \nu)$ можем определить интеграл как предел в пространстве L_2:

$$\int_{\mathcal{R}} f \mathrm{d}\mathcal{W} := \lim_{n\to\infty} \int_{\mathcal{R}} f_n \mathrm{d}\mathcal{W},$$

где (f_n) – любая последовательность, сходящаяся к f в L_2. Благодаря изометричности такие пределы существуют и не зависят от последовательности (f_n). Упомянутые выше свойства интеграла без труда переносятся с класса ступенчатых функций на всё пространство L_2.

Комплексный гауссовский белый шум и интеграл по нему

Совершенно аналогично предыдущему определяется гауссовский комплексный белый шум и интеграл по нему. Ковариация комплексного шума определяется формулой

$$\mathbb{E}\mathcal{W}(A)\overline{\mathcal{W}(B)} = \nu(A \cap B).$$

Здесь величины $\mathcal{W}(A) \in \mathbf{C}$ и интегрируются комплекснозначные функции $f \in L_{2,\mathbf{C}}(\mathcal{R}, \mathcal{A}, \nu)$. Свойство изометричности имеет вид

$$\mathbb{E}\left(\int_{\mathcal{R}} f \mathrm{d}\mathcal{W} \cdot \overline{\int_{\mathcal{R}} g \mathrm{d}\mathcal{W}}\right) = \int_{\mathcal{R}} f\, \overline{g}\, \mathrm{d}\nu.$$

Комплексная конструкция необходима для спектрального представления стационарного процесса

$$X(t) = \int_{-\infty}^{\infty} e^{itu} \mathrm{d}\mathcal{W}(u),$$

причём соответствующая мера контроля ν – это спектральная мера процесса X. Даже если X – вещественный процесс, соответствующий шум $\mathcal{W}$ будет комплексным.

3.2 Интегральные представления

Мы знаем, что свойства любого центрированного гауссовского процесса $X(t), t \in T$, полностью определяются его ковариационной функцией $K(s,t) = \mathbb{E}X(s)X(t)$. Чтобы построить такой процесс, достаточно располагать гауссовским белым шумом $\mathcal{W}$ на некотором пространстве $(\mathcal{R}, \mathcal{A}, \nu)$ и такой системой функций $\{m_t, t \in T\} \subset L_2(\mathcal{R}, \mathcal{A}, \nu)$, что

$$(m_s, m_t)_2 = \int_{\mathcal{R}} m_s(u) m_t(u)\, \mathrm{d}\nu(u) = K(s,t), \qquad s,t \in T.$$

В этом случае процесс

$$\tilde{X}(t) = \int_{\mathcal{R}} m_t\, \mathrm{d}\mathcal{W}, \qquad t \in T, \tag{3.1}$$

будет иметь нужную ковариационную функцию $K(s,t)$. Выражение (3.1) будем называть *интегральным представлением* процесса X.

Пример 3.1 *(Винеровский процесс).* Здесь достаточно положить $(\mathcal{R}, \mathcal{A}, \nu) = (\mathbb{R}_+, \mathcal{B}, \lambda)$, где $\mathcal{B}$ – борелевская σ-алгебра, λ – мера Лебега, и

$$m_t(u) = \mathbf{1}_{[0,t]}(u). \tag{3.2}$$

Очевидно, что

$$\tilde{X}(t) = \int \mathbf{1}_{[0,t]}(u)\, \mathrm{d}\mathcal{W}(u) = \mathcal{W}([0,t])$$

будет винеровским процессом.

Пример 3.2 *(Броуновский мост).* Для броуновского моста W^0 можно предложить два представления. Одно строится исходя из представления винеровского процесса и линейного соотношения

$$W^0(t) = W(t) - tW(1). \tag{3.3}$$

Очевидно, функции

$$m_t^0(u) = (m_t - tm_1)(u) = (1-t)\mathbf{1}_{[0,t]}(u) - t\,\mathbf{1}_{(t,1]}(u)$$

обеспечивают интегральное представление для W^0.

Альтернативное представление строится на квадрате $[0,1]^2$ с двумерной мерой Лебега. Здесь можно положить

$$\tilde{m}_t^0(u) = \mathbf{1}_{[0,t]\times[0,1-t]}(u).$$

Тогда (нарисуйте картинку!)

$$\begin{aligned}
(\tilde m_s, \tilde m_t)_2 &= \lambda^2 \left([0,s]\times[0,1-s]\bigcap[0,t]\times[0,1-t]\right)\\
&= \min(s,t)\cdot\min(1-s,1-t)\\
&= \min(s,t)\cdot(1-\max(s,t))\\
&= \min(s,t)-st,
\end{aligned}$$

как и требуется для представления W^0.

Пример 3.3 *(Дробное броуновское движение).* Интегральное представление дробного броуновского движения $W^{(\alpha)}(t)$ строится следующим образом [148]. Положим $\mathcal{R}=\mathbb{R}$, пусть $\nu=\lambda$ – мера Лебега и $\mathcal{W}$ – соответствующий белый шум. Тогда рассмотрим процесс

$$\begin{aligned}
&W^{(\alpha)}(t)\\
=\ &\int c_\alpha\left((t-u)^{\frac{\alpha-1}{2}}\mathbf{1}_{u\le t}-(-u)^{\frac{\alpha-1}{2}}\mathbf{1}_{u\le 0}\right)\mathrm{d}\mathcal{W}(u). \qquad (3.4)
\end{aligned}$$

Интеграл корректно определен именно при $0<\alpha<2$. При подходящей нормировке c_α это действительно будет дробное броуновское движение, так как для всех $t\ge s$

$$\begin{aligned}
&\mathbb{E}\left(W^{(\alpha)}(t)-W^{(\alpha)}(s)\right)^2\\
=\ & c_\alpha^2\int\left((t-u)^{\frac{\alpha-1}{2}}\mathbf{1}_{u\le t}-(s-u)^{\frac{\alpha-1}{2}}\mathbf{1}_{u\le s}\right)^2\mathrm{d}u\\
=\ & c_\alpha^2\int\left((t-s-v)^{\frac{\alpha-1}{2}}\mathbf{1}_{v\le t-s}-(-v)^{\frac{\alpha-1}{2}}\mathbf{1}_{v\le 0}\right)^2\mathrm{d}v\\
=\ & c_\alpha^2(t-s)^\alpha\int\left((1-w)^{\frac{\alpha-1}{2}}\mathbf{1}_{w\le 1}-(-w)^{\frac{\alpha-1}{2}}\mathbf{1}_{w\le 0}\right)^2\mathrm{d}w\\
=\ & const\cdot(t-s)^\alpha.
\end{aligned}$$

Вычисляя интеграл, можно показать, что нужное соотношение

$$\mathbb{E}\left(W^{(\alpha)}(t)-W^{(\alpha)}(s)\right)^2=(t-s)^\alpha$$

достигается при

$$c_\alpha = \frac{\left[\sin(\frac{\pi\alpha}{2})\Gamma(\alpha+1)\right]^{1/2}}{\Gamma(\frac{\alpha+1}{2})} \ . \tag{3.5}$$

Упражнение 3.2 (см. [50, 127]) Докажите следующие соотношения между дробными броуновскими движениями с одним и двумя параметрами $W^{(\alpha K/2)}$ и $W^{\alpha,K}$.

$$W^{(\alpha K/2)}(t) = c_1 W^{\alpha,K}(t) + c_2 \int_0^\infty (1 - e^{-ut^\alpha}) u^{-\frac{K+1}{2}} \mathrm{d}\mathcal{W}(u)$$

при $0 < K < 1, 0 < \alpha < 2$ и

$$W^{\alpha,K}(t) = c_1 W^{(\alpha K/2)}(t) + c_2 \int_0^\infty (1 - e^{-ut^\alpha}) u^{-\frac{K+1}{2}} \mathrm{d}\mathcal{W}(u)$$

при $1 < K < 2, 0 < \alpha < \frac{2}{K}$, где c_1 и c_2 некоторые положительные константы, зависящие от α, K, а $\mathcal{W}$ – стандартный белый шум; слагаемые в правой части равенств независимы.

Пример 3.4 *(Процессы и операторы Римана–Лиувилля[1]).* В отличие от предыдущих и последующих примеров, процессы Римана–Лиувилля принято определять именно через их интегральное представление, а не через ковариационные функции.

Напомним, что оператор дробного интегрирования Римана–Лиувилля определяется соотношением

$$R_\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} f(u) du, \qquad \alpha > 0. \tag{3.6}$$

При этом

$$R_\alpha : L_2[0,1] \mapsto \begin{cases} L_p[0,1], & \alpha > \frac{1}{2} - \frac{1}{p} \ , \\ \mathbb{C}[0,1], & \alpha > \frac{1}{2} \ . \end{cases}$$

При $\alpha = 1$ мы получаем обычный оператор интегрирования. Замечательным свойством оператора Римана–Лиувилля является полугрупповое свойство вида $R_\alpha R_\beta = R_{\alpha+\beta}$.

[1] Riemann–Liouville processes.

По аналогии с (3.6) *процесс Римана–Лиувилля* порядка α определяется как интеграл по белому шуму на прямой (с мерой Лебега в качестве меры контроля)

$$R^\alpha(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} \mathrm{d}\mathcal{W}(u), \qquad \alpha > 1/2.$$

При $\alpha = 1$ он совпадает с винеровским процессом. Ограничение $\alpha > 1/2$ требуется для того, чтобы ядро интегрирования $(t-\cdot)^{\alpha-1}$ принадлежало пространству $L_2[0,t]$, т.е. чтобы интеграл по белому шуму был корректно определен.

Процесс R^α является H-самоподобным с индексом самоподобия $H = 2\alpha - 1$.

Полугрупповое свойство даёт $R_\alpha R^\beta = R^{\alpha+\beta}$.

Как можно усмотреть из (3.4), при $\alpha \in (1/2, 3/2)$ процесс Римана–Лиувилля порядка α по своим локальным свойствам близок к дробному броуновскому движению с показателем $\alpha' = 2\alpha - 1$.

Заметим, однако, что в отличие от ДБД семейство процессов Римана–Лиувилля не имеет ограничений по гладкости траекторий, т.е. нет границы сверху для индекса α.

Ещё одно отличие от ДБД состоит в том, что процесс R^α не имеет стационарных приращений, но взамен он обладает свойством однородности ошибки прогноза [140]:

$$R^\alpha(t_0+\cdot) - \mathbb{E}\left(R^\alpha(t_0+\cdot)\middle|\mathcal{F}_{t_0}\right) = R^\alpha(\cdot), \qquad \forall t_0 \geq 0,$$

по распределению. Здесь $\mathcal{F}_{t_0}$ обозначает σ-алгебру прошлого для исходного белого шума $\mathcal{W}$ до момента t_0, т.е.

$$\mathcal{F}_{t_0} = \sigma\{\mathcal{W}(A), A \subset [0, t_0]\}.$$

Замечание 3.1 В эконометрической литературе, где процесс Римана–Лиувилля возникает как предел дискретных схем, его часто называют „дробным броуновским движением“, открывая возможность путаницы с „настоящим“ ДБД из примера 2.5. О дальнейшем сопоставлении этих двух процессов см. [150].

Пример 3.5 *(Броуновский лист).* Этот пример вполне аналогичен примеру 3.1 и служит его обобщением на многомерный случай. Здесь нужно положить $(\mathcal{R}, \mathcal{A}, \nu) = (\mathbb{R}_+^d, \mathcal{B}^d, \lambda^d)$, где $\mathcal{B}^d$ – борелевская σ-алгебра в $\mathbb{R}^d$, λ^d – d-мерная мера Лебега, и определить „прямоугольники“

$$[0, t] := \{u \in \mathcal{R} : 0 \leq \ u_j \leq t_j,\ 1 \leq j \leq d\}.$$

Тогда пересечение прямоугольников тоже будет прямоугольником (нарисуйте картинку):

$$[0, s] \cap [0, t] = \{u \in \mathcal{R} : 0 \leq \ u_j \leq \min(s_j, t_j),\ 1 \leq j \leq d\}.$$

Поэтому функции

$$m_t(u) = \mathbf{1}_{[0,t]}(u)$$

обладают свойством

$$(m_s, m_t)_2 = \lambda^d([0, s] \cap [0, t]) = \prod_{j=1}^{d} \min(s_j, t_j),$$

и поэтому

$$\tilde{X}(t) = \int \mathbf{1}_{[0,t]}(u) \mathrm{d}\mathcal{W}(u) = \mathcal{W}([0, t])$$

будет броуновским листом.

Пример 3.6 *(Броуновская функция Леви на $\mathbb{R}^d$).* Напомним, что БФЛ на $\mathbb{R}^d$ определяется соотношениями (2.7). Построим её представление через белый шум, называемое *интегрально-геометрической конструкцией Ченцова* [40].

Пусть $\mathcal{R}$ – пространство всех гиперплоскостей в $\mathbb{R}^d$. На $\mathcal{R}$ существует единственная с точностью до константы мера ν, инвариантная относительно всех унитарных преобразований $\mathbb{R}^d$. Для $t \in \mathbb{R}^d$ обозначим A_t множество всех гиперплоскостей, пересекающих отрезок $\overline{0, t} := \{rt, 0 \leq r \leq 1\}$. Мера этого множества, как легко понять, пропорциональна длине отрезка $\overline{0,t}$. Отнормируем меру ν так, чтобы было верно $\nu(A_t) = ||t||$. Пусть теперь $\mathcal{W}$ – белый шум на $\mathcal{R}$ с мерой контроля ν. Тогда

$$W^L(t) := \mathcal{W}(A_t) = \int_{\mathcal{R}} \mathbf{1}_{A_t} \mathrm{d}\mathcal{W}, \qquad t \in \mathbb{R}^d,$$

будет БФЛ на $\mathbb{R}^d$. Действительно соотношения $\mathbb{E}W^L(t) = 0$ и $W^L(0) = 0$ очевидны. Кроме того, для любых $s, t \in \mathbb{R}^d$ симметрическая разность $A_s \Delta A_t$ состоит в точности из тех гиперплоскостей, которые пресекают отрезок

$$\overline{s,t} := \{s + r(t-s), 0 \le r \le 1\}$$

(мы игнорируем множество ν-меры нуль гиперплоскостей, проходящих непосредственно через одну из точек $0, s, t$). Поэтому

$$\begin{aligned} \mathbb{E}(W^L(s) - W^L(t))^2 &= \mathbb{E}(\mathcal{W}(A_s) - \mathcal{W}(A_t))^2 \\ &= \nu(A_s \Delta A_t) = ||s-t||, \end{aligned}$$

что и требовалось.

Интегрально-геометрическую конструкцию можно „упаковать“ в само пространство $\mathbb{R}^d$, причём она станет более элементарной, но гораздо менее понятной. Действительно, пусть $\mathcal{R}_0$ – множество гиперплоскостей, проходящих через ноль. Тогда между множествами $\mathcal{R} \backslash \mathcal{R}_0$ и $\mathbb{R}^d \backslash \{0\}$ существует естественная биекция: каждой гиперплоскости соответствует её точка, ближайшая к началу координат. При такой биекции множество A_t перейдет в шар $\tilde{A}_t$ радиуса $||t||/2$ с центром в $\frac{t}{2}$ (т.е. в шар, построенный на отрезке $\overline{0,t}$ как на диаметре; проверьте это!). Мера ν перейдет в сферически инвариантную меру $\tilde{\nu} = \mathrm{d}r\mu(\mathrm{d}\theta)$, где $\mu(\mathrm{d}\theta)$ – подходящим образом нормированная равномерная мера на единичной сфере (докажите!). Очевидно, мера $\tilde{\nu}$ отлична от меры Лебега в $\mathbb{R}^d$. Если $\tilde{\mathcal{W}}$ – белый шум на $\mathbb{R}^d$ с мерой контроля $\tilde{\nu}$, то $\tilde{\mathcal{W}}(\tilde{A}_t)$ также будет БФЛ.

Интегрально-геометрическую конструкцию БФЛ, аналогичную конструкции Ченцова, можно реализовать на сфере[2],

[2]Вместо гиперплоскостей здесь нужно использовать круги максимального радиуса.

в гиперболическом пространстве и некоторых других случаях.

Пример 3.7 *(Броуновская функция Леви на L_1, [20]). Броуновской функцией Леви* (БФЛ) на метрическом пространстве (T, ρ) с отмеченной точкой ϑ называется центрированная гауссовская случайная функция $\{W^L(t), t \in T\}$, для которой $W^L(\vartheta) = 0$ и

$$\mathbb{E}(W^L(t) - W^L(s))^2 = \rho(s,t), \qquad s, t \in T.$$

Одним из наиболее интересных классов пространств, на которых БФЛ существуют, являются пространства L_1. Итак, пусть $T = L_1(U, \mathcal{U}, \mu)$ с отмеченной точкой $\vartheta = 0$, и расстояние в T задано L_1-метрикой

$$\rho(f,g) = \int_U |f(u) - g(u)| \mu(\mathrm{d}u).$$

Положим $\mathcal{R} = U \times \mathbb{R}$, $\nu = \mu \times \lambda$ и рассмотрим белый шум $\mathcal{W}$ на $\mathcal{R}$ с мерой контроля ν. Для функции $f \in T$ определим её подграфик

$$A_f = \{(u,r) : |r| \le |f(u)|, rf(u) \ge 0,\ u \in U, r \in \mathbb{R}\} \subset \mathcal{R}.$$

Тогда

$$W^L(f) := \mathcal{W}(A_f) = \int_{\mathcal{R}} \mathbf{1}_{A_f} \mathrm{d}\mathcal{W}, \qquad f \in T,$$

является БФЛ. Действительно,

$$\begin{aligned}
\mathbb{E}(W_f^L - W_g^L)^2 &= \int_{\mathcal{R}} (\mathbf{1}_{A_f} - \mathbf{1}_{A_g})^2 \mathrm{d}\nu = \nu(A_f \Delta A_g) \\
&= \int_U \mu(\mathrm{d}u) \lambda\left((A_f \Delta A_g) \cap (u \times \mathbb{R})\right) \\
&= \int_U \mu(\mathrm{d}u) |f(u) - g(u)| \\
&= ||f - g||_1 = \rho(f,g).
\end{aligned}$$

Упражнение 3.3 Пусть S – единичная сфера в $\mathbb{R}^d$ и μ – равномерная мера на S, нормированная таким образом, чтобы для любого $x \in \mathbb{R}^d$ было верно

$$||x|| = \int_S |(x,u)|\mu(\mathrm{d}u).$$

Тогда $\mathbb{R}^d$ изометрически вкладывется в $L_1(S,\mu)$ по правилу

$$x \mapsto f_x(\cdot) := (x,\cdot).$$

Установите при помощи этого вложения связь между интегральными представлениями БФЛ на $\mathbb{R}^d$ и на L_1-пространствах, описанными в примерах 3.6 и 3.7.

Интегральные представления стационарных процессов

Мы будем рассматривать вещественный стационарный гауссовский центрированный процесс $X(t), t \in \mathbb{R}$.

В силу стационарности его ковариационная функция имеет вид $\mathbb{E}X(t)X(s) = K(t-s)$, причём функция K является вещественной и неотрицательно определённой. Поэтому она допускает *спектральное представление*[3]

$$K(\tau) = \int_{-\infty}^{\infty} e^{i\tau u}\nu(\mathrm{d}u), \qquad \tau \in \mathbb{R},$$

где ν – конечная симметричная мера на $\mathbb{R}$. В частности, верно $\mathbb{D}X(t) = \nu(\mathbb{R})$ при всех t. Сам процесс X также допускает спектральное представление вида

$$X(t) = \int_{-\infty}^{\infty} e^{itu}\mathcal{W}(\mathrm{d}u), \qquad t \in \mathbb{R}, \tag{3.7}$$

где $\mathcal{W}$ - комплексный гауссовский шум с некоррелированными, но зависимыми значениями. Процесс X может быть выражен через обычный вещественный гауссовский белый шум

[3] Здесь предполагается, что основы спектральной теории стационарных процессов читателю известны; подробности см. в учебнике [10].

следующим образом. Пусть $\mathcal{W}^{(re)}, \mathcal{W}^{(im)}$ – две независимых копии вещественного гауссовского белого шума на $(0,\infty)$ с мерой контроля $\nu/2$, и пусть $\mathcal{W}_0$ – независимая с $\mathcal{W}^{(re)}, \mathcal{W}^{(im)}$ нормальная центрированная случайная величина с распределением $N(0,\nu(\{0\}))$. Тогда $\mathcal{W}(\{0\}) = \mathcal{W}_0$ и

$$\mathcal{W}(A) = \begin{cases} \mathcal{W}^{(re)}(A) + i\mathcal{W}^{(im)}(A), & A \subset (0,\infty), \\ \mathcal{W}^{(re)}(-A) - i\mathcal{W}^{(im)}(-A), & A \subset (-\infty, 0). \end{cases}$$

Здесь, конечно, имеется зависимость вида $\mathcal{W}(-A) = \overline{\mathcal{W}(A)}$, которая и обеспечивает вещественность X в интеграле (3.7). В этом случае наш процесс записывается в виде суммы независимых слагаемых

$$X(t) = \mathcal{W}_0 + 2\int_0^\infty \cos(tu)\mathcal{W}^{(re)}(\mathrm{d}u) + 2\int_0^\infty \sin(tu)\mathcal{W}^{(im)}(\mathrm{d}u).$$

Приведённая конструкция может быть распространена на случай случайных полей $(t, u \in \mathbb{R}^d)$, случайных последовательностей $(t \in \mathbb{Z}, u \in \mathbb{S}^1)$, а также периодических процессов $(t \in \mathbb{S}^1, u \in \mathbb{Z})$. Её также можно реализовать для случайных процессов и полей со стационарными приращениями.

Рассмотрим теперь ещё два представления стационарных процессов. Предположим, что спектральная мера ν имеет плотность f и возьмём произвольную измеримую функцию $\theta : \mathbb{R} \mapsto \mathbb{S}^1$. Определим с её помощью семейство функций $\{m_t, t \in \mathbb{R}\} \subset L_{2,\mathbf{C}}(\mathbb{R})$ формулой

$$m_t(u) := \theta(u)e^{itu}\sqrt{f(u)}, \qquad u \in \mathbb{R}.$$

Очевидно,

$$\begin{aligned} (m_t, m_s)_2 &= \int_{\mathbb{R}} m_t(u)\overline{m_s(u)}\mathrm{d}u \\ &= \int_{\mathbb{R}} |\theta(u)|^2 e^{i(t-s)u} f(u)\mathrm{d}u \\ &= \int_{\mathbb{R}} e^{i(t-s)u}\nu(\mathrm{d}u) \\ &= K(t,s) = \mathbb{E}X(t)X(s). \end{aligned}$$

Другое, более наглядное представление, можно получить из (m_t), используя преобразование Фурье

$$\mathcal{F} : L_{2,\mathbf{C}}(\mathbb{R}) \mapsto L_{2,\mathbf{C}}(\mathbb{R}).$$

Обозначим $h = \mathcal{F}\left(\theta(\cdot)\sqrt{f(\cdot)}\right)$ и рассмотрим семейство

$$\tilde{m}_t(\cdot) = \mathcal{F}(m_t)(\cdot) = h(\cdot - t).$$

В силу изометричности преобразования Фурье мы имеем

$$(\tilde{m}_t, \tilde{m}_s)_2 = (m_t, m_s)_2 = \mathbb{E}X(t)X(s).$$

Если вспомогательная функция θ удачно подобрана так, что $h(\cdot)$ вещественна, то мы получаем *сдвиговое представление* процесса X через белый шум с лебеговой мерой контроля

$$X(t) = \int_{\mathbb{R}} h(u - t)\mathcal{W}(\mathrm{d}u).$$

Пример 3.8 *(Процесс Орнштейна–Уленбека).* Мы определили процесс Орнштейна–Уленбека в примере 2.6. Он имеет ковариационную функцию

$$K(\tau) = e^{-|\tau|/2} = \int_{\mathbb{R}} e^{i\tau u} \frac{2\mathrm{d}u}{\pi(1 + 4u^2)} \ .$$

Поэтому спектральная плотность имеет вид $f(u) = \frac{2}{\pi(1+4u^2)}$. Полагая $\theta(u) = \frac{1+2iu}{(1+4u^2)^{1/2}}$, найдём

$$\begin{aligned} m_t(u) &= e^{itu} \left(\frac{2}{\pi(1 + 4u^2)}\right)^{1/2} \frac{1 + 2iu}{(1 + 4u^2)^{1/2}} \\ &= e^{itu} \ (2\pi)^{-1/2} \ (\frac{1}{2} - iu)^{-1} \end{aligned}$$

и

$$h(v) = (\mathcal{F}m_0)(v) = e^{-v/2} \ \mathbf{1}_{v>0} \ .$$

Таким образом, сдвиговое представление процесса Орнштейна–Уленбека имеет вид

$$X(t) = \int_t^{\infty} e^{-(v-t)/2} \ \mathcal{W}(\mathrm{d}v).$$

Глава 4

Измеримые функционалы и ядро

4.1 Основные определения

После того, как мы запаслись изрядным количеством примеров, можно продолжить построение общей теории, начатое в главе 1. Мы по-прежнему рассматриваем гауссовский вектор X со значениями в линейном пространстве $\mathcal{X}$, удовлетворяющем обычным условиям. Будем предполагать $\mathbb{E}X = 0$ и обозначать $K : \mathcal{X}^* \mapsto \mathcal{X}$ ковариационный оператор X, а $P = N(0, K)$ – распределение X в $\mathcal{X}$.

Рассмотрим произвольный линейный непрерывный функционал $f \in \mathcal{X}^*$. Поскольку случайная величина (f, X) имеет нормальное распределение, то её второй момент конечен,

$$\mathbb{E}(f, X)^2 = \int_{\mathcal{X}} |f|^2 \mathrm{d}P < \infty.$$

Таким образом, определено каноническое вложение I^* пространства $\mathcal{X}^*$ в гильбертово пространство $L_2(\mathcal{X}, P)$. Замыкание образа $I^*(\mathcal{X}^*)$ в $L_2(\mathcal{X}, P)$ называется пространством *измеримых линейных функционалов* и обозначается $\mathcal{X}_P^*$. Оно наследует из $L_2(\mathcal{X}, P)$ скалярное произведение

$$(z_1, z_2)_{\mathcal{X}_P^*} = (z_1, z_2)_2 = \int_{\mathcal{X}} z_1 z_2 \mathrm{d}P = \mathbb{E}(z_1(X)\, z_2(X)).$$

В частности,

$$||z||^2_{\mathcal{X}_P^*} = \mathbb{E}\, z(X)^2.$$

Каждый измеримый линейный функционал является пределом, в смысле пространства $L_2(X, P)$ и P-почти наверное, последовательности линейных непрерывных функционалов.

В дальнейшем нам будет удобнее рассматривать I^* именно как вложение

$$I^* : \mathcal{X}^* \mapsto \mathcal{X}_P^*.$$

Сопряжённый оператор $I : \mathcal{X}_P^* \mapsto \mathcal{X}$ определяется естественным соотношением

$$(f, Iz) = (I^* f, z)_{\mathcal{X}_P^*} = \mathbb{E}(f, X) z(X), \qquad \forall f \in \mathcal{X}^*, z \in \mathcal{X}_P^*.$$

Существование сопряжённого оператора не очевидно, так как мы ничего не сказали здесь о непрерывности исходного оператора I^*. Однако можно доказать, что при обычных условиях, сформулированных в главе 1, он действительно существует.

Оператор I является линейным и взаимно-однозначным. Действительно, если при некотором $z \in \mathcal{X}_P^*$ верно $Iz = 0$, то для любого $f \in \mathcal{X}^*$ имеем

$$(I^* f, z) = (f, Iz) = 0.$$

Рассматривая последовательность $I^* f_n$, сходящуюся к z, получим $||z||_{\mathcal{X}_P^*} = 0$, т.е. $z = 0$.

Ещё нам важно отметить, что ковариационный оператор допускает факторизацию

$$K = II^*. \tag{4.1}$$

Действительно,

$$(f, II^* g) = (I^* f, I^* g) = \mathbb{E}(f, X)(g, X) = (f, Kg) \qquad \forall f, g \in \mathcal{X}^*.$$

Теперь можно дать фундаментальное определение: множество $H_P := I(\mathcal{X}_P^*) \subset \mathcal{X}$, снабжённое структурой скалярного произведения

$$(h_1, h_2)_{H_P} := \left(I^{-1}h_1, I^{-1}h_2\right)_{\mathcal{X}_P^*}, \qquad h_1, h_2 \in H_P,$$

и соответствующей нормой

$$|h|_{H_P}^2 := (h, h)_{H_P}, \qquad h \in H_P,$$

называется *ядром* распределения P.

Единичный шар $\{h \in H_P : |h|_{H_P} \leq 1\}$ иногда называют *эллипсоидом рассеяния* меры P.

Корректность определения гарантируется взаимной однозначностью оператора I. Можно также говорить об H_P как о ядре вектора X, но в этом случае следует помнить, что ядро полностью определяется распределением этого вектора.

Всё дальнейшее изложение покажет, что ядро содержит всю важную информацию о P и X; решение любой важной задачи выражается в терминах ядра. Пока что упомянем некоторые простые свойства ядра.

- В силу (4.1) имеем $K(\mathcal{X}^*) \subset H_P \subset \mathcal{X}$. Если H_P конечномерно, то в невырожденном случае все три пространства совпадают. В противном случае все они различны.

- Если H_P бесконечномерно, то $P(H_P) = 0$ (несмотря на всю важность H_P для описания P!).

- Топологический носитель меры P совпадает с замыканием H_P в $\mathcal{X}$.

- Пространство H_P сепарабельно.

- Шары $\{h \in H_P : |h|_{H_P} \leq r\}$ являются компактными подмножествами $\mathcal{X}$.

Пример 4.1 *(Стандартная гауссовская мера в $\mathbb{R}^\infty$).* Рассмотрим стандартный гауссовский вектор X, определённый в примере 2.1. Для любых $f, g \in \mathbf{c_0} = \mathcal{X}^*$ имеем

$$(I^* f, I^* g)_{X_P^*} = \mathbb{E} \left(\sum_{j=1}^{\infty} f_j X_j \right) \left(\sum_{j=1}^{\infty} g_j X_j \right) = \sum_{j=1}^{\infty} f_j g_j = (f, g)_2.$$

Отсюда легко выводится, что линейные измеримые функционалы $z \in \mathcal{X}_P^*$ имеют вид

$$z(x) = \sum_{j=1}^{\infty} z_j x_j, \qquad (z_j) \in \ell_2.$$

Далее, используя координатные функционалы δ_k, найдём

$$(Iz)_k = (\delta_k, Iz) = (I^* \delta_k, z)_{\mathcal{X}_P^*} = \mathbb{E} \left[X_k \left(\sum_{j=1}^{\infty} z_j X_j \right) \right] = z_k.$$

Это означает, что $H_P = \ell_2$, $(h_1, h_2)_{H_P} = (h_1, h_2)_2$ и, в частности, $|h|_{H_P} = ||h||_2$.

Пример 4.2 *(Гильбертово пространство).* Гауссовский вектор X в гильбертовом пространстве был определён в примере 2.2. Для любых элементов $f, g \in \mathcal{X} = \mathcal{X}^*$ имеем

$$(I^* f, I^* g)_{X_P^*} = \mathbb{E} \left(\sum_{j=1}^{\infty} f_j \sigma_j \xi_j \right) \left(\sum_{j=1}^{\infty} g_j \sigma_j \xi_j \right) = \sum_{j=1}^{\infty} f_j g_j \sigma_j^2.$$

Отсюда легко выводится, что

$$\mathcal{X}_P^* = \left\{ z : z(x) = \sum_{j=1}^{\infty} z_j x_j, \quad \sum_{j=1}^{\infty} z_j^2 \sigma_j^2 < \infty \right\},$$

$$||z||_{\mathcal{X}_P^*}^2 = \sum_{j=1}^{\infty} z_j^2 \sigma_j^2, \qquad z \in \mathcal{X}_P^*.$$

Снова прибегая к координатным функционалам, найдём

$$(e_k, Iz) = (I^* e_k, z) = \mathbb{E}\left(\sigma_k \xi_k \left(\sum_{j=1}^{\infty} z_j \sigma_j \xi_j\right)\right) = \sigma_k^2 z_k,$$

т.е.

$$Iz = \sum_{j=1}^{\infty} \sigma_j^2 z_j e_j \ .$$

Если $h = Iz \in H_P$, то

$$||h||_{H_P}^2 = ||z||_{X_P}^2 = \sum_{j=1}^{\infty} z_j^2 \sigma_j^2 = \sum_{j=1}^{\infty} \frac{h_j^2}{\sigma_j^2}.$$

Это означает, что

$$H_P = \left\{ h \in \mathcal{X} : \ \sum_{j=1}^{\infty} \frac{h_j^2}{\sigma_j^2} < \infty \right\}$$

и

$$(h_1, h_2)_{H_P} = \sum_{j=1}^{\infty} \frac{(h_1)_j (h_2)_j}{\sigma_j^2} \ ,$$

где $h_j = (h, e_j)$ – координаты h в базисе (e_j).

Пример 4.3 *(Конечномерный гауссовский вектор).*

Пусть $\mathcal{X} = \mathbb{R}^n$ и вектор X имеет гауссовское распределение $P = N(0, K)$, причём его ковариационный оператор $K : \mathbb{R}^n \mapsto \mathbb{R}^n$ невырожден, т.е. $K(\mathbb{R}^n) = \mathbb{R}^n$. Тогда в некотором базисе K имеет диагональную форму

$$K\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} \sigma_j^2 x_j e_j,$$

причём все $\sigma_j > 0$. Отсюда следует, что X записывается в виде $X = \sum_{j=1}^{n} \sigma_j \xi_j e_j$ из предыдущего примера. Применяя полученные там результаты, видим, что $H_P = \mathbb{R}^n$ и

$$(h_1, h_2)_{H_P} = \sum_{j=1}^{\infty} \frac{(h_1)_j (h_2)_j}{\sigma_j^2} = (K^{-1} h_1, h_2),$$

и в частности,

$$|h|^2_{H_P} = (K^{-1}h, h) = ||K^{-1/2}h||^2.$$

4.2 Теорема о факторизации

В большинстве случаев не существует элементарных рассуждений, подобных описанным выше, позволяющих мгновенно находить ядро гауссовского распределения. Здесь будет полезным следующий результат:

Теорема 4.1 *(теорема о факторизации). Пусть $\mathcal{H}$ – некоторое гильбертово пространство, и $J : \mathcal{H} \mapsto \mathcal{X}$ взаимно-однозначное отображение, для которого верно*

$$K = JJ^*.$$

Тогда ядро распределения P находится как $H_P = J(\mathcal{H})$, а скалярное произведение и норма в H_P вычисляются по формулам

$$(h_1, h_2)_{H_P} = (J^{-1}h_1, J^{-1}h_2)_{\mathcal{H}} \ , \qquad \forall h_1, h_2 \in H_P,$$
$$|h|_{H_P} = ||J^{-1}h||_{\mathcal{H}} \ , \qquad \forall h \in H_P. \tag{4.2}$$

Замечание 4.1 Если оператор J не является взаимно-однозначным, то равенство $H_P = J(\mathcal{H})$ всё же имеет место, но вместо (4.2) будет верно

$$|h|_{H_P} = \inf_{\ell : J\ell = h} \{||\ell||_{\mathcal{H}}\}, \qquad \forall h \in H_P.$$

Замечание 4.2 Мы уже видели, что оператор I порождает нужную факторизацию, так что он формально может рассматриваться в рамках теоремы. Однако её суть как раз заключается в том, что мы можем выбирать $\mathcal{H}$ и J по своему усмотрению для удобства вычислений.

Доказательство теоремы. Определим линейную изометрию U между пространствами $I^*(\mathcal{X}^*) \subset X_P^*$ и $J^*(\mathcal{X}^*) \subset \mathcal{H}$ соотношением

$$UI^*f := J^*f, \qquad f \in \mathcal{X}^*.$$

Имеем

$$\begin{aligned}(UI^*f, UI^*g)_{\mathcal{H}} &= (J^*f, J^*g)_{\mathcal{H}} = (f, JJ^*g)\\ &= (f, Kg) = (f, II^*g) = (I^*f, I^*g)_{\mathcal{X}_P^*},\end{aligned}$$

так что скалярное произведение (а значит, нормы и расстояния) сохраняется. Очевидно, изометрия U может быть продолжена до соответствующих замыканий. По определению, замыканием $I^*(\mathcal{X}^*)$ служит всё пространство X_P^*. Аналогично доказывается, что замыканием $J^*(\mathcal{X}^*)$ является всё пространство $\mathcal{H}$. Действительно, если элемент $\ell \in \mathcal{H}$ ортогонален $J^*(\mathcal{X}^*)$, то при всех $f \in \mathcal{X}^*$

$$0 = (J^*f, \ell)_{\mathcal{H}} = (f, J\ell).$$

Отсюда следует $J\ell = 0$ и из взаимной однозначности J получаем $\ell = 0$. Итак, продолжение U будет изометрией пространств X_P^* и $\mathcal{H}$. Нам остаётся проверить операторное тождество

$$I = JU. \tag{4.3}$$

В этом случае будем иметь

$$H_P = I(X_P^*) = JU(X_P^*) = J(\mathcal{H}),$$

$$|h|_{H_P} = ||I^{-1}h||_{X_P^*} = ||U^{-1}J^{-1}h||_{X_P^*} = ||J^{-1}h||_{\mathcal{H}}.$$

Итак, обратимся к (4.3). Для любых $z \in X_P^*$, $f \in \mathcal{X}^*$ имеем

$$(f, JUz) = (J^*f, Uz)_{\mathcal{H}} = (U^{-1}J^*f, z)_{X_P^*} = (I^*f, z) = (f, Iz).$$

Отсюда следует $(JU)z = Iz$. $\square$

Пример 4.4 *(Винеровский процесс).* Найдём ядро распределения винеровского процесса, определённого в примере 2.3. Пусть $\mathcal{X} = \mathbb{C}[0,1]$, $X = W$, $\mathcal{H} = L_2[0,1]$. Оператор $J : \mathcal{H} \mapsto \mathcal{X}$ определим как оператор интегрирования

$$(J\ell)(t) = \int_0^t \ell(s)\mathrm{d}s.$$

Напомним, что $\mathcal{X}^* = \mathbb{M}[0,1]$ – пространство зарядов конечной вариации (знакопеременных мер) на $[0,1]$. Легко проверить, что $J^* : \mathbb{M}[0,1] \mapsto \mathcal{H}$ задаётся формулой

$$(J^*\mu)(s) = \mu[s,1].$$

Отсюда

$$\begin{aligned}
(JJ^*\mu)(t) &= \int_0^t (J^*\mu)(s)\mathrm{d}s = \int_0^t \mu[s,1]\mathrm{d}s \\
&= \int_0^1 \int_0^1 \mathbf{1}_{s\le t}\mathbf{1}_{s\le u}\mu(\mathrm{d}u)\mathrm{d}s \\
&= \int_0^1 \min\{t,u\}\mu(\mathrm{d}u) = K\mu(t).
\end{aligned}$$

Таким образом, условие факторизации выполнено. Кроме того, оператор J взаимно-однозначен. Применяя теорему 4.1, мы приходим к выводу, что ядро винеровской меры имеет вид

$$\begin{aligned}
H_P &= \left\{h : h(t) = \int_0^t \ell(s)\mathrm{d}s, \ell \in L_2[0,1]\right\} \\
&= \{h : \ h \in AC[0,1], h(0) = 0,\ h' \in L_2[0,1]\}\,. \qquad (4.4)
\end{aligned}$$

Здесь $AC[0,1]$ – класс абсолютно непрерывных функций. Норма и скалярное произведение в H_P задаются формулами

$$|h|^2_{H_P} = \int_0^1 h'(s)^2\mathrm{d}s,$$

$$(h_1,h_2)_{H_P} = \int_0^1 h_1'(s)h_2'(s)\mathrm{d}s.$$

Фактически это соболевское пространство типа W_2^1 с односторонним краевым условием. Ядро винеровской меры появилось в работах Камерона и Мартина [66] и было первым среди изученных ядер. Поэтому иногда его и даже ядра других гауссовских распределений называют *пространством Камерона–Мартина.*

Замечание 4.3 Ядро (как множество и как гильбертово пространство) на самом деле не зависит от того, в каком пространстве мы рассматриваем гауссовский вектор. Например, если рассматривать винеровский процесс как случайный элемент пространства $\tilde{\mathcal{X}} = L_2[0,1]$, ядро окажется тем же самым пространством Камерона–Мартина.

Пример 4.5 *(Ядро процесса с интегральным представлением).*

Рассмотрим гауссовский процесс $X(t), t \in T$, допускающий интегральное представление по белому шуму на некотором пространстве $(\mathcal{R}, \mathcal{A}, \nu)$,

$$X(t) = \int_{\mathcal{R}} m_t(u)\ \mathcal{W}(\mathrm{d}u), \qquad t \in T, m_t \in L_2(\mathcal{R}, \mathcal{A}, \nu),$$

что равносильно возможности представить ковариационную функцию X в виде

$$K(s,t) := \mathbb{E}X(s)X(t) = \int_{\mathcal{R}} m_s\, m_t\ \mathrm{d}\nu, \qquad s,t \in T.$$

Чтобы избежать топологических тонкостей, будем считать, что T – компактное топологическое пространство и что траектории X непрерывны на T. Тогда можно считать, что X – случайный вектор в $\mathcal{X} = \mathbb{C}(T)$ с интегральным ковариационным оператором (см. (2.2), пример 2.4).

Положим $\mathcal{H} = L_2(\mathcal{R}, \mathcal{A}, \nu)$ и определим факторизующий оператор $J : \mathcal{H} \mapsto \mathcal{X}$ формулой

$$(J\ell)(t) := (\ell, m_t)_{\mathcal{H}} = \int_{\mathcal{R}} \ell\, m_t\, \mathrm{d}\nu, \qquad t \in T.$$

Сопряжённый оператор $J^* : \mathbb{M}(T) \mapsto \mathcal{H}$ действует по формуле

$$(J^*\mu)(u) = \int_T m_s(u)\mu(ds)$$

и мы получаем

$$
\begin{aligned}
(JJ^*\mu)(t) &= \int_{\mathcal{R}} (J^*\mu)(u) m_t(u)\nu(\mathrm{d}u) \\
&= \int_{\mathcal{R}} \int_T m_s(u)\mu(\mathrm{d}s) m_t(u)\nu(\mathrm{d}u) \\
&= \int_T \int_{\mathcal{R}} m_s(u) m_t(u)\nu(\mathrm{d}u)\mu(\mathrm{d}s) \\
&= \int_T K(s,t)\mu(\mathrm{d}s) = (K\mu)(t).
\end{aligned}
$$

Таким образом, условие факторизации выполнено. Применяя теорему 4.1, мы приходим к выводу, что

$$
H_P = \left\{ h:\, h(t) = \int_{\mathcal{R}} \ell(u) m_t(u)\nu(\mathrm{d}u), \ell \in L_2(\mathcal{R}, \mathcal{A}, \nu) \right\}. \quad (4.5)
$$

Если оператор J ещё и взаимно-однозначен, то можно найти и норму

$$
|h|_P^2 = \int l(u)^2 \nu(\mathrm{d}u),
$$

и соответствующее скалярное произведение.

Ядро винеровского процесса (4.4), конечно, может быть получено как частный случай этой конструкции с помощью представления (3.2). Приведём ещё два примера.

Пример 4.6 *(Ядро дробного броуновского движения* [12]*).*

Напомним, что процесс дробного броуновского движения $W^{(\alpha)}(t), t \in \mathbb{R}$, был определен в примере 2.5, а его интегральное представление найдено в примере 3.3. Подставляя это представление в (4.5), найдём, что ядро имеет вид

$$
\left\{ h : h(t) = \int_{-\infty}^{\infty} c_\alpha \left((t-u)^{\frac{\alpha-1}{2}} \mathbf{1}_{u \le t} - (-u)^{\frac{\alpha-1}{2}} \mathbf{1}_{u \le 0} \right) \ell(u) \mathrm{d}u \right\},
$$

$\ell \in L_2(\mathbb{R})$. Иными словами, функция принадлежит ядру ДБД, если её дробная производная порядка $\frac{\alpha+1}{2}$ квадратично интегрируема.

Упражнение 4.1 Найдите ядро процесса Римана–Лиувилля, определённого в примере 3.4.

Пример 4.7 *(Ядро броуновского листа).* Напомним, что броуновский лист $W(t), t \in \mathbb{R}_+^d$, был определен в примере 2.7, а его интегральное представление найдено в примере 3.5. Подставляя это представление в (4.5), найдём, что ядро имеет вид

$$\left\{ h : h(t) = \int_{[0,t]} \ell(u)\lambda^d(\mathrm{d}u), \quad \ell \in L_2(\mathbb{R}_+^d) \right\}.$$

Иными словами, функция $h : \mathbb{R}_+^d \mapsto \mathbb{R}$ принадлежит ядру броуновского листа, если её смешанная производная $\frac{\partial^d}{\partial t_1 \ldots \partial t_d} h$ квадратично интегрируема.

Полезно будет узнать, как ведёт себя ядро при линейном преобразовании гауссовского вектора. Итак, пусть $X \in \mathcal{X}$ – центрированный гауссовский вектор с ковариационным оператором K_X и распределением $P = N(0, K_X)$, а $L : \mathcal{X} \to \mathcal{Y}$ – непрерывное линейное отображение. Рассмотрим $Y := LX$ – образ X под действием L. Очевидно, Y – гауссовский вектор в $\mathcal{Y}$ с ковариационным оператором $K_Y = LK_X L^*$ и распределением $Q := PL^{-1}$.

Предложение 4.1 *Верно равенство*

$$H_Q = L(H_P)$$

и

$$|v|_{H_Q} = \inf_{h \in L^{-1}v} |h|_{H_P}, \qquad \forall v \in H_Q.$$

Доказательство. Пусть I, I^* – канонические операторы, связанные с вектором X. Тогда мы находим факторизацию

$$K_Y = LK_X L^* = LII^*L^* = (LI)(LI)^*.$$

По теореме 4.1 имеем $H_Q = (LI)(\mathcal{X}_P^*) = L(H_P)$, а также для любого $v \in H_Q$

$$|v|_{H_Q} = \inf_{z \in (LI)^{-1}v} ||z||_{\mathcal{X}_P^*} = \inf_{z \in (LI)^{-1}v} |Iz|_{H_P} = \inf_{h \in L^{-1}v} |h|_{H_P}. \qquad \square$$

Проиллюстрируем действие этого предложения двумя примерами.

Пример 4.8 *(Ядро процесса Орнштейна–Уленбека).* Напомним, что процесс Орнштейна–Уленбека был определен в примере 2.6 именно как линейное преобразование винеровского процесса, $X(t) = e^{-t/2}W(e^t)$. Положим $\mathcal{X} = \mathbb{C}[0,e]$, $\mathcal{Y} = \mathbb{C}[0,1]$, а оператор L зададим как $(Lw)(t) = e^{-t/2}w(e^t)$. Согласно предложению 4.1, достаточно найти образ винеровского ядра (4.4) при отображении L. Пусть $v = Lh$. Тогда h можно частично выразить через v по формуле

$$h(s) = \sqrt{s}\; v(\ln s), \qquad 1 \le s \le e.$$

Отсюда $h(1) = v(0)$ и

$$\begin{aligned}
\int_1^e h'(s)^2 \mathrm{d}s &= \int_1^e \left(\frac{v(\ln s)}{2\sqrt{s}} + \frac{\sqrt{s}v'(\ln s)}{s}\right)^2 \mathrm{d}s \\
&= \int_1^e \left(\frac{v(\ln s)}{2} + v'(\ln s)\right)^2 \frac{\mathrm{d}s}{s} \\
&= \int_0^1 \left(\frac{v(t)}{2} + v'(t)\right)^2 \mathrm{d}t \\
&= \frac{1}{4}\int_0^1 v(t)^2 \mathrm{d}t + \int_0^1 v(t)v'(t)\mathrm{d}t + \int_0^1 v'(t)^2 \mathrm{d}t \\
&= \frac{1}{4}\int_0^1 v(t)^2 \mathrm{d}t + \frac{v(1)^2 - v(0)^2}{2} + \int_0^1 v'(t)^2 \mathrm{d}t.
\end{aligned}$$

Функция h не определена условием $Lh = v$ на интервале $[0,1]$ однозначно; известно лишь, что $h(0) = 0$ и $h(1) = v(0)$. При этих условиях минимум интеграла квадрата производной достигается на линейной функции и равен $v(0)^2$, поскольку

$$\int_0^1 h'(s)^2 \mathrm{d}s \ge \left(\int_0^1 h'(s)\mathrm{d}s\right)^2 = (h(1) - h(0))^2 = h(1)^2 = v(0)^2.$$

Получаем, что ядро H^U процесса Орнштейна–Уленбека в $\mathbb{C}[0,1]$ имеет вид

$$H^U = \{v:\ v \in AC[0,1],\ v' \in L_2[0,1]\}$$

и

$$|v|^2_{H^U} = \frac{v(1)^2 + v(0)^2}{2} + \frac{1}{4}\int_0^1 v(t)^2 \mathrm{d}t + \int_0^1 v'(t)^2 \mathrm{d}t, \qquad v \in H^U.$$

Стационарность процесса выражается в том, что все моменты времени входят в интегральное выражение для нормы равноправно и имеется симметрия по отношению к инверсии времени.

Упражнение 4.2 Найдите ядро процесса Орнштейна–Уленбека в $\mathbb{C}[a,b]$ для произвольного интервала $[a,b] \subset \mathbb{R}$. Решите аналогичную задачу для стационарного гауссовского процесса X с ковариационной функцией несколько более общего вида $\mathbb{E}X(s)X(t) = \alpha e^{-\beta|s-t|}$.

Пример 4.9 *(Ядро броуновского моста).* Напомним, что одно из возможных определений броуновского моста имеет вид

$$W^0(t) = W(t) - tW(1), \qquad 0 \le t \le 1,$$

где W – винеровский процесс. Получается, что $W^0 = LW$, где линейный оператор $L : \mathbb{C}[0,1] \mapsto \mathbb{C}[0,1]$ задан соотношением $(Lw)(t) = w(t) - tw(1)$. Следовательно, ядро броуновского моста H^0 получается из ядра винеровского процесса (4.4) применением L. Легко понять, что

$$H^0 = \{v : \ v \in AC[0,1], v(0) = v(1) = 0, \, v' \in L_2[0,1]\}.$$

Действительно, любая функция из правой части равенства принадлежит винеровскому ядру и удовлетворяет условию $Lh = h$. Поэтому она принадлежит H^0. Обратно, если функция h принадлежит образу ядра винеровской меры, то она удовлетворяет всем условиям из правой части равенства.

Остаётся найти гильбертову структуру в найденном ядре. Пусть $v \in H^0$ и $Lh = v$. Тогда $h(t) = v(t) + th(1)$, откуда $h'(t) = v'(t) + h(1)$ и

$$\begin{aligned} |h|^2 &= \int_0^1 h'(t)^2 \mathrm{d}t = \int_0^1 v'(t)^2 \mathrm{d}t + h(1)^2 + 2h(1)\int_0^1 v'(t)\mathrm{d}t \\ &= \int_0^1 v'(t)^2 \mathrm{d}t + h(1)^2, \end{aligned}$$

поскольку
$$\int_0^1 v'(t)\mathrm{d}t = v(1) - v(0) = 0.$$
Минимум указанного выражения по $h \in L^{-1}v$ достигается, очевидно, при $h(1) = 0$, т.е. при $h = v$. Мы приходим к выводу, что
$$|v|^2_{H^0} = \int_0^1 v'(t)^2\mathrm{d}t,$$
т.е. норма и скалярное произведение в H^0 – те же самые, что в ядре винеровской меры.

4.3 Альтернативные подходы к определению ядра

Воспроизводящее ядро (RKHS)

Одним из распространенных подходов к построению ядра является концепция *воспроизводящего ядра*[1] [45]. Пусть T – произвольное множество, и $K : T \times T \to \mathbb{R}$ – неотрицательно определённая функция (ядро). Пространство H, воспроизводящее ядро K, состоит из функций $f : T \mapsto \mathbb{R}$. Оно строится следующим образом. Берется линейная оболочка функций $K(t, \cdot), t \in T$; в ней вводится скалярное произведение исходя из соотношения
$$\langle K(s, \cdot), K(t, \cdot)\rangle := K(s, t), \qquad s, t \in T.$$
H получается как пополнение этой оболочки относительно гильбертова расстояния.

Установим связь этой конструкции с нашим понятием ядра. Пусть T – метрический компакт, $X \in \mathbb{C}(T)$ – центрированный гауссовский случайный процесс с непрерывными траекториями и ковариационной функцией $K(s, t)$. Для любого $t \in T$ рассмотрим функционал $(\delta_t, x) := x(t)$ и положим $h_t := II^*\delta_t$. Тогда
$$h_t(s) = (\delta_s, II^*\delta_t) = (I^*\delta_s, I^*\delta_t)_{\mathcal{X}^*_P} = \mathbb{E}X(s)X(t) = K(s, t),$$

[1] Reproducing kernel Hilbert space, RKHS.

т.е. $h_t = K(\cdot, t)$. Скалярное произведение находится по знакомой формуле

$$(h_s, h_t)_{H_P} = (II^*\delta_s, II^*\delta_t)_{H_P} = (I^*\delta_s, I^*\delta_t)_{\mathcal{X}_P^*} = K(s,t).$$

Таким образом, воспроизводящее ядро является замкнутым подпространством H_P. Приближая произвольные линейные функционалы линейными комбинациями δ-функционалов, нетрудно показать, что оба пространства в действительности совпадают.

Векторные интегралы

Представим себе, что нам удалось корректно определить векторнозначные интегралы, а именно, в обычных обозначениях,

$$\mathbb{E}\,[z(X)\,X] = \int_{\mathcal{X}} z(x)xP(\mathrm{d}x), \qquad z \in \mathcal{X}_P^*.$$

Тогда при всех $z \in \mathcal{X}_P^*, f \in \mathcal{X}^*$ верно

$$(f, \mathbb{E}z(X)X) = \mathbb{E}z(X)(f,X) = (z, I^*f)_{\mathcal{X}_P^*} = (f, Iz).$$

Получаем представление ядра меры P через векторные интегралы [61, 124]

$$H_P = \{Iz, z \in \mathcal{X}_P^*\} = \left\{ \int_{\mathcal{X}} z(x)xP(\mathrm{d}x),\ z \in \mathcal{X}_P^* \right\}.$$

Полярное представление

Рассмотрим множество линейных непрерывных функционалов

$$B = \left\{ f \in \mathcal{X}^* : \mathbb{E}(f,X)^2 = \int_{\mathcal{X}} (f,x)^2 P(\mathrm{d}x) = ||I^*f||^2_{\mathcal{X}_P^*} \le 1 \right\}.$$

Для любого $h = Iz \in H_P$ имеем

$$\sup_{f\in B} |(f,h)| = \sup_{f\in B} |(f,Iz)| = \sup_{f\in B} |(I^*f, z)_{\mathcal{X}_P^*}| \le ||z||_{\mathcal{X}_P^*} = |h|_{H_P}.$$

Приближая z непрерывными функционалами, нетрудно показать, что здесь имеет место равенство и

$$H_P = \left\{ h \in \mathcal{X} : \sup_{f \in B} |(f, h)| < \infty \right\}.$$

Глава 5

Теорема Камерона–Мартина

Пусть X – случайный вектор в линейном пространстве $\mathcal{X}$, а h – вектор в $\mathcal{X}$. Обозначим через P распределение X. Тогда распределение P_h вектора $X + h$ определяется формулой

$$P_h(A) = P(A - h)$$

на всех измеримых множествах $A \subset \mathcal{X}$ и называется *сдвигом* P на вектор h. Нас интересует вопрос об абсолютной непрерывности P_h относительно P. Напомним, что мера Q *абсолютно непрерывна* относительно P (записывается $Q \ll P$), если $P(A) = 0$ влечёт $Q(A) = 0$. Эквивалентное свойство: существует такая *плотность* $g := \frac{\mathrm{d}Q}{\mathrm{d}P} \in L_1(\mathcal{X}, P)$, что для всех измеримых множеств $A \subset \mathcal{X}$ верно

$$Q(A) = \int_A g dP.$$

Если $P_h \ll P$, то h называется *допустимым сдвигом* для P. Если ch – допустимый сдвиг для P при всех $c \in \mathbb{R}$, то говорят, что h определяет *допустимое направление* для P.

Нас интересует случай гауссовских мер, т.е. $P = N(a, K)$ и $P_h = N(a + h, K)$.

Суть происходящего можно усмотреть уже в одномерном случае.

Пример 5.1 Пусть $\mathcal{X} = \mathbb{R}$ и $P = N(0,1), h \in \mathbb{R}$. Тогда плотность меры P относительно меры Лебега есть

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\},$$

а для меры $P_h = N(h,1)$ имеем плотность

$$p_h(x) = \frac{1}{\sqrt{2\pi}} \exp\{-(x-h)^2/2\}.$$

Поэтому $P_h \ll P$ и

$$\frac{\mathrm{d}P_h}{\mathrm{d}P}(x) = \frac{p_h(x)}{p(x)} = \exp\left\{hx - \frac{h^2}{2}\right\}.$$

В общем случае не каждый сдвиг будет допустимым, но для допустимых сдвигов форма плотности будет именно такой: экспонента от линейного функционала с нормирующей константой квадратичного вида.

Теорема 5.1 *(Теорема Камерона–Мартина.) Пусть P – центрированная гауссовская мера в линейном пространстве $\mathcal{X}$. Тогда $P_h \ll P$ в том и только в том случае, когда верно $h \in H_P$. Если $h \in H_P$, то плотность $\frac{\mathrm{d}P_h}{\mathrm{d}P}$ имеет вид*

$$\frac{\mathrm{d}P_h}{\mathrm{d}P}(x) = \exp\left\{z(x) - \frac{|h|^2_{H_P}}{2}\right\}, \tag{5.1}$$

где $z \in \mathcal{X}_P^*$ – такой линейный измеримый функционал, что $Iz = h$.

Из теоремы следует, что любой допустимый сдвиг гауссовской меры определяет допустимое направление. Формула для плотности (5.1) называется *формулой Камерона–Мартина.*

Собственно говоря, американские математики Камерон и Мартин в 40-х годах XX века рассматривали не общий случай, а только винеровский процесс, для которого мы получаем следующее следствие. Пусть $\mathcal{X} = \mathbb{C}[0,1]$, $X = W$ – винеровский процесс, P и H_P – соответственно его распределение (винеровская мера) и ядро. Напомним, что

$$H_P = \{h \in AC[0,1], h(0) = 0,\ h' \in L_2[0,1]\},$$

$$|h|^2_{H_P} = \int_0^1 h'(s)^2 ds, \qquad h \in H_P.$$

Что касается функционала z, ассоциированного с вектором h формулой $Iz = h$, то нетрудно показать, что он имеет вид винеровского интеграла

$$z(w) = \int_0^1 h'(s) dw(s).$$

Действительно, с учётом изометрического свойства винеровских интегралов для любого $t \in [0,1]$ найдём

$$\begin{aligned} Iz(t) &= \delta_t(Iz) = (I^*\delta_t, z) = \mathbb{E}W(t)z(W) \\ &= \mathbb{E}\left(\int_0^1 \mathbf{1}_{[0,t]}(s)\mathrm{d}w(s) \cdot \int_0^1 h'(s)\mathrm{d}w(s)\right) \\ &= \int_0^t h'(s)\mathrm{d}s = h(t). \end{aligned}$$

Таким образом, формула Камерона–Мартина для винеровского процесса имеет вид

$$\frac{\mathrm{d}P_h}{\mathrm{d}P}(w) = \exp\left\{\int_0^1 h'(s)\mathrm{d}w(s) - \frac{1}{2}\int_0^1 h'(s)^2\mathrm{d}s\right\}.$$

Доказательство теоремы Камерона–Мартина.

1) *Достаточность.* Пусть $h \in H_P$. Рассмотрим меру

$$Q(dx) = \exp\left\{ z(x) - \frac{|h|^2_{H_P}}{2} \right\} \; P(dx).$$

Нам достаточно доказать, что $Q = P_h$. Для этого мы проверим совпадение характеристических функций. Пусть $f \in \mathcal{X}^*$. Тогда

$$\int_{\mathcal{X}} e^{i(f,x)} P_h(dx) = \mathbb{E} e^{i(f,X+h)} = e^{i(f,h)-(f,Kf)/2},$$

где X – вектор с распределением P. С другой стороны,

$$\begin{aligned} \int_{\mathcal{X}} e^{i(f,x)} Q(dx) &= \int_{\mathcal{X}} \exp\left\{ i(f,x) + z(x) - \frac{|h|^2_{H_P}}{2} \right\} \; P(dx) \\ &= \exp\left\{ -\frac{|h|^2_{H_P}}{2} \right\} \; \mathbb{E} \exp\left\{ i(f,X) + z(X) \right\}. \end{aligned}$$

Заметим, что двумерный случайный вектор $Y = ((f,X), z(X))$ является центрированным гауссовским с ковариационной матрицей

$$K^Y = \begin{pmatrix} (f,Kf) & (f,h) \\ (f,h) & |h|^2_{H_P} \end{pmatrix}.$$

Действительно, $\mathbb{E}(f,X)^2 = (f,Kf)$ по определению ковариационного оператора,

$$\mathbb{E}(f,X)z(X) = (I^* f, z)_{\mathcal{X}^*_P} = (f, Iz) = (f,h)$$

и

$$\mathbb{E} z(X)^2 = ||z||^2_{\mathcal{X}^*_P} = |h|^2_{H_P}.$$

По известной формуле для конечномерных гауссовских векторов верно

$$\mathbb{E} e^{v_1 Y_1 + v_2 Y_2} = \exp\left\{ \frac{K^Y_{11} v_1^2 + 2K^Y_{12} v_1 v_2 + K^Y_{22} v_2^2}{2} \right\}.$$

Подставляя сюда $v_1 = i, v_2 = 1$, получим

$$\mathbb{E} \exp\left\{ i(f,X) + z(X) \right\} = \exp\left\{ \frac{-(f,Kf) + 2i(f,h) + |h|^2_{H_P}}{2} \right\}.$$

Отсюда

$$\int_{\mathcal{X}} e^{i(f,x)} Q(dx) = \exp\left\{\frac{-(f,Kf)+2i(f,h)}{2}\right\},$$

и совпадение характеристических функций проверено.

2) *Необходимость*. Пусть h – допустимый сдвиг, покажем, что $h \in H_P$. Для этого мы должны найти такой линейный измеримый функционал $z \in \mathcal{X}_P^*$, что $h = Iz$.

Для каждого $f \in \mathcal{X}^*$ обозначим μ_f и ν_f распределения случайных величин (f, X) и $(f, X + h)$. Оба они являются нормальными: если $P = N(0, K)$, то $\mu_f = N(0, (f, Kf))$ и $\nu_f = N((f, h), (f, Kf))$.

Так как $P_h \ll P$, то для любого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для измеримого $A \subset \mathcal{X}$ из $P(A) \le \delta$ вытекает $P_h(A) \le \varepsilon$. В частности, для любого борелевского множества $B \subset \mathbb{R}$ из $\mu_f(B) \le \delta$ вытекает $\nu_f(B) \le \varepsilon$.

Это, в свою очередь, возможно только если разность барицентров распределений μ_f и ν_f соизмерима с их общим среднеквадратическим уклонением $(f, Kf)^{-1/2}$:

$$\sup_{f\in\mathcal{X}^*, f\neq 0} \frac{|(f,h)|}{(f,Kf)^{1/2}} < \infty. \tag{5.2}$$

Действительно, положим $\varepsilon = 1/2$ и выберем соответствующее δ. Возьмём

$$B = \{r \in \mathbb{R} : \ |r| \le \Phi^{-1}(1-\delta/2)\ (f,Kf)^{1/2}\}.$$

Тогда $\mu_f(\mathbb{R}\backslash B) = \delta$, откуда $\nu_f(\mathbb{R}\backslash B) \le 1/2$ и $\nu_f(B) \ge 1/2$.

С другой стороны, полагая

$$B' = \{r \in \mathbb{R} : \ |r - (f,h)| \le 2(f,Kf)^{1/2})\},$$

имеем $\nu_f(B') > 1/2$. Поэтому множества B и B' имеют непустое пересечение, т.е.

$$|(f,h)| \le \left(2 + \Phi^{-1}(1-\delta/2)\right)(f,Kf)^{1/2},$$

что и подтверждает (5.2). Учитывая, что $(f, Kf) = |I^*f|^2$, получим из (5.2)

$$\sup_{f\in\mathcal{X}^*, f\neq 0} \frac{|(f,h)|}{|I^*f|} < \infty.$$

Это означает, что линейный функционал $\mathcal{A}$, заданный соотношением $\mathcal{A}(I^* f) = (f, h)$, на плотном подпространстве $I^*\mathcal{X}^* \subset \mathcal{X}_P^*$ допускает продолжение до линейного непрерывного функционала на $\mathcal{X}_P^*$. По теореме Рисса об общей форме линейного непрерывного функционала на гильбертовом пространстве, найдётся такой элемент $z \in \mathcal{X}_P^*$, что

$$(f, h) = \mathcal{A}(I^* f) = (I^* f, z) = (f, Iz), \qquad f \in \mathcal{X}^*.$$

Отсюда следует, что $h = Iz$. $\square$

Интересной иллюстрацией полезности теоремы Камерона–Мартина является следующая оценка меры множества сдвинутой меры, принадлежащая шведскому математику К. Бореллю[1].

Предложение 5.1 *Неравенство Борелля для сдвинутых множеств* [62]. *Пусть P – центрированная гауссовская мера на линейном пространстве $\mathcal{X}$. Пусть $A \subset \mathcal{X}$ – симметричное множество, т.е. $A = -A$. Тогда для любого $h \in H_P$ верно*

$$P(A+h) \geq P(A)\ \exp\left\{-\frac{|h|_{H_P}^2}{2}\right\}. \tag{5.3}$$

Эта оценка замечательна своей простотой и особенно общностью.

Доказательство предложения 5.1. В силу симметрии P и A мы имеем $P(A+h) = P(-A+h) = P(A-h)$. Поэтому

$$\begin{aligned}
P(A+h) &= \frac{1}{2}\left(P(A+h) + P(A-h)\right) \\
&= \frac{1}{2}\left(P_{-h}(A) + P_h(A)\right) \\
&= \exp\left(-\frac{|h|_{H_P}^2}{2}\right) \int_A \frac{e^{-z(x)} + e^{z(x)}}{2}\, P(dx) \\
&\geq \exp\left(-\frac{|h|_{H_P}^2}{2}\right) P(A). \quad \square
\end{aligned}$$

[1] Christer Borell.

Замечание 5.1 *Позднее мы покажем, что для симметричных выпуклых множеств верно противоположное неравенство* $P(A) \geq P(A + h)$.

Глава 6

Изопериметрические неравенства

6.1 Евклидово пространство

Изопериметрические задачи знакомы нам из школьных курсов геометрии или физики в следующей форме: „среди всех тел заданного объёма найти то, которое имеет наименьшую площадь поверхности“ или „среди всех тел заданной площади поверхности найти то, которое имеет наибольший объём“. В случае евклидова объёма решением обеих задач является шар. Соответствующее изопериметрическое неравенство можно сформулировать следующим образом: если $A \subset \mathbb{R}^n$ – „гладкое“ множество и B - шар в $\mathbb{R}^n$, то

$$\lambda^n(A) = \lambda^n(B) \curvearrowright |\partial A| \geq |\partial B|,$$

где в правой части символ ∂ обозначает границу множества, а $|\cdot|$ – поверхностную меру. Эта формулировка не представляется удобной для перехода к бесконечномерным пространствам – слишком „хрупко“ понятие поверхностной меры. Поэтому перепишем изопериметрическое неравенство в несколько ином виде. Пусть B_r – замкнутый шар радиуса r с центром в нуле. Определим *r-расширение* множества A формулой

$$A^r := A + B_r.$$

Заменяя площадь поверхности на объём маленькой r-полоски вокруг множества, получим новую формулировку изопериметрического неравенства

$$\lambda^n(A) = \lambda^n(B) \curvearrowright \lambda^n(A^r \backslash A) \geq \lambda^n(B^r \backslash B),$$

что равносильно соотношению

$$\lambda^n(A) = \lambda^n(B) \curvearrowright \lambda^n(A^r) \geq \lambda^n(B^r). \tag{6.1}$$

Замечательно, что это верно *для всех* $r > 0$, а не только для малых.

Замечание 6.1 Изопериметрическое неравенство в евклидовом пространстве известно с античных времён. Первые доказательства на приемлемом для современной математики уровне строгости восходят к Я. Штейнеру [165]. Подробное изложение см. в [8].

Геометрическую формулировку (6.1) можно легко превратить в алгебраическую, где шар вообще не будет упоминаться. Учитывая, что объём евклидова шара вычисляется по формуле

$$\lambda^n(B_R) = c_n R^n, \qquad c_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}\ ,$$

мы видим, что радиус шара B, равновеликого множеству A, равен $\left(\frac{\lambda^n(A)}{c_n}\right)^{1/n}$, а радиус шара B^r равен $\left(\frac{\lambda^n(A)}{c_n}\right)^{1/n} + r$. Соответственно получаем

$$\lambda^n(A^r) \geq c_n \left(\left(\frac{\lambda^n(A)}{c_n}\right)^{1/n} + r\right)^n,$$

что равносильно

$$\left(\frac{\lambda^n(A^r)}{c_n}\right)^{1/n} \geq \left(\frac{\lambda^n(A)}{c_n}\right)^{1/n} + r.$$

Функцию $\varphi(v) := \left(\frac{v}{c_n}\right)^{1/n}$ можно назвать *изопериметрической функцией* пространства $(\mathbb{R}^n, \lambda^n)$.

6.2 Евклидова сфера

Следующим важным примером пространства с изопериметрическим свойством является евклидова сфера $\mathbb{S}^n$. Как известно, на ней определены естественное (геодезическое) расстояние $\rho_n(\cdot, \cdot)$ и единственная конечная мера σ^n, инвариантная относительно всех вращений сферы. Поскольку сфера не является векторным пространством, расширение множества $A \subset \mathbb{S}^n$ следует понимать как

$$A^r := \{x \in \mathbb{S}^n : \inf_{y \in A} \rho(x, y) \leq r\}$$

(в евклидовом пространстве такое определение равносильно приведённому выше). Изопериметрическое неравенство, доказанное П. Леви, утверждает (см. [8]), что, подобно евклидовому случаю, решением изопериметрической задачи является шар B, т.е.

$$\sigma^n(A) = \sigma^n(B) \curvearrowright \sigma^n(A^r) \geq \sigma^n(B^r).$$

Поскольку расширением шара на сфере является шар, то изопериметрическому неравенству на сфере также можно придать алгебраическую форму вида $\widetilde{\varphi}(\sigma^n(A^r)) \geq \widetilde{\varphi}(\sigma^n(A)) + r$ с подходящей изопериметрической функцией $\widetilde{\varphi}$.

Очень важно заметить, что шарами в $\mathbb{S}^n$ являются „шапочки“, которые можно получить как пересечения $\mathbb{S}^n$ с евклидовыми полупространствами. В дальнейшем мы рассмотрим предельный переход от $(\mathbb{S}^n, \sigma^n)$ к гауссовским распределениям, и не так уж удивительно, что решением изопериметрической задачи в гауссовском случае окажется полупространство.

6.3 Конструкция Пуанкаре

Теперь мы покажем, как можно перейти от равномерных распределений на подходящим образом выбранных сферах к стандартному гауссовскому распределению в $\mathbb{R}^n$ (эту конструкцию приписывают великому французскому математику А. Пуанкаре[1]). В частности, удаётся перенести на гауссовский случай изопериметрическое неравенство.

Пусть $m \geq n$. Обозначим $\pi_{m,n}$ естественную проекцию $\mathbb{R}^m \mapsto \mathbb{R}^n$, т.е.

$$\pi_{m,n}(x_1, \ldots, x_m) = (x_1, \ldots, x_n).$$

Пусть σ^m – единичная инвариантная мера на сфере $\widetilde{\mathbb{S}}^m$ радиуса $\sqrt{m}$ с центром в нуле в пространстве $\mathbb{R}^m$. Определим проекцию $\nu_{m,n} = \sigma^m \pi_{m,n}^{-1}$ меры σ^m на пространство $\mathbb{R}^n$ соотношением

$$\nu_{m,n}(A) = \sigma^m \left(\pi_{m,n}^{-1}(A)\right).$$

Предложение 6.1 *(Лемма Пуанкаре). При $m \to \infty$ меры $\nu_{m,n}$ слабо сходятся в $\mathbb{R}^n$ к стандартному гауссовскому распределению $N(0, E_n)$.*

Доказательство. Пусть $X_1, \ldots, X_n, \ldots, X_m$ – независимые стандартные нормальные случайные величины, образующие вектор $\mathbf{X^m} = (X_1, \ldots, X_m) \in \mathbb{R}^m$. Тогда σ^m будет распределением случайного вектора $\left(\frac{X_j \sqrt{m}}{||\mathbf{X^m}||}\right)_{1 \leq j \leq m}$ в $\mathbb{R}^m$, а $\nu_{m,n}$ – распределением вектора $\left(\frac{X_j \sqrt{m}}{||\mathbf{X^m}||}\right)_{1 \leq j \leq n}$ в $\mathbb{R}^n$. Остаётся заметить, что вектор $(X_j)_{1 \leq j \leq n}$ имеет нужное распределение $N(0, E_n)$, а поправочный множитель по закону больших чисел сходится к единице:

$$\frac{\sqrt{m}}{||\mathbf{X^m}||} = \sqrt{\frac{m}{X_1^2 + \cdots + X_m^2}} \Rightarrow 1. \quad \square$$

[1] H. Poincaré.

6.4 Евклидово пространство с гауссовской мерой

Рассмотрим теперь изопериметрическую задачу в пространстве $\mathbb{R}^n$ со стандартной гауссовской мерой $P = N(0, E_n)$. Понятие r-расширения множества остаётся таким же, как в случае лебеговой меры. Как независимо (и почти одновременно) показали К. Борелль [59] и В.Н. Судаков с Б.С. Цирельсоном [31], решением изопериметрической задачи является *полупространство* Π, т.е.

$$P(A) = P(\Pi) \curvearrowright P(A^r) \geq P(\Pi^r). \tag{6.2}$$

Придадим этому неравенству алгебраическую форму. Полупространство Π можно выбрать в виде

$$\Pi = \left\{x \in \mathbb{R}^n : \ x_1 \leq \Phi^{-1}(P(A))\right\},$$

где $\Phi^{-1}(\cdot)$ – функция, обратная к функции распределения стандартного нормального закона $\Phi(\cdot)$. Очевидно,

$$\Pi^r = \left\{x \in \mathbb{R}^n : \ x_1 \leq \Phi^{-1}(P(A)) + r\right\},$$

откуда $P(\Pi^r) = \Phi\left(\Phi^{-1}(P(A)) + r\right)$, и мы можем переписать (6.2) в виде

$$\Phi^{-1}(P(A^r)) \geq \Phi^{-1}(P(A)) + r. \tag{6.3}$$

Видим, что изопериметрической функцией для стандартной гауссовской меры является $\Phi^{-1}(\cdot)$.

Доказательство изопериметрического неравенства (6.3). Воспользуемся конструкцией Пуанкаре и соответствующими обозначениями. Пусть $A \subset \mathbb{R}^n$ и

$$\Pi_{\tilde{q}} = \{x \in \mathbb{R}^n : x_1 \leq \tilde{q}\}$$

– семейство полупространств. Выбирая значение параметра $q = \Phi^{-1}(P(A))$, получим $P(\Pi_q) = P(A)$. На сфере $\widetilde{\mathbb{S}}^m$ мы рассмотрим множество $A_m = \pi_{m,n}^{-1}(A)$ и „шапочки“ (шары)

$$B_{m,\tilde{q}} = \{x \in \widetilde{\mathbb{S}}^m : x_1 \leq \tilde{q}\} = \pi_{m,n}^{-1}(\Pi_{\tilde{q}}),$$

среди которых выделим

$$B_m := \{x \in \widetilde{\mathbb{S}}^m : x_1 \leq q_m\},$$

причём параметр q_m выбираем так, чтобы $\sigma^m(B_m) = \sigma^m(A_m)$.

По лемме Пуанкаре имеем

$$\lim_{m\to\infty} \sigma^m\left(B_{m,\tilde{q}}\right) = \Phi(\tilde{q}) < P(A), \qquad \forall \tilde{q} < q.$$

Сравнивая это с соотношением

$$\lim_{m\to\infty} \sigma^m\left(B_m\right) = \lim_{m\to\infty} \sigma^m\left(A_m\right) = \lim_{m\to\infty} \sigma^m\left(\pi_{m,n}^{-1}(A)\right) = P(A),$$

видим, что для любого $\tilde{q} < q$ при больших m верно

$$B_m \supset B_{m,\tilde{q}}. \tag{6.4}$$

Наш следующий аргумент – сжатие расстояний при проекции. Обозначая $\rho_m(\cdot,\cdot)$ геодезическое расстояние на сфере $\widetilde{\mathbb{S}}^m$, мы имеем

$$\rho_m(x,y) \geq ||x-y|| \geq ||\pi_{m,n}x - \pi_{m,n}y||, \qquad \forall x,y \in \widetilde{\mathbb{S}}^m.$$

Отсюда следует соотношение между расширениями множеств

$$(A_m)^r \subset \pi_{m,n}^{-1}\left(A^r\right). \tag{6.5}$$

Последовательно используя лемму Пуанкаре, включение (6.5), изопериметрическое неравенство Леви и (6.4), получим

$$\begin{aligned} P(A^r) &= \lim_{m\to\infty} \sigma^m \pi_{m,n}^{-1}(A^r) \geq \liminf_{m\to\infty} \sigma^m\left((A_m)^r\right) \\ &\geq \liminf_{m\to\infty} \sigma^m\left((B_m)^r\right) \geq \liminf_{m\to\infty} \sigma^m\left((B_{m,\tilde{q}})^r\right). \end{aligned}$$

Воспользуемся ещё таким свойством (сфера большого радиуса локально плоская): для любых $r,\varepsilon > 0$ и $\tilde{q} \in \mathbb{R}$ при достаточно больших m верно

$$(B_{m,\tilde{q}})^r \supset B_{m,\tilde{q}+r-\varepsilon}\ .$$

Отсюда

$$\begin{aligned}\liminf_{m\to\infty}\sigma^m\left((B_{m,\tilde q})^r\right) &\geq \liminf_{m\to\infty}\sigma^m\left(B_{m,\tilde q+r-\varepsilon}\right)\\ &= \liminf_{m\to\infty}\sigma^m\pi_{m,n}^{-1}(\Pi_{\tilde q+r-\varepsilon})\\ &= P\left(\Pi_{\tilde q+r-\varepsilon}\right)=\Phi(\tilde q+r-\varepsilon).\end{aligned}$$

Таким образом,

$$P(A^r)\geq\Phi(\tilde q+r-\varepsilon).$$

Устремляя $\varepsilon\to 0$, $\tilde q\to q$, получим (6.3). $\square$

Пусть теперь $P=N(0,K)$ – произвольная гауссовская мера в $\mathbb{R}^n$. Будем считать оператор K невырожденным. Проявим „здоровый конформизм“ и вместо того, чтобы доказывать что-то новое, видоизменим понятия так, чтобы воспользоваться результатом для стандартной меры. Напомним, что вектор с распределением $N(0,K)$ можно получить преобразованием $Y=LX$ из вектора X со стандартным гауссовским распределением, используя оператор $L=K^{1/2}$. Обозначим $D=\{x\in R^n:||L^{-1}x||\leq 1\}$ – L-образ единичного евклидова шара. Напомним также, что D совпадает с единичным шаром ядра (эллипсоидом рассеяния) меры $N(0,K)$ (см. пример 4.3).

Переопределим теперь понятие r-расширения, сообразно ситуации, как $A^r:=A+rD$. Докажем, что при таком определении неравенство (6.3) остаётся верным без изменений.

Действительно,

$$\begin{aligned}&\Phi^{-1}(P(A^r))=\Phi^{-1}(\mathbb{P}(Y\in A^r))=\Phi^{-1}(\mathbb{P}(LX\in A^r))\\ =\;&\Phi^{-1}(\mathbb{P}(X\in L^{-1}(A^r)))\\ =\;&\Phi^{-1}(\mathbb{P}(X\in L^{-1}(A)+L^{-1}(rD)))\\ =\;&\Phi^{-1}(\mathbb{P}(X\in L^{-1}(A)+B_r))=\Phi^{-1}(\mathbb{P}(X\in(L^{-1}(A))^r))\\ \geq\;&\Phi^{-1}(\mathbb{P}(X\in(L^{-1}(A))))+r=\Phi^{-1}(\mathbb{P}(Y=LX\in A))+r\\ =\;&\Phi^{-1}(P(A))+r.\end{aligned}$$

6.5 Общий случай

Пусть теперь $X \in \mathcal{X}$ – произвольный центрированный гауссовский вектор в линейном пространстве с распределением P и ядром H_P. Пусть $D = \{h \in H : \|h|_{H_P} \le 1\}$ – единичный шар ядра. Определим расширение множества $A \subset \mathcal{X}$ формулой $A^r := A + rD$. Мы можем в этой общей ситуации воспроизвести результат, полученный ранее для $\mathcal{X} = \mathbb{R}^n$.

Теорема 6.1 *Гауссовское изопериметрическое неравенство. Для любого измеримого $A \subset \mathcal{X}$ верно*

$$\Phi^{-1}(P_*(A^r)) \ge \Phi^{-1}(P(A)) + r, \tag{6.6}$$

где внутренняя мера $P_*(C)$ определяется для любого $C \subset \mathcal{X}$ формулой

$$P_*(C) := \sup_{B \subset C,\, B \text{ компактно}} P(B). \tag{6.7}$$

Равенство в (6.6) *достигается на полупространствах.*

Замечание 6.2 Можно избавиться от внутренней меры, записав, что для любого измеримого B верно

$$B \cap A^r = \emptyset \curvearrowright P(B) \le \widehat{\Phi}\left(\Phi^{-1}(P(A)) + r\right), \tag{6.8}$$

где $\widehat{\Phi}(\cdot) = 1 - \Phi(\cdot)$ – хвост распределения стандартного нормального закона.

6.6 Принцип концентрации

Одно из замечательных следствий изопериметрического неравенства – *принцип концентрации*, утверждающий, что липшицевы функционалы от гауссовского вектора имеют распределения, чья концентрация не уступает гауссовской.

Функционал $f : \mathcal{X} \mapsto \mathbb{R}$ называется H_P-*липшицевым* с константой σ (записываем $f \in Lip_{H_P}(\sigma)$), если

$$|f(x+h) - f(x)| \le \sigma |h|_{H_P}, \qquad \forall x \in \mathcal{X}, h \in H_P\ .$$

Пример 6.1 *(Супремум как липшицев функционал).* Пусть $\mathcal{X} = \mathbb{C}(T)$ и $X = (X(t))_{t\in T} \in \mathcal{X}$ – центрированный гауссовский процесс. Рассмотрим функционал

$$f(x) = \sup_{t\in T} x(t).$$

Тогда для любых $x \in \mathcal{X}, h = Iz \in H_P$ верно

$$\begin{aligned}
|f(x+h) - f(x)| &= \left|\sup_{t\in T} x(t) - \sup_{t\in T}(x(t)+h(t))\right| \le \sup_{t\in T}|h(t)| \\
&= \sup_{t\in T}|(\delta_t, Iz)| = \sup_{t\in T}|(I^*\delta_t, z)_{\mathcal{X}_P^*}| \\
&\le \sup_{t\in T}||I^*\delta_t||_{\mathcal{X}_P^*}||z||_{\mathcal{X}_P^*} \\
&= \sup_{t\in T}||I^*\delta_t||_{\mathcal{X}_P^*}\ |h|_{H_P} = \sigma|h|_{H_P},
\end{aligned}$$

где

$$\sigma^2 = \sup_{t\in T}||I^*\delta_t||^2_{\mathcal{X}_P^*} = \sup_{t\in T}\mathbb{E}X(t)^2.$$

Упражнение 6.1 Пусть X – центрированный гауссовский вектор в сепарабельном банаховом пространстве $(\mathcal{X}, ||\cdot||)$, а $K : \mathcal{X}^* \mapsto \mathcal{X}$ – ковариационный оператор X. Докажите, что функционал $f(x) = ||x||$ принадлежит классу $Lip_{H_P}\left(\sqrt{||K||}\right)$.

Напомним, что число m называется *медианой* распределения функционала f, если

$$\mathbb{P}(f(X) \le m) \ge \frac{1}{2} \text{ и } \mathbb{P}(f(X) \ge m) \ge \frac{1}{2}\ .$$

Медиана может быть не единственной, но это не сказывается на следующей оценке.

Теорема 6.2 *(Принцип концентрации). Если $f \in Lip_{H_P}(\sigma)$ и m – медиана функционала f, то для любого $r > 0$ верно*

$$\mathbb{P}(f(X) \ge m + r) \le \widehat{\Phi}\left(\frac{r}{\sigma}\right), \tag{6.9}$$

$$\mathbb{P}(f(X) \le m - r) \le \widehat{\Phi}\left(\frac{r}{\sigma}\right). \tag{6.10}$$

Равенство достигается, если f – линейный функционал.

Доказательство. Положим $A = \{x \in \mathcal{X} : f(x) \leq m\}$ и подставим соотношение $P(A) \geq \frac{1}{2}$ в изопериметрическое неравенство (6.6). Получим

$$\Phi^{-1}(P_*(A^{\frac{r}{\sigma}})) \geq \Phi^{-1}(P(A)) + \frac{r}{\sigma} \geq \frac{r}{\sigma} \ . \qquad (6.11)$$

С другой стороны, для любого $y \in A^{\frac{r}{\sigma}}$ имеется представление $y = x + h$ при $x \in A, |h|_{H_P} \leq \frac{r}{\sigma}$. Поэтому верно

$$f(y) \leq f(x) + \sigma|h|_{H_P} \leq m + r.$$

Это можно записать в виде

$$\{x \in \mathcal{X} : f(x) > m + r\} \cap A^{\frac{r}{\sigma}} = \emptyset,$$

откуда в силу (6.11)

$$\mathbb{P}(f(X) > m + r) \leq 1 - P_* \left(A^{\frac{r}{\sigma}}\right) \leq 1 - \Phi\left(\frac{r}{\sigma}\right) = \widehat{\Phi}\left(\frac{r}{\sigma}\right).$$

Для доказательства (6.10) достаточно применить (6.9) к функционалу $-f$. $\square$

Теорема 6.2 показывает, что распределения функционалов можно контролировать с помощью всего лишь двух числовых параметров, m и σ. Имея оценки вероятностей уклонений, нетрудно оценить и моментные характеристики функционала. Приведём только два простых примера.

Следствие 6.1 *Если* $f \in Lip_{H_P}(\sigma)$, *то*

$$\mathbb{E}\exp\{\alpha|f(X)|^2\} < \infty, \qquad 0 \leq \alpha < \frac{1}{2\sigma^2} \ .$$

Следствие 6.2 *Если* $f \in Lip_{H_P}(\sigma)$, *то*

$$\mathbb{D}f(X) \leq \sigma^2.$$

Действительно, в силу стандартных свойств дисперсии

$$\begin{aligned}
\mathbb{D}f(X) \quad &\leq \quad \mathbb{E}(f(X) - m)^2 = \int_0^\infty \mathbb{P}\left(|f(X) - m| \geq r^{1/2}\right) \mathrm{d}r \\
&\leq \quad 2\int_0^\infty \widehat{\Phi}\left(\frac{r^{1/2}}{\sigma}\right) \mathrm{d}r = \sigma^2.
\end{aligned}$$

Существует несколько альтернативных подходов к изопериметрическим неравенствам для гауссовских мер. Об одном из них (А. Эрхард [89]), основанном на преобразованиях симметризации, можно прочитать в [19]. С.Г. Бобков предложил подход на базе функциональных неравенств [57] и ещё один, более элементарный, см. [58].

Имеется обширная литература по изопериметрическим неравенствам и неравенствам концентрации для *негауссовских* мер, см. например, [124, 125, 174, 175].

Глава 7

Выпуклость мер и другие неравенства

7.1 Выпуклость мер

Функция $\varphi : \mathcal{X} \mapsto \mathbb{R}$, заданная на линейном пространстве $\mathcal{X}$, называется выпуклой вверх, если

$$\varphi(\alpha x + (1-\alpha)y) \geq \alpha\varphi(x) + (1-\alpha)\varphi(y)$$

для всех $x, y \in \mathcal{X}$ и $\alpha \in [0,1]$. Аналогичное понятие можно ввести и для мер. Самым естественным обобщением было бы свойство

$$\mu(\alpha A + (1-\alpha)B) \geq \alpha\mu(A) + (1-\alpha)\mu(B),$$

но, к сожалению, интересных мер μ с таким свойством просто нет. Однако возможно ввести очень близкое понятие выпуклости, зависящее от некоторой функции Q:

$$\mu(\alpha A + (1-\alpha)B) \geq Q(\mu(A), \mu(B), \alpha)$$

для всех измеримых $A, B \subset \mathcal{X}$ и всех $\alpha \in [0,1]$. Например, знаменитое неравенство Брунна–Минковского [8, 99] как раз утверждает, что для меры Лебега в $\mathbb{R}^n$ верно

$$\begin{aligned} & \left[\lambda^n(\alpha A + (1-\alpha)B)\right]^{1/n} \\ \geq \quad & \alpha \left[\lambda^n(A)\right]^{1/n} + (1-\alpha)\left[\lambda^n(B)\right]^{1/n}. \end{aligned} \tag{7.1}$$

Рассмотрим ещё две разновидности выпуклости мер – логарифмическую выпуклость и выпуклость по Эрхарду.

Мера μ называется *логарифмически выпуклой*, если

$$\ln \mu(\alpha A + (1-\alpha)B) \geq \alpha \ln \mu(A) + (1-\alpha) \ln \mu(B) \tag{7.2}$$

или, что то же самое,

$$\mu(\alpha A + (1-\alpha)B) \geq \mu(A)^\alpha \mu(B)^{1-\alpha}.$$

Комбинируя неравенство Брунна–Минковского (7.1) с выпуклостью вверх логарифмической функции, видим, что для меры Лебега в $\mathbb{R}^n$ верно (7.2). Преимущество логарифмической выпуклости (7.2) перед (7.1) состоит в том, что (7.2) инвариантно по отношению к размерности пространства и потому может быть распространено на бесконечномерный случай.

К. Борелль [60] доказал, что любая мера в $\mathbb{R}^n$ с плотностью вида $p(x) = e^{\varphi(x)}$, где $\varphi(\cdot)$ – выпуклая вверх функция, является логарифмически выпуклой. В частности, для гауссовской меры $N(a,K)$ с невырожденным оператором K

$$\varphi(x) = \ln\left((2\pi)^{-n/2}(detK)^{-1/2}\right) - \frac{1}{2}\left(K^{-1}(x-a),(x-a)\right)$$

– выпуклая вверх квадратичная функция. Несложные предельные переходы показывают, что любая гауссовская мера в линейном пространстве логарифмически выпукла.

Поскольку свойство гауссовости не является чем-то особенным с точки зрения определения логарифмической выпуклости, то неудивительно, что нет интересных множеств, для которых в (7.2) для гауссовской меры достигалось бы *равенство.*

Мера μ называется *выпуклой по Эрхарду*, если

$$\begin{aligned} & \Phi^{-1}\left(\mu_*(\alpha A + (1-\alpha)B)\right) \\ \geq \quad & \alpha\Phi^{-1}\left(\mu(A)\right) + (1-\alpha)\Phi^{-1}\left(\mu(B)\right). \end{aligned} \tag{7.3}$$

Здесь $\Phi^{-1}(\cdot)$, как обычно, – функция, обратная к функции распределения стандартного нормального закона, а μ_* в левой части – внутренняя мера, определённая в (6.7). Любая гауссовская мера является выпуклой по Эрхарду[1], причём равенство в (7.3) достигается на полупространствах с параллельными границами.

Из любого вида выпуклости можно извлечь такое полезное следствие.

Следствие 7.1 *(Неравенство Андерсона* [43]*). Пусть P – центрированная гауссовская мера в линейном пространстве $\mathcal{X}$ и A – измеримое выпуклое центрально-симметричное подмножество $\mathcal{X}$. Тогда для любого $h \in \mathcal{X}$ верно*

$$P(A+h) \leq P(A).$$

Доказательство. Положим $A' = A + h$ и $B' = A - h$. Поскольку A симметрично, то

$$B' = A - h = -A - h = -(A+h) = -A'.$$

В силу симметрии меры P имеем $P(A') = P(B')$.

Применим свойство (7.2) к множествам A' и B'. Получим

$$\ln P\left(\frac{A'+B'}{2}\right) \geq \frac{1}{2}\ln P(A') + \frac{1}{2}\ln P(B') = \ln P(A').$$

Наконец, из выпуклости A следует

$$\frac{A'+B'}{2} = \frac{A+h+A-h}{2} = \frac{A+A}{2} = A.$$

Поэтому $P(A) \geq P(A') = P(A+h)$. □

[1] Это очень непростой результат. Французский математик А. Эрхард (A. Ehrhard [89]) доказал (7.3) в 1983 г. для выпуклых множеств. Позднее в 1996 г. Р. Латала [121] доказал это соотношение для случая, когда только одно из множеств является выпуклым. Наконец, К. Борелль [63] исследовал общий случай в 2003 г. О дальнейших улучшениях см. [52, 64, 100].

Пример 7.1 *(Прохождение процесса вблизи кривой).* Пусть $X(t), t \in T$, – гауссовский процесс с нулевым средним и задана переменная ширина „коридора“ $\varepsilon : T \mapsto \mathbb{R}_+$. Рассмотрим симметричное выпуклое множество функций

$$A = \{x : T \mapsto \mathbb{R}, |x(t)| \leq \varepsilon(t), \forall t \in T\}.$$

По неравенству Андерсона для любой функции $h : T \mapsto \mathbb{R}$ верно

$$\begin{aligned} \mathbb{P}(X \in A) &= \mathbb{P}\{|X(t)| \leq \varepsilon(t), \forall t \in T\} \\ &\geq \mathbb{P}\{|X(t) - h(t)| \leq \varepsilon(t), \forall t \in T\}. \end{aligned}$$

Таким образом, график гауссовского процесса с большей вероятностью проходит по полосе вокруг его математического ожидания, чем по полосе той же формы, но с другим центром.

Другое интересное приложение выпуклости – исследование распределения выпуклого функционала от гауссовского вектора X. Пусть $\varphi : \mathcal{X} \mapsto \mathbb{R}$ – функция, выпуклая вниз, т.е.

$$\varphi(\alpha x + (1-\alpha)y) \leq \alpha\varphi(x) + (1-\alpha)\varphi(y)$$

для всех $x, y \in \mathcal{X}$ и $\alpha \in [0,1]$. Обозначим $F(r) = \mathbb{P}\{\varphi(X) \leq r\}$ функцию распределения функционала $\varphi(X)$.

Предложение 7.1 *Функция*

$$r \mapsto \Phi^{-1}(F(r))$$

выпукла вверх.

Доказательство. Для любых $r_1, r_2 \in \mathbb{R}$ положим

$$A = \{x \in \mathcal{X} : \varphi(x) \leq r_1\}, \qquad B = \{x \in \mathcal{X} : \varphi(x) \leq r_2\}.$$

Тогда для любого $\alpha \in [0,1]$ верно

$$\alpha A + (1-\alpha)B \subset \{x \in \mathcal{X} : \varphi(x) \leq \alpha r_1 + (1-\alpha) r_2\}.$$

Из неравенства Эрхарда (7.3) получаем

$$\begin{aligned}
& \Phi^{-1}\left(F(\alpha r_1 + (1-\alpha) r_2)\right) \\
= \; & \Phi^{-1}\left(\mathbb{P}\{\varphi(X) \le \alpha r_1 + (1-\alpha) r_2\}\right) \\
\ge \; & \Phi^{-1}\left(\mathbb{P}\{X \in \alpha A + (1-\alpha) B\}\right) \\
\ge \; & \alpha \Phi^{-1}\left(\mathbb{P}\{X \in A\}\right) + (1-\alpha)\Phi^{-1}\left(\mathbb{P}\{X \in B\}\right) \\
= \; & \alpha \Phi^{-1}\left(F(r_1)\right) + (1-\alpha)\Phi^{-1}\left(F(r_2)\right). \qquad \square
\end{aligned}$$

Из этого предложения следует, что распределение случайной величины $\varphi(X)$ обладает примерно той же степенью регулярности, что и выпуклые функции общего вида на прямой. А именно, если Q – распределение $\varphi(X)$, то Q абсолютно непрерывно относительно меры Лебега за исключением возможного атома в точке отрыва $r_* = \inf\{r : F(r) > 0\}$. На интервале (r_*, ∞) мера Q имеет плотность, непрерывную всюду, кроме не более чем счётного числа точек, где она имеет скачки вниз.

Следствие 7.2 *Пусть m - медиана распределения случайной величины $\varphi(X)$. Тогда*

$$m \le \mathbb{E}\,\varphi(X). \tag{7.4}$$

Доказательство. Будем для простоты считать, что атома в распределении $\varphi(X)$ нет. Тогда $F(m) = \frac{1}{2}$ и величина $F(\varphi(X))$ равномерно распределена на отрезке $[0,1]$. По определению, $\Phi^{-1}\left(F(m)\right) = \Phi^{-1}\left(\frac{1}{2}\right) = 0$. С другой стороны, используя выпуклость функции $\Phi^{-1}\left(F(\cdot)\right)$ и неравенство Йенсена, получим

$$\Phi^{-1}\left(F(\mathbb{E}\varphi(X))\right) \ge \mathbb{E}\Phi^{-1}\left(F(\varphi(X))\right) = \int_0^1 \Phi^{-1}(p)\mathrm{d}p = 0.$$

Утверждение следствия вытекает из монотонности функции $\Phi^{-1}\left(F(\cdot)\right)$. $\square$

7.2 Растяжения

Большое число гауссовских неравенств посвящено сравнению гауссовских мер множеств A и rA, где A - центрально-симметричное выпуклое множество. Наиболее известно здесь так называемое S-свойство, которое сформулировали как гипотезу[2] Квапень и Сава [120], а позднее доказали Латала и Олешкевич [123][3]. Оно говорит об экстремальном свойстве полос вида

$$S_{r,f} := \{x \in \mathcal{X} : |(f,x)| \le r\}, \qquad r > 0, f \in \mathcal{X}^*.$$

Пусть P – распределение центрированного гауссовского вектора X. Не теряя общности, мы можем предположить, что $\mathbb{E}(f,X)^2 = 1$. Тогда

$$P\left(S_{r,f}\right) = 2\Phi(r) - 1 := G(r).$$

S-свойство утверждает, что среди центрально-симметричных выпуклых множеств полоса наиболее медленно теряет массу при сжатии и наиболее медленно увеличивает её при растяжении. Аналитически это выражается неравенствами

$$P(A) = G(\rho) \curvearrowright P(rA) \ge G(r\rho), \qquad r \ge 1;$$

$$P(A) = G(\rho) \curvearrowright P(rA) \le G(r\rho), \qquad r \le 1.$$

Свойства такого типа широко применяются в геометрических приложениях теории вероятностей, однако для теории случайных процессов они пока не используются, потому что оценка $G(r\rho)$ убывает при $r \to 0$ степенным образом, в то время как для типичных множеств A, связанных с процессами, $P(rA)$ при $r \to 0$ убывает экспоненциально.

Упомянем, наконец, более специальное свойство – *B-выпуклость*, введённое как гипотеза В. Банащиком и окончательно обоснованное позднее Д. Кордеро-Эроскином, М. Фраделизи и Б. Морэ[4]. По сравнению с рассмотренными выше видами выпуклости оно относится к более узкому классу множеств и может быть сформулировано следующим образом.

Теорема 7.1 *Пусть P – центрированное гауссовское распределение в $\mathcal{X}$, и A центрально-симметричное выпуклое*

---

[2] S-conjecture.

[3] S. Kwapień и J. Sawa; R. Latała и K. Oleszkiewicz.

[4] W. Banaszczyk; D. Cordero-Erausquin, M. Fradelizi, B. Maurey [71].

подмножество $\mathcal{X}$. Тогда функция

$$t \mapsto \ln P(e^t A)$$

выпукла вверх при $t \in (0, \infty)$.

Из этой теоремы следует, например, что

$$P\left(e^{\frac{t_1+t_2}{2}} A\right) \geq P\left(e^{t_1} A\right)^{1/2} P\left(e^{t_2} A\right)^{1/2}.$$

Полагая $c_1 = e^{t_1}$, $c_2 = e^{t_2}$, получим

$$P\left(\sqrt{c_1 c_2} A\right) \geq P\left(c_1 A\right)^{1/2} P\left(c_2 A\right)^{1/2},$$

в то время как обычная логарифмическая выпуклость (7.2) даёт более слабое неравенство

$$\begin{aligned} P\left(\frac{c_1+c_2}{2} A\right) &= P\left(\frac{1}{2}(c_1 A) + \frac{1}{2}(c_2 A)\right) \\ &\geq P\left(c_1 A\right)^{1/2} P\left(c_2 A\right)^{1/2}. \end{aligned}$$

7.3 Корреляционная гипотеза

Если X_1 и X_2 – два независимых вектора в пространстве $\mathcal{X}$, то для любых измеримых множеств $A_1, A_2 \subset \mathcal{X}$ верно

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1)\ \mathbb{P}(X_2 \in A_2).$$

Имея дело с зависимыми *гауссовскими* векторами, мы всё же можем утверждать, что для множеств A_1, A_2 из некоторого класса верно

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) \geq \mathbb{P}(X_1 \in A_1)\ \mathbb{P}(X_2 \in A_2). \tag{7.5}$$

Знаменитая *корреляционная гипотеза*, едва ли не самая привлекательная открытая проблема теории гауссовских процессов, утверждает, что (7.5) верно, если X_1, X_2 – центрированные гауссовские векторы, а A_1, A_2 – симметричные выпуклые множества. О её истории и различных формулировках см. [162].

Корреляционную гипотезу можно записать и в несколько ином виде. Пусть X – центрированный гауссовский вектор, а $\mathcal{A}_1, \mathcal{A}_2$ – симметричные выпуклые множества. Тогда

$$\mathbb{P}(X \in \mathcal{A}_1 \cap \mathcal{A}_2) \geq \mathbb{P}(X \in \mathcal{A}_1)\ \mathbb{P}(X \in \mathcal{A}_2), \tag{7.6}$$

что с первого взгляда выглядит как парадоксальная „независимость X от самого себя“. Связь между двумя видами неравенств проясняется, если положить $X = (X_1, X_2) \in \mathcal{X} \times \mathcal{X}$ и определить симметричные выпуклые множества формулами

$$\begin{aligned} \mathcal{A}_1 &= \{(x_1, x_2) \in \mathcal{X} \times \mathcal{X} :\ x_1 \in A_1\}, \\ \mathcal{A}_2 &= \{(x_1, x_2) \in \mathcal{X} \times \mathcal{X} :\ x_2 \in A_2\}. \end{aligned}$$

Тогда (7.6) переходит в (7.5). С другой стороны, полагая $X = X_1 = X_2$ в (7.5), получим (7.6).

Корреляционная гипотеза (7.6) доказана в некоторых важных частных случаях:

- для $\mathcal{X} = \mathbb{R}^2$ (Л. Питт [159]).

- в случае, когда одно из множеств $\mathcal{A}_1$ или $\mathcal{A}_2$ является полосой, например,

  $$\mathcal{A}_2 = \{x \in \mathcal{X} : |(f, x)| \leq r\}, \qquad r \geq 0, f \in \mathcal{X}^*$$

  (результат получен независимо Ч.Г. Кхатри и З. Шидаком [111, 164]).

- в случае, когда одно из множеств $\mathcal{A}_1$ или $\mathcal{A}_2$ является симметричным эллипсоидом (Ж. Арже [103]).

- в некоторых других частных случаях ... (см. ссылки в [122]).

Используя утверждение для полосы шаг за шагом, получим для полос вида $\mathcal{A}_j = \{x \in \mathcal{X} : |(f_j, x)| \leq r_j\}$, $j = 2, 3, \ldots$

$$\begin{aligned} &\mathbb{P}(X \in \mathcal{A}_1 \cap (\mathcal{A}_2 \cap \mathcal{A}_3)) \geq \mathbb{P}\,(X \in (\mathcal{A}_1 \cap \mathcal{A}_2))\ \mathbb{P}\,(X \in \mathcal{A}_3) \\ \geq\ &\mathbb{P}\,(X \in \mathcal{A}_1)\ \mathbb{P}\,(X \in \mathcal{A}_2)\ \mathbb{P}\,(X \in \mathcal{A}_3)\,, \end{aligned}$$

и, продолжая в том же духе, приходим к следующему неравенству

$$\mathbb{P}\left(X \in \mathcal{A}_1 \bigcap \left(\bigcap_{j=2}^{n} \mathcal{A}_j\right)\right) \geq \mathbb{P}\left(X \in (\mathcal{A}_1)\right) \prod_{j=2}^{n} \mathbb{P}(X \in \mathcal{A}_j).$$

В частности, при $A_1 = \mathcal{X}$ находим

$$\mathbb{P}\left(X \in \bigcap_{j=2}^{n} \mathcal{A}_j\right) \geq \prod_{j=2}^{n} \mathbb{P}\left(X \in \mathcal{A}_j\right). \tag{7.7}$$

Последнее соотношение известно как *неравенство Кхатри–Шидака* [111, 164][5]. Оно, разумеется, верно и для пересечения счётного числа полос.

Хотя в полном объёме корреляционная гипотеза не доказана, очень часто оказывается достаточным следующий более слабый вариант, авторами которого являются Г. Шехтман, Т. Шлумпрехт и Дж. Зинн [162] (для случая $1 - \varepsilon = \sqrt{\frac{1}{2}}$) и Венбо Ли [129] (для произвольного ε)[6].

Теорема 7.2 *(Слабое корреляционное неравенство). Для любого $\varepsilon \in (0,1)$ существует такое $K_\varepsilon > 0$, что для всех центрированных гауссовских векторов X и всех симметричных выпуклых $\mathcal{A}_1, \mathcal{A}_2$ верно*

$$\mathbb{P}\left(X \in \mathcal{A}_1 \cap \mathcal{A}_2\right) \geq \mathbb{P}\left(X \in (1-\varepsilon)\mathcal{A}_1\right)\ \mathbb{P}\left(X \in \frac{\mathcal{A}_2}{K_\varepsilon}\right). \tag{7.8}$$

Доказательство теоремы 7.2. Пусть $\tilde{X}$ – вектор, имеющий то же распределение, что X, и независимый с X. Нетрудно проверить, что для любого $a > 0$ векторы $X - a\tilde{X}$ и $X + a^{-1}\tilde{X}$ независимы. Для этого достаточно проверить, что значения любых линейных функционалов от этих векторов

[5] C. Khatri–Z. Šidák inequality.

[6] G. Schechtman, T. Schlumprecht, J. Zinn, Wenbo V. Li.

некоррелированы, а мы действительно имеем

$$\begin{aligned} & \mathbb{E}(f, X - a\tilde{X})(g, X + a^{-1}\tilde{X}) \\ = \; & \mathbb{E}(f,X)(g,X) + a^{-1}\mathbb{E}(f,X)(g,\tilde{X}) \\ & -a\,\mathbb{E}(f,\tilde{X})(g,X) - \mathbb{E}(f,\tilde{X})(g,\tilde{X}) \\ = \; & \mathbb{E}(f,X)(g,X) + a^{-1}\mathbb{E}(f,X)\mathbb{E}(g,\tilde{X}) \\ & -a\,\mathbb{E}(f,\tilde{X})\mathbb{E}(g,X) - \mathbb{E}(f,\tilde{X})(g,\tilde{X}) \\ = \; & 0. \end{aligned}$$

Заметим, что в силу устойчивости гауссовских распределений вектор $X - a\tilde{X}$ равнораспределен с $(1+a^2)^{1/2}X$, а вектор $X + a^{-1}\tilde{X}$ равнораспределен с вектором $(1+a^{-2})^{1/2}X$.

Далее, используя неравенство Андерсона применительно к вектору $(X, X) \in \mathcal{X} \times \mathcal{X}$, множеству $\mathcal{A}_1 \times \mathcal{A}_2 \subset \mathcal{X} \times \mathcal{X}$ и случайному сдвигу $(a\tilde{X}, -a^{-1}\tilde{X}) \in \mathcal{X} \times \mathcal{X}$, найдём

$$\begin{aligned} & \mathbb{P}(X \in \mathcal{A}_1, X \in \mathcal{A}_2) \\ \geq \; & \int_{\mathcal{X}} \mathbb{P}(X \in \mathcal{A}_1 + ah, X \in \mathcal{A}_2 - a^{-1}h) P(\mathrm{d}h) \\ = \; & \mathbb{P}(X - a\tilde{X} \in \mathcal{A}_1, X + a^{-1}\tilde{X} \in \mathcal{A}_2) \\ = \; & \mathbb{P}(X - a\tilde{X} \in \mathcal{A}_1)\; \mathbb{P}(X + a^{-1}\tilde{X} \in \mathcal{A}_2) \\ = \; & \mathbb{P}((1+a^2)^{1/2}X \in \mathcal{A}_1)\; \mathbb{P}((1+a^{-2})^{1/2}X \in \mathcal{A}_2). \end{aligned}$$

Подставляя сюда $a = \frac{(2\varepsilon-\varepsilon^2)^{1/2}}{1-\varepsilon}$, получим (7.8) с константой $K_\varepsilon = (2\varepsilon - \varepsilon^2)^{1/2}$. $\square$

Приведём типичное приложение слабого корреляционного неравенства – к теории малых уклонений. Пусть X – центрированный гауссовский вектор в нормированном пространстве $(\mathcal{X}, ||\cdot||)$, допускающий представление в виде суммы $X = X_1 + X_2$, где X_1, X_2 – гауссовские, возможно зависимые, причём X_1 хорошо изучен, а X_2 сравнительно мал. Тогда, используя неравенство треугольника и (7.8), можем получить

$$\begin{aligned} & \mathbb{P}(||X|| \leq r) = \mathbb{P}(||X_1 + X_2|| \leq r) \\ \geq \; & \mathbb{P}(||X_1|| \leq (1-\varepsilon)r, ||X_2|| \leq \varepsilon r) \\ \geq \; & \mathbb{P}\left(||X_1|| \leq (1-\varepsilon)^2 r\right)\; \mathbb{P}\left(||X_2|| \leq \frac{\varepsilon r}{K_\varepsilon}\right). \end{aligned} \tag{7.9}$$

Если вектор X_2 существенно меньше, чем X_1, то второй множитель будет менее существенным, чем первый, для асимптотического поведения вероятности при $r \to 0$, и малые уклонения суммы сводятся к малым уклонениям одного из слагаемых. Из найденной оценки снизу легко получается и оценка сверху. Записывая $X_1 = (X_1 + X_2) + (-X_2)$ и заменяя r на $(1-\varepsilon)^{-2} r$, получим

$$
\begin{aligned}
&\mathbb{P}(||X_1|| \le (1-\varepsilon)^{-2} r) \\
\ge\ &\mathbb{P}\left(||X_1 + X_2|| \le r\right)\ \mathbb{P}\left(||X_2|| \le \frac{\varepsilon(1-\varepsilon)^{-2} r}{K_\varepsilon}\right),
\end{aligned}
$$

т.е.

$$
\mathbb{P}\left(||X|| \le r\right) \le \frac{\mathbb{P}(||X_1|| \le (1-\varepsilon)^{-2} r)}{\mathbb{P}\left(||X_2|| \le \frac{\varepsilon(1-\varepsilon)^{-2} r}{K_\varepsilon}\right)}\,. \tag{7.10}
$$

7.4 Экстремальное изменение меры при сдвигах

Изопериметрическое неравенство позволяет определить, насколько может увеличиться мера множества при переходе к его ε-расширению. В этом параграфе мы посмотрим, насколько может измениться мера множества при его *сдвиге*. Это оказывается сравнительно простой задачей. Она решена Дж. Кэлбсом и Венбо Ли[7] в [117].

Теорема 7.3 *Пусть X – центрированный гауссовский вектор в $\mathcal{X}$ с распределением P. Тогда для любого измеримого $A \subset \mathcal{X}$ и для любого $h \in H_P$ верно*

$$
\Phi^{-1}\left(P(A)\right) - |h|_{H_P} \le \Phi^{-1}\left(P(A+h)\right) \le \Phi^{-1}\left(P(A)\right) + |h|_{H_P}.
$$

Равенство в обоих неравенствах достигается на соответствующих полупространствах.

Доказательство теоремы 7.3. Чтобы не погрязнуть в деталях, рассмотрим только случай $\mathcal{X} = \mathbb{R}^n$, $P = N(0,E)$ и $|h|_{H_P} = 1$. Будем доказывать только оценку сверху.

[7] J. Kuelbs, Wenbo V. Li.

Рассмотрим полупространство

$$\Pi = \{x \in \mathbb{R}^n : (h,x) \le r\}, \qquad \text{где } r = \Phi^{-1}(P(A)).$$

Тогда $P(\Pi) = \mathbb{P}((h,X) \le r) = \Phi(r) = P(A)$ и

$$\Pi + h = \{x \in \mathbb{R}^n : (h,x) \le r+1\}.$$

Мы проверим, что

$$P(A+h) \le P(\Pi + h). \tag{7.11}$$

Тогда получится требуемое

$$\begin{aligned} \Phi^{-1}\left(P(A+h)\right) &\le \Phi^{-1}\left(P(\Pi+h)\right) \\ &= \Phi^{-1}\left(\mathbb{P}((h,X) \le r+1)\right) = r+1 \\ &= \Phi^{-1}\left(P(A)\right) + 1 = \Phi^{-1}\left(P(A)\right) + |h|_{H_P}. \end{aligned}$$

По формуле Камерона–Мартина, мы имеем

$$\begin{aligned} P(A+h) &= P_{-h}(A) = \int_A e^{-\frac{1}{2}-(h,x)} P(dx) \\ &= \int_{A\cap\Pi} e^{-\frac{1}{2}-(h,x)} P(dx) + \int_{A\setminus\Pi} e^{-\frac{1}{2}-(h,x)} P(dx) \end{aligned}$$

и

$$\begin{aligned} P(\Pi+h) &= P_{-h}(\Pi) = \int_\Pi e^{-\frac{1}{2}-(h,x)} P(dx) \\ &= \int_{A\cap\Pi} e^{-\frac{1}{2}-(h,x)} P(dx) + \int_{\Pi\setminus A} e^{-\frac{1}{2}-(h,x)} P(dx). \end{aligned}$$

Первые слагаемые двух выражений совпадают, а для вторых, учитывая

$$P(A\setminus\Pi) = P(A) - P(A\cap\Pi) = P(\Pi) - P(A\cap\Pi) = P(\Pi\setminus A)$$

и определение Π, мы находим нужное неравенство

$$\begin{aligned} &\int_{A\setminus\Pi} e^{-(h,x)} P(dx) \le e^{-r} P(A\setminus\Pi) \\ &= e^{-r} P(\Pi\setminus A) \le \int_{\Pi\setminus A} e^{-(h,x)} P(dx). \end{aligned}$$

Отсюда следует (7.11). $\square$

Упражнение 7.1 Повторите доказательство для $h \in \mathbb{R}^n$ с произвольной нормой. Затем выведите теорему для произвольного гауссовского вектора в $\mathbb{R}^n$.

Дальнейшую информацию о гауссовских и связанных с ними неравенствах можно найти в обзорах [51, 122].

Глава 8

Принцип больших уклонений

8.1 Общий принцип больших уклонений

В этой главе рассматривается версия принципа больших уклонений применительно к гауссовским распределениям. В полной общности принцип больших уклонений представляет собой методику оценивания вероятности того, что случайный объект находится вдали от своего „типичного положения“. Он часто используется в математической статистике для сопоставления теоретических моделей и практических данных, в статистической механике и т.д. Наиболее простым и типичным примером принципа больших уклонений является следующий результат.

Теорема 8.1 *(Теорема Крамера–Чернова* [69, 72]*). Пусть $X_1, X_2 \ldots$ – последовательность независимых одинаково распределённых случайных величин, удовлетворяющих условию*

$$\mathbb{E}\exp\{\gamma X_j\} < \infty, \qquad |\gamma| < \gamma_0$$

при некотором $\gamma_0 > 0$. Обозначим $S_n = \sum_{j=1}^n X_j$. Тогда для *любого $r > \mathbb{E}X_1$ верно*

$$\lim_{n\to\infty} \frac{\ln \mathbb{P}\left\{\frac{S_n}{n} \geq r\right\}}{n} = -I(r),$$

где функция

$$I(r) = \sup_{\gamma}\{\gamma r - \ln \mathbb{E}e^{\gamma X_1}\}$$

называется функцией уклонений.

Верхняя оценка в теореме Крамера–Чернова получается простым применением неравенства Чебышева

$$\begin{aligned}\mathbb{P}\left\{\frac{S_n}{n} \geq r\right\} &= \mathbb{P}\{S_n \geq rn\} \leq \frac{\mathbb{E}e^{\gamma S_n}}{e^{\gamma rn}} \\ &\leq \frac{\left(\mathbb{E}e^{\gamma X_1}\right)^n}{e^{\gamma rn}} = \exp\left\{-n\left(\gamma r - \ln \mathbb{E}e^{\gamma X_1}\right)\right\},\end{aligned}$$

с последующей оптимизацией по параметру γ. Однако нижняя оценка требует более тонких и асимптотических вычислений.

При попытке обобщить теорему Крамера–Чернова на случайные вектора результат несколько усложняется.

Теорема 8.2 *(Принцип больших уклонений в $\mathbb{R}^n$). Пусть $X_1, X_2, \ldots \in \mathbb{R}^n$ – последовательность независимых одинаково распределённых случайных векторов, удовлетворяющих условию*

$$\mathbb{E}\exp\{(\gamma, X_j)\} < \infty, \qquad \gamma \in \mathbb{R}^n,\ |\gamma| < \gamma_0$$

при некотором $\gamma_0 > 0$. Обозначим $S_n = \sum_{j=1}^n X_j$. Тогда для любого открытого множества $G \subset \mathbb{R}^n$ верно

$$\liminf_{n\to\infty} \frac{\ln \mathbb{P}\left\{\frac{S_n}{n} \in G\right\}}{n} \geq -J(G),$$

а для любого замкнутого множества $F \subset \mathbb{R}^n$ верно

$$\limsup_{n\to\infty} \frac{\ln \mathbb{P}\left\{\frac{S_n}{n} \in F\right\}}{n} \leq -J(F),$$

где $J(A) = \inf_A I(\cdot)$, *а функция* $I : \mathbb{R}^n \mapsto [0, \infty]$ *определяется формулой*

$$I(h) = \sup_{\gamma \in \mathbb{R}^n} \{(\gamma, h) - \ln \mathbb{E} e^{(\gamma, X_1)}\}$$

и называется функцией уклонений.

Подставим сюда векторы со стандартным гауссовским распределением. Очевидно, $\ln \mathbb{E} e^{(\gamma, X_1)} = \frac{(\gamma,\gamma)}{2}$, откуда легко следует $I(h) = \frac{|h|^2}{2}$.

Заметим также, что в гауссовском случае $\frac{S_n}{n} = \frac{X_1}{\sqrt{n}}$ по распределению, поэтому результат можно сформулировать прямо в терминах X_1. Далее, после того, как мы избавились от сумм, можно заменить натуральный параметр n на вещественный $R = \sqrt{n}$. Получится, что для любого открытого $G \subset \mathbb{R}^n$ верно

$$\liminf_{R \to \infty} \frac{\ln \mathbb{P}\{X \in RG\}}{R^2} \geq -J(G),$$

а для любого замкнутого $F \subset \mathbb{R}^n$ верно

$$\limsup_{R \to \infty} \frac{\ln \mathbb{P}\{X \in RF\}}{R^2} \leq -J(F)$$

при $J(A) = \inf_{h \in A} \frac{|h|^2}{2}$.

Наконец, заметим, что для любого множества $A \subset \mathbb{R}^n$ верно

$$\inf_{Cl(A)} I(\cdot) \leq \inf_A I(\cdot) \leq \inf_{Int(A)} I(\cdot).$$

Если же

$$\inf_{Cl(A)} I(\cdot) = \inf_A I(\cdot) = \inf_{Int(A)} I(\cdot),$$

то множество A называется *регулярным*. В условиях теоремы 8.2 для регулярного множества мы имеем

$$\lim_{n \to \infty} \frac{\ln \mathbb{P}\left\{\frac{S_n}{n} \in A\right\}}{n} = -J(A),$$

а для гауссовского случая

$$\lim_{R \to \infty} \frac{\ln \mathbb{P}\{X \in RA\}}{R^2} = -J(A).$$

После знакомства с теоремой 8.2 выглядит естественным такое определение [180]: говорят, что семейство распределений (P_n) в топологическом пространстве $\mathcal{X}$ удовлетворяет *принципу больших уклонений* со скоростью v_n и функцией уклонений $I : \mathcal{X} \mapsto [0, \infty]$, если для любого открытого множества $G \subset \mathcal{X}$ верно

$$\liminf_{n\to\infty} \frac{\ln P_n(G)}{v_n} \geq -J(G),$$

а для любого замкнутого множества $F \subset \mathcal{X}$ верно

$$\limsup_{n\to\infty} \frac{\ln P_n(F)}{v_n} \leq -J(F),$$

где $J(A) = \inf_A I(\cdot)$.

Общую теорию больших уклонений можно изучать по книге [77].

8.2 Гауссовский принцип больших уклонений

Пусть теперь X – центрированный гауссовский вектор с распределением P в линейном пространстве $\mathcal{X}$. Определим функцию уклонений формулой, аналогичной конечномерному случаю:

$$I(h) = \begin{cases} \frac{|h|^2_{H_P}}{2}, & h \in H_P, \\ +\infty, & h \notin H_P. \end{cases}$$

Тогда верна следующая теорема:

Теорема 8.3 *(Гауссовский принцип больших уклонений). Для любого открытого множества $G \subset \mathcal{X}$ верно*

$$\liminf_{R\to\infty} \frac{\ln \mathbb{P}\{X \in RG\}}{R^2} \geq -J(G), \tag{8.1}$$

а для любого замкнутого $F \subset \mathcal{X}$ верно

$$\limsup_{R\to\infty} \frac{\ln \mathbb{P}\{X \in RF\}}{R^2} \leq -J(F), \tag{8.2}$$

где $J(A) = \inf_A I(\cdot)$.

Множество $A \subset \mathcal{X}$ называется *регулярным*, если

$$\inf_{Cl(A)} I(\cdot) = \inf_{A} I(\cdot) = \inf_{Int(A)} I(\cdot).$$

Из теоремы 8.3 следует, что для таких множеств верно

$$\lim_{R\to\infty} \frac{\ln \mathbb{P}\{X \in RA\}}{R^2} = -J(A).$$

Доказательство теоремы 8.3.

Нижняя оценка. Пусть G – открытое множество и вектор h принадлежит $G \cap H_P$. Тогда найдётся такая симметричная окрестность нуля V, что $h + V \subset G$. Используя предложение 5.1, получим

$$P(RG) \geq P(Rh + RV) \geq P(RV) \exp\{-R^2 |h|^2_{H_P}/2\},$$

причём

$$\lim_{R\to\infty} P(RV) = P(\mathcal{X}) = 1. \tag{8.3}$$

Отсюда

$$\liminf_{R\to\infty} \frac{\ln \mathbb{P}\{X \in RG\}}{R^2} \geq \frac{-|h|^2_{H_P}}{2} .$$

Максимизируя правую часть по h, получим (8.1).

Верхняя оценка. Пусть F – замкнутое множество. Фиксируем $\rho < \inf_{h\in F} |h|_{H_P}$. Тогда шар $B := \{h \in H_P : |h|_{H_P} \leq \rho\}$ и множество F не пересекаются. Воспользуемся тем, что шар B компактен. Поскольку F замкнуто, то для каждой точки $h \in B$ найдётся такая выпуклая окрестность нуля V_h, что $h + V_h \cap F = \emptyset$. Выделим из покрытия $\{h + V_h/2\}$ множества B конечное подпокрытие $\{h_i + V_{h_i}/2\}$ и положим

$$V = \bigcap_i (V_{h_i})/2 .$$

Тогда

$$(B + V) \cap F = \emptyset. \tag{8.4}$$

Действительно, для любого $h \in B$ найдётся такое i, что верно $h \in h_i + V_{h_i}/2$. Поэтому для любого $v \in V$ имеем

$$h + v \in h_i + V_{h_i}/2 + V \subset h_i + V_{h_i}/2 + V_{h_i}/2 = h_i + V_{h_i} \ .$$

Следовательно, $h+v \notin F$ и (8.4) доказано. Из него, применяя изопериметрическое неравенство, находим

$$P(RF) \leq P(\mathcal{X} \backslash (RB + RV)) \leq 1 - \Phi\left(\Phi^{-1}(P(RV)) + R\rho\right).$$

Из (8.3) находим, что при достаточно больших R верно

$$\Phi^{-1}(P(RV)) \geq \Phi^{-1}(1/2) = 0,$$

откуда

$$P(RF) \leq 1 - \Phi\left(R\rho\right).$$

Следовательно,

$$\limsup_{R \to \infty} \frac{\ln \mathbb{P}\left\{X \in RF\right\}}{R^2} \leq \limsup_{R \to \infty} \frac{\ln\left(1 - \Phi\left(R\rho\right)\right)}{R^2} = \frac{-\rho^2}{2} \ .$$

Остаётся перейти к пределу $\rho \nearrow \inf_{h \in F} |h|_{H_P}$, что равносильно $\frac{\rho^2}{2} \nearrow J(F)$, и (8.2) доказано. $\square$

Замечание 8.1 Гауссовская версия принципа больших уклонений была независимо установлена Вентцелем и Фрейдлиным [9, 38] для гильбертова пространства. Ещё раньше частный случай винеровской меры был рассмотрен Шильдером [163].

8.3 Следствия принципа больших уклонений

Пример 8.1 *(Большие уклонения максимума гауссовского процесса).* Пусть $X(t), t \in T$, – центрированный гауссовский процесс с непрерывными траекториями на компакте T. Тогда

$$\lim_{R \to \infty} \frac{\ln \mathbb{P}\left(\max_T X \geq R\right)}{R^2} = \frac{-1}{2\sigma^2}, \tag{8.5}$$

$$\lim_{R\to\infty} \frac{\ln \mathbb{P}\left(\max_T |X| \geq R\right)}{R^2} = \frac{-1}{2\sigma^2}, \tag{8.6}$$

где

$$\sigma^2 = \max_T \mathbb{E}\, X(t)^2.$$

Применим принцип больших уклонений к вектору $X = (X(t))$ в пространстве $\mathbb{C}(T)$ и множеству $A = \{x : \max_T x \geq 1\}$. Заметим, что A – регулярное множество. Действительно, A – замкнутое множество, поэтому $J(A) = J(Cl(A))$. С другой стороны, $Int(A) = \{x : \max_T x > 1\}$. Поэтому для любого $h \in A$ и любого $\varepsilon > 0$ имеем $(1+\varepsilon)h \in Int(A)$ и

$$I(h) = \lim_{\varepsilon\to 0}(1+\varepsilon)^2 I(h) = \lim_{\varepsilon\to 0} I((1+\varepsilon)h) \geq J(Int(A)).$$

Следовательно,

$$J(A) = \inf_{h\in A} I(h) \geq J(Int(A)).$$

Обратное неравенство $J(A) \leq J(Int(A))$ очевидно, и регулярность A доказана. Значит, предел в (8.5) существует и равен $-J(A)$. Вычислим его.

Для любого $t \in T$ положим $\sigma_t^2 = \mathbb{E}X(t)^2$, $\delta_t x = x(t)$, и

$$h_t = \sigma_t^{-2} K\delta_t = \sigma_t^{-2} II^*\delta_t.$$

Тогда

$$|h_t|^2_{H_P} = \sigma_t^{-4} ||I^*\delta_t||^2_{\mathcal{X}_P^*} = \sigma_t^{-4}\sigma_t^2 = \sigma_t^{-2}$$

и

$$h_t(t) = (\delta_t, h) = (\delta_t, \sigma_t^{-2} II^*\delta_t) = \sigma_t^{-2}(I^*\delta_t, I^*\delta_t) = 1.$$

Поэтому $h_t \in A$ и

$$J(A) = \inf_{h\in A} I(h) \leq \inf_{t\in T} I(h_t) = \inf_{t\in T} \frac{|h_t|^2_{H_P}}{2} = \inf_{t\in T} \frac{\sigma_t^{-2}}{2} = \frac{1}{2\sigma^2}\ .$$

С другой стороны, пусть $h = Iz \in A$. Тогда при некотором $t \in T$

$$1 \leq h(t) = (\delta_t, h) = (\delta_t, Iz) = (I^*\delta_t, z) \leq \sigma_t ||z||_{\mathcal{X}_P^*}.$$

Поэтому
$$|h|_{H_P} = ||z||_{\mathcal{X}_P^*} \geq \inf_{t\in T} \sigma_t^{-1} = \sigma^{-1}.$$
Значит, $J(A) \geq \frac{1}{2\sigma^2}$. Получаем $J(A) = \frac{1}{2\sigma^2}$, и (8.5) доказано.

Упражнение 8.1 Доказите (8.6) аналогичным рассуждением.

Упражнение 8.2 Для функции $f \in \mathbb{C}[0,1]$ определим её модуль непрерывности

$$\omega(f,u) := \sup_{|s-t|\leq u} |f(s) - f(t)|.$$

Пусть W – винеровский процесс. Докажите, что

$$\lim_{R\to\infty} \frac{\ln \mathbb{P}\,(\omega(W,u) \geq R)}{R^2} = \frac{-1}{2u}\;, \qquad 0 < u \leq 1. \tag{8.7}$$

Пример 8.2 *(Распределение нормы гауссовского вектора).* Пусть X – центрированный гауссовский вектор в сепарабельном банаховом пространстве $(\mathcal{X}, ||\cdot||)$. Пусть P – распределение X, а $K : \mathcal{X}^* \mapsto \mathcal{X}$ – ковариационный оператор X. Тогда

$$\lim_{R\to\infty} \frac{\ln \mathbb{P}\,(||X|| \geq R)}{R^2} = \frac{-1}{2||K||}\,. \tag{8.8}$$

Как известно, норма вектора в банаховом пространстве записывается в виде

$$||x|| = \sup_{f\in S^*} (f,x), \qquad x \in \mathcal{X}, \tag{8.9}$$

где S^* – единичная сфера в сопряжённом пространстве. Поэтому мы имеем
$$||X|| = \sup_{f\in S^*} (f,X),$$
и можно применять рассуждения из предыдущего примера с функционалами $f \in S^*$ вместо моментов времени $t \in T$. Получим
$$\lim_{R\to\infty} \frac{\ln \mathbb{P}\,(||X|| \geq R)}{R^2} = \frac{-1}{2\sigma^2},$$

где
$$\sigma^2 = \sup_{f\in S^*} \mathbb{E}(f,X)^2 = \sup_{f\in S^*} (f,Kf).$$
Очевидно, что
$$\sigma^2 \le \sup_{f\in S^*} ||f||\,||Kf|| \le ||K||.$$

С другой стороны,
$$\begin{aligned} ||K|| &= \sup_{f\in S^*} ||Kf|| = \sup_{f,g\in S^*} (g,Kf) = \sup_{f,g\in S^*} \mathbb{E}(f,X)(g,X) \\ &\le \left\{ \sup_{f\in S^*} \mathbb{E}(f,X)^2 \sup_{g\in S^*} \mathbb{E}(g,X)^2 \right\}^{1/2} = \sigma^2. \end{aligned}$$

Таким образом, $\sigma^2 = ||K||$, и (8.8) доказано.

Упражнение 8.3 (Эквивалентность моментов нормы гауссовского вектора.) Используя принцип концентрации (применительно к супремум-функционалу, см. представление (8.9) и пример 6.1), а также неравенство (7.4), докажите, что для любого $p \ge 1$ найдётся такая положительная константа c_p, что для любого центрированного гауссовского вектора в нормированном пространстве верно
$$\mathbb{E}||X|| \le (\mathbb{E}||X||^p)^{1/p} \le c_p\, \mathbb{E}||X||. \qquad (8.10)$$

Глава 9

Функциональный закон повторного логарифма

9.1 Классический ЗПЛ

Напомним две классические формы закона повторного логарифма (ЗПЛ):

- *для сумм независимых случайных величин.*

  Пусть (X_j) – последовательность центрированных независимых случайных величин с единичной дисперсией. Положим
  $$S_n = \sum_{j=1}^{n} X_j \ .$$
  Тогда ЗПЛ Хартмана–Винтнера [104][1] утверждает, что почти наверное
  $$\begin{aligned} \limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} &= 1, \\ \liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} &= -1. \end{aligned} \tag{9.1}$$

[1] P. Hartmann, A. Wintner.

- *для винеровского процесса.* ЗПЛ Хинчина [7, 39] утверждает, что почти наверное

$$\begin{aligned}\limsup_{T\to\infty}\frac{W(T)}{\sqrt{2T\ln\ln T}} &= 1,\\ \liminf_{T\to\infty}\frac{W(T)}{\sqrt{2T\ln\ln T}} &= -1.\end{aligned}$$

Сходство формулировок объясняется тем, что при $T=n$ можно представить $W(T)$ в виде суммы независимых одинаково распределённых приращений $W(j)-W(j-1)$.

Подчеркнем, что ЗПЛ является формой изучения *больших уклонений* фигурирующих в нём величин. Например, типичное значение S_n согласно центральной предельной теореме имеет порядок $\sqrt{n}$, в то время как ЗПЛ имеет дело с величинами большего порядка $\sqrt{n\ln\ln n}$. То же самое можно сказать и о ЗПЛ для винеровского процесса.

Закон повторного логарифма был открыт Хинчиным [112] для числа успехов в схеме Бернулли. Важный результат Колмогорова даёт ЗПЛ для сумм не одинаково распределённых ограниченных независимых случайных величин. Последующая история ЗПЛ, отмеченная работами Колмогорова, Леви, Эрдёша, Феллера, Петрова, Леду и др., и литература по этой теме очень обширны, см. [91, 56, 126, 24] и дальнейшие ссылки в этих работах.

9.2 Функциональный ЗПЛ

Теперь мы переходим к изучению функционального ЗПЛ (далее обозначаемого ФЗПЛ) для винеровского процесса. В отличие от классического ЗПЛ функциональный закон принимает во внимание не только значение $W(T)$, но и всю траекторию $W(t), 0\le t\le T$. Для каждого $T>3$ определим случайные элементы пространства $\mathbb{C}[0,1]$ формулами

$$X_T(s)=\frac{W(sT)}{\sqrt{T}}\ ,\qquad 0\le s\le 1,$$

$$Y_T(s) = \frac{W(sT)}{\sqrt{2T\ln\ln T}} = \frac{X_T(s)}{\sqrt{2\ln\ln T}}\,, \qquad 0 \le s \le 1. \tag{9.2}$$

Заметим, что обе эти функции содержат в сжатом виде всю траекторию процесса $W(t), 0 \le t \le T$. Кроме того,

$$Y_T(1) = \frac{W(T)}{\sqrt{2T\ln\ln T}}\,,$$

так что изучать предельное поведение $Y(\cdot)$ – естественная задача в свете ЗПЛ Хинчина.

Отметим на будущее, что в силу самоподобия винеровского процесса процесс X_T при любом T сам является винеровским.

Дальше пойдет речь о *сходимости к множеству*, а не к предельной точке. Поэтому имеет смысл напомнить соответствующее определение.

Пусть $(Y_T)_{T\ge T_0}$ – семейство элементов некоторого метрического пространства $(\mathcal{X},\rho)$. Говорят, что компактное множество $K\subset\mathcal{X}$ является *предельным* для Y_T при $T\to\infty$ и пишут $Y_T \hookrightarrow K$, если выполнены два условия

1. $\lim_{T\to\infty}\inf_{h\in K}\rho(Y_T,h)=0$.

2. для любого $h\in K$ верно $\liminf_{T\to\infty}\rho(Y_T,h)=0$.

Первое условие говорит, что при больших T элемент Y_T близок к *некоторому* элементу компакта K. Второе условие говорит, что любая окрестность *любого* элемента K посещается семейством (Y_T) при сколь угодно больших T.

Соль приведённого определения в том, что из $Y_T \hookrightarrow K$ для любого непрерывного функционала $g:\mathcal{X}\mapsto\mathbb{R}$, следует

$$\limsup_{T\to\infty} g(Y_T) = \sup_{h\in K} g(h), \qquad \liminf_{T\to\infty} g(Y_T) = \inf_{h\in K} g(h). \tag{9.3}$$

Таким образом, анализ предельного поведения функционала сводится к решению экстремальной задачи на предельном компакте. Для доказательства (9.3) заметим, что из второго свойства сходимости и непрерывности g следует, что для любого $h\in K$

$$L := \limsup_{T\to\infty} g(Y_T) \ge g(h),$$

и, оптимизируя по h, находим

$$L \geq \sup_{h \in K} g(h).$$

Обратно, выделим такую последовательность $T_n \to \infty$, что

$$\lim_{n\to\infty} g(Y_{T_n}) = L.$$

По первому свойству сходимости, найдётся такая последовательность $h_n \in K$, что

$$\lim_{n\to\infty} \rho(Y_{T_n}, h_n) = 0.$$

Пользуясь компактностью K, можно выделить сходящуюся подпоследовательность $h_{n_j} \to h \in K$. Тогда $Y_{T_{n_j}} \to h$ и

$$L = \lim_{j\to\infty} g(Y_{T_{n_j}}) = g(h),$$

откуда следует $L \leq \sup_{h\in K} g(h)$. Первое утверждение в (9.3) доказано. Второе получается из первого заменой g на $-g$. □

В качестве предельного компакта в ФЗПЛ будет выступать единичный шар ядра (эллипсоид рассеяния) винеровской меры, т.е. множество

$$\begin{aligned} K &= \{h : |h|_{H_P} \leq 1\} \\ &= \left\{h \in AC[0,1] : h(0) = 0,\ \int_0^1 h'(s)^2 \mathrm{d}s \leq 1\right\}. \end{aligned} \tag{9.4}$$

В контексте ФЗПЛ K часто называют *шаром Штрассена.*

Теорема 9.1 *(Функциональный закон повторного логарифма (закон Штрассена*[2]*) для винеровского процесса). Пусть семейство* $(Y_T)_{T\geq 3}$ *задано формулой* (9.2), *а шар* K *– формулой* (9.4). *Тогда с вероятностью единица верно*

$$Y_T \hookrightarrow K.$$

[2]V. Strassen [168].

Доказательство теоремы 9.1 достаточно техническое, но очень поучительное. В его основе лежит идея экспоненциального блокирования: мы будем главным образом следить за поведением Y_T вдоль экспоненциально растущих последовательностей $T_n = \gamma^n$, причём параметр $\gamma > 1$ будет варьироваться в зависимости от поставленной цели.

Введём подходящие обозначения. Пусть $\gamma > 1$. Представим Y_T в виде

$$Y_T = \widehat{Y}_T + Y_T^0,$$

где $\widehat{Y}_T(s) := Y_T\left(\min(s, \gamma^{-1})\right)$ – это функция $Y_T(\cdot)$, остановленная в момент времени $\gamma^{-1} < 1$. Соответственно определяется $Y_T^0 := Y_T - \widehat{Y}_T$. Нетрудно видеть, что функция

$$Y_T^0(s) = \begin{cases} 0, & 0 \le s \le \gamma^{-1}, \\ (2T \ln\ln T)^{-1/2}\left(W(sT) - W(\gamma^{-1}T)\right), & \gamma^{-1} \le s \le 1, \end{cases}$$

полностью определяется приращениями процесса W на интервале $[\gamma^{-1}T, T]$. Поэтому, когда мы рассматриваем последовательность моментов $T_n = \gamma^n$, случайные функции $Y_{T_n}^0$ будут независимы. Обозначим $\rho(x, y) = ||x - y||$ равномерное расстояние между функциями.

Техническая часть доказательства состоит в обосновании четырех утверждений о сходимости рядов.

1. Для любых $\gamma > 1$ и $\varepsilon > 0$ верно

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{h \in K} \rho(Y_{T_n}, h) > \varepsilon\right) < \infty. \tag{9.5}$$

2. Для любого $\gamma > 1$ можно найти такое $\varepsilon_1 = \varepsilon_1(\gamma) > 0$, что

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{T_n \le T \le T_{n+1}} \rho(Y_T, Y_{T_n}) > \varepsilon_1\right) < \infty \tag{9.6}$$

и $\lim_{\gamma \searrow 0} \varepsilon_1(\gamma) = 0$.

3. Для любых $h \in K$, $\gamma > 1$ и $\varepsilon > 0$ верно

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\rho(Y_{T_n}, h) < \varepsilon\right) = \infty. \tag{9.7}$$

4. Для любого $\gamma > 1$ можно найти такое $\varepsilon_2 = \varepsilon_2(\gamma) > 0$, что

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left\|\widehat{Y}_{T_n}\right\| > \varepsilon_2\right) < \infty \tag{9.8}$$

и $\lim_{\gamma\to\infty} \varepsilon_2(\gamma) = 0$.

С помощью двух первых утверждений мы покажем, что при больших T функции Y_T приближаются к множеству K, а с помощью двух последних, – что каждый элемент K является предельной точкой.

По лемме Бореля–Кантелли из (9.5) следует[3], что при любом $\gamma > 1$

$$\limsup_{n\to\infty} \inf_{h\in K} \rho(Y_{T_n}, h) = 0,$$

а из (9.6) следует, что

$$\limsup_{n\to\infty} \sup_{T_n \le T \le T_{n+1}} \rho(Y_T, Y_{T_n}) \le \varepsilon_1.$$

По неравенству треугольника

$$\inf_{h\in K} \rho(Y_T, h) \le \inf_{h\in K} \rho(Y_{T_n}, h) + \rho(Y_T, Y_{T_n}),$$

откуда

$$\limsup_{T\to\infty} \inf_{h\in K} \rho(Y_T, h) \le \varepsilon_1.$$

Наконец, устремляя $\gamma \searrow 1$, получим

$$\limsup_{T\to\infty} \inf_{h\in K} \rho(Y_T, h) \le \lim_{\gamma\searrow 1} \varepsilon_1(\gamma) = 0.$$

---

[3]Здесь и далее утверждения по ходу доказательства справедливы „с вероятностью единица“.

Зафиксируем $h \in K$ и $\gamma > 1$. Вычитая из расходящегося ряда сходящийся, мы получаем расходящийся ряд. Поэтому из (9.7) и (9.8) следует

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\rho(Y_{T_n}, h) \le \varepsilon, \left\|\widehat{Y}_{T_n}\right\| \le \varepsilon_2\right)$$
$$\ge \sum_{n=1}^{\infty} \left(\mathbb{P}\left(\rho(Y_{T_n}, h) \le \varepsilon\right) - \mathbb{P}\left(\left\|\widehat{Y}_{T_n}\right\| > \varepsilon_2\right)\right) = \infty.$$

Положим здесь $\varepsilon = \varepsilon_2$ и, предполагая выполнение неравенств $\rho(Y_{T_n}, h) < \varepsilon_2, \left\|\widehat{Y}_{T_n}\right\| \le \varepsilon_2$, выведем из неравенства треугольника

$$\rho(Y^0_{T_n}, h) \le \rho(Y_{T_n}, h) + \rho(Y^0_{T_n}, Y_{T_n}) = \rho(Y_{T_n}, h) + \left\|\widehat{Y}_{T_n}\right\| \le 2\varepsilon_2.$$

Таким образом,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\rho(Y^0_{T_n}, h) \le 2\varepsilon_2\right) = \infty.$$

Поскольку элементы $Y^0_{T_n}$ независимы, из леммы Бореля–Кантелли следует

$$\liminf_{n\to\infty} \rho(Y^0 T_n, h) \le 2\varepsilon_2.$$

С другой стороны, из (9.8) по той же лемме Бореля–Кантелли мы получаем

$$\limsup_{n\to\infty} \left\|\widehat{Y}_{T_n}\right\| \le \varepsilon_2.$$

По неравенству треугольника

$$\rho(Y_{T_n}, h) \le \rho(Y^0_{T_n}, h) + \rho(Y^0_{T_n}, Y_{T_n}),$$

поэтому

$$\liminf_{n\to\infty} \rho(Y_{T_n}, h) \le \liminf_{n\to\infty} \rho(Y^0_{T_n}, h) + \limsup_{n\to\infty} \left\|\widehat{Y}_{T_n}\right\| \le 3\varepsilon_2.$$

Наконец,

$$\liminf_{T\to\infty} \rho(Y_T, h) \le \liminf_{n\to\infty} \rho(Y_{T_n}, h) \le 3\varepsilon_2$$

и, устремляя $\gamma \to \infty$, получим

$$\liminf_{T\to\infty} \rho(Y_T, h) \le 3 \lim_{\gamma\to\infty} \varepsilon_2(\gamma) = 0.$$

Остаётся проверить, что справедливы утверждения о рядах (9.5) – (9.8). Докажем (9.5). Положим $L_T = 2\ln\ln T$. Обозначим

$$\mathbb{U} := \{x \in \mathbb{C}[0,1] : ||x|| \le 1\}$$

– единичный шар в $\mathbb{C}[0,1]$. Учитывая, что

$$\lim_{r\to\infty} \mathbb{P}(W \in r\mathbb{U}) = 1 \tag{9.9}$$

и пользуясь изопериметрическим неравенством (6.8), найдём при достаточно больших n

$$\begin{aligned}
& \mathbb{P}\left(\inf_{h\in K} \rho(Y_{T_n}, h) > \varepsilon\right) = \mathbb{P}\left(Y_{T_n} \notin K + \varepsilon\mathbb{U}\right) \\
= \;& \mathbb{P}\left(\frac{W}{\sqrt{L_{T_n}}} \notin K + \varepsilon\mathbb{U}\right) \\
= \;& \mathbb{P}\left(W \notin \sqrt{L_{T_n}}K + \varepsilon\sqrt{L_{T_n}}\mathbb{U}\right) \\
\le \;& \widehat{\Phi}\left(\Phi^{-1}\left(\mathbb{P}(W \in \varepsilon\sqrt{L_{T_n}}\mathbb{U})\right) + \sqrt{L_{T_n}}\right) \\
\le \;& \widehat{\Phi}\left(1 + \sqrt{L_{T_n}}\right) \\
\le \;& \exp\left\{\frac{-\left(1+\sqrt{L_{T_n}}\right)^2}{2}\right\} \\
\le \;& \exp\left\{\frac{-L_{T_n}}{2} - \sqrt{L_{T_n}}\right\}.
\end{aligned}$$

Используя грубую оценку $\exp(-x) \le cx^{-4}$ применительно к

$x = \sqrt{L_{T_n}}$, получим

$$\begin{aligned}
\mathbb{P}\left(\inf_{h\in K}\rho(Y_{T_n},h) > \varepsilon\right) &\leq \exp\left\{\frac{-L_{T_n}}{2} - \sqrt{L_{T_n}}\right\} \\
&\leq c\,L_{T_n}^{-2}\ \exp\left\{\frac{-L_{T_n}}{2}\right\} \\
&= \frac{c}{(2(\ln\ln\gamma + \ln n))^2}\ \frac{1}{\ln T_n} \\
&= \frac{c}{(2(\ln\ln\gamma + \ln n))^2}\ \frac{1}{n\,\ln\gamma},
\end{aligned}$$

и это приводит к сходящемуся ряду. Соотношение (9.5) доказано.

Докажем (9.7). Пусть $\varepsilon > 0$ и $h \in K$. Используя (9.9) и неравенство Борелля для сдвигов (5.3)), получим при больших n

$$\begin{aligned}
\mathbb{P}\left(\rho(Y_{T_n},h) \leq \varepsilon\right) &= \mathbb{P}\left(Y_{T_n} \in h + \varepsilon\mathbb{U}\right) \\
&= \mathbb{P}\left(\frac{W}{\sqrt{L_{T_n}}} \in h + \varepsilon\mathbb{U}\right) \\
&= \mathbb{P}\left(W \in h\sqrt{L_{T_n}} + \varepsilon\sqrt{L_{T_n}}\mathbb{U}\right) \\
&\geq \exp\left\{-\frac{L_{T_n}^2|h|_{H_P}^2}{2}\right\}\mathbb{P}\left(W \in \varepsilon\sqrt{L_{T_n}}\mathbb{U}\right) \\
&\geq (\ln(T_n))^{-|h|_{H_P}^2}\ \frac{1}{2} = \frac{1}{2}\,(n\,\ln\gamma)^{-|h|_{H_P}^2}\ .
\end{aligned}$$

Поскольку $|h|_{H_P} \leq 1$, то ряд из последних выражений расходится и оценка (9.7) доказана.

Докажем (9.8). Пусть $\varepsilon > 0$ и $\gamma > 1$. Используя самоподобие винеровского процесса, найдём

$$\begin{aligned}
\mathbb{P}\left(\left\|\widehat{Y}_{T_n}\right\| > \varepsilon\right) &= \mathbb{P}\left(\frac{\left\|W(\min(\cdot,\gamma^{-1}))\right\|}{\sqrt{L_{T_n}}} > \varepsilon\right) \\
&= \mathbb{P}\left(\max_{0\le s\le\gamma^{-1}} |W(s)| > \varepsilon\sqrt{L_{T_n}}\right) \\
&= \mathbb{P}\left(\gamma^{-1/2}\max_{0\le s\le 1} |W(s)| > \varepsilon\sqrt{L_{T_n}}\right) \\
&= \mathbb{P}\left(\max_{0\le s\le 1} |W(s)| > \varepsilon\sqrt{\gamma L_{T_n}}\right).
\end{aligned}$$

Применяя принцип больших уклонений для максимумов (8.6), найдём при больших n

$$\begin{aligned}
&\mathbb{P}\left(\max_{0\le s\le 1} |W(s)| > \varepsilon\sqrt{\gamma L_{T_n}}\right) \le \exp\left\{\frac{-\varepsilon^2\gamma L_{T_n}}{3}\right\} \\
= \quad &(\ln(T_n))^{-2\varepsilon^2\gamma/3} = (n\,\ln\gamma)^{-2\varepsilon^2\gamma/3}.
\end{aligned}$$

Ряд из таких величин будет сходиться, если мы выберем $\varepsilon := \varepsilon_2(\gamma) = 2\gamma^{-1/2}$. Очевидно, что при таком выборе верно и условие (9.8), и требуемое соотношение $\lim_{\gamma\to\infty}\varepsilon_2(\gamma) = 0$.

Докажем (9.6). Пусть $\varepsilon > 0$, $\gamma > 1$ и $T \in [T_n, T_{n+1}]$. Из определения Y_T следует

$$Y_T(s) = \frac{W(sT)}{\sqrt{TL_T}} = \frac{\sqrt{T_{n+1}}}{\sqrt{TL_T}}\, X_{T_{n+1}}\left(\frac{sT}{T_{n+1}}\right).$$

Поэтому

$$\begin{aligned}
&\rho(Y_T, Y_{T_{n+1}}) \\
= \quad &\sup_{0\le s\le 1}\left|\frac{\sqrt{T_{n+1}}}{\sqrt{TL_T}}\, X_{T_{n+1}}\left(\frac{sT}{T_{n+1}}\right) - \frac{1}{\sqrt{L_{T_{n+1}}}}\, X_{T_{n+1}}(s)\right| \\
\le \quad &\sup_{0\le s\le 1}\Delta_1(s) + \sup_{0\le s\le 1}\Delta_2(s),
\end{aligned}$$

где

$$\begin{aligned}
\Delta_1(s) &= \frac{\sqrt{T_{n+1}}}{\sqrt{TL_T}} \left| X_{T_{n+1}} \left(\frac{sT}{T_{n+1}} \right) - X_{T_{n+1}}(s) \right| ; \\
\Delta_2(s) &= \left| \frac{\sqrt{T_{n+1}}}{\sqrt{TL_T}} - \frac{1}{\sqrt{L_{T_{n+1}}}} \right| \cdot \left| X_{T_{n+1}}(s) \right| .
\end{aligned}$$

Для первого слагаемого воспользуемся оценкой больших уклонений модуля непрерывности винеровского процесса (8.7). Получим, что при больших n верно

$$\begin{aligned}
&\mathbb{P}\left\{ \sup_{0\le s\le 1} \Delta_1(s) \ge \varepsilon \right\} \\
&= \mathbb{P}\left\{ \frac{\sqrt{T_{n+1}}}{\sqrt{TL_T}} \sup_{0\le s\le 1} \left| W\left(\frac{sT}{T_{n+1}} \right) - W(s) \right| \ge \varepsilon \right\} \\
&\le \mathbb{P}\left\{ \omega\left(W, 1-\gamma^{-1}\right) \ge \frac{\varepsilon\sqrt{TL_T}}{\sqrt{T_{n+1}}} \right\} \\
&= \mathbb{P}\left\{ \omega\left(W, 1-\gamma^{-1}\right) \ge \frac{\varepsilon\sqrt{L_{T_n}}}{\sqrt{\gamma}} \right\} \\
&\le \exp\left\{ -\frac{\varepsilon^2 L_{T_n}}{3\gamma} \left(1-\gamma^{-1}\right)^{-1} \right\} \\
&= (n \ln \gamma)^{-\frac{2\varepsilon^2}{3\gamma}\left(1-\gamma^{-1}\right)^{-1}} .
\end{aligned}$$

Соответствующий ряд сходится, если $\varepsilon \ge 2(\gamma-1)^{1/2}$. Аналогично, но даже проще, рассматривается и второе слагаемое. Для него получаем оценку вида

$$\begin{aligned}
\mathbb{P}\left\{ \sup_{0\le s\le 1} \Delta_2(s) \ge \varepsilon \right\} &\le \mathbb{P}\left\{ ||W|| \ge \frac{c\varepsilon\sqrt{L_{T_n}}}{1-\gamma^{-1}} \right\} \\
&\le \exp\left\{ -\frac{c^2\varepsilon^2}{3} L_{T_n} \left(1-\gamma^{-1}\right)^{-2} \right\} \\
&= (n \ln \gamma)^{-\frac{2c^2\varepsilon^2}{3}\left(1-\frac{1}{\gamma}\right)^{-2}} .
\end{aligned}$$

Соответствующий ряд сходится, если $\varepsilon \ge 2c^{-1}\left(1-\gamma^{-1}\right)$. Остаётся выбрать

$$\varepsilon_1(\gamma) = 2\max\left\{2(\gamma-1)^{1/2}; 2c^{-1}\left(1-\gamma^{-1}\right)\right\},$$

и (9.6) доказано при одновременном выполнении требования $\lim_{\gamma\searrow 1}\varepsilon_1(\gamma) = 0$. □

Упражнение 9.1 Пусть W – винеровский процесс. Докажите, что верхний предел

$$\limsup_{T\to\infty} \frac{\int_0^T W(s)^2 \mathrm{d}s}{T^2 \ln\ln T}$$

не случаен (с вероятностью единица) и найдите его.

9.3 Некоторые уточнения и обобщения ФЗПЛ

Скорость сходимости

Интересно узнать, насколько быстра сходимость в ФЗПЛ. Поскольку ФЗПЛ представляет собой пару утверждений, то и скорость сходимости нужно вычислять для каждого из них.

Что касается сближения Y_T с множеством K, то К. Грилл и М. Талагран[4] независимо доказали, что для некоторых c_1, c_2 почти наверное при достаточно больших n

$$\frac{c_1}{(\ln\ln T)^{2/3}} \le \inf_{h\in K}\rho(Y_T, h) \le \frac{c_2}{(\ln\ln T)^{2/3}}\ .$$

О сближении с конкретным элементом $h \in K$ можно сказать следующее. Если h находится *внутри* K, т.е. $|h|_{H_P} < 1$, то скорость сближения имеет порядок $(\ln\ln T)^{-1}$. Точнее из результатов А. де Акоста и Э.Чаки[5] следует, что

$$\liminf_{T\to\infty}\left(||Y_T - h||\ln\ln T\right) = \frac{\pi}{4\sqrt{1-|h|^2_{H_P}}}\ .$$

[4]K. Grill, M. Talagrand [102, 170].
[5]A. de Acosta, E. Csáki [74, 73].

Для *пограничных* элементов h с единичной нормой порядок скорости сближения зависит от h. Скорость всегда медленнее $(\ln\ln T)^{-1}$, но не хуже, чем $(\ln\ln T)^{-2/3}$.

Упражнение 9.2 Выведите из сказанного, что почти наверное

$$\liminf_{T\to\infty}\left(\frac{\ln\ln T}{T}\right)^{1/2}\sup_{0\le t\le T}|W(t)| = \frac{\pi}{\sqrt{8}}\ .$$

Это – закон повторного логарифма *в форме Чжуна*[6].

Другие нормы

Пространство $\mathbb{C}[0,1]$ можно заменить на другие нормированные пространства, в которые с вероятностью единица попадают траектории винеровского процесса, например, пространство $L_p[0,1]$ или пространство α-гёльдеровских функций с нормой

$$||x||_\alpha = \sup_{\substack{s,t\in[0,1]\\ s\neq t}}\frac{|x(s)-x(t)|}{|s-t|^\alpha}, \qquad 0<\alpha<\frac{1}{2}\ .$$

Чем сильнее норма, тем больше соответствующее расстояние и сильнее результат ФЗПЛ: можно применять следствия к большему классу функционалов, см. [49, 75, 76].

Выбор нормы влияет на скорость сходимости в ФЗПЛ.

Многомерный процесс

Вместо обычного винеровского процесса W можно рассмотреть его многомерный аналог, то есть векторный процесс

$$W(t) = \left(W^{(1)}(t),\dots,W^{(n)}(t)\right)\in\mathbb{R}^n,$$

где $(W^{(j)}(t))_{j=1}^n$ – независимые одномерные винеровские процессы. Сохраняя определение Y_T, данное в (9.2), получим сходимость к множеству K^n функций $h\in\mathbb{C}([0,1],\mathbb{R}^n)$, удовлетворяющих условиям

$$h_j(0)=0, \quad h_j(\cdot)\in AC[0,1]$$

[6] K. L. Chung [68, 107, 108].

и

$$\int_0^1 \left(\sum_{j=1}^{n} h_j'(s)^2 \right) \mathrm{d}s \le 1.$$

Упражнение 9.3 С помощью (9.3) найдите для n-мерного винеровского процесса предел

$$\limsup_{T\to\infty} \frac{||W(t)||}{\sqrt{T \ln\ln T}} \ .$$

Многопараметрический процесс

Рассмотрим броуновский лист $W(t), t \in \mathbb{R}_+^d$, определённый в примере 2.7. С учётом его самоподобия, можно рассмотреть случайные элементы

$$Y_T(s) = \frac{W(Ts)}{\sqrt{2T^d \ln\ln T}} \ , \qquad T > 3, s \in [0,1]^d,$$

в пространстве $\mathbb{C}([0,1]^d)$ и доказать, что предельным множеством для них будет единичный шар ядра распределения W, найденного в примере 4.7, см. [53, 154].

Дробное броуновское движение

Пусть $W^{(\alpha)}(t)$ – дробное броуновское движение с показателем $\alpha \in (0,2)$. Самоподобие $W^{(\alpha)}$ означает, что процесс

$$X_T(s) = \frac{W^{(\alpha)}(sT)}{T^{\alpha/2}}$$

также будет дробным броуновским движением. Соответственно, в ФЗПЛ будет фигурировать

$$Y_T(s) = \frac{W^{(\alpha)}(sT)}{\sqrt{2T^\alpha \ln\ln T}} = \frac{X_T(s)}{\sqrt{2\ln\ln T}} \ , \qquad 0 \le s \le 1.$$

ФЗПЛ в этом случае утверждает, что $Y_T \hookrightarrow K^{(\alpha)}$, где предельное множество $K^{(\alpha)}$ будет единичным шаром ядра распределения процесса $W^{(\alpha)}$, найденного в примере 4.6.

9.4 Сильный принцип инвариантности

Для того, чтобы доказать ФЗПЛ для случайных блужданий, содержащий теорему Хартмана–Винтнера, достаточно показать, что значения блуждания близки к значениям винеровского процесса. Такой тип результатов называется *сильным принципом инвариантности* (СПИ).

Мы приведём здесь несколько наиболее значимых результатов. Общая схема такова. Пусть (X_j) – последовательность независимых одинаково распределённых случайных величин с единичной дисперсией, заданных на некотором вероятностном пространстве. Предположим, что на некотором другом вероятностном пространстве совместно построены последовательность величин $(\widetilde{X}_j)$, одинаково распределённая с (X_j), и винеровский процесс W. Положим

$$\widetilde{S}_k = \sum_{j=1}^{k} \widetilde{X}_j$$

и

$$\Delta_n = \max_{1\le k\le n} |\widetilde{S}_k - W(k)|.$$

СПИ утверждает, что при определённых моментных ограничениях на общее распределение величин (X_j) возможно построение с определённым образом убывающими разностями Δ_n. Приведём несколько результатов, двигаясь от более сильных предположений к более слабым.

- СПИ Комлоша–Майора–Тушнади (известный как КМТ-конструкция[7]). Если для некоторого $c > 0$ конечен экспоненциальный момент: $\mathbb{E}\exp(c|X_j|) < \infty$, то почти наверное
$$\limsup_{n\to\infty} \frac{\Delta_n}{\ln n} < \infty.$$
- СПИ Саханенко [26]. Если $\mathbb{E}|X_j|^p < \infty$ для некоторого $p > 2$, то почти наверное
$$\lim_{n\to\infty} \frac{\Delta_n}{n^{1/p}} = 0.$$

[7] J. Komlós, P. Major, G. Tusnády [115, 116].

- СПИ Штрассена [168]. Из одного (уже сделанного ранее) предположения $\mathbb{E}|X_j|^2 = 1$ следует, что почти наверное
$$\lim_{n\to\infty} \frac{\Delta_n}{(n \ln\ln n)^{1/2}} = 0. \tag{9.10}$$
Эту оценку невозможно улучшить без дальнейших моментных предположений. Однако следующий результат показывает, как чуть лучшую скорость аппроксимации можно получить, если слегка изменить способ аппроксимации.

- СПИ Майора [146]. В тех же условиях имеем почти наверное
$$\lim_{n\to\infty} \frac{\widetilde{\Delta}_n}{n^{1/2}} = 0,$$
где
$$\widetilde{\Delta}_n = \max_{1\le k\le n} |\widetilde{S}_k - W(\widetilde{k})|,$$
а
$$\widetilde{k} = \sum_{j=1}^{k} \mathbb{D}\left(X_j \cdot \mathbf{1}_{|X_j|\le j}\right)$$
удовлетворяет соотношениям $\widetilde{k} \le k$ и
$$\lim_{k\to\infty} \frac{\widetilde{k}}{k} = 1.$$

9.5 ФЗПЛ для случайного блуждания

Пусть (X_j) и S_k – такие же, как в теореме Хартмана–Винтнера (9.1). Нормированные траектории случайного блуждания $Z_n(s), 0 \le s \le 1$, определим формулами
$$Z_n\left(\frac{k}{n}\right) = \begin{cases} 0, & k = 0, \\ \frac{S_k}{\sqrt{2n\ln\ln n}}, & 1 \le k \le n, \end{cases} \tag{9.11}$$
и линейной интерполяцией между точками вида $\frac{k}{n}$.

Теорема 9.2 *(Функциональный закон повторного логарифма (закон Штрассена) для случайных блужданий* [168]*.)*

Пусть семейство $(Z_n)_{n\geq 3}$ *задано формулой* (9.11)*, а множество* K *– формулой* (9.4)*. Тогда с вероятностью единица верно*

$$Z_n \hookrightarrow K.$$

Доказательство теоремы 9.2.

Построим последовательность величин $(\widetilde{X}_j)$, одинаково распределённую с (X_j), и винеровский процесс W, о которых говорится в СПИ Штрассена. Пусть $\widetilde{Z}_n$ – последовательность, построенная аналогично Z_n, но исходя из $(\widetilde{X}_j)$. Из одинаковой распределённости (X_j) и $(\widetilde{X}_j)$ следует, что предельные множества у (Z_n) и $(\widetilde{Z}_n)$ одинаковы. Поэтому нам достаточно показать, что $\widetilde{Z}_n \hookrightarrow K$. Пусть Y_n – нормированные траектории процесса W, построенные по (9.2) при $T = n$, а $\widetilde{Y}_n$ – их линейные интерполяции по точкам $\frac{k}{n}, 0 \leq k \leq n$. Тогда

$$\begin{aligned}
\rho(\widetilde{Z}_n, \widetilde{Y}_n) &= \max_{0\leq s\leq 1} |\widetilde{Z}_n(s) - \widetilde{Y}_n(s)| = \max_{0\leq k\leq n} |\widetilde{Z}_n(\frac{k}{n}) - \widetilde{Y}_n(\frac{k}{n})| \\
&= \max_{0\leq k\leq n} |\widetilde{Z}_n(\frac{k}{n}) - Y_n(\frac{k}{n})| = \max_{0\leq k\leq n} \frac{|\widetilde{S}_k - W(k)|}{\sqrt{2n \ln\ln n}} \\
&= \frac{\Delta_n}{\sqrt{2n \ln\ln n}} .
\end{aligned}$$

Из (9.10) следует, что $\lim_{n\to\infty} \rho(\widetilde{Z}_n, \widetilde{Y}_n) = 0$. С другой стороны, приглядевшись к процессу $Y_n - \widetilde{Y}_n$ (ошибке интерполяции), можно заметить, что на каждом отрезке вида $[\frac{k}{n}, \frac{k+1}{n}]$ он представляет собой нормированную копию броуновского моста, к тому же эти копии независимы между собой. Отсюда легко заключить, что $\lim_{n\to\infty} \rho(Y_n, \widetilde{Y}_n) = 0$. По неравенству треугольника имеем

$$\lim_{n\to\infty} \rho(Y_n, \widetilde{Z}_n) = 0.$$

Следовательно, из $Y_n \hookrightarrow K$ следует $\widetilde{Z}_n \hookrightarrow K$. □

Замечание 9.1 ЗПЛ Хартмана–Винтнера (9.1) получается из теоремы 9.2 применением (9.3) к функционалу $g(x) = x(1)$.

Глава 10

Метрическая энтропия и свойства траекторий

10.1 Основные определения

Пусть (T, ρ) – метрическое пространство. Определим энтропийное число[1] $N(\varepsilon)$ как минимальное количество множеств диаметра не более ε, достаточное, чтобы покрыть T. Тогда $N(\cdot)$ – невозрастающая функция и $N(0+) < \infty$ в том и только в том случае, когда T – конечное множество и $N(\varepsilon) < \infty$ при всех $\varepsilon > 0$ в том и только в том случае, когда T относительно компактно. Величина $H(\varepsilon) = \ln N(\varepsilon)$ называется *метрической энтропией* пространства T.

Определим ёмкостное число[2] $M(\varepsilon)$ как максимально возможное количество точек, которые можно так разместить в T, чтобы все попарные расстояния между ними были больше ε. Тогда $M(\cdot)$ – невозрастающая функция и $M(0+) < \infty$ в том и только в том случае, когда T – конечное множество и $M(\varepsilon) < \infty$ при всех $\varepsilon > 0$ в том и только в том случае, когда T относительно компактно. Величина $\mathcal{C}(\varepsilon) = \ln M(\varepsilon)$ называется *метрической ёмкостью* пространства T.

---

[1] Covering number.

[2] Packing number.

Следующее предложение показывает, что N и M в общем-то описывают одно и то же свойство.

Предложение 10.1 *Для любого пространства* (T,ρ) *и любого* $\varepsilon > 0$ *верны оценки*

$$N(2\varepsilon) \le M(\varepsilon) \le N(\varepsilon) \tag{10.1}$$

и

$$H(2\varepsilon) \le \mathcal{C}(\varepsilon) \le H(\varepsilon). \tag{10.2}$$

Доказательство. Возьмём такую конфигурацию из $M(\varepsilon)$ точек, что все попарные расстояния между ними больше ε. Очевидно, к ней нельзя добавить ни одной точки с сохранением этого свойства. Это означает, что шары радиуса ε с центрами в точках конфигурации образуют покрытие T. Поскольку диаметр каждого шара не превосходит 2ε, мы получаем, что $N(2\varepsilon) \le M(\varepsilon)$. С другой стороны, рассмотрим покрытие T, состоящее из $N(\varepsilon)$ множеств диаметра не более ε и такую систему точек, что все попарные расстояния между ними больше ε. Очевидно, любой элемент покрытия содержит не более одной точки системы. Поэтому число точек в системе не превосходит $N(\varepsilon)$. Таким образом, $M(\varepsilon) \le N(\varepsilon)$. Неравенство (10.1) доказано. Неравенство (10.2) получается логарифмированием неравенства (10.1). □

Пусть теперь $X(t), t \in T$, – гауссовский случайный процесс. Тогда на T можно ввести „*естественное расстояние*“, порождённое процессом X,

$$\rho(s,t)^2 := \mathbb{E}|X(s) - X(t)|^2.$$

Строго говоря, ρ не является расстоянием в том смысле, что для некоторых $s \ne t$ возможно $\rho(s,t) = 0$. Однако это обстоятельство никак не сказывается на данных выше определениях и свойствах введённых понятий. Любитель строгости может свести дело к настоящему метрическому пространству, отождествляя точки, находящиеся на нулевом расстоянии друг от друга.

В дальнейшем $N, M, H, \mathcal{C}$ обозначают соответствующие понятия, относящиеся к пространству (T,ρ). С их помощью мы будем изучать свойства процесса X.

10.2 Верхние оценки

Основной оценке предпошлём важный технический результат.

Лемма 10.1 *Пусть* $(X_j)_{1\le j\le N}$ – *центрированные нормальные случайные величины и* $\max_{j\le N}\mathbb{E}X_j^2\le\sigma^2$. *Тогда*

$$\mathbb{E}\max_{1\le j\le N} X_j \le \sqrt{2\ln N}\;\sigma. \tag{10.3}$$

Доказательство. Оценим преобразование Лапласа: для любого λ верно

$$\begin{aligned}\mathbb{E}\exp\left\{\lambda\max_{1\le j\le N} X_j\right\} &\le \mathbb{E}\left(\sum_{1\le j\le N}\exp\{\lambda X_j\}\right)\\ &\le \sum_{1\le j\le N}\exp\left\{\lambda^2(\mathbb{E}X_j^2)/2\right\}\\ &\le N\;\exp\left\{\lambda^2\sigma^2/2\right\}.\end{aligned}$$

По неравенству Йенсена

$$\lambda\,\mathbb{E}\max_{1\le j\le N} X_j \le \ln\mathbb{E}\exp\left\{\lambda\max_{1\le j\le N} X_j\right\}\le \ln N+\lambda^2\sigma^2/2.$$

Полагая здесь $\lambda=\sqrt{2\ln N}\,\sigma^{-1}$, получим (10.3). □

Интегралом Дадли процесса X назовем интеграл, построенный по его метрической энтропии,

$$\mathcal{D}(u)=\int_0^u \sqrt{H(\varepsilon)}\;\mathrm{d}\varepsilon.$$

Теорема 10.1 *(Верхняя оценка по Дадли*[3]*). Для любого центрированного гауссовского процесса* $X(t), t\in T$, *верна оценка*

$$\mathbb{E}\sup_{t\in T} X(t)\le 4\sqrt{2}\mathcal{D}(\sigma/2),$$

где $\sigma^2=\sup_{t\in T}\mathbb{E}X(t)^2$.

---

[3] R. Dudley, [82].

Доказательство. Следующие рассуждения часто называются *цепным методом* оценивания. Положим

$$\varepsilon_j = \sigma \cdot 2^{-j}, \qquad j = 1, 2, \ldots$$

Для каждого ε_j выберем минимальное покрытие T множествами диаметра, не превосходящего ε_j. В каждом из элементов покрытия выберем по одной точке. Полученное множество, состоящее из $N(\varepsilon_j)$ точек, обозначим S_j. Определим отображение $\pi_j : T \mapsto S_j$ так, чтобы выполнялось

$$\rho(x, \pi_j(x)) \le \varepsilon_j\,, \qquad \forall x \in T.$$

Тогда справедливо

$$\sup_{t\in S_j} X(t) \le \sup_{t\in S_{j-1}} X(t) + \sup_{t\in S_j}(X(t) - X(\pi_{j-1}(t))).$$

Переходя к математическим ожиданиям и применяя к последнему слагаемому лемму 10.1, получим

$$\mathbb{E}\sup_{t\in S_j} X(t) \le \mathbb{E}\sup_{t\in S_{j-1}} X(t) + \sqrt{2H(\varepsilon_j)}\;\varepsilon_{j-1}.$$

Кроме того, по той же лемме для исходного слагаемого верно

$$\mathbb{E}\sup_{t\in S_1} X(t) \le \sqrt{2H(\varepsilon_1)}\;\sigma.$$

Рассуждая по индукции и используя монотонность функции $H(\cdot)$, видим, что для любого $j \ge 1$ верно

$$\begin{aligned}
\mathbb{E}\sup_{t\in S_j} X(t) \quad &\le \quad \sqrt{2H(\varepsilon_1)}\;\sigma + \sum_{k=2}^{j}\sqrt{2H(\varepsilon_k)}\;\varepsilon_{k-1} \\
&= \quad 2\sqrt{2H(\varepsilon_1)}\;\varepsilon_1 + \sum_{k=2}^{j} 2\sqrt{2H(\varepsilon_k)}\;\varepsilon_k \\
&= \quad \sum_{k=1}^{j} 2\sqrt{2H(\varepsilon_k)}\;\varepsilon_k \\
&\le \quad 4\sqrt{2}\sum_{k=1}^{j}\int_{\varepsilon_{k+1}}^{\varepsilon_k}\sqrt{H(\varepsilon)}\;\mathrm{d}\varepsilon \\
&\le \quad 4\sqrt{2}\int_0^{\varepsilon_1}\sqrt{H(\varepsilon)}\;\mathrm{d}\varepsilon = 4\sqrt{2}\mathcal{D}(\sigma/2).
\end{aligned}$$

Остаётся перейти от S_j ко всему пространству T. Это делается в несколько этапов.

1. Если T – конечное множество, то при некотором j будет верно $T = S_j$ и утверждение теоремы уже доказано.

2. Если $T = \{t_i\}_{i=1}^{\infty}$ – бесконечное счётное множество, то можно применить доказанное к конечным множествам $T_n = \{t_i\}_{i=1}^{n}$ и воспользоваться тем, что

$$\mathbb{E} \sup_{t\in T_n} X(t) \nearrow \mathbb{E} \sup_{t\in T} X(t),$$

   а также тем, что все энтропийные характеристики пространства (T_n, ρ) не превосходят соответствующих характеристик (T, ρ).

3. Если T – произвольное множество, то в нём можно выделить счётное всюду плотное множество $T_\#$ (в противном случае интеграл Дадли бесконечен и нам нечего доказывать). Применим доказанное к $T_\#$ и рассмотрим версию процесса, для которой $\sup_T X = \sup_{T_\#} X$. Для неё утверждение теоремы будет выполнено.

□

Замечание 10.1 Согласно (7.4) медиана случайной величины $\sup_T X$ мажорируется её математическим ожиданием. Поэтому оценку Дадли можно использовать и для оценки медианы, например, комбинировать её с принципом концентрации (6.9).

Упражнение 10.1 (Теорема Пизье[4]). Пусть $X(t), t \in T$, – центрированный случайный процесс (необязательно гауссовский), для которого верно $\sigma^2 = \sup_{t\in T} \mathbb{E}X(t)^2 < \infty$. Докажите, что

$$\mathbb{E} \sup_{t\in T} X(t) \le 4 \int_0^{\sigma} \sqrt{N(\varepsilon)}\, \mathrm{d}\varepsilon.$$

[4] G. Pisier, [157].

Часто приходится оценивать максимум *абсолютных величин* значений процесса. Однако это не требует новых оценок благодаря следующему результату.

Предложение 10.2 *Для любого центрированного гауссовского процесса $X(t), t \in T$, верна оценка*

$$\mathbb{E} \sup_{t \in T} |X(t)| \le 2 \left(\mathbb{E} \sup_{t \in T} X(t) + \inf_{t \in T} (\mathbb{D}X(t))^{1/2} \right). \tag{10.4}$$

Доказательство предложения 10.2. будем использовать стандартные обозначения

$$x_+ = \max\{x, 0\}, \quad x_- = \max\{-x, 0\}.$$

Тогда верно $|x| = x_+ + x_-$, $(-x)_- = x_+$. Сразу заметим, что

$$\begin{aligned}
\mathbb{E} \sup_{t \in T} (X(t))_- &= \mathbb{E} \sup_{t \in T} (-X(t))_- = \mathbb{E} \sup_{t \in T} (X(t))_+ \ , \\
\mathbb{E} \sup_{t \in T} |X(t)| &= \mathbb{E} \sup_{t \in T} ((X(t))_+ + (X(t))_-) \\
&\le \mathbb{E} \sup_{t \in T} (X(t))_+ + \mathbb{E} \sup_{t \in T} (X(t))_- \\
&= 2\,\mathbb{E} \sup_{t \in T} (X(t))_+ \ .
\end{aligned} \tag{10.5}$$

С другой стороны, всегда верно

$$\sup_{t \in T} (X(t))_+ \le \sup_{t \in T} X(t) + \inf_{t \in T} |X(t)|. \tag{10.6}$$

Действительно, если $\sup_{t \in T} X(t) \ge 0$, то

$$\sup_{t \in T} (X(t))_+ = \sup_{t \in T} X(t)$$

и неравенство (10.6) очевидно, а если $\sup_{t \in T} X(t) < 0$, то и левая и правая часть (10.6) равны нулю. Из (10.6) находим, что

$$\begin{aligned}
\mathbb{E} \sup_{t \in T} (X(t))_+ &\le \mathbb{E} \sup_{t \in T} X(t) + \inf_{t \in T} \mathbb{E}|X(t)| \\
&\le \mathbb{E} \sup_{t \in T} X(t) + \inf_{t \in T} (\mathbb{D}X(t))^{1/2}.
\end{aligned}$$

Комбинируя это неравенство с (10.5), получим (10.4). $\square$

Замечание 10.2 Второе слагаемое в оценке (10.4) необходимо: рассмотрим одноточечное множество $T = \{t\}$. Тогда $\mathbb{E}\sup_T X = \mathbb{E}X(t) = 0$, и первого слагаемого недостаточно для корректной оценки.

Предложение 10.3 *(Достаточное энтропийное условие непрерывности). Пусть $X(t)$, $t \in T$, – такой центрированный гауссовский процесс, что $\mathcal{D}(u) < \infty$ при $u > 0$. Тогда для любого $t \in T$ с вероятностью единица процесс X непрерывен в t, т.е.*

$$\lim_{\rho(s,t)\to 0} X(s) = X(t). \qquad (10.7)$$

Замечание 10.3 Можно доказать и нечто большее: если верно $\mathcal{D}(u) < \infty$, то с вероятностью единица (10.7) справедливо для всех $t \in T$ одновременно.

Доказательство предложения 10.3. Зафиксируем точку $t \in T$ и положим $T_n = \{s \in T : \rho(s,t) \leq \frac{1}{n}\}$.

Пусть $H_n(\cdot)$, $\mathcal{D}_n(\cdot)$ – метрическая энтропия и интеграл Дадли случайного процесса $\{X(s)-X(t), s \in T_n\}$. Обозначим $S_n = \sup_{s\in T_n}(X(s) - X(t))$. Очевидно, что $H_n(\varepsilon) \leq H(\varepsilon)$. Поэтому из теоремы 10.1 следует оценка

$$\mathbb{E}S_n \leq 4\sqrt{2}\mathcal{D}_n\left(n^{-1}\right) \leq 4\sqrt{2}\mathcal{D}\left(n^{-1}\right).$$

Последовательность случайных величин S_n неотрицательна и монотонно убывает к некоторому пределу S. По лемме Фату $\mathbb{E}S \leq \lim_n \mathbb{E}S_n = 0$. Отсюда следует, что почти наверное $S = 0$. □

Отдельно стоит упомянуть о регулярности *стационарных* процессов. Здесь есть два красивых результата. Первый из них показывает, что стационарный гауссовский процесс либо непрерывен, либо катастрофически разрывен.

Теорема 10.2 *(Альтернатива Беляева [3]). Для любого стационарного гауссовского процесса $X(t), t \in \mathbb{R}$, верно одно из двух утверждений*

а) Процесс X имеет непрерывные траектории.

б) Для любого невырожденного интервала $T \subset \mathbb{R}$ с вероятностью единица верно

$$\sup_{t\in T} X(t) = \infty.$$

Второй результат даёт необходимое и достаточное условие непрерывности стационарного процесса.

Теорема 10.3 *(Критерий Ферника[5]). Пусть $X(t), t \in \mathbb{R}$, – стационарный гауссовский процесс. Тогда X имеет непрерывные траектории (относительно естественной метрики ρ) в том и только в том случае, когда его интеграл Дадли конечен.*

Для нестационарных процессов интеграл Дадли не даёт критерия непрерывности: существуют непрерывные процессы с бесконечным интегралом Дадли. Более того, известно, что необходимые и достаточные условия непрерывности или ограниченности гауссовского процесса не могут быть сформулированы в энтропийных терминах. Для формулировки таких условий нужны более тонкие средства, например, мажорирующие меры [19, 92, 93, 126, 182] или усовершенствованный цепной метод[6] [176, 177]. Дальнейшие исторические комментарии и библиография по энтропийным методам приведены в [19]. Интересные приложения энтропийного метода к негауссовским процессам, а также задачам анализа и эргодической теории можно найти в [181, 182].

10.3 Нижние оценки

В основе нижних энтропийных оценок лежит следующий результат, который мы приведём здесь без доказательства.

Теорема 10.4 *(Теорема сравнения Судакова–Ферника* [28, 29, 30, 92]*). Пусть $X(t), Y(t)$ – два центрированных гауссовских*

[5] X. Fernique, [92].
[6] Generic chaining.

процесса, заданных на одном и том же параметрическом множестве T. Пусть

$$\mathbb{E}(X(t)-X(s))^2 \geq \mathbb{E}(Y(t)-Y(s))^2 \qquad \forall s,t\in T.$$

Тогда

$$\mathbb{E}\sup_{t\in T} X(t) \geq \mathbb{E}\sup_{t\in T} Y(t).$$

Иными словами, процесс с бо́льшими разбросами значений имеет бо́льшее ожидание максимума.

Теорему сравнения удобно комбинировать со следующей элементарной оценкой, противоположной лемме 10.1. Важное отличие состоит в предположении независимости величин.

Лемма 10.2 *Пусть $(X_j)_{1\leq j\leq N}$ – независимые центрированные нормальные случайные величины и $\min_{j\leq N}\mathbb{E}X_j^2 \geq \sigma^2$. Тогда при $c_* = 0.64$ верно*

$$\mathbb{E}\max_{1\leq j\leq N} X_j \geq c_*\sqrt{\ln N}\,\sigma. \tag{10.8}$$

Доказательство. Пусть $c<\sqrt{2}$. Покажем, что

$$\lim_{N\to\infty}\mathbb{P}\left\{\max_{1\leq j\leq N} X_j \leq c\sqrt{\ln N}\,\sigma\right\} = 0.$$

Действительно,

$$\begin{aligned}
\mathbb{P}\left\{\max_{1\leq j\leq N} X_j \leq c\sqrt{\ln N}\,\sigma\right\} &= \prod_{j=1}^{N}\mathbb{P}\left\{X_j \leq c\sqrt{\ln N}\sigma\right\}\\
&\leq \left(1-\widehat{\Phi}\left(c\sqrt{\ln N}\right)\right)^N\\
&\leq \exp\left\{-N\widehat{\Phi}\left(c\sqrt{\ln N}\right)\right\},
\end{aligned}$$

где $\widehat{\Phi}$ – хвост стандартного нормального закона. Воспользуемся неравенством

$$\widehat{\Phi}(u) \geq \frac{1}{\sqrt{2\pi}}\left(\frac{1}{u}-\frac{1}{u^3}\right)e^{-u^2/2}.$$

При больших $u = c\sqrt{\ln N}$ получаем

$$\widehat{\Phi}\left(c\sqrt{\ln N}\right) \geq \frac{N^{-c^2/2}}{3c\sqrt{\ln N}} \ ,$$

откуда при $N \to \infty$

$$\mathbb{P}\left\{\max_{1\leq j\leq N} X_j \leq c\sqrt{\ln N}\,\sigma\right\} \leq \exp\left\{-\frac{N^{1-c^2/2}}{3c\sqrt{\ln N}}\right\} \to 0.$$

Значит, при больших N медиана величины $\max_{1\leq j\leq N} X_j$ превосходит $c\sqrt{\ln N}\,\sigma$. При больших N неравенство (7.4) даёт

$$\mathbb{E}\max_{1\leq j\leq N} X_j \geq c\sqrt{\ln N}\,\sigma.$$

Несколько первых значений N могут только повлиять на константу в (10.8). В конечном счёте, оказывается, что значение $c_* = 0.64$ достаточно для справедливости (10.8). $\square$

Теорема 10.5 *(Нижняя оценка по Судакову). Для любого центрированного гауссовского процесса $X(t), t \in T$ и для любого $\varepsilon > 0$ верна оценка*

$$\mathbb{E}\sup_{t\in T} X(t) \geq \frac{c_*}{\sqrt{2}}\ \sqrt{\mathcal{C}(\varepsilon)}\ \varepsilon. \tag{10.9}$$

Доказательство. Зафиксируем $\varepsilon > 0$. Выберем точки $\{t_1, \ldots, t_m\}$ в T так, что $m = M(\varepsilon)$ и

$$\rho^2(t_i, t_j) = \mathbb{E}|X(t_i) - X(t_j)|^2 > \varepsilon^2, \qquad i \neq j.$$

Пусть $Y_1, \ldots, Y_m$ – независимые величины с одинаковым нормальным распределением $N(0, \frac{\varepsilon^2}{2})$. Применяя теорему сравнения 10.4 к величинам $(X(t_j))_{j\leq m}$ и $(Y_j)_{j\leq m}$, находим

$$\mathbb{E}\sup_{t\in T} X(t) \geq \mathbb{E}\sup_{j\leq m} X(t_j) \geq \mathbb{E}\sup_{j\leq m} Y_j\ .$$

Но по лемме 10.2 верно

$$\mathbb{E}\max_{1\leq j\leq m} Y_j \geq c_*\sqrt{\ln m}\ \frac{\varepsilon}{\sqrt{2}} = \frac{c_*}{\sqrt{2}}\ \sqrt{\mathcal{C}(\varepsilon)}\ \varepsilon.$$

Комбинируя две оценки, получим (10.9). $\square$

Упражнение 10.2 Пусть $X(t), t \in T$, – центрированный гауссовский процесс и связанное с ним пространство (T,ρ) компактно. Пусть $c > 0, \varepsilon_0 > 0$ и $\delta \in (0,2)$. Докажите следующие утверждения:

а) Если $N(\varepsilon) \le \exp\left(-c\,\varepsilon^{-(2-\delta)}\right)$ при $\varepsilon \in (0,\varepsilon_0)$, то процесс X с вероятностью единица непрерывен на (T,ρ).

б) Если $N(\varepsilon) \ge \exp\left(-c\,\varepsilon^{-(2+\delta)}\right)$ при $\varepsilon \in (0,\varepsilon_0)$, то процесс X с вероятностью единица не ограничен (и следовательно, разрывен) на (T,ρ).

Упражнение 10.3 Пусть $X(t), t \in \mathbb{R}$, – стационарный гауссовский процесс со спектральной плотностью

$$f(u) = |u|^{-1}(\ln|u|)^{-1-\beta}, \qquad \beta > 0, |u| > 2.$$

При каких значениях параметра β процесс X будет непрерывным?

Упражнение 10.4 Пусть $X(t), t \in [0,1]$, – дробное броуновское движение, определённое в примере 2.5. Найдите такие числа C_1 и C_2, зависящие от параметра α, что

$$C_1 \le \mathbb{E} \sup_{t\in[0,1]} X(t) \le C_2\,.$$

10.4 GB-множества и GC-множества

Содержимое данной главы может быть выражено и на чисто геометрическом языке. Пусть $\mathcal{H}$ – гильбертово пространство; обозначим $(\cdot,\cdot)$ и $||\cdot||$ соответствующее скалярное произведение и норму. Центрированная гауссовская случайная функция $X(h), h \in \mathcal{H}$, называется *изонормальной*, если

$$cov(X(h), X(h')) = (h,h'), \qquad h,h' \in \mathcal{H}.$$

Заметим, что естественное расстояние, порождённое X, совпадает с гильбертовой метрикой,

$$\rho(h,h') = ||h-h'||, \qquad h,h' \in \mathcal{H}. \tag{10.10}$$

Пусть $T \subset \mathcal{H}$. Назовём T *GB-множеством* (соответственно, *GC-множеством*), если сужение изонормальной функции $X(h), h \in T$, имеет ограниченную (соответственно, непрерывную) версию.

Все GC-множества являются GB-множествами, а все GB-множества вполне ограничены (докажите, что оба противоположных утверждения не верны!). Поэтому классы замкнутых GB- и GC-множеств являются собственными подклассами класса компактов. Легко видеть, что GB и GC классы инвариантны относительно сдвигов и унитарных вращений $\mathcal{H}$.

В силу (10.10) можно забыть о процессе X и формулировать результаты этой главы, рассматривая T как метрическое подпространство $\mathcal{H}$ и используя обозначения $H_T(\cdot)$ и $\mathcal{C}_T(\cdot)$ для соответствующих метрической энтропии и метрической ёмкости. В частности, по теореме 10.5 необходимое условие

$$\sup_{\varepsilon>0} \mathcal{C}_T(\varepsilon)\,\varepsilon^2 < \infty$$

выполнено для любого GB-множества T, а по теореме 10.1 достаточное условие

$$\int_0^\infty \sqrt{H_T(\varepsilon)}\,\mathrm{d}\varepsilon < \infty$$

обеспечивает принадлежность T классу GC-множеств.

Глава 11

Малые уклонения

11.1 Определения и первые примеры

В этой главе предполагается, что задан центрированный гауссовский $\mathcal{X}$-значный случайный вектор как измеримое отображение $X : (\Omega, \mathbb{P}) \mapsto \mathcal{X}$ со значениями в сепарабельном банаховом пространстве $(\mathcal{X}, ||\cdot||)$.

Как обычно, обозначаем через P распределение вектора X. Соответственно, пространство H_P – это ядро меры P, а $D := \{h \in H_P : |h|_{H_P} \leq 1\}$ – эллипсоид рассеяния. Определим ещё единичный шар пространства $\mathcal{X}$,

$$U := \{x \in \mathcal{X} : ||x|| \leq 1\}.$$

Задача о вероятностях малых уклонений (*вероятностях малых шаров*) состоит в исследовании поведения

$$\mathbb{P}\left(||X|| \leq \varepsilon\right) = P\left(\varepsilon U\right), \qquad \varepsilon \to 0.$$

Типичным является ответ вида

$$\mathbb{P}\left(||X|| \leq \varepsilon\right) \sim c_1 \varepsilon^a \exp\{-c_2 \varepsilon^{-b}\}, \qquad \varepsilon \to 0. \tag{11.1}$$

Назовем b *порядком малых уклонений*[1], а c_2 – *постоянной малых уклонений*. Найти ответ типа (11.1) удаётся лишь в

[1] Small deviation rate.

небольшом числе случаев, главным образом для марковских процессов типа винеровского процесса или броуновского моста. Например, для винеровского процесса [68, 156]

$$\mathbb{P}\left(\sup_{0\le t\le 1}|W(t)|\le\varepsilon\right)\sim\frac{4}{\pi}\ \exp\{-\frac{\pi^2}{8}\,\varepsilon^{-2}\}\,,\qquad \varepsilon\to 0,\quad (11.2)$$

и [67]

$$\mathbb{P}\left(\int_0^1|W(t)|^2\mathrm{d}t\le\varepsilon^2\right)\sim\frac{4\varepsilon}{\sqrt{\pi}}\ \exp\{-\frac{1}{8}\,\varepsilon^{-2}\}\,,\quad \varepsilon\to 0,\quad (11.3)$$

а для броуновского моста [114]

$$\mathbb{P}\left(\sup_{0\le t\le 1}|W^0(t)|\le\varepsilon\right)\sim\frac{\sqrt{2\pi}}{\varepsilon}\ \exp\{-\frac{\pi^2}{8}\,\varepsilon^{-2}\}\,,\qquad \varepsilon\to 0,$$

и [44]

$$\mathbb{P}\left(\int_0^1|W^0(t)|^2\mathrm{d}t\le\varepsilon^2\right)\sim\frac{\sqrt{8}}{\sqrt{\pi}}\ \exp\{-\frac{1}{8}\,\varepsilon^{-2}\}\,,\qquad \varepsilon\to 0.$$

Можно заметить, что маленькое отличие (процесс ранга 1) между винеровским процессом и броуновским мостом влияет на степенной член асимптотики, но не на экспоненциальный.

11.2 Марковский случай

Проиллюстрируем подход к малым уклонениям для марковских процессов на примере асимптотики (11.2). Воспользуемся самоподобием винеровского процесса, чтобы записать

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le 1}|W(t)|\le\varepsilon\right) &= \mathbb{P}\left(\sup_{0\le t\le 1}|\varepsilon W(\varepsilon^{-2}t)|\le\varepsilon\right)\\ &= \mathbb{P}\left(\sup_{0\le t\le \varepsilon^{-2}}|W(t)|\le 1\right)=f(0,\varepsilon^{-2}),\end{aligned}$$

где

$$f(x,T):=\mathbb{P}\left(\sup_{0\le t\le T}|x+W(t)|\le 1\right).$$

Если $p(\cdot)$ – плотность распределения стандартного нормального закона $N(0,1)$, то марковское свойство винеровского процесса и разложение Тейлора дают, без претензии на строгость, для малых δ

$$\begin{aligned} f(x,T) &= o(\delta) + \int_{\frac{x-1}{\sqrt{\delta}}}^{\frac{x+1}{\sqrt{\delta}}} f(x+\sqrt{\delta}y, T-\delta)p(y)\mathrm{d}y \\ \approx & \int_{-\infty}^{\infty} \left(f + f'_x\sqrt{\delta}y + \frac{1}{2}f''_{xx}\delta y^2 - f'_T\delta\right)p(y)\mathrm{d}y \\ = & \ f(x,T) + \frac{1}{2}f''_{xx}(x,T)\delta - f'_T(x,T)\delta. \end{aligned}$$

После сокращения получим классическое уравнение теплопроводности

$$\frac{1}{2}f''_{xx}(x,T) = f'_T(x,T)$$

с граничными условиями

$$f(1,T) = f(-1,T) = 0, f(x,0) \equiv 1.$$

Решения с разделяющимися переменными этого уравнения имеют форму $f_k(x,T) = g_k(x)e^{-\lambda_k T}$, с граничными условиями $g_k(\pm 1) = 0$. Подставляя в уравнение теплопроводности, придем к обыкновенному дифференциальному уравнению

$$\begin{cases} g''_k(x) + 2\lambda_k g_k(x) = 0, & |x| < 1, \\ g_k(\pm 1) = 0, \end{cases}$$

откуда $g_k(x) = \cos(\pi(k+1)/2)x)$ и $2\lambda_k = \pi^2(k+1/2)^2$ при $k = 0, 1, 2, \ldots$. Поэтому искомое решение можно записать в виде ряда

$$f(x,T) = \sum_{k=0}^{\infty} c_k \cos(\pi(k+1/2)x)\exp\{-\pi^2(k+1/2)^2 T/2\},$$

а коэффициенты c_k подобрать для выполнения последнего краевого условия $f(x,0) \equiv 1$, т.е. (с учётом ортогональности косинусов)

$$c_k = \frac{\int_{-1}^{1}\cos(\pi(k+1/2)x)\mathrm{d}x}{\int_{-1}^{1}\cos^2(\pi(k+1/2)x)\mathrm{d}x} = \frac{2(-1)^k}{\pi(k+1/2)}\ .$$

На асимптотику решения при $T \to \infty$ влияет только первый член ряда,

$$f(x,T) \sim c_0 \cos(\pi x/2) \exp\{-\pi^2 T/8\}, \qquad T \to \infty.$$

Подставляя сюда $x = 0, T = \varepsilon^{-2}$, придем к (11.2).

Анализ малых уклонений аддитивных норм типа (11.3) производится аналогично, но вместо уравнения для вероятностей $f(x,T)$ выводится уравнение теплопроводности для преобразования Лапласа [6], например,

$$\widetilde{f}(x,T) := \mathbb{E} \exp\left\{ -\int_0^T |x + W(t)|^2 \mathrm{d}t \right\}.$$

С некоторой потерей точности формулы (11.2), (11.3) могут быть распространены на L_q-нормы с весом, например,

$$\ln \mathbb{P}\left(\sup_{0 \le t \le 1} |\rho(t) W(t)| \le \varepsilon \right) \sim -\frac{\pi^2}{8} \, ||\rho||^2_{L_2[0,1]} \, \varepsilon^{-2}, \qquad \varepsilon \to 0,$$

если функция ρ^2 интегрируема по Риману, и

$$\ln \mathbb{P}\left(\int_0^1 |\rho(t) W(t)|^q \mathrm{d}t \le \varepsilon^q \right) \sim -c(q) \, ||\rho||^2_{L_m[0,1]} \, \varepsilon^{-2}, \; \varepsilon \to 0,$$

если $1 \le q < \infty, m = \frac{2q}{q+2}$ и функция ρ^m интегрируема по Риману [130, 131, 137].

11.3 Прямой энтропийный метод

В частном случае, когда $\mathcal{X}$ – это пространство непрерывных функций, а $||\cdot||$ – супремум-норма, можно получить содержательные оценки малых уклонений в терминах энтропийных характеристик из главы 10.

Итак, пусть $X(t), t \in T$, – центрированный гауссовский случайный процесс, а $N(\varepsilon)$ – его энтропийные числа. Обозначим, как и ранее, $\sigma^2 := \sup_T \mathbb{E} X(t)^2$ и

$$\rho^2(s,t) := \mathbb{E}(X(s) - X(t))^2.$$

Теорема 11.1 [48] *Предположим, что*

$$N(\varepsilon) \leq \Psi(\varepsilon), \qquad \forall \varepsilon > 0, \tag{11.4}$$

причём мажоранта Ψ непрерывна, не возрастает и удовлетворяет условию регулярности

$$\Psi(\varepsilon/2) \leq C\,\Psi(\varepsilon), \qquad \forall\, \varepsilon > 0. \tag{11.5}$$

Тогда для всех $0 < \varepsilon < \sigma/2$ верно

$$\ln \mathbb{P}\left\{\sup_{s,t\in T} |X(s) - X(t)| \leq C_0\varepsilon\right\} \geq -C_1\,\widetilde{\Psi}(\varepsilon), \tag{11.6}$$

где C_0 – абсолютная постоянная, $C_1 = C_1(C)$ и

$$\widetilde{\Psi}(\varepsilon) = \int_\varepsilon^\sigma \frac{\Psi(u)}{u}\,\mathrm{d}u, \qquad 0 < \varepsilon < \sigma/2.$$

Замечание 11.1 Условие регулярности запрещает использовать в качестве мажоранты экспоненциальные функции, но допускает степенные и логарифмические.

Замечание 11.2 Формула (11.6) оценивает малые уклонения для несколько необычной нормы $\sup_{s,t\in T}|X(s) - X(t)|$, называемой *размахом* случайного процесса X. Однако из неё несложно вывести оценку и для классической супремум-нормы.

Воспользовавшись слабым корреляционным неравенством (7.8), найдём, что для любого $t \in T$ и любого $\delta \in (0,1)$ верно

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{s\in T}|X(s)| \leq \varepsilon\right\}\\
\geq\;& \mathbb{P}\left\{\sup_{s\in T}|X(s) - X(t)| \leq (1-\delta)\varepsilon; |X(t)| \leq \delta\varepsilon\right\}\\
\geq\;& \mathbb{P}\left\{\sup_{s\in T}|X(s) - X(t)| \leq (1-\delta)^2\varepsilon\right\}\mathbb{P}\left\{|X(t)| \leq \frac{\delta\varepsilon}{K_\delta}\right\}.
\end{aligned}$$

Второй множитель имеет порядок ε и практически никогда не сказывается на логарифмической асимптотике малых уклонений, т.е. с учётом (11.6) и (11.5)

$$\ln \mathbb{P}\left\{\sup_{s\in T}|X(s)| \leq \varepsilon\right\} \succeq -\widetilde{\Psi}(\varepsilon/C_0) \approx -\widetilde{\Psi}(\varepsilon). \tag{11.7}$$

Замечание 11.3 Если условие (11.5) заменить на более сильное двустороннее условие

$$C'\Psi(\varepsilon) \le \Psi(\varepsilon/2) \le C\,\Psi(\varepsilon)$$

с некоторым $C' > 1$ (это исключает логарифмические мажоранты и фактически означает, что Ψ – это функция, похожая на регулярно меняющуюся функцию с отрицательным индексом), то несложно проверить, что $\widetilde{\Psi}(\varepsilon) \le \frac{C'\ln 2}{C'-1}\,\Psi(\varepsilon)$, и мы приходим к *оценке Талаграна*[2] для вероятностей малых уклонений

$$\ln \mathbb{P}\left\{\sup_{s,t\in T} |X(s) - X(t)| \le C_0\varepsilon\right\} \ge -C_1'\,\Psi(\varepsilon), \tag{11.8}$$

при $0 < \varepsilon < \sigma/2$.

Например, если $T = [0,1]$ и $X = W^{(\alpha)}$ – дробное броуновское движение с показателем α, то $N(\varepsilon) \sim \varepsilon^{-2/\alpha}$ и из оценки Талаграна следует

$$\ln \mathbb{P}\left\{\sup_{0\le t\le 1} |W^{(\alpha)}(t)| \le \varepsilon\right\} \ge -c(\alpha)\varepsilon^{-2/\alpha}. \tag{11.9}$$

Впрочем, следующий пример показывает, что функцию $\widetilde{\Psi}$ не всегда можно заменить на Ψ как в (11.8).

Пример 11.1 Пусть $T = \mathbb{N}$ и $X(s), s \in \mathbb{N},$ – центрированные независимые нормальные величины с дисперсиями $\sigma_s^2 = e^{-2s}$. Легко проверить, что $N(\varepsilon) \approx |\ln\varepsilon|$, т.е. допустима мажоранта $\Psi(\varepsilon) \asymp |\ln\varepsilon|$. При этом $\widetilde{\Psi}(\varepsilon) \asymp |\ln\varepsilon|^2$ и оценка (11.7) даёт

$$\ln \mathbb{P}\left\{\sup_{s\in\mathbb{N}} |X(s)| \le \varepsilon\right\} \succeq -|\ln\varepsilon|^2.$$

Нетрудно проверить, что эта оценка точна по порядку, т.е.

$$\ln \mathbb{P}\left\{\sup_{s\in\mathbb{N}} |X(s)| \le \varepsilon\right\} \asymp -|\ln\varepsilon|^2. \tag{11.10}$$

[2] M. Talagrand, [124, 172].

Упражнение 11.1 Докажите оценку (11.10).

Доказательство теоремы 11.1 основано на следующей оценке.

Лемма 11.1 *Пусть* $(\varepsilon_k)_{k\geq 0}$ *– убывающая последовательность положительных чисел, причём верно* $\varepsilon_0 \geq \sigma$. *Пусть* $(b_k)_{k\geq 0}$ *– суммируемая последовательность неотрицательных чисел и* $b = \sum_{k\geq 0} b_k$. *Тогда*

$$\mathbb{P}\left\{\sup_{s,t\in T} |X(s) - X(t)| \leq 2b\right\} \geq \prod_{k=0}^{\infty} \mathbb{P}\{\varepsilon_k|\xi| \leq b_k\}^{N(\varepsilon_{k+1})}, \tag{11.11}$$

где ξ *– стандартная нормальная величина.*

Доказательство леммы 11.1. Воспользуемся цепным методом оценивания из доказательства теоремы 10.1. Для каждого ε_k выберем минимальное покрытие T множествами, диаметры которых не превосходят ε_k. В каждом из элементов покрытия выберем по одной точке. Полученное множество, состоящее из $N(\varepsilon_k)$ точек, обозначим S_k. Определим отображение $\pi_k : T \mapsto S_k$ так, чтобы выполнялось

$$\rho(x, \pi_k(x)) \leq \varepsilon_k\,, \qquad \forall x \in T.$$

Заметим, что при $k = 0$ верно $N(\varepsilon_0) = 1$, т.е. множество S_0 состоит из единственного элемента и

$$\sup_{s,t\in S_0} |X(s) - X(t)| = 0.$$

Далее, запишем оценку

$$\begin{aligned} &|X(s) - X(t)| \leq |X(s) - X(\pi_k(s))| \\ &+|X(\pi_k(s)) - X(\pi_k(t))| + |X(\pi_k(t)) - X(t)|, \end{aligned}$$

из которой следует

$$\begin{aligned} &\sup_{s,t\in S_{k+1}} |X(s) - X(t)| \\ \leq\ &\sup_{s,t\in S_k} |X(s) - X(t)| + 2\sup_{t\in S_{k+1}} |X(t) - X(\pi_k(t))|. \end{aligned}$$

По индукции получаем

$$\sup_{s,t\in S_{n+1}} |X(s)-X(t)| \le 2\sum_{k=0}^{n} \sup_{t\in S_{k+1}} |X(t)-X(\pi_k(t))|.$$

Ключевой шаг доказательства – использование неравенства Кхатри–Шидака (7.7), которое даёт

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{s,t\in S_{n+1}} |X(s)-X(t)| \le 2b\right\}\\
\ge\ &\mathbb{P}\left\{\bigcap_{k=0}^{n}\left\{\sup_{t\in S_{k+1}} |X(t)-X(\pi_k(t))| \le b_k\right\}\right\}\\
=\ &\mathbb{P}\left\{\bigcap_{k=0}^{n}\bigcap_{t\in S_{k+1}} \{|X(t)-X(\pi_k(t))| \le b_k\}\right\}\\
\ge\ &\prod_{k=0}^{n}\mathbb{P}\{\varepsilon_k|\xi|\le b_k\}^{N(\varepsilon_{k+1})}.
\end{aligned}$$

Заменяя в правой части конечное произведение на бесконечное, получаем оценку, не зависящую от n. Переход от множеств S_n к множеству T повторяет формальные рассуждения из доказательства теоремы 10.1. $\square$

Доказательство теоремы 11.1. Пусть $\varepsilon\in(0,\sigma/2)$. Чтобы воспользоваться оценкой (11.11), мы должны построить подходящие последовательности (ε_k) и (b_k). Построение проведём отдельно для $\varepsilon_k\le\varepsilon$ и $\varepsilon_k\ge\varepsilon$. Фиксируем некоторое $r\in(\frac{1}{2},1)$, значение которого скажется только на возникающих константах.

Начнем с зоны $\varepsilon_k\le\varepsilon$. Положим $\varepsilon_k=2^{-k}\varepsilon$, $b_k=r^k\varepsilon$, $k=0,1,\dots$. Очевидно,

$$b := \sum_{k=0}^{\infty} b_k = \frac{\varepsilon}{1-r}$$

является подходящей оценкой. Используя соотношение (11.4) и итерируя условие (11.5), получим

$$N(\varepsilon_{k+1}) = N(2^{-k-1}\varepsilon) \le \Psi(2^{-k-1}\varepsilon) \le C^{k+1}\Psi(\varepsilon).$$

Поскольку $2r > 1$, то в силу стандартных оценок хвостов нормального закона

$$\begin{aligned}\mathbb{P}\{\varepsilon_k|\xi| \le b_k\} &= \mathbb{P}\{|\xi| \le (2r)^k\} = 1 - \mathbb{P}\{|\xi| \ge (2r)^k\} \\ &\ge \exp\left\{-2\mathbb{P}\{|\xi| \ge (2r)^k\}\right\} \\ &\ge \exp\left\{-4\exp[-(2r)^{2k}/2]\right\}.\end{aligned}$$

Таким образом,

$$\begin{aligned}& \prod_{k=0}^{\infty} \mathbb{P}\{\varepsilon_k|\xi| \le b_k\}^{N(\varepsilon_{k+1})} \\ \ge\ & \prod_{k=0}^{\infty} \exp\{-4\exp[-(2r)^{2k}/2]C^{k+1}\Psi(\varepsilon)\} \\ =\ & \exp\Big\{-4\sum_{k=0}^{\infty} \exp[-(2r)^{2k}/2]C^{k+1}\Psi(\varepsilon)\Big\} \\ :=\ & \exp\{-c(r)\Psi(\varepsilon)\}.\end{aligned}$$

Учитывая монотонность функции Ψ и условие $\varepsilon < \sigma/2$, мы имеем

$$\widetilde{\Psi}(\varepsilon) \ge \int_{\varepsilon}^{2\varepsilon} \frac{\Psi(u)}{u}\,\mathrm{d}u \ge \ln 2 \cdot \Psi(2\varepsilon) \ge \frac{\ln 2}{C}\ \Psi(\varepsilon), \qquad (11.12)$$

так что мы получаем необходимую оценку

$$\prod_{k=0}^{\infty} \mathbb{P}\{\varepsilon_k|\xi| \le b_k\}^{N(\varepsilon_{k+1})} \ge \exp\left\{-\frac{c(r)C}{\ln 2}\ \widetilde{\Psi}(\varepsilon)\right\}.$$

Теперь рассмотрим зону $\varepsilon_k \ge \varepsilon$. Здесь мы построим *конечную* систему уровней. Выберем $n = n(\varepsilon)$ из соотношения

$$r^n\Psi(\varepsilon) \le \Psi(\sigma) < r^{n-1}\Psi(\varepsilon).$$

Положим $\varepsilon_0 = \sigma$, а далее выберем ε_k из соотношения

$$\Psi(\varepsilon_k) = r^{n-k}\Psi(\varepsilon), \qquad 1 \le k \le n.$$

В частности, $\varepsilon_n = \varepsilon$.

При любом $k \geq 1$ мы имеем $\Psi(\varepsilon_k) \leq r^{-1}\Psi(\varepsilon_{k-1})$. Заметим, что

$$\begin{aligned}\ln\left(\frac{\varepsilon_{k-1}}{\varepsilon_k}\right)\ \Psi(\varepsilon_k) \quad &\leq \quad \int_{\varepsilon_k}^{\varepsilon_{k-1}} \frac{\mathrm{d}u}{u}\ r^{-1}\Psi(\varepsilon_{k-1}) \\ &\leq \quad r^{-1}\int_{\varepsilon_k}^{\varepsilon_{k-1}} \frac{\Psi(u)\mathrm{d}u}{u}\ . \qquad (11.13)\end{aligned}$$

Наконец, назначим

$$b_k = r^{n-k}\varepsilon, \qquad 0 \leq k < n.$$

Так как

$$\frac{b_k}{\varepsilon_k} = \frac{r^{n-k}\varepsilon}{\varepsilon_k} = r^{n-k}\ \frac{\varepsilon_n}{\varepsilon_k} \leq 1,$$

то при оценке вероятностей будем исходить из простейшей оценки

$$\mathbb{P}\left(\varepsilon_k|\xi| \leq b_k\right) \geq c\ \frac{b_k}{\varepsilon_k} = c\ \frac{r^{n-k}\varepsilon_n}{\varepsilon_k}\ ,$$

где $c = (2/\pi)^{1/2}$. Отсюда

$$\prod_{k=0}^{n-1} \mathbb{P}\left(\varepsilon_k|\xi| \leq b_k\right)^{\Psi(\varepsilon_{k+1})} \geq \prod_{k=0}^{n-1}\left(c\ \frac{r^{n-k}\varepsilon_n}{\varepsilon_k}\right)^{\Psi(\varepsilon_{k+1})} := \Pi_1 \cdot \Pi_2,$$

где

$$\Pi_1 := \prod_{k=0}^{n-1}\left(c\, r^{n-k}\right)^{\Psi(\varepsilon_{k+1})}, \qquad \Pi_2 := \prod_{k=0}^{n-1}\left(\frac{\varepsilon_n}{\varepsilon_k}\right)^{\Psi(\varepsilon_{k+1})}.$$

Для Π_1 мы легко получаем

$$|\ln \Pi_1| \leq \sum_{k=0}^{n-1}\left(|\ln c| + |\ln r|(n-k)\right) r^{n-k-1}\Psi(\varepsilon) \leq c(r)\Psi(\varepsilon),$$

а эта величина уже была оценена в (11.12). Для Π_2 делаем суммирование по частям и с учётом (11.13) приходим к

$$\begin{aligned}
|\ln \Pi_2| &\le \sum_{k=0}^{n-1} \sum_{l=k+1}^{n} \ln\left(\frac{\varepsilon_{l-1}}{\varepsilon_l}\right) \Psi(\varepsilon_{k+1}) \\
&\le (1-r)^{-1} \sum_{l=1}^{n} \ln\left(\frac{\varepsilon_{l-1}}{\varepsilon_l}\right) \Psi(\varepsilon_l) \\
&\le r^{-1} \int_{\varepsilon_n}^{\varepsilon_0} \frac{\Psi(u)\mathrm{d}u}{u} = r^{-1}\widetilde{\Psi}(\varepsilon).
\end{aligned}$$

Остаётся соединить две последовательности ε_k, перенумеровать их и применить лемму 11.1. □

Оценки, получаемые прямым энтропийным методом просты в применении и часто, хотя и не всегда, достаточно эффективны. Однако им не хватает общности в том, что касается рассматриваемых норм, они не позволяют оценивать вероятности малых уклонений *сверху* и дают недостаточно точные результаты в более сложных ситуациях. Поэтому мы сейчас перейдём к менее элементарному, но гораздо более общему *двойственному* энтропийному методу.

11.4 Двойственный энтропийный метод

Пусть X – центрированный гауссовский вектор со значениями в сепарабельном банаховом пространстве $(\mathcal{X}, ||\cdot||)$. В отсутствие марковского свойства, как правило, удаётся найти только порядок малых уклонений. Поэтому мы сосредоточимся на изучении *функции малых уклонений*

$$\phi(\varepsilon) := -\ln \mathbb{P}(||X|| \le \varepsilon),$$

которая при $\varepsilon \to 0$ для многих векторов удовлетворяет асимптотическому соотношению $\phi(\varepsilon) \asymp \varepsilon^{-b}$.

Следуя Дж. Кэлбсу и Венбо Ли, авторам двойственного энтропийного метода [118, 133], рассмотрим $M_D(\varepsilon)$ – ёмкостные числа эллипсоида рассеяния D относительно расстояния, заданного нормой $||\cdot||$, и $\mathcal{C}_D(\varepsilon) = \ln M_D(\varepsilon)$ – соответствующую метрическую ёмкость. Слово „двойственный“ в заголовке подчеркивает, что речь идет о характеристиках энтропийной природы, *не совпадающих* с теми, что рассматривались в главе 10. О соотношении между двумя видами энтропии будет сказано в п. 11.5.

Верхняя оценка вероятностей малых уклонений (она же будет нижней оценкой функции малых уклонений) выглядит следующим образом.

Предложение 11.1 *Для любых* $r > 0, \lambda > 0$ *верно*

$$\phi(r) \geq \mathcal{C}_D\left(\frac{2r}{\lambda}\right) - \frac{\lambda^2}{2}\ . \tag{11.14}$$

Доказательство. Пусть $U := \{x \in \mathcal{X} : ||x|| \leq 1\}$ и P – распределение вектора X в $\mathcal{X}$. Пусть $\varepsilon > 0$. Положим $n = M_D(\varepsilon)$.

Рассмотрим конфигурацию точек $\{h_j, 1 \leq j \leq n\} \subset D$ такую, что $||h_i - h_j|| > \varepsilon$ при $j \neq i$. Тогда все шары вида $h_j + \frac{\varepsilon}{2}U$ не пересекаются, и то же самое верно для растянутых шаров $\lambda(h_j + \frac{\varepsilon}{2}U) = \lambda h_j + \frac{\lambda\varepsilon}{2}U$. Используя неравенство Борелля для сдвинутых множеств (5.3), получим

$$\begin{aligned}
1 &\geq \sum_{j=1}^{n} P\left(\lambda h_j + \frac{\lambda\varepsilon}{2}U\right) \\
&\geq n \min_{1\leq j\leq n} P\left(\frac{\lambda\varepsilon}{2}U\right)\ \exp\left\{-\frac{\lambda^2|h_j|^2_{H_P}}{2}\right\} \\
&\geq n\ P\left(\frac{\lambda\varepsilon}{2}U\right)\ \exp\left\{-\frac{\lambda^2}{2}\right\}.
\end{aligned}$$

Отсюда

$$P\left(\frac{\lambda\varepsilon}{2}U\right) \leq \exp\left\{\frac{\lambda^2}{2}\right\}\ n^{-1}$$

и

$$\ln P\left(\frac{\lambda\varepsilon}{2}U\right) \leq \frac{\lambda^2}{2} - \ln n.$$

Наконец, сменим обозначения, положив $r := \frac{\lambda\varepsilon}{2}$, тогда $\varepsilon = \frac{2r}{\lambda}$, и получим

$$\ln P(rU) \le \frac{\lambda^2}{2} - \ln M_D\left(\frac{2r}{\lambda}\right),$$

что равносильно неравенству (11.14). $\square$

Следствие 11.1 *Пусть при некоторых $\beta \in (0,2), c > 0$ при малых $\varepsilon > 0$ верно $\mathcal{C}_D(\varepsilon) \ge c\,\varepsilon^{-\beta}$. Тогда при некотором $\widetilde{c} > 0$ при малых $r > 0$ верно $\phi(r) \ge \widetilde{c}\, r^{-\frac{2\beta}{2-\beta}}$.*

Доказательство следствия. Фиксируем $\delta > 0$. Подставим в (11.14) значение $\lambda := \delta\, r^{-\frac{\beta}{2-\beta}}$. Получим

$$\begin{aligned}
\ln P(rU) &\le \frac{\delta^2}{2r^{\frac{2\beta}{2-\beta}}} - \mathcal{C}_D\left(\frac{2r\cdot r^{\frac{\beta}{2-\beta}}}{\delta}\right), \\
&\le \frac{\delta^2}{2r^{\frac{2\beta}{2-\beta}}} - c\left(\frac{2r^{\frac{2}{2-\beta}}}{\delta}\right)^{-\beta} \\
&= \left(\frac{\delta^2}{2} - c\,2^{-\beta}\delta^\beta\right) r^{\frac{-2\beta}{2-\beta}} := -\widetilde{c}\; r^{\frac{-2\beta}{2-\beta}},
\end{aligned}$$

причём $\widetilde{c} > 0$, если δ достаточно мало. $\square$

Упражнение 11.2 Выведите с помощью оценки (11.14), что для любого гауссовского вектора в сепарабельном банаховом пространстве верно соотношение $\mathcal{C}_D(\varepsilon) = o\left(\varepsilon^{-2}\right)$. Поэтому значение параметра $\beta \ge 2$ в следствии 11.1 невозможно.

Перейдём к противоположным оценкам. Пусть $N_D(\varepsilon)$ – энтропийные числа эллипсоида рассеяния D относительно расстояния, заданного нормой $||\cdot||$, и $H_D(\varepsilon) = \ln N_D(\varepsilon)$ – соответствующая метрическая энтропия. Нижняя оценка вероятностей малых уклонений (она же будет верхней оценкой функции малых уклонений) выглядит следующим образом.

Предложение 11.2 *Для любого $r > 0$ верно*

$$\phi(2r) \le \ln 2 + H_D\left(\frac{r}{\sqrt{2\phi(r)}}\right). \tag{11.15}$$

Доказательство. Пусть $r, \lambda > 0$ и $n = N_D\left(\frac{r}{\lambda}\right)$. Покроем эллипсоид D минимальным числом множеств диаметра не более $\frac{\varepsilon}{\lambda}$, а затем впишем каждое из этих множеств в шар того же радиуса. Получится покрытие

$$D \subset \bigcup_{j=1}^{n} \left\{h_j + \frac{r}{\lambda}\,U\right\}.$$

Умножая на λ, получим

$$\lambda D \subset \bigcup_{j=1}^{n} \{\lambda h_j + r\,U\}.$$

Отсюда следует

$$\lambda D + rU \subset \bigcup_{j=1}^{n} \{\lambda h_j + 2r\,U\}$$

и по неравенству Андерсона

$$P\,(\lambda D + rU) \le \sum_{j=1}^{n} P\,(\lambda h_j + 2r\,U) \le n\ P(2rU).$$

С другой стороны, изопериметрическое неравенство даёт

$$P\,(\lambda D + rU) \ge \Phi\left(\Phi^{-1}\left(P(rU)\right) + \lambda\right).$$

Сравнивая полученные оценки, находим

$$\Phi\left(\Phi^{-1}\left(P(rU)\right) + \lambda\right) \le n\ P(2rU).$$

Чтобы упростить левую часть, положим $\lambda = \sqrt{2\phi(r)}$. Тогда

$$\Phi(-\lambda) \le \exp\{-\lambda^2/2\} = \exp\{-\phi(r)\} = P(rU),$$

откуда

$$-\lambda \le \Phi^{-1}\left(P(rU)\right),$$

и мы приходим к оценке

$$n\ P(2rU) \ge \Phi(0) = \frac{1}{2}\ .$$

Логарифмируем и находим

$$\ln P(2rU) \geq -\ln 2 - \ln n,$$

что равносильно искомой оценке (11.15). □

Замечание 11.4 Недостатком неравенства (11.15) является, конечно, его итеративность - функция $\phi(\cdot)$ присутствует и в левой, и в правой части. Однако для практически интересных примеров $\phi(\cdot)$ входит в правую часть „в меньшей степени“, что и обеспечивает реальную полезность (11.15). В частности, удаётся доказать оценки, противоположные полученным выше.

Следствие 11.2 *Пусть при некоторых $\beta \in (0,2), c > 0$ при малых $\varepsilon > 0$ верно $N_D(\varepsilon) \leq c\,\varepsilon^{-\beta}$. Тогда при некотором $\widetilde{c} > 0$ при малых $r > 0$ верно неравенство $\phi(r) \leq \widetilde{c}\, r^{-\frac{2\beta}{2-\beta}}$.*

К сожалению, известное доказательство этого факта [133], хотя и основано на (11.15), но также включает ссылки на понятия и результаты, не рассматриваемые в этом курсе. Было бы очень интересно найти короткое „самодостаточное“ доказательство.

Пример 11.2 *(Малые уклонения процесса Римана–Лиувилля* [132]*).*

Для процесса Римана–Лиувилля, определённого в примере 3.4, согласно (4.5) эллипсоид рассеяния имеет вид

$$D = \left\{ h(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \ell(s) \mathrm{d}s, \ ||\ell||_{L_2[0,1]} \leq 1 \right\}.$$

Энтропийные характеристики этого класса функций в $\mathbb{C}[0,1]$ достаточно хорошо изучены: известно, что $H_D(\varepsilon) \approx \varepsilon^{-1/\alpha}$. Отсюда находим $\phi(r) \approx r^{-\frac{2}{2\alpha-1}}$. Можно даже показать, что для некоторой константы $Q(\alpha)$ верно более точное утверждение $\phi(r) \sim Q(\alpha)\, r^{-\frac{2}{2\alpha-1}}$.

Пример 11.3 *(Малые уклонения дробного броуновского движения* [132]*).*

Сравнивая имеющиеся интегральные представления для дробного броуновского движения с показателем α и процесса Римана–Лиувилля с показателем $\frac{\alpha+1}{2}$, см. (3.4) и (3.6), можно заметить, что разница между ними (с точностью до нормирующих констант) есть гладкий процесс со сравнительно большими вероятностями малых уклонений. Используя оценки сравнения (7.9), можно показать, что функции малых уклонений двух процессов эквивалентны. С учётом результатов предыдущего примера, находим, что для ДБД с показателем α верно

$$\phi(r) \sim Q\left(\frac{\alpha+1}{2}\right) \left(\frac{r}{c_\alpha \Gamma(\frac{\alpha+1}{2})}\right)^{-2/\alpha},$$

где c_α – константа из формулы (3.5). Это согласуется с односторонней оценкой (11.9), полученной совсем другим методом.

11.5 Двойственность метрических энтропий

С первого взгляда может показаться, что две энтропии (параметрического множества процесса в его естественной метрике и эллипсоида рассеяния гауссовского вектора в нормированном пространстве), рассмотренные в предыдущих параграфах, не имеют ничего общего между собой. На самом же деле между ними есть глубокая связь, идущая от *гипотезы двойственности*[3] из теории линейных операторов. Забудем на короткое время о теории вероятностей и малых уклонениях и рассмотрим проблему на языке линейных операторов. Пусть $V : (\mathcal{X}_1, ||\cdot||_1) \to (\mathcal{X}_2, ||\cdot||_2)$ – компактный линейный оператор, действующий из одного банахова пространства в другое. Обозначим через

$$B_1 = \{x \in \mathcal{X}_1 : ||x||_1 \le 1\}, \quad B_2 = \{x \in \mathcal{X}_2 : ||x||_2 \le 1\}$$

[3] Duality conjecture.

соответствующие единичные шары. Компактность V означает, что множество $V(B_1)$ компактно в $\mathcal{X}_2$. Следовательно, мы можем рассмотреть энтропийные числа $N_{V(B_1)}(\varepsilon)$ этого множества в метрике пространства $\mathcal{X}_2$ и назвать их энтропийными числами оператора V,

$$N_V(\varepsilon) := N_{V(B_1)}(\varepsilon).$$

Функция $N_V(\cdot)$ характеризует сложность устройства оператора V. Наряду с оператором V рассмотрим сопряжённый (двойственный) оператор

$$V^* : (\mathcal{X}_2^*, ||\cdot||_{*,2}) \to (\mathcal{X}_1^*, ||\cdot||_{*,1}).$$

Свойства операторов V и V^* тесно связаны: их нормы равны, и если V компактен, то и V^* компактен. Поэтому возникает естественный вопрос: как связаны между собой энтропийные числа операторов V и V^*? Ещё в 1972 г. А. Пич[4] высказал *гипотезу двойственности* для энтропийных чисел, которая в полном объёме всё ещё не доказана (хотя, конечно, и не опровергнута). Эта гипотеза утверждает, что существуют две такие универсальные константы a и b, что для любого линейного оператора V и всех $\varepsilon > 0$ верно

$$b^{-1} \ln N_{V^*}(a\varepsilon) \le \ln N_V(\varepsilon) \le b \ln N_{V^*}(a^{-1}\varepsilon).$$

Суть утверждения заключена уже в первом неравенстве; второе следует автоматически применением первого к V^*. Гипотеза двойственности доказана для случая, когда хотя бы одно из пространств является гильбертовым [46].

Вернемся к интересующей нас ситуации с гауссовским вектором $X \in \mathcal{X}$. Рассмотрим канонические операторы вложения $I^* : \mathcal{X}^* \to \mathcal{X}_P^*$ и $I : \mathcal{X}_P^* \to \mathcal{X}$. Образ единичного шара под действием оператора I это в точности эллипсоид рассеяния. Поэтому энтропийные числа $N_I(\varepsilon)$ – те самые, на которых основан двойственный энтропийный подход к малым уклонениям. Запишем норму вектора X в виде супремума,

$$||X|| = \sup_{f \in B^*} (f, X),$$

[4]A. Pietsch.

где B^* – единичный шар пространства $\mathcal{X}^*$. Применяя к множеству $T = B^*$ и случайному процессу $\widetilde{X}(f) := (X, f)$ технику прямого энтропийного подхода, мы сталкиваемся с необходимостью изучать энтропию множества B^* с метрикой

$$\rho(f,g)^2 = \mathbb{E}(\widetilde{X}(f) - \widetilde{X}(g))^2 = \mathbb{E}(f-g,X)^2 = ||I^*f - I^*g||^2_{\mathcal{X}_P^*}.$$

Поэтому в прямом подходе нам нужна именно энтропия оператора I^*. Мы пришли к выводу, что энтропийные числа двух наших подходов к малым уклонениям относятся к сопряжённым операторам! Поскольку пространство $\mathcal{X}_P^*$ является гильбертовым, то в интересующем нас случае неравенства двойственности имеют место и мы можем при необходимости заменять одну энтропию на другую.

11.6 Гильбертово пространство

Если $\mathcal{X}$ – сепарабельное гильбертово пространство, то задачу малых уклонений удаётся решить значительно точнее, чем в общем случае. Зная разложение Кархунена–Лоэва (2.1), можно записать норму центрированного гауссовского вектора в виде

$$||X||^2 = \sum_{j=1}^{\infty} \sigma_j^2\, \xi_j^2 \tag{11.16}$$

(вспомним, что здесь ξ_j – независимые стандартные нормальные величины), так что проблема малых уклонений сводится к изучению суммы независимых случайных величин. Из (11.16) следует явная формула для преобразования Лапласа,

$$\mathbb{E}\exp\{-\gamma||X||^2\} = \prod_{j=1}^{\infty}(1+2\gamma\sigma_j^2)^{-1/2}.$$

Поскольку поведение распределения нормы около нуля (малые уклонения) связано с поведением преобразования Лапласа при больших γ, то в данном случае всё определяется асимптотикой последовательности (σ_j). Аккуратными вычислениями, см. [110], можно получить, например, такой результат.

Предложение 11.3 *Если $\alpha > 1$, $\theta \in \mathbb{R}$, и $\sigma_j^2 \sim C\,\frac{(\ln j)^\theta}{j^\alpha}$, то*

$$\phi(\varepsilon) \sim C^{\frac{1}{\alpha-1}} \left(\frac{\alpha-1}{2}\right)^{1-\frac{\theta}{\alpha-1}} \beta^{\frac{\alpha}{\alpha-1}}\, |\ln\varepsilon|^{\frac{\theta}{\alpha-1}}\, \varepsilon^{\frac{-2}{\alpha-1}},$$

где $\beta := \frac{\pi}{\alpha\sin(\pi/\alpha)}$. Если же известна более точная (двучленная) асимптотика,

$$\sigma_j^2 = C\,\left(j + \delta + O\left(j^{-1}\right)\right)^{-\alpha},$$

то

$$P(||X|| \le \varepsilon) \sim M\varepsilon^\gamma \exp\left\{-C^{\frac{1}{\alpha-1}}\,\frac{\alpha-1}{2}\,\beta^{\frac{\alpha}{\alpha-1}}\,\varepsilon^{\frac{-2}{\alpha-1}}\right\},$$

где $\gamma = \frac{2-\alpha-2\delta\alpha}{2(\alpha-1)}$, а постоянная M зависит от всей последовательности (σ_j).

Проблема однако заключается в том, что гауссовский вектор в гильбертовом пространстве обычно задан не в форме (2.1), а иначе, например, $\mathcal{X} = L_2(T,\mu)$ и известна ковариационная функция $K(s,t) = \mathbb{E}X(s)X(t)$. В этом случае, прежде, чем применять предложение 11.3, нужно исследовать асимптотическое поведение неизвестной последовательности (σ_j). В этом направлении недавно появилось множество результатов и интересных примеров, см. [94, 95, 96, 97, 110, 152], [2, 21, 22, 23, 98], [33, 151].

Теория вероятностей малых уклонений в гильбертовом пространстве допускает два совершенно различных естественных обобщения: на случай сумм взвешенных независимых величин [47, 65, 86, 135] и на случай L_p-пространств с произвольным конечным p, см. [34]-[37], [47], [131, 137, 138, 139].

11.7 Другие результаты

В следствиях 11.1 и 11.2 рассмотрены критерии степенного поведения функции малых уклонений. Однако в некоторых важных случаях приходится иметь дело и с логарифмическими множителями. Поэтому полезно такое уточнение [118, 133].

Предложение 11.4 *Пусть $\beta \in (0,2), \gamma \in \mathbb{R}$. Тогда соотношения $H_D(\varepsilon) \approx \varepsilon^{-\beta} |\ln \varepsilon|^{\gamma}$ и $\phi(r) \approx r^{-\frac{2\beta}{2-\beta}} |\ln r|^{\frac{2\gamma}{2-\beta}}$ равносильны.*

Например, из результатов об энтропии классов гладких функций [32, 54] с помощью предложения 11.4 можно вывести такой результат. Рассмотрим d-параметрический броуновский лист $X = W$ как случайный элемент пространства $L_p[0,1]^d$, $1 \leq p < \infty$. Тогда $H_D(\varepsilon) \approx \varepsilon^{-1} |\ln \varepsilon|^{d-1}$ и $\phi(r) \approx r^{-2} |\ln r|^{2d-2}$. Поведение малых уклонений броуновского листа в равномерной норме (случай $p = \infty$) является трудной и интересной нерешенной задачей. Здесь известно, что

$$\phi(r) \begin{cases} \sim \frac{\pi^2}{8}\, r^{-2}, & d = 1, \\ \approx r^{-2} |\ln r|^3, & d = 2, [173], \\ \preceq r^{-2} |\ln r|^{2d-1} & d > 2, [85], \\ \succeq r^{-2} |\ln r|^{2d-2+q}, & \text{при некотором } q = q(d),\ d > 2, [55]. \end{cases}$$

При $d > 2$ между верхней и нижней оценками виден логарифмический зазор.

Для очень гладких процессов функция малых уклонений может вообще иметь логарифмическое поведение. В таких случаях полезен другой результат [1].

Предложение 11.5 *Пусть $\gamma > 0, \zeta \in \mathbb{R}$. Тогда равносильны соотношения $H_D(\varepsilon) \approx |\ln \varepsilon|^{\gamma} \ln |\ln \varepsilon|^{\zeta}$ и $\phi(r) \approx |\ln r|^{\gamma} \ln |\ln \varepsilon|^{\zeta}$.*

В байесовской статистике для моделирования случайных плотностей распределения часто используется стационарный гауссовский случайный процесс с нормальной спектральной плотностью $f(u) = \exp(-u^2)$, и для анализа скорости сходимости байесовских оценок оказывается критически важно знать поведение его вероятностей малых уклонений [178]. Рассмотрим этот процесс как элемент семейства процессов X_ν со спектральными плотностями

$$f_\nu(u) = \begin{cases} \exp(-|u|^\nu), & 0 < \nu < \infty, \\ \mathbf{1}_{[-1,1]}(u), & \nu = \infty. \end{cases}$$

Тогда для процесса X_ν с помощью результатов об энтропии классов аналитических функций [15] и предложения 11.5 можно (см. [1]) вывести, что

$$\phi(r) \approx H_D(r) \approx \begin{cases} \frac{|\ln r|^2}{\ln|\ln r|}, & 1 < \nu \le \infty, \\ \\ |\ln r|^{1+\frac{1}{\nu}}, & 0 < \nu \le 1. \end{cases}$$

Все приведённые примеры указывают на общее правило: *чем глаже траектории процесса, тем больше вероятности малых уклонений, т.е. меньше функция малых уклонений.*

Много дополнительной информации о теории малых уклонений можно найти в обзорах [134, 136], а актуальную литературу по теме – в интернет-библиографии [18].

Глава 12

Разложения гауссовских векторов

12.1 Постановка задачи

Задачу о разложении гауссовского вектора в ряд можно ставить двумя способами.

Сильная форма. Задан гауссовский $\mathcal{X}$-значный случайный вектор как измеримое отображение $X : (\Omega, \mathbb{P}) \mapsto \mathcal{X}$. Найти разложение вида

$$X(\omega) = \sum_{j=1}^{\infty} \xi_j(\omega) x_j,$$

которое справедливо $\mathbb{P}$-почти наверное, причём ξ_j – независимые нормальные величины, заданные на $(\Omega, \mathbb{P})$ и (x_j) – последовательность неслучайных векторов в $\mathcal{X}$.

Слабая форма. Задана гауссовская мера P на $\mathcal{X}$. Построить такой случайный вектор X вида

$$X(\omega) = \sum_{j=1}^{\infty} \xi_j(\omega) x_j,$$

что X имеет распределение P, а ξ_j – независимые нормальные величины и ряд сходится $\mathbb{P}$-почти наверное.

12.2 Ряды из независимых случайных векторов

Для простоты изложения предположим далее, что на $\mathcal{X}$ задана норма $||\cdot||$ и $(\mathcal{X}, ||\cdot||)$ – сепарабельное банахово пространство. Рассмотрим ряд из частных сумм

$$S_n = \sum_{j=1}^{n} X_j,$$

где $X_j \in \mathcal{X}$ – независимые случайные векторы, заданные на общем вероятностном пространстве $(\Omega, \mathbb{P})$. Можно рассмотреть три типа сходимости S_n:

- распределения S_n слабо сходятся к некоторому распределению P на $\mathcal{X}$.
- Существует предельный случайный вектор $S \in \mathcal{X}$, такой, что $S_n \overset{\mathbb{P}}{\Rightarrow} S$, то есть

$$\lim_{n\to\infty} \mathbb{P}\{||S_n - S|| > \varepsilon\} \qquad \forall \varepsilon > 0.$$

- Справедливо равенство

$$\lim_{n\to\infty} S_n = S \qquad \mathbb{P}\text{-почти наверное.} \tag{12.1}$$

Из общей теории сходимости известно, что из третьего свойства следует второе, а из второго – первое. Однако в случае независимых слагаемых все три свойства эквивалентны.

Теорема 12.1 *Если векторы X_j – центрированные гауссовские, то из сходимости распределений S_n следует* (12.1).

Замечание 12.1 Гауссовость в этой теореме не так важна как независимость. Утверждение остаётся верным, например, для рядов, составленных из симметрично распределённых независимых векторов.

Доказательство теоремы 12.1. Мы проведём его в несколько этапов, постепенно расширяя класс рассматриваемых пространств.

Этап 1. $\mathcal{X} = \mathbb{R}^1$. Здесь X_j – обыкновенные случайные величины. Пусть $\sigma_j^2 = \mathbb{D}X_j$. Тогда S_n имеет распределение $N\left(0, \sum_{j=1}^n \sigma_j^2\right)$. Из сходимости распределений S_n следует, что $\sum_{j=1}^\infty \sigma_j^2 < \infty$. Поскольку сходимость рядов, составленных из математических ожиданий и дисперсий, гарантирует сходимость почти наверное ряда из независимых случайных величин (классическая „теорема о двух рядах“ Колмогорова–Хинчина), то (12.1) имеет место.

Этап 2. $\mathcal{X} = \mathbb{R}^n$. Утверждение получается применением результата этапа 1 к координатам векторов X_j. Напомним, что они тоже независимы, центрированы и для каждой координаты распределения сумм сходятся к распределению этой координаты относительно предельного закона. С другой стороны, очевидно, что если отдельные координаты сумм S_n сходятся почти наверное, то и векторы S_n тоже имеют это свойство.

Этап 3. $\mathcal{X}$ – произвольное конечномерное пространство. Утверждение следует из доказанного на предыдущем этапе и существования линейного изоморфизма между $\mathcal{X}$ и соответствующим $\mathbb{R}^n$, причём этот изоморфизм и обратное к нему отображение являются ограниченными операторами.

Этап 4. $\mathcal{X}$ – произвольное сепарабельное банахово пространство. Сначала переформулируем понятие сходимости почти наверное в терминах, близких к сходимости по вероятности. Будем говорить, что последовательность случайных элементов S_n нормированного пространства $\mathcal{X}$ *фундаментальна по вероятности*, если

$$\lim_{n\to\infty} \mathbb{P}\left\{ \sup_{n_1,n_2\geq n} ||S_{n_1} - S_{n_2}|| \geq \varepsilon \right\} = 0, \qquad \forall \varepsilon > 0.$$

Покажем, что если последовательность S_n фундаментальна по вероятности и пространство $\mathcal{X}$ полно, то найдётся предельный случайный элемент $S \in \mathcal{X}$, для которого $\lim_n S_n = S$

почти наверное. Действительно, положим

$$M_n = \sup_{n_1, n_2 \geq n} ||S_{n_1} - S_{n_2}||.$$

Поскольку M_n – убывающая последовательность, то существует предел $M = \lim_n M_n$. Для любых $n \geq 1$, $\varepsilon > 0$ имеем $\{M \geq \varepsilon\} \subset \{M_n \geq \varepsilon\}$. Отсюда получаем

$$\mathbb{P}\{M \geq \varepsilon\} \leq \lim_{n\to\infty} \mathbb{P}\{M_n \geq \varepsilon\} = 0.$$

Ввиду произвольности ε видим, что $M = 0$ почти наверное. Таким образом, последовательность S_n фундаментальна и предел S существует.

Вернемся к рядам. Предположим, что последовательность частных сумм сходится по распределению. Покажем, что она фундаментальна по вероятности. Пусть P – предельное распределение S_n. Фиксируем малые числа $\varepsilon, \delta > 0$. Воспользуемся тем, что любая конечная мера в сепарабельном банаховом пространстве плотна, т.е. найдётся компактное множество $K \subset \mathcal{X}$, для которого $P(K) \geq 1 - \delta$. Выберем в K конечную ε-сеть $x_1, \ldots, x_m$. Пусть $\widetilde{\mathcal{X}}$ – линейная оболочка этой сети. Как обычно, расстояние между точкой и множеством определяется формулой

$$\rho(x, A) = \inf_{y\in A} ||x - y||, \qquad x \in \mathcal{X}, A \subset \mathcal{X}.$$

Тогда

$$\sup_{x\in K} \rho(x, \widetilde{\mathcal{X}}) \leq \sup_{x\in K} \inf_{1\leq j\leq m} ||x - x_j|| \leq \varepsilon.$$

Поэтому из определения сходимости распределений

$$\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}\{\rho(S_n, \widetilde{\mathcal{X}}) \geq 2\varepsilon\} &\leq P\{x : \rho(x, \widetilde{\mathcal{X}}) > 2\varepsilon\} \\
&\leq P(\mathcal{X} \backslash K) \leq \delta.
\end{aligned}$$

Итак, для достаточно больших n верно

$$\mathbb{P}\{\rho(S_n, \widetilde{\mathcal{X}}) \geq 2\varepsilon\} \leq 2\delta.$$

Из функционального анализа известно, что существует линейный проектор $L : \mathcal{X} \mapsto \widetilde{\mathcal{X}}$, удовлетворяющий условию $||L|| \leq 2$. Запишем

$$S_n = (S_n - LS_n) + LS_n := S'_n + \widetilde{S}_n.$$

Выведем теперь оценки, позволяющие свести дело к конечномерному случаю. Найдём элемент $y_n \in \widetilde{\mathcal{X}}$, реализующий минимум

$$||S_n - y_n|| = \rho(S_n, \widetilde{\mathcal{X}}).$$

Тогда

$$\begin{array}{rcl} ||S'_n|| & = & ||S_n - LS_n|| \\ & \leq & ||S_n - y_n|| + ||y_n - Ly_n|| + ||Ly_n - LS_n|| \\ & \leq & (1 + ||L||)\ ||S_n - y_n|| \leq 3\rho(S_n, \widetilde{\mathcal{X}}). \end{array}$$

Для получения равномерной оценки воспользуемся следующей леммой.

Лемма 12.1 *(Неравенство Леви). Если случайные векторы X_j независимы и симметрично распределены, то для любого $r > 0$ их суммы допускают оценку*

$$\mathbb{P}\left\{\max_{k\leq n} ||S_k|| \geq r\right\} \leq 2\ \mathbb{P}\left\{||S_n|| \geq r\right\}. \qquad (12.2)$$

Применяя (12.2) к суммам S'_n, найдём

$$\begin{array}{rcl} \mathbb{P}\left\{\max\limits_{k\leq n} ||S'_k|| \geq 6\varepsilon\right\} & \leq & 2\mathbb{P}\left\{||S'_n|| \geq 6\varepsilon\right\} \\ & \leq & 2\mathbb{P}\left\{\rho(S_n, \widetilde{\mathcal{X}}) \geq 2\varepsilon\right\} \leq 2\delta. \end{array}$$

Следовательно,

$$\mathbb{P}\left\{\max_{k} ||S'_k|| \geq 6\varepsilon\right\} = \lim_{n\to\infty} \mathbb{P}\left\{\max_{k\leq n} ||S'_k|| \geq 6\varepsilon\right\} \leq 2\delta.$$

С другой стороны, конечномерная последовательность $\widetilde{S}_n$ сходится по распределению, будучи линейной проекцией сходящейся последовательности S_n. В силу результата этапа 3 суммы $\widetilde{S}_n$ будут фундаментальны по вероятности. Поэтому при больших n мы получим

$$\begin{aligned}
& \mathbb{P}\left\{\sup_{n_1,n_2\ge n}||S_{n_1}-S_{n_2}||\ge 13\varepsilon\right\} \\
\le\ & \mathbb{P}\{\sup_{n_1,n_2\ge n}||S'_{n_1}-S'_{n_2}||\ge 12\varepsilon\} \\
& +\mathbb{P}\{\sup_{n_1,n_2\ge n}||\widetilde{S}_{n_1}-\widetilde{S}_{n_2}||\ge \varepsilon\} \\
\le\ & \mathbb{P}\{\sup_k||S'_k||\ge 6\varepsilon\}+\delta\le 3\delta.
\end{aligned}$$

Поскольку δ произвольно, мы проверили фундаментальность по вероятности последовательности S_n, что и требовалось доказать. □

12.3 Построение вектора с заданным распределением

Пусть в $\mathcal{X}$ задано центрированное гауссовское распределение $P = N(0,K)$ и пусть H_P – соответствующее ядро. Далее, пусть (ξ_j) – последовательность независимых $N(0,1)$-распределённых случайных величин, а (h_j) – ортонормальный базис в $\mathcal{X}_P^*$. Тогда $(h_j) = (Iz_j)$ – ортонормальный базис в H_P. (Здесь I – канонический изоморфизм пространств $\mathcal{X}_P^*$ и H_P, см. главу 4.) Покажем, что ряд

$$X(\omega)=\sum_{j=1}^{\infty}\xi_j(\omega)h_j$$

сходится и что распределение суммы X совпадает с P, [161]. В соответствии с теоремой 12.1 достаточно проверить, что частные суммы $S_n=\sum_{j=1}^{n}\xi_j(\omega)h_j$ сходятся по распределению к P. Проверим сходимость характеристических функций. Действительно, для $f\in\mathcal{X}^*$ имеем

$$(f,S_n)=\sum_{j=1}^{n}\xi_j(\omega)(f,h_j)=\sum_{j=1}^{n}\xi_j(\omega)(I^*f,z_j), \tag{12.3}$$

$$
\begin{aligned}
\mathbb{E}e^{i(f,S_n)} &= \mathbb{E}\exp\left\{ i\sum_{j=1}^{n} \xi_j(\omega)(I^*f,z_j)\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(I^*f,z_j)^2\right\} \\
&\to \exp\left\{-\frac{1}{2}\sum_{j=1}^{\infty}(I^*f,z_j)^2\right\} \\
&= \exp\left\{-\frac{1}{2}||I^*f||_2^2\right\} = \int_{\mathcal{X}} e^{i(f,x)}P(\mathrm{d}x). \ \square
\end{aligned}
$$

12.4 Разложение заданного вектора

Теперь решим задачу о разложении в сильной форме. Пусть задан вектор X с распределением P. Воспользуемся предыдущей конструкцией, но уточним выбор случайных величин (ξ_j), выбирая $\xi_j = z_j(X)$. Напомним, что линейные измеримые функционалы образуют базис в $\mathcal{X}_P^*$, поэтому

$$
\mathbb{E}z_i(X)z_j(X) = \int_{\mathcal{X}} z_i z_j \mathrm{d}P = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}
$$

Мы уже знаем, что сумма

$$
Y = \sum_{j=1}^{\infty} z_j(X)h_j
$$

определена $\mathbb{P}$-почти наверное. Остаётся доказать, что $Y = X$ с вероятностью единица. Для любого функционала $f \in \mathcal{X}^*$ имеем $\mathbb{P}$-почти наверное в силу (12.3)

$$
(f,Y) = \lim_{n\to\infty}\sum_{j=1}^{n} z_j(X)(I^*f,z_j)
$$

С другой стороны, в пространстве $\mathcal{X}_P^* \subset L_2(\mathcal{X},P)$

$$
\lim_{n\to\infty}\sum_{j=1}^{n} z_j(\cdot)(I^*f,z_j) = (I^*f)(\cdot),
$$

то есть

$$\mathbb{E}\left|\sum_{j=1}^{n} z_j(X)(I^* f, z_j) - (f, X)\right|^2 = 0.$$

Отсюда следует, что $(f, X) = (f, Y)$ почти наверное. Следовательно, для любого $f \in \mathcal{X}^*$ имеем $\mathbb{E}e^{i(f,X-Y)} = 1$ и $X = Y$ с вероятностью единица. □

12.5 Примеры разложений: винеровский процесс

Рассмотрим несколько примеров разложений, взяв в качестве гауссовского вектора винеровский процесс W, а в качестве пространства $\mathbb{C}[0,1]$. Напомним, что интегрированием любого базиса в $L_2[0,1]$ можно получить некоторый базис в ядре H_P для винеровской меры.

Пример 12.1 *(Базис косинусов).* Возьмём в $L_2[0,1]$ базис

$$\begin{cases} \varphi_0(s) = 1, \\ \varphi_j(s) = \sqrt{2}\cos(\pi j s), \quad j \geq 1. \end{cases}$$

Интегрируя, получим базис в ядре

$$\begin{cases} h_0(t) = t, \\ h_j(t) = \sqrt{2}\ \frac{\sin(\pi j t)}{\pi j}, \quad j \geq 1. \end{cases}$$

Приходим к разложению

$$W(t) = \xi_0 t + \sqrt{2}\sum_{j=1}^{\infty} \xi_j\ \frac{\sin(\pi j t)}{\pi j}\ .$$

Заметим, что $W(1) = \xi_0$, и воспользуемся представлением (3.3) для броуновского моста. Видим, что если отбросить первое (линейное) слагаемое, то оставшаяся сумма даст разложение броуновского моста.

Пример 12.2 *(Базис синусов).* Возьмём в $L_2[0,1]$ базис

$$\varphi_j(s) = \sqrt{2}\sin(\pi j s), \qquad j \geq 1.$$

Интегрируя, получим базис в ядре

$$h_j(t) = \sqrt{2}\ \frac{1-\cos(\pi j t)}{\pi j}, \qquad j \geq 1.$$

Приходим к разложению

$$W(t) = \sqrt{2}\sum_{j=1}^{\infty} \xi_j\ \frac{1-\cos(\pi j t)}{\pi j}\ .$$

Это разложение фигурировало ещё в работах самого Н. Винера.

Пример 12.3 *(разложение Кархунена–Лоэва)*[1]. Выберем в $L_2[0,1]$ базис

$$\varphi_j(s) = \sqrt{2}\cos(\pi(j-1/2)s), \qquad j \geq 1.$$

Интегрируя, получим базис в ядре

$$h_j(t) = \sqrt{2}\ \frac{\sin(\pi(j-\frac{1}{2})t)}{\pi(j-\frac{1}{2})}\ , \qquad j \geq 1.$$

Приходим к разложению

$$W(t) = \sqrt{2}\sum_{j=1}^{\infty} \xi_j\ \frac{\sin(\pi(j-\frac{1}{2})t)}{\pi(j-\frac{1}{2})}\ .$$

Это разложение интересно тем, что функции h_j сами ортогональны в $L_2[0,1]$. Если рассмотреть W как элемент $L_2[0,1]$, а не $\mathbb{C}[0,1]$, то (h_j) будет ортогональной системой собственных функций ковариационного оператора винеровской меры в $L_2[0,1]$.

[1] K. Karhunen, M. Loève.

Пример 12.4 *(Разложение Пэли–Винера[2]).* Выберем в пространстве $L_2[0,1]$ базис

$$\begin{cases} \varphi_0(s) = 1, \\ \varphi_{2j}(s) = \sqrt{2}\cos(2\pi j s), & j \geq 1, \\ \varphi_{2j-1}(s) = \sqrt{2}\sin(2\pi j s), & j \geq 1. \end{cases}$$

Интегрируя, получим базис в ядре

$$\begin{cases} h_0(t) = t, \\ h_{2j}(t) = \frac{\sin(2\pi j t)}{\sqrt{2}\pi j}, & j \geq 1, \\ h_{2j-1}(t) = \frac{(1-\cos(2\pi j t))}{\sqrt{2}\pi j}, & j \geq 1. \end{cases}$$

Приходим к разложению

$$W(t) = \xi_0 t + \sum_{j=1}^{\infty} \xi_{2j} \, \frac{\sin(2\pi j t)}{\sqrt{2}\pi j} + \sum_{j=1}^{\infty} \xi_{2j-1} \, \frac{(1-\cos(2\pi j t))}{\sqrt{2}\pi j} \; .$$

Пример 12.5 *(Разложение Хаара–Шаудера[3]).* Обозначим

$$\psi(s) = \begin{cases} 1, & 0 \leq s \leq 1/2, \\ -1, & 1/2 < s \leq 1, \\ 0, & s < 0 \text{ или } s > 1, \end{cases}$$

и

$$h(t) = \int_0^t \psi(s) \mathrm{d}s = \begin{cases} t, & 0 \leq t \leq \frac{1}{2} \; , \\ 1 - t, & \frac{1}{2} \leq t \leq 1, \\ 0, & t < 0 \text{ или } t > 1. \end{cases}$$

Возьмём в $L_2[0,1]$ базис Хаара

$$\begin{cases} \varphi_0(s) = 1, \\ \varphi_{j,k}(s) = 2^{j/2}\psi\left(2^j(t - \frac{k}{2^j})\right), & j \geq 0, \quad 0 \leq k \leq 2^j - 1. \end{cases}$$

[2] R. E. A. C. Paley, N. Wiener, [153].

[3] A. Haar, Yu. Schauder.

Интегрируя, получим базис в ядре

$$\begin{cases} h_0(t) = t, \\ h_{j,k}(t) = 2^{-j/2} h\left(2^j(t - \frac{k}{2^j})\right), \qquad j \geq 0, \quad 0 \leq k \leq 2^j - 1. \end{cases}$$

„Треугольные" функции $h_{j,k}$ называются *функциями Шаудера.* Приходим к разложению

$$W(t) = \xi_0 t + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \xi_{j,k} h_{j,k}(t) \ . \qquad (12.4)$$

Представление винеровского процесса (12.4) часто называют *конструкцией Леви*, [16]. Аналогичные представления возможны и с использованием других вейвлет-базисов в $L_2[0,1]$.

Разложения с помощью базиса Шаудера, подобные конструкции Леви, можно строить для практически произвольных гауссовских процессов [70]. Они оказались очень удобным инструментом для исследования свойств процессов, в том числе для оценок вероятностей малых уклонений [166, 167].

Интересно, что аналогичное разложение W на полуоси $[0,\infty)$ будет выглядеть более однородно:

$$W(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \xi_{j,k} h_{j,k}(t) \ .$$

На интервале $[0,1]$ все слагаемые с отрицательными индексами j будут линейными функциями. В сумме они дают первое слагаемое из (12.4).

Пример 12.6 *(Разложение комплексного винеровского процесса, [153]).* Рассмотрим комплексный винеровский процесс $W(t) = W_1(t) + iW_2(t)$, где W_1, W_2 – независимые вещественные винеровские процессы. Будем рассматривать W как гауссовский вектор со значениями в пространстве комплексных непрерывных функций $\mathbb{C}_{\mathbf{C}}([0,1])$.

Возьмём в $L_{2,\mathbf{C}}[0,1]$ базис

$$\varphi_j(s) = \exp\{2\pi i j s\}, \qquad j \in \mathbb{Z}.$$

Интегрируя, получим базис в ядре

$$h_j(s) = \frac{\exp\{2\pi i j t\} - 1}{2\pi i j}, \qquad j \in \mathbb{Z}.$$

Приходим к разложению

$$W(t) = \sum_{j=-\infty}^{\infty} \xi_j \, \frac{\exp\{2\pi i j t\} - 1}{2\pi i j}, \tag{12.5}$$

где (ξ_j) – независимые *комплексные* стандартные гауссовские величины.

Пример 12.7 *(Разложение комплексного дробного броуновского движения).* Предыдущий пример можно обобщить на случай комплексного дробного броуновского движения, т.е. процесса $W^{(\alpha)}(t) = W_1^{(\alpha)}(t) + iW_2^{(\alpha)}(t)$, где $W_1^{(\alpha)}, W_2^{(\alpha)}$ – независимые вещественные ДБД (см. пример 2.5). По аналогии с разложением (12.5) К. Джапаридзе и Х. ван Зантен[4] нашли разложение [87]

$$W^{(\alpha)}(t) = \sum_{j=-\infty}^{\infty} \sigma_j \, \xi_j \, \frac{\exp\{2i\omega_j t\} - 1}{2i\omega_j}, \tag{12.6}$$

где (ξ_j) – снова независимые комплексные стандартные гауссовские величины, ω_j – вещественные нули функции Бесселя $J_{1-\alpha/2}(\cdot)$, а дисперсии коэффициентов разложения вычисляются по формулам

$$\sigma_j^2 = \left[(2-\alpha)\Gamma(1-\alpha/2)^2(\omega_j/2)^\alpha J_{-\alpha/2}(\omega_j)V\right]^{-1},$$

$$V = \frac{\Gamma(\frac{3-\alpha}{2})}{\alpha\Gamma(\frac{\alpha+1}{2})\Gamma(3-\alpha)}.$$

Заметим, что при $\alpha = 1$ верно $J_{1/2}(z) = (2\pi/z)^{1/2} \sin z$. Поэтому $\omega_j = \pi j$ и (12.6) превращается в (12.5).

Варианты разложения (12.6) для других гауссовских процессов и полей можно найти в [87, 88, 147].

[4] K. Dzhaparidze, J. H. van Zanten.

Упражнение 12.1 Рассмотрим броуновский лист (он же – поле Винера–Ченцова) W на квадрате $[0,1]^2$, определённый в примере 2.7, как случайный элемент пространства $\mathbb{C}([0,1]^2)$. Пусть H – ядро распределения W. Постройте ортонормальный базис в H и разложение W в ряд по образцу винеровского процесса.

12.6 Линейные операторы и разложения

Пусть $\mathcal{X}$ – линейное пространство, удовлетворяющее обычным условиям, $\mathcal{H}$ – некоторое гильбертово пространство и $J : \mathcal{H} \mapsto \mathcal{X}$ – линейный оператор. Любому ортонормированному базису (e_j) в $\mathcal{H}$ мы можем сопоставить формальный ряд

$$X = \sum_{j=1}^{\infty} \xi_j \, Je_j \ , \tag{12.7}$$

где (ξ_j) – последовательность независимых $N(0,1)$-распределённых случайных величин.

Можно доказать, что ряд (12.7) сходится с вероятностью единица или расходится с вероятностью единица, причём сам факт сходимости не зависит от выбора базиса (e_j). В дальнейшем будем предполагать, что ряд (12.7) сходится. Тогда X будет гауссовским вектором с нулевым средним и ковариационным оператором $K = JJ^*$. Поэтому распределение $P = N(0, JJ^*)$ вектора X не зависит от выбора базиса в $\mathcal{H}$. Более того, по теореме о факторизации ядро H_P меры P совпадает с $J(\mathcal{H})$. Будем говорить, что гауссовский случайный вектор X и мера P *ассоциированы* с оператором J.

Если $(\mathcal{X}, ||\cdot||)$ – нормированное пространство, то в терминах X можно сформулировать важные характеристики компактности оператора J. Так, ***ℓ-нормой*** оператора J называется (см. [17, 158]) величина

$$||J||_\ell := \left(\mathbb{E}||X||^2\right)^{1/2}.$$

Выбор второго момента для определения нормы здесь не важен, так как согласно (8.10) все моменты нормы гауссовского вектора эквивалентны.

Стохастические аппроксимационные числа $\ell_n(J)$ характеризуют качество возможной аппроксимации оператора J операторами конечного ранга [133, 158]:

$$\ell_n(J) = \inf \left\{ ||J - F||_\ell \ ; \ F : \mathcal{H} \mapsto \mathcal{X}, rank(F) < n \right\}.$$

Можно показать, что $\ell_n(J)$ эквивалентны качеству аппроксимации ассоциированного вектора гауссовскими случайными векторами конечного ранга:

$$\ell_n(J)^2 = \inf_{\substack{x_1,\dots,x_{n-1} \\ \xi_1,\dots,\xi_{n-1}}} \left\{ \mathbb{E} \Big\| X - \sum_{j=1}^{n-1} \xi_j x_j \Big\|^2 \right\}.$$

Заметим, что скорость сходимости ряда (12.7) может, вообще говоря, зависеть от выбора базиса (e_j). Однако базис с правильным порядком скорости сходимости всегда существует [119]:

Упражнение 12.2 Пусть $\alpha > 0, \beta \in \mathbb{R}$. Докажите, что найдётся такая константа $C_{\alpha,\beta}$, что для любого оператора J и ассоциированного с ним гауссовского вектора X в нормированном пространстве из

$$\ell_n(J)^2 \le n^{-\alpha}(1 + \ln n)^\beta, \qquad n \ge 1,$$

следует, что найдётся такой базис (e_j), для которого верно

$$\mathbb{E} \Big\| X - \sum_{j=1}^{n-1} \xi_j \, J e_j \Big\|^2 \le C_{\alpha,\beta} \; n^{-\alpha}(1 + \ln n)^\beta, \qquad n \ge 1.$$

Глава 13

Квантование гауссовских векторов

13.1 Постановка задачи

Задача квантования (дискретизации)[1] случайных векторов происходит из теории передачи информации. Представим себе, что по каналу связи должен быть передан некоторый сигнал (картинка, запись звука и т.д.). Множество всех возможных сигналов образует метрическое пространство $(\mathcal{X}, \rho)$. Одна из возможных идей организации передачи состоит в использовании „словарей". Словарь $Y = \{y_j\}_{1 \le j \le n}$ – это конечное подмножество пространства $\mathcal{X}$. Как узел приема, так и узел передачи информации располагают копией словаря. Когда нужно передать сигнал $x \in \mathcal{X}$, то узел передачи определяет ближайший к x элемент словаря y_j и передаёт по каналу *номер* j аппроксимирующего элемента. Узел приема по словарю восстанавливает значение y_j. Разумеется, передача номера происходит гораздо быстрее, чем передача всего сигнала. За ускорение передачи мы платим ошибкой, так как на выходе процедуры вместо сигнала x оказывается его аппроксимация y_j.

[1] Quantization.

К анализу описанной процедуры приме́ним байесовский подход, то есть будем считать, что на $\mathcal{X}$ задана вероятностная мера P, характеризующая распределение вероятности того, что потребуется передавать сигнал x. Пусть X – случайный элемент $\mathcal{X}$ с распределением P. Тогда средние ошибки квантования по словарю Y можно определить формулой

$$d(Y,p) = \left(\mathbb{E} \min_{1\le j\le n} \rho(X,y_j)^p\right)^{1/p}.$$

Обычно изучается квантование *высокой разрешимости*[2], т.е. асимптотическое поведение процедуры при размере словаря, стремящемся к бесконечности.

Но как построить разумный словарь большого размера в сложном пространстве сигналов? Один из вариантов – рассмотреть случайный словарь, элементы которого Y_j – независимые случайные элементы с тем же распределением P, что у передаваемого сигнала. Эту схему мы и будем изучать. Сделав экспоненциальную замену переменной $n = e^\lambda$, чтобы работать со степенными функциями, придём к выражению для ошибки квантования

$$D(\lambda,p) = \left(\mathbb{E} \min_{1\le j\le e^\lambda} \rho(X,Y_j)^p\right)^{1/p}.$$

В дальнейшем, следуя обширной литературе, будем считать, что $(\mathcal{X}, ||\cdot||)$ – нормированное пространство, соответствующее расстояние $\rho(x,y) = ||x-y||$, а общее распределение P всех векторов – центрированное гауссовское.

13.2 Квантование и малые уклонения

При сделанных предположениях оказывается, что ошибки квантования гауссовского вектора тесно связаны с поведением его вероятностей малых уклонений, см. [79, 90].

Напомним, что функция малых уклонений случайного вектора X определяется соотношением

$$\phi(\varepsilon) := -\ln \mathbb{P}(||X|| \le \varepsilon).$$

[2]High resolution.

В качестве наиболее наглядного примера приведём следующий результат.

Теорема 13.1 *Пусть* $M_\lambda = \min_{j \le e^\lambda} ||Y_j - X||$. *Тогда*

$$\lim_{\lambda\to\infty} \mathbb{P}\left(M_\lambda \ge 2\phi^{-1}(\lambda/6)\right) = 0.$$

При минимальных дополнительных предположениях можно оценить и моменты уклонений $D(\lambda,p) = (\mathbb{E}M_\lambda^p)^{1/p}$.

Теорема 13.2 *Предположим, что функция малых уклонений вектора* X *удовлетворяет условию регулярности: при некотором* $c > 0$

$$\phi(c\,r) \ge 2\,\phi(r), \qquad r < r_0. \tag{13.1}$$

Тогда для любого $p > 0$ *верно*

$$\limsup_{\lambda\to\infty} \frac{D(\lambda,p)}{\phi^{-1}(\lambda/2)} \le 2.$$

По поводу последней теоремы заметим, что благодаря высокой концентрации гауссовских распределений значение анализируемого момента p не играет особой роли.

Суть обеих теорем в том, что если вероятности малых уклонений не слишком малы, то есть функции ϕ, ϕ^{-1} оцениваются сверху, то можно оценить и ошибку квантования.

Доказательство теоремы 13.1. Для $x \in \mathcal{X}$, $r > 0$ обозначим

$$\nu(x,r) = \inf\{|h|_{H_P}, ||h - x|| \le r\}.$$

Тогда верно

$$\mathbb{P}\left(||Y - x|| \le 2r\right) \ge \exp(-\phi(r) - \nu(x,r)^2/2). \tag{13.2}$$

Действительно, для любого h, такого, что $||x - h|| \le r$, имеем включение

$$\{y : ||y - x|| \le 2r\} \supset \{y : ||y - h|| \le r\}.$$

По неравенству Борелля для сдвинутых множеств (5.3) имеем

$$\begin{aligned}\mathbb{P}\left(||Y-x|| \le 2r\right) &\ge \mathbb{P}\left(||Y-h|| \le r\right)\\ &\ge \mathbb{P}\left(||Y|| \le r\right)\exp(-|h|^2_{H_P}/2)\\ &= \exp(-\phi(r)-|h|^2_{H_P}/2).\end{aligned}$$

Минимизируя по h, приходим к (13.2).

Теперь докажем, что для любого $\delta>0$ при достаточно малых r верно неравенство

$$\mathbb{P}\left(\nu(X,r)\ge 2\sqrt{(2+\delta)\phi(r)}\right) \le \exp(-\phi(r)). \tag{13.3}$$

Пусть $D=\{h : |h|_{H_P}\le 1\}$ – эллипсоид рассеяния меры P, а $U=\{x : ||x||\le 1\}$ – единичный шар пространства $\mathcal{X}$. Тогда изопериметрическое неравенство для любого $u>0$ даёт

$$\begin{aligned}\mathbb{P}\left(\nu(X,r)\ge u\right) &= \mathbb{P}\left((X+rU)\cap uD=\emptyset\right)\\ &= \mathbb{P}\left(X\notin rU+uD\right)\\ &\le \widehat{\Phi}\left(\Phi^{-1}\left(P(rU)\right)+u\right)\\ &= \widehat{\Phi}\left(\Phi^{-1}\left(\exp(-\phi(r)\right)+u\right).\end{aligned}$$

Учитывая, что при достаточно малых $p>0$ верно

$$\Phi^{-1}\left(p\right)\ge -\sqrt{(2+\delta)|\ln p|},$$

получаем

$$\mathbb{P}\left(\nu(X,r)\ge u\right)\le \widehat{\Phi}\left(-\sqrt{(2+\delta)\phi(r)}+u\right).$$

Полагая $u=2\sqrt{(2+\delta)\phi(r)}$, имеем

$$\begin{aligned}\mathbb{P}\left(\nu(X,r)\ge 2\sqrt{(2+\delta)\phi(r)}\right) &\le \widehat{\Phi}\left(\sqrt{(2+\delta)\phi(r)}\right)\\ &\le \exp\left(-(2+\delta)\phi(r)/2\right)\\ &\le \exp(-\phi(r)),\end{aligned}$$

так что мы в конце концов приходим к (13.3).

Приступим теперь к оценке вероятностей. Для любых положительных λ, r и любого $x \in \mathcal{X}$ в силу (13.2) имеем оценку

$$\begin{aligned} \mathbb{P}\left(\min_{j\le e^\lambda} ||Y_j - x|| \ge 2r\right) &= \mathbb{P}\left(||Y - x|| \ge 2r\right)^{[e^\lambda]} \\ &= \left(1 - \mathbb{P}\left(||Y - x|| \le 2r\right)\right)^{[e^\lambda]} \\ &\le \exp\left\{-\mathbb{P}\left(||Y - x|| \le 2r\right)[e^\lambda]\right\} \\ &\le \exp\left\{-\exp(-\phi(r) - \nu(x,r)^2/2)[e^\lambda]\right\}. \end{aligned}$$

Воспользуемся неравенством (13.3) и получим

$$\begin{aligned} \mathbb{P}\left(M_\lambda \ge 2r\right) \le\ & \exp\left\{-\exp(-\phi(r) - 2(2+\delta)\phi(r))[e^\lambda]\right\} \\ & + \exp(-\phi(r)). \end{aligned}$$

Полагая в этом неравенстве $\delta = 1/3$ и $r = \phi^{-1}(\lambda/6)$, приходим к

$$\mathbb{P}\left(M_\lambda \ge 2\phi^{-1}(\lambda/6)\right) \to 0 \qquad \text{при } \lambda \to \infty. \qquad \square$$

Еще более тесной оказывается связь между ошибкой квантования и вероятностями малых уклонений *со случайными центрами*, см. [78, 80, 81]. Последнее понятие определяется следующим образом. Пусть X – гауссовский вектор с распределением P. Определим функцию малых уклонений со случайными центрами формулой

$$\psi(\omega, r) = -\ln P\{x : ||x - X(\omega)|| \le r\}.$$

Иными словами, мы выбираем случайный центр шара по распределению P, а затем измеряем уклонения относительно этого центра. По неравенству Андерсона верно

$$P\{x : ||x - X(\omega)|| \le r\} \le P\{x : ||x|| \le r\}.$$

Поэтому

$$\psi(\omega, r) \ge \phi(r).$$

С другой стороны, можно показать, что при почти всех ω верно

$$\lim_{r\to 0} \frac{\psi(\omega, r)}{2\phi(r/2)} \le 1.$$

Можно доказать, что при выполнении естественного условия регулярности $\phi(r/2) \le C\phi(r)$ для вероятностей малых уклонений со случайными центрами имеется детерминированный эквивалент, т.е. существует такая невозрастающая функция $\phi_*(r)$, что

$$\lim_{r\to 0} \frac{\psi(\omega, r)}{\phi_*(r)} = 1 \qquad \text{по вероятности.}$$

Из сказанного ясно, что функции $\phi(r)$ и $\phi_*(r)$, как правило, имеют одинаковый порядок роста при $r \searrow 0$. Тем не менее, точное соотношение между ними неизвестно. Даже для винеровского процесса в пространстве $\mathbb{C}[0,1]$ мы не знаем постоянную a в формуле $\phi_*(r) \sim a\, r^{-2}$. Известно лишь, что такое a существует и $\frac{\pi^2}{4} \le a \le \pi^2$.

Ключевая роль функции $\phi_*(r)$ в теории квантования гауссовских векторов видна из следующей теоремы [80].

Теорема 13.3 *Предположим, что функция ϕ_* удовлетворяет условию* (13.1)*. Тогда для любого $p > 0$ верно*

$$\lim_{\lambda\to\infty} \frac{D(\lambda, p)}{\phi_*^{-1}(\lambda)} = 1.$$

Заинтересованный читатель найдёт много других интересных результатов и алгоритмов квантования гауссовских процессов в работах З. Графа, Г. Лушги и Ж. Пажеса, см. [101], [141]–[145] и др.

Глава 14

Что читать дальше

Этот краткий курс никоим образом не претендует на то, чтобы дать полное представление об огромном поле исследований в чистой и прикладной математике, связанных с гауссовскими процессами, но он должен помочь войти в этот интересный мир.

Среди теоретических объектов для дальнейшего изучения в первую очередь стоит упомянуть мажорирующие меры и усовершенствованный цепной метод[1], мощный инструмент, разработанный Ферником и Талаграном для описания условий ограниченности и непрерывности траекторий гауссовских процессов. Эта техника впечатляет своей способностью разрешать критически сложные случаи. Основанная на неявных характеристиках, она долго считалась трудной и сложной для понимания, но в конечном счёте были найдены более доступные способы изложения. На эту тему есть обширная литература „из первых рук“ [92, 93], [171, 176, 177]. См. также [19] и особенно [124, Глава 6].

Большая и важная литература посвящена исследованию функционалов от гауссовских процессов, в том числе дифференциальному исчислению (исчисление Маллявэна) [4, 5], [13], и полиномиальным разложениям (винеровский хаос), см. [126], [124, Глава 5], [155].

[1] Majorizing measures, generic chaining.

Монографии Адлера и Тейлора [42] и Кошневизана [113] можно порекомендовать тем, кого интересуют геометрия и другие свойства случайных полей.

О разнообразных новейших приложениях гауссовских процессов можно узнать из книг и обзорных статей Мандьеса [149], Виллинджера и др. [183] (модели коммуникационных сетей), Расмуссена и Вильямса [160] (машинное обучение), ван дер Ваарта и ван Зантена [179] (модели априорных распределений в байесовской статистике).

Среди изданий на близкую тему на русском языке укажем монографии Богачёва [4], Го [11], Лифшица [19], Розанова [25], а среди зарубежной литературы – курсы лекций Адлера [41], Леду [124] и монографию Ферника [93]. В более специальной литературе близкие темы затрагиваются в книгах Адлера и Тейлора [42], Ибрагимова и Розанова [14], Хида и Хитсуда [105], Янсона [109].

Предметный указатель

Литература

[1] Аурзада Ф., Ибрагимов И. А., Лифшиц М.А., ван Зантен Х. Малые уклонения гладких стационарных гауссовских процессов. Теор. вероятн. и её примен., **53** (2008), 788–798.

[2] Бегин Л., Никитин Я. Ю., Орсингер Е. Точные константы в оценках вероятностей малых L_2-шаров для некоторых гауссовских процессов. Записки научн. семин. ПОМИ, **298** (2003), 5–21.

[3] Беляев Ю. К. Локальные свойства выборочных функций гауссовских стационарных процессов. Теор. вероятн. и её примен., **5** (1960), 128–131.

[4] Богачёв В. И. Гауссовские меры. Наука-Физматлит, Москва, 1997.

[5] Богачёв В. И. Дифференцируемые меры и исчисление Маллявэна. НИЦ Регулярная и Хаотическая Динамика, Москва – Ижевск, 2008.

[6] Боровков А. А., Могульский А. А. О вероятностях малых уклонений для случайных процессов. Труды Института Математики СО АН СССР, **13** (1989), 147–168.

[7] Булинский А. В., Ширяев А. Н. Теория случайных процессов. Физматлит, Москва, 2003.

[8] Бураго Ю. Д., Залгаллер В. А. Геометрические неравенства. Наука, Москва, 1980.

[9] Вентцель А. Д. Теоремы, касающиеся функционала действия для гауссовских случайных функций. Теор. вероятн. и её примен., **17** (1972), 542–544.

[10] Вентцель А. Д. Курс теории случайных процессов. Наука, Москва, 1975.

[11] Го, Х.-С. Гауссовские меры в банаховых пространствах. Мир, Москва, 1979.

[12] Голосов Ю. Н., Молчан Г. М. Гауссовские стационарные процессы с асимптотически степенным спектром. Доклады АН СССР, **184** (1969), 546–549.

[13] Давыдов Ю.А., Лифшиц М.А., Смородина Н.В. Распределения стохастических функционалов. Наука, Москва, 1995.

[14] Ибрагимов И. А., Розанов Ю. А. Гауссовские случайные процессы. Наука, Москва, 1970.

[15] Колмогоров А. Н., Тихомиров В. М. ε-энтропия и ε-ёмкость множеств в функциональных пространствах. Успехи матем. наук, **14** (1959), 3–86; см. также в кн.: Колмогоров А. Н. Теория информации и теория алгоритмов. Наука, Москва, 1987, 119–198.

[16] Леви П. Стохастические процессы и броуновское движение. Наука, Москва, 1972.

[17] Линде В., Пич А. Отображения гауссовских циллиндрических мер в банаховых пространствах. Теор. вероятн. и её примен., **19** (1974), 472–487.

[18] Лифшиц М. А. Библиография по вероятностям малых уклонений. `www.proba.jussieu.fr/pageperso/smalldev/biblio.html`

[19] Лифшиц М. А. Гауссовские случайные функции. Издательство ТВиМС, Киев, 1995.

[20] Лифшиц М. А. О представлении полей Леви индикаторами. Теор. вероятн. и её примен., **29** (1979), 624–628.

[21] Назаров А. И. О точной константе в асимптотике малых уклонений в L_2-норме некоторых гауссовских процессов. В сб.: Нелинейные уравнения и математический

анализ (сер. Проблемы математического анализа, вып. 26 (2003)), 179–214.

[22] Назаров А. И., Никитин Я. Ю. Логарифмическая асимптотика малых шаров в L_2-норме для некоторых дробных гауссовских процессов. Теор. вероятн. и её примен., **49** (2004), 695–711.

[23] Назаров А. И., Пусев Р. С. Точная асимптотика малых уклонений в L_2-норме для некоторых взвешенных гауссовских процессов. Записки научн. семин. ПОМИ, **364** (2009), 166–199.

[24] Петров В. В. Предельные теоремы для сумм независимых случайных величин. Наука, Москва, 1987.

[25] Розанов Ю. А. Гауссовские бесконечномерные распределения. Труды математического института им. В.А. Стеклова, **108** (1968), Наука, Москва.

[26] Саханенко А. И. Скорость сходимости в принципе инвариантности для разнораспределённых случайных величин с экспоненциальными моментами. Труды института математики СО АН СССР, **3** (1984), 4–49.

[27] Скороход А. В. Интегрирование в гильбертовом пространстве. Наука, Москва, 1975.

[28] Судаков В. Н. Меры Гаусса, Коши и ϵ-энтропия. Доклады АН СССР, **185** (1969), 51–53.

[29] Судаков В. Н. Гауссовские случайные процессы и меры телесных углов в гильбертовом пространстве, Доклады АН СССР, **197** (1971), 43–45.

[30] Судаков В. Н. Геометрические проблемы теории бесконечномерных вероятностных распределений. Труды математического института АН СССР им. В.А. Стеклова, **141** (1976), Наука, Ленинград.

[31] Судаков В. Н., Цирельсон Б. С. Экстремальные свойства полупространств для сферически инвариантных мер. Записки научн. семин. ЛОМИ, **41** (1974), 14–41.

[32] Темляков В. Н. Оценки асимптотических характеристик классов функций с ограниченной смешанной производ-

ной или разностью. Труды математического института АН СССР им В. А. Стеклова, **189** (1989), 138–167.

[33] Фаталов В. Р. Константы в асимптотиках вероятностей малых уклонений для гауссовских процессов и полей. Успехи матем. наук, **58** (2003), 89–134.

[34] Фаталов В. Р. Времена пребывания и точные асимптотики малых уклонений бесселевских процессов для L_p-норм, $p > 0$. Известия РАН. Сер. матем., **71**, вып. 4 (2007), 69–102.

[35] Фаталов В. Р. Точная асимптотика малых уклонений для нестационарного процесса Орнштейна–Уленбека в L_p-норме, $p \geq 2$. Вестник МГУ. Сер. матем. мех., **62** (2007), 3–8.

[36] Фаталов В.Р. Точные асимптотики малых уклонений для стационарного процесса Орнштейна–Уленбека и некоторых гауссовских диффузий в L_p-норме, $2 \leq p \leq \infty$. Проблемы передачи информации, **44**, вып. 2 (2008), 75–95.

[37] Фаталов В. Р. Малые уклонения для двух классов гауссовских стационарных процессов и L_p-функционалов, $0 < p \leq \infty$. Проблемы передачи информации, **46** (2010), 68–93.

[38] Фрейдлин М. И. Функционал действия для одного класса случайных процессов. Теор. вероятн. и её примен., **17** (1972), 536–541.

[39] Хинчин А. Я. Асимптотические законы теории вероятностей. ОНТИ, Москва – Ленинград, 1936.

[40] Ченцов Н. Н. Многопараметрическое броуновское движение и обобщённый белый шум. Теор. вероятн. и её примен., **2** (1957), 281–282.

[41] Adler R. J. An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes. Lect. Notes Inst. Math. Stat., **12**, IMS, Hayword, 1990.

[42] Adler R. J., Tailor J. E. Random Fields and Their Geometry. Springer, New York, 2007.

[43] Anderson T. W. The integral of symmetric unimodal function. Proc. Amer. Math. Soc., **6** (1955), 170–176.

[44] Anderson T. W., Darling D. A. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Ann. Math. Stat., **23** (1952), 193–212.

[45] Aronszajn N. Theory of reproducing kernels. Trans. Amer. Math. Soc., **68** (1950), 337–404.

[46] Artstein S., Milman V. D., Szarek S. J. Duality of metric entropy. Ann. Math., **159** (2004), 1313–1328.

[47] Aurzada F. Lower tail probabilities of some random sequences in l_p. J. Theor. Probab., **20** (2007), 843–858.

[48] Aurzada F., Lifshits M. A. Small deviations of Gaussian processes via chaining. Stoch. Proc. Appl., **118** (2008), 2344–2368.

[49] Baldi P., Ben Arous G., Kerkacharian G. Large deviations and Strassen theorem in Hölder norm. Stoch. Proc. Appl., **42** (1992), 171–180.

[50] Bardina X., Es-Sebaiy K. An extension of bifractional Brownian motion. Commun. Stochast. Analysis, **5** (2011), 333–340.

[51] Barthe F. The Brunn–Minkowskii theorem and related geometric and functional inequalities, in: "Proc. Intern. Congr. Math. Madrid, 2006, **II**", Eur. Math. Soc., Zürich, 2006, 1529–1546.

[52] Barthe F., Huet N. On Gaussian Brunn–Minkowskii inequalities. Studia Math., **191** (2009), 283–304.

[53] Bass R. F., Pyke R. Functional law of the iterated logarithm and uniform central limit theorem for processes indexed by sets. Ann. Probab., **12** (1984), 13–34.

[54] Belinsky E. S. Estimates of entropy numbers and Gaussian measures for classes of functions with bounded mixed derivative. J. Approx. Theory, **93** (1998), 114–127.

[55] Bilyk D., Lacey M., Vagarshakyan A. On the small ball inequality in all dimensions. J. Funct. Anal., **254** (2008), 2470–2502.

[56] Bingham N. H. Variants on the law of the iterated logarithm. Bull. Lond. Math. Soc., **18** (1986), 433–467.

[57] Bobkov S. G. Extremal properties of half-spaces for log-concave distributions. Ann. Probab., **24** (1996), 35–48.

[58] Bobkov S. G. A functional form of isoperimetric inequality for the Gaussian measure. J. Funct. Anal., **135** (1996), 39–49.

[59] Borell C. The Brunn–Minkowski inequality in Gauss space. Invent. Math., **30** (1975), 207–216.

[60] Borell C. Convex measures on locally convex spaces. Ark. Mat., **12** (1974), 239–252.

[61] Borell C. Gaussian Radon measures on locally convex spaces. Math. Scand., **38** (1976), 265–284.

[62] Borell C. A note on Gaussian measures which agree on small balls. Ann. Inst. H. Poincaré, Ser. B, **13** (1977), 231–238.

[63] Borell C. The Ehrhard inequality. Comptes Rendus Math. Acad. Sci. Paris, **337** (2003), 663–666.

[64] Borell C. Inequalities of the Brunn–Minkowski type for Gaussian measure. Probab. Theory Rel. Fields, **140** (2008), 195–205.

[65] Borovkov A. A., Ruzankin P. S. On small deviations of series of weighted random variables. J. Theor. Probab., **21** (2008), 628–649.

[66] Cameron R. H., Martin W. T. Transformations of Wiener integrals under translations. Ann. Math., **45** (1944), 386–396.

[67] Cameron R. H., Martin W. T. The Wiener measure of Hilbert neighborhoods in the space of real continuous functions. J. Math. Phys., **23** (1944), 195–209.

[68] Chung K. L. On maximum of partial sums of sequences of independent random variables. Trans. Amer. Math. Soc., **64** (1948), 205–233.

[69] Chernoff H. A measure of asymptotic efficiency for tests of hypothesis based on sums of observations. Ann. Math. Stat., **23** (1952), 493–507.

[70] Ciesielski Z., Kerkyacharian G., Roynette B. Quelques espaces fonctionnels associés à des processus gaussiens. Studia Math., **107** (1993), 173–203.

[71] Cordero-Erausquin D., Fradelizi M., Maurey B. The B-conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. J. Funct. Anal., **44** (2004), 410–427.

[72] Cramér H. Sur un nouveau théorème-limite de la théorie de probabilités. Act. Sci. Indust., **736** (1938), 5–23.

[73] Csáki E. A relation between Chung's and Strassen's law of the iterated logarithm. Z. Wahrsch. verw. Geb., **54** (1980), 287–301.

[74] de Acosta A. Small deviations in the functional central limit theorem with applications to functional laws of the iterated logarithm. Ann. Probab., **11** (1983), 78–101.

[75] Deheuvels P., Lifshits M. A. Strassen-type functional laws for strong topologies. Probab. Theory Rel. Fields, **97** (1993), 151–167.

[76] Deheuvels P., Lifshits M. A. Necessary and sufficient condition for the Strassen law of the iterated logarithm in non-uniform topologies. Ann. Probab., **22** (1994), 1838–1856.

[77] Dembo A., Zeitouni O. Large Deviation Techniques and Applications (2-nd edition). Springer, New York, 1998.

[78] Dereich S. Small ball probabilities around random centers of Gaussian measures and applications to quantization. J. Theor. Probab., **16** (2003), 427–449.

[79] Dereich S., Fehringer F., Matoussi A., Scheutzow M. On the link between small ball probabilities and the quantization problem for Gaussian measures on Banach spaces. J. Theor. Probab., **16** (2003), 249–265.

[80] Dereich S., Lifshits M. A. Probabilities of randomly centered small balls and quantization in Banach spaces. Ann. Probab., **33** (2005), 1397–1421.

[81] Dereich S., Scheutzow M. High-resolution quantization and entropy coding for fractional Brownian motion. Electron. J. Probab., **11** (2006), No. 28, 700–722.

[82] Dudley R. M. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. J. Funct. Anal., **1** (1967), 290–330.

[83] Dudley R. M. Sample functions of Gaussian processes. Ann. Probab., **1** (1973), 66–103.

[84] Dudley R. M., Feldman J., Le Cam, L. On seminorms, and probabilities, and abstract Wiener spaces. Ann. Math., **93** (1971), 390–408.

[85] Dunker T., Kühn T., Lifshits M. A., Linde W. Metric entropy of integration operators and small ball probabilities for the Brownian sheet. J. Approx. Theory, **101** (1999), 63–77.

[86] Dunker T., Lifshits M. A., Linde W. Small deviations of sums of independent variables, in: ”High Dimensional Probability”, Ser. Progress in Probability, **43**, Birkhäuser, Basel, 1998, 59–74.

[87] Dzhaparidze K., van Zanten J. H. Krein's spectral theory and the Paley–Wiener expansion for fractional Brownian motion. Ann. Probab., **33** (2005), 620–644.

[88] Dzhaparidze K., van Zanten J. H., Zareba P. Representations of isotropic Gaussian random fields with homogeneous increments. J. Appl. Math. Stoch. Anal., **7731** (2006), 1–25.

[89] Ehrhard A. Symetrisation dans l'espace de Gauss. Math. Scand., **53** (1983), 281–301.

[90] Fehringer, F. Kodierung von Gaußmassen. Ph.D. Thesis. TU Berlin, Berlin, 2001.

[91] Feller W. The general form of the so-called law of the iterated logarithm. Trans. Amer. Math. Soc., **54** (1943), 373–402.

[92] Fernique X. Régularité des trajectoires des fonctions aléatoires gaussiennes, in: ”Ecole d'Été de Probabilités de

Saint-Flour, IV-1974". Lect. Notes Math., **480**, Springer, Berlin, 1975, 1–96.

[93] Fernique X. Fonctions Aléatoires Gaussiennes. Vecteurs Aléatoirs Gaussiens. CRM, Montreal, 1997.

[94] Fill J. A., Torcaso F. Asymptotic analysis via Mellin transforms for small deviations in L_2-norm of integrated Brownian sheets. Probab. Theory Rel. Fields, **130** (2003), 259–288.

[95] Gao F., Hannig J., Lee T.-Y., Torcaso F. Laplace transforms via Hadamard factorization with applications to small ball probabilities. Electron. J. Probab., **8** (2003), No. 13, 1–20.

[96] Gao F., Hannig J., Torcaso F. Integrated Brownian motions and exact l_2-small balls, Ann. Probab., **31** (2003), 1320–1337.

[97] Gao F., Hannig J., Lee T.-Y., Torcaso F. Exact L^2-small balls of Gaussian processes. J. Theor. Probab., **17** (2004), 503–520.

[98] Gao F., Li W. V. Small ball probabilities for the Slepian Gaussian fields. Trans. Amer. Math. Soc., **359** (2005), 1339–1350.

[99] Gardner R. J. The Brunn–Minkowskii inequality. Bull. Amer. Math. Soc., **39** (2002), 355–405.

[100] Gardner R. J., Zvavitch A. Gaussian Brunn–Minkowskii inequalities. Trans. Amer. Math. Soc., **362** (2010), 5333–5353.

[101] Graf S., Luschgy H. Foundations of Quantization for Probability Distributions. Lect. Notes Math., **1730**, Springer, Berlin, 2000.

[102] Grill K. Exact rate of convergence in Strassen's law of iterated logarithm. J. Theor. Probab., **5** (1992), 197–205.

[103] Hargé G. A particular case of correlation inequality for the Gaussian measure. Ann. Probab., **27** (1999), 1939–1951.

[104] Hartman P., Wintner A. On the law of the iterated logarithm. Amer. J. Math., **63** (1941), 169–176.

[105] Hida T., Hitsuda M. Gaussian Processes. AMS, Providence, 1993.

[106] Houdré C., Villa J. An example of infinite dimensional quasi-helix, in: "Contemporary Mathematics", **336**, AMS, Providence, 2003, 195–201.

[107] Jain N. C., Pruitt W. E. Maximum of partial sums of independent random variables. Z. Wahrsch. verw. Geb., **27** (1973), 141–151.

[108] Jain N. C., Pruitt W. E. The other law of the iterated logarithm. Ann. Probab., **3** (1975), 1046–1049.

[109] Janson S. Gaussian Hilbert Spaces. Cambridge University Press, Cambridge, 1997.

[110] Karol' A., Nazarov A., Nikitin Ya. Small ball probabilities for Gaussian random fields and tenzor products of compact operators. Trans. Amer. Math. Soc., **360** (2008), 1443–1474.

[111] Khatri C. G. On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. Ann. Math. Stat., **38** (1967), 1853–1867.

[112] Khintchine A. Über einen Satz der Wahrscheinlichkeitsrechnung. Fund. Math., **6** (1924), 9–12.

[113] Khoshnevisan D. Multiparameter Processes: an Introduction to Random Fields. Springer, New York, 2002.

[114] Kolmogorov A. N. Sulla determinazione empirica di una legge di distribuzione. G. Inst. Ital. Attuari, **4** (1933), 83–91.

[115] Komlós J., Major P., Tusnády G. An approximation of partial sums of independent RV'-s and the sample DF. I. Z. Wahrsch. verw. Geb., **32** (1975), 111–131.

[116] Komlós J., Major P., Tusnády G. An approximation of partial sums of independent RV'-s and the sample DF. II. Z. Wahrsch. verw. Geb., **34** (1976), 34–58.

[117] Kuelbs J., Li W. V. Some shift inequalities for Gaussian measures, in: "High Dimensional Probability. Proceedings of the conference, Oberwolfach, Germany, August 1996". Ser.

Progress in Probability, **43**. Birkhäuser, Basel, 1998, 233–243.

[118] Kuelbs J., Li W. V. Metric entropy and the small ball problem for Gaussian measures. J. Funct. Anal., **116** (1993), 133–157.

[119] Kühn T., Linde W. Optimal series representation of fractional Brownian sheet. Bernoulli, **8** (2002), 669–696.

[120] Kwapień S., Sawa J. On some conjecture concerning Gaussian measures of dilatations of convex symmetric sets. Studia Math., **105** (1993), 173–187.

[121] Latała R. A note on Ehrhard inequality. Studia Math., **118** (1996), 169–174.

[122] Latała R. On some inequalities for Gaussian measures, in: "Proc. Intern. Congr. Math. Beijing, **II**". Higher Education Press, 2002, 813–821.

[123] Latała R., Oleszkiewitcz K. Gaussian measures of dilatations of convex symmetric sets. Ann. Probab., **27** (1999), 1922–1938.

[124] Ledoux M. Isoperimetry and Gaussian Analysis. Lect. Notes Math., **1648**, Springer, Berlin, 1996, 165–294.

[125] Ledoux M. Concentration of Measure Phenomenon. Ser. Math. Surveys and Monographs, **89** AMS, Providence, 2001.

[126] Ledoux M., Talagrand M. Probability in Banach Spaces, Springer, Berlin, 1991.

[127] Lei P., Nualart D. A decomposition of the bifractional Brownian motion and some applications, Statist. Probab. Letters, **79** (2009), 619–624.

[128] Lévy P. Problèmes Concrets d'Analyse Fonctionnelle. Gautier-Villars, Paris, 1951.

[129] Li W. V. A Gaussian correlation inequality and its applications to small ball probabilities. Electron. Commun. Probab., **4** (1999), 111–118.

[130] Li W. V. Small deviations for Gaussian Markov processes under the sup-norm. J. Theor. Probab., **12** (1999), 971–984.

[131] Li W. V. Small ball probabilities for Gaussian Markov processes under the L_p-norm. Stoch. Proc. Appl., **92** (2001), 87–102.

[132] Li W. V., Linde W. Existence of small ball constants for fractional Brownian motions, Comptes Rendus Math. Acad. Sci. Paris, **326** (1998), 1329–1334.

[133] Li W. V., Linde W. Approximation, metric entropy and small ball estimates for Gaussian measures. Ann. Probab., **27** (1999), 1556–1578.

[134] Li W. V., Shao Q.-M. Gaussian processes: inequalities, small ball probabilities and applications, in: ”Stochastic Processes: Theory and Methods, Handbook of Statistics, **19**”. North-Holland/Elsevier, Amsterdam, 2001, 533–597.

[135] Lifshits M. A. On the lower tail probabilities of some random series. Ann. Probab., **25** (1997), 424–442.

[136] Lifshits M. A. Asymptotic behavior of small ball probabilities, in: ”Probab. Theory and Math. Statist. Proc. VII International Vilnius Conference”, VSP/TEV, Vilnius, 1999, 453–468.

[137] Lifshits M. A., Linde W. Approximation and entropy numbers of Volterra operators with application to Brownian Motion. Mem. Amer. Math. Soc., **157**, No. 745 (2002), 1–87.

[138] Lifshits M. A., Linde W., Shi Z. Small deviations of Riemann–Liouville processes in L_q-norms with respect to fractal measures. Proc. Lond. Math. Soc., **92** (2006), 224–250.

[139] Lifshits M. A., Linde W., Shi Z. Small deviations of Gaussian random fields in L_q-spaces. Electron. J. Probab., **11**, No. 46 (2006), 1204–1223.

[140] Lifshits M. A., Simon T. Small deviations for fractional stable processes. Ann. Inst. H. Poincaré, Ser. B, **41** (2005), 725–752.

[141] Luschgy H., Pagès G. Functional quantization of Gaussian processes. J. Funct. Anal., **196** (2002), 486–531.

[142] Luschgy H., Pagès G. Sharp asymptotics of the functional quantization problem for Gaussian processes. Ann. Probab., **32** (2004), 1574–1599.

[143] Luschgy H., Pagès G. High-resolution product quantization for Gaussian processes under sup-norm distortion. Bernoulli, **13** (2007), 653–671.

[144] Luschgy H., Pagès G. Expansion for Gaussian processes and Parseval schemes. Electron. J. Probab., **14** (2009), 1198–1221.

[145] Luschgy H., Pagès G., Wilbertz B. Asymptotically optimal quantization schemes for Gaussian processes. ESAIM: Probab. Stat., **14** (2010), 93–116.

[146] Major P. An improvement of Strassen's invariance principle. Ann. Probab., **7** (1979), 55–61.

[147] Malyarenko A. An optimal series expansion of the multiparameter fractional Brownian motion. J. Theor. Probab., **21** (2008), 459–475.

[148] Mandelbrot B.B.,van Ness J. Fractional Brownian motions, fractional noises and applications. SIAM Rev., **10** (1968), 422–437.

[149] Mandjes M. Large Deviations for Gaussian Queues: Modelling Communication Networks. Wiley, Chichester, 2007.

[150] Marinucci D., Robinson P. M. Alternative forms of fractional Brownian motion. J. Stat. Plan. Infer., **80** (1999), 111–122.

[151] Nazarov A. I. Exact L_2-small ball asymptotics of Gaussian processes and the spectrum of boundary-value problems. J. Theor. Probab., **22** (2009), 640–665.

[152] Nazarov A. I., Nikitin Ya. Yu. Exact L_2-small ball behavior of integrated Gaussian processes and spectral asymptotics of boundary value problems. Probab. Theory Rel. Fields, **129** (2004), 469–494.

[153] Paley R. E. A. C., Wiener N. Fourier Transforms in the Complex Domain. Amer. Math. Soc. Colloq. Publ., **19**, AMS, New York, 1934. Reprint: AMS, Providence, 1987.

[154] Park W. J. On Strassen version of the law of the iterated logarithm for two-parameter Gaussian processes. J. Multivariate Anal., **4** (1974), 479–485.

[155] Peccati G., Taqqu M. S. Wiener Chaos: Moments, Cumulants and Diagrams. Springer, Milano, 2011.

[156] Petrovskii I. G. Über das Irrfahrtproblem. Math. Ann., **109** (1934), 425–444.

[157] Pisier G. Conditions d'entropie assurant la continuité de certains processus et applications à l'analyse harmonique. Seminaires d'analyse fonctionnelle, Exp. 13–14 (1980), 1–43.

[158] Pisier G. The Volume of Convex Bodies and Banach Space Geometry. Cambridge University Press, Cambridge, 1989.

[159] Pitt L. A Gaussian correlation inequality for symmetric convex sets. Ann. Probab., **5** (1977), 470–474.

[160] Rasmussen C. E., Williams C. K. I. Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA, 2006.

[161] Sato H. Souslin support and Fourier expansion of a Gaussian Radon measure, in: "Probability in Banach Spaces III. Proc. Intern. Conf., Medford, USA, 1980". Lect. Notes Math., **860**, Springer, Berlin, 1981, 299–313.

[162] Schechtman G., Schlumprecht T., Zinn J. On the Gaussian measure of intersection. Ann. Probab., **26** (1998), 346–357.

[163] Schilder M. Some asymptotic formulas for Wiener integrals. Trans. Amer. Math. Soc., **125** (1966), 63–85.

[164] Šidák Z. Rectangular confidence regions for the means of multivariate normal distributions. J. Amer. Statist. Assoc., **62** (1967), 626–633.

[165] Steiner J. Einfacher Beweis der isoperimetrischen Hauptsätze. J. Reine Angew. Math., **18** (1838), 281–296.

[166] Stolz W. Une méthode élémentaire pour l'évaluation des petites boules browniennes. Comptes Rendus Math. Acad. Sci. Paris, **316** (1993), 1217–1220.

[167] Stolz W. Small ball probabilities for Gaussian processes under non-uniform norms. J. Theor. Probab., **9** (1996), 613–630.

[168] Strassen V. An invariance principle for the law of the iterated logarithm. Z. Wahrsch. verv. Geb., **3** (1964), 211–226.

[169] Talagrand M. Regularity of Gaussian processes. Acta Math., **159** (1987), 99–149.

[170] Talagrand M. On the rate of clustering in Strassen's law of the iterated logarithm for Brownian motion, in: "Probability in Banach space. VIII". Ser. Progress in Probability, **30**. Birkhäuser, Basel, 1992, 339–347.

[171] Talagrand M. Simple proof of the majorizing measure theorem. Geom. Funct. Anal., **2** (1992), 118–125.

[172] Talagrand M. New Gaussian estimates for enlarged balls. Geom. Funct. Anal., **3** (1993), 502–526.

[173] Talagrand M. The small ball problem for the Brownian sheet. Ann. Probab., **22** (1994), 1331–1354.

[174] Talagrand M. Concentration of measure and isoperimetric inequalities in product spaces. Publ. Math. de l'I.H.E.S., **81** (1995), 73–205.

[175] Talagrand M. A new look at independence. Ann. Probab., **24** (1996), 1–34.

[176] Talagrand M. Majorizing measures: the generic chaining. Ann. Probab., **24** (1996), 1049–1103.

[177] Talagrand M. The Generic Chaining: Upper and Lower Bounds for Stochastic Processes, Springer, Berlin, 2005.

[178] van der Vaart A. W., van Zanten J. H. Rates of contraction of posterior distributions based on Gaussian process priors. Ann. Statist., **36** (2008), 1435–1463.

[179] van der Vaart A. W., van Zanten J. H. Bayesian inference with rescaled Gaussian process priors. Electron. J. Statist., **1** (2007), 433–448.

[180] Varadhan S. R. S. Asymptotic probabilities and differential equations. Comm. Pure Appl. Math., **19** (1966), 261–286.

[181] Weber M. Entropie Métrique et Convergence Presque Partout. Ser. Travaux en Cours. **58**, Hermann, Paris, 1998.

[182] Weber M. Dynamical Systems and Processes. Ser. IRMA Lect. in Math. and Theor. Phys., **14**, EMS, Zürich, 2009.

[183] Willinger W., Paxson V., Riedli R. H., Taqqu M. S . Long range dependence and data network traffic, in: ”Theory and Applications of Long Range Dependence”, Birkhäuser, Basel, 2003, 373–408.

Михаил Анатольевич ЛИФШИЦ

ЛЕКЦИИ ПО ГАУССОВСКИМ ПРОЦЕССАМ

Учебное пособие

Зав. редакцией физико-математической
литературы *Н. Р. Крамор*
Выпускающие *Т. С. Симонова, Н. А. Крылова*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
196105, Санкт-Петербург, пр. Ю. Гагарина, д. 1, лит. А.
Тел./факс: (812) 336-25-09, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 12.11.15.
Бумага офсетная. Гарнитура Школьная. Формат 84×108 $^{1}/_{32}$.
Печать офсетная. Усл. п. л. 10,08. Тираж 100 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленного оригинал-макета
в ПАО «Т8 Издательские Технологии».
109316, г. Москва, Волгоградский пр., д. 42, к. 5.