

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ

На правах рукописи



Палионная Софья Игоревна

**Асимптотические свойства оценок риска в задачах
множественной проверки гипотез**

1.1.4 — теория вероятностей и математическая статистика

ДИССЕРТАЦИЯ

на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:

доктор физико-математических наук, доцент

Шестаков Олег Владимирович

Москва — 2023

Оглавление

	Стр.
Введение	3
Глава 1. Сведение задачи множественной проверки гипотез к задаче пороговой обработки	10
1.1 Постановка задачи множественной проверки гипотез	10
1.2 Пороговая обработка	13
1.3 Оценка среднеквадратичного риска при пороговой обработке . . .	17
1.4 Сильная состоятельность оценки риска	24
1.5 Асимптотическая нормальность оценки риска	29
1.6 Оценка скорости сходимости распределения оценки риска	37
Глава 2. Обращение линейных однородных операторов	43
2.1 Линейные однородные операторы	43
2.2 Вейвлет-разложение и вейвлет-коэффициенты	45
2.3 Вейвлет-разложение в случае обращения линейного однородного оператора	51
2.4 Пороговая обработки и оценка среднеквадратичного риска в случае обращения линейных однородных операторов	53
2.5 Сильная состоятельность оценки риска	55
2.6 Асимптотическая нормальность оценки риска	62
2.7 Оценка скорости сходимости	69
Заключение	74
Список литературы	76

Введение

Актуальность темы. В современном мире, где всё большую значимость приобретают задачи обработки больших данных, количество компонент-предикторов в моделях данных может быть очень велико, что существенно затрудняет работу с ними. Поэтому прежде чем перейти к задаче фильтрации шума в такой модели, необходимо преобразовать исходные данные так, чтобы выделить значимые признаки и удалить незначимые. Такое преобразование позволит привести наблюдаемые данные к «экономному» представлению. При этом важную роль при решении данной задачи играет универсальность выбранного метода, т.е. метод должен быть адаптивен к различным входным данным и моделям и в каждом случае давать результат, близкий к оптимальному.

Для того, чтобы осуществить сжатие данных можно прибегнуть к пороговой обработке, которая по сути эквивалентна задаче множественной проверки гипотез для набора коэффициентов разложения исходных данных: если нулевая гипотеза отклоняется, то соответствующая компонента разложения данных зануляется. В случае построения семейства статистических выводов возникает эффект множественных сравнений. Существуют различные методы решения задачи множественной проверки гипотез, устраняющие эффект множественных сравнений [48]. Суть этих методов заключается в контроле меры, обобщающей ошибку первого рода при множественной проверке гипотез. К этим мерам относятся FWER (Family-Wise Error Rate), FDR (False Discovery Rate), pFDR (positive False Discovery Rate), HMP (harmonic mean p-value) (см. работы [15; 59; 64; 65]). В данной диссертации будет использоваться FDR-мера, которая представляет собой долю ложных отклонений нулевых гипотез. Широкое распространение FDR-мера получила в случаях, когда количество проверяемых гипотез настолько велико, что предпочтительнее совершить небольшое количество ошибок первого рода для увеличения статистической мощности.

В рамках задачи пороговой обработки ключевую роль играет выбор порогового значения, который будет гарантировать построение в некотором смысле наилучшей модели, что приводит к необходимости вычислять погрешность модели после применения пороговой обработки. На практике получить значение самой погрешности (или риска) не представляется возможным, т.к. она зависит от неизвестных значений исходных данных. Однако можно рассмотреть несмещенные оценки риска, которые носят название SURE-оценок (Stein Unbiased Risk Estimator). Для вычисления таких оценок требуются только наблюдаемые данные, что позволяет вычислять эти оценки на практике. В работах [24], [26], [28] были предложены способы вычисления порогового значения, направленные на минимизацию среднеквадратичной погрешности (риска). В данной диссертации задача сжатия данных будет сведена к задаче пороговой обработки с использованием FDR-порога, контроль над которым будет осуществляться с помощью алгоритма Бенжамини-Хочберга [15].

Во многих областях знаний после применения пороговой обработки сигнала оказывается, что лишь небольшое количество компонент-предикторов исходной модели значимо отличны от нуля. В частности, такие модели рассматриваются в задачах из различных областей компьютерного зрения, аудио- и видео- обработки данных, обработки электроэнцефалограмм и т.д. Эти наблюдения приводят к необходимости рассматривать в некотором смысле разреженный сигнал исходных данных. В данной диссертации будут рассмотрены несколько вариантов определения разреженности сигнала.

Кроме того, зачастую информация, доступная для наблюдения, представляет собой некоторое преобразование исходных данных. Такие ситуации возникают, например, в астрофизических и томографических приложениях, физике плазмы и др. [39—41]. В этом случае дополнительно возникает задача обращения преобразования. Во второй главе данной диссертации будет рассмотрена постановка задачи в случае, когда исходный вектор данных подвергается действию линейного однородного оператора. При этом существование и ограниченность обратного оператора, вообще говоря, не гарантируется, что делает

невозможным применение обратного оператора к наблюдаемым данным для получения исходного сигнала. Для решения этой задачи будет использоваться аппарат вейвлет-анализа, который представляет собой нелинейный метод обработки сигнала. Выбор вейвлет-анализа обусловлен рядом преимуществ этого метода. Во-первых, вейвлет-разложение приводит исходные данные к разреженному представлению, что является ключевым предположением в постановке задачи диссертации. Во-вторых, вейвлет-анализ, в отличие, например, от анализа Фурье, позволяет эффективнее работать с нестационарными сигналами. Наконец, реализация вейвлет-анализа в различных библиотеках программного обеспечения позволяет получать практические результаты с хорошей вычислительной скоростью.

В частности, в диссертации рассмотрен метод вейвлет-вейвлет-разложения, который был предложен в статье [10]. Этот метод предполагает представление преобразованных оператором исходных данных в виде ряда из сдвигов и растяжений некоторой выбранной вейвлет-функции. Основная идея применения вейвлет-вейвлет-метода заключается в том, чтобы представить исходные данные в виде разложения по вейвлет-базису и выразить исходные данные через коэффициенты разложения их преобразования. Также стоит отметить, что при применении вейвлет-вейвлет-метода преобразованные исходные данные раскладываются в ряд по ортонормированному базису, что сохраняет независимость шумовых коэффициентов разложения. Однако при применении этого метода возникает необходимость вводить корректировку значения среднеквадратичного риска для каждого масштаба разложения, что приводит к необходимости вычисления значения порога на каждом уровне.

Цель работы. Целью данной диссертации является исследование асимптотических свойств оценки среднеквадратичной погрешности множественной проверки гипотез с использованием FDR-порога в задаче оценивания математического ожидания гауссова вектора в случае рассмотрения векторов большой размерности и в случае, когда исходный сигнал данных был преобразован действием линейного однородного оператора.

Научная новизна. Результаты, изложенные в диссертации, являются новыми и заключаются в следующем.

Доказана сильная состоятельность и асимптотическая нормальность оценки риска в задаче множественной проверки гипотез с использованием FDR-порога. Также в диссертации оценивается скорость сходимости распределения данной оценки риска к нормальному закону.

Рассмотрен случай, когда исходный сигнал был подвержен линейному однородному преобразованию. Для такой постановки задачи доказывается сильная состоятельность и асимптотическая нормальность оценки риска, а также оценивается скорость сходимости распределения данной оценки риска к нормальному закону.

Методы исследования. В диссертации используются методы теории вероятностей и математической статистики, а также математического анализа. Основным методом является сведение задачи пороговой обработки к задаче множественной проверки гипотез с FDR-мерой, для контроля над которой используется алгоритм Бенжамини-Хочберга. Рассматриваются случаи использования мягкой и жесткой пороговой обработки. Также рассматривается задача обращения линейного однородного оператора, для решения которой используется аппарат вейвлет-анализа.

Основные положения, выносимые на защиту:

1. Сильная состоятельность оценки риска при множественной проверке гипотез с FDR-порогом.
2. Асимптотическая нормальность оценки риска при множественной проверке гипотез с FDR-порогом.
3. Оценка скорости сходимости распределения оценки риска к нормальному закону с использованием FDR-метода множественной проверки гипотез.
4. Сильная состоятельность оценки риска при множественной проверке гипотез с FDR-порогом при обращении линейных однородных операторов.

5. Асимптотическая нормальность оценки риска при множественной проверке гипотез с FDR-порогом при обращении линейных однородных операторов.
6. Оценка скорости сходимости распределения оценки риска к нормальному закону с использованием FDR-метода множественной проверки гипотез при обращении линейных однородных операторов.

Соответствие паспорту научной специальности. В диссертации исследуются асимптотические свойства оценки риска и скорость сходимости распределения оценки риска в задаче множественной проверки гипотез, в силу чего диссертация соответствует паспорту специальности 1.1.4 «Теория вероятностей и математическая статистика».

Апробация работы. Основные результаты работы докладывались автором на следующих конференциях:

1. Научная конференция «Тихоновские чтения 2020», МГУ имени М.В.Ломоносова, факультет ВМК, Москва, Россия, 26-31 октября 2020 г.

Тема доклада: Свойства оценки риска при множественной проверке гипотез с использованием FDR-метода

2. Ломоносовские чтения 2021. Секция вычислительная математика и кибернетика, Москва, Россия, 20-29 апреля 2021 г.

Тема доклада: Асимптотическое поведение оценки риска FDR-метода в задаче множественной проверки гипотез

3. XXXVI International Seminar on Stability Problems for Stochastic Models, г.Петрозаводск, Россия, 21-25 июня 2021 г.

Тема доклада: Asymptotic behavior of a risk estimate for the FDR-method in the problem of multiple hypothesis testing

4. Ломоносовские чтения - 2022, Секция вычислительная математика и кибернетика, МГУ имени М.В.Ломоносова, факультет ВМК, Москва, Россия, 14-22 апреля 2022 г.

Тема доклада: Свойства оценок риска в задачах обращения линейных

операторов при использовании FDR-метода множественной проверки гипотез

Публикации. Основные результаты по теме диссертации изложены в 5 статьях [66–70] автора. Статья [70] (в соавторстве) опубликована в рецензируемом научном журнале, входящем в базы Scopus и Web of Science, статья [69] (в соавторстве) опубликована в рецензируемом научном журнале, входящем в базу Scopus, статьи [66–68] опубликованы в рецензируемых научных журналах, входящих в базу RSCI. Также работы автора представлены в материалах конференций [71–74]. Список работ автора приведен в конце автореферата и диссертации. В этих работах постановки задач принадлежат О.В. Шестакову, а все основные результаты, приведенные в статьях и диссертации, получены С.И. Палионной самостоятельно.

Объем и структура работы. Диссертация состоит из введения, двух глав, заключения и списка литературы, включающего в себя 74 наименований. Общий объем диссертации составляет 83 страниц. В диссертацию вошли результаты, полученные в рамках программы Московского центра фундаментальной и прикладной математики.

В первой главе рассматривается задача оценивания математического ожидания гауссова вектора в случае, когда наблюдаемые данные зашумлены, а неизвестный вектор исходных данных принадлежит некоторому классу разреженности. Приводятся условия, при которых имеет место сильная состоятельность и асимптотическая нормальность оценки риска. Оценивается скорость сходимости распределения оценки риска к нормальному закону.

Во второй главе рассматривается задача оценивания математического ожидания гауссова вектора в случае, когда наблюдаемые данные зашумлены и представляют собой преобразованные исходные данные, подвергнутые действию линейного однородного оператора, а неизвестный вектор исходных данных принадлежит некоторому классу разреженности. Рассматривается задача обращения линейного однородного оператора, приводятся условия, при которых имеет место сильная состоятельность и асимптотическая нормальность

оценки риска. Оценивается скорость сходимости распределения оценки риска к нормальному закону.

В заключении кратко приведены основные результаты диссертации.

Глава 1. Сведение задачи множественной проверки гипотез к задаче пороговой обработки

1.1 Постановка задачи множественной проверки гипотез

Рассмотрим линейную регрессионную модель, у которой количество компонент-предикторов значительно велико. Классической задачей в таком случае будет являться выделение значимых признаков для модели и удаление из рассмотрения шумовых. Существует множество различных методов для решения проблемы выбора статистической модели, к примеру, скорректированный R^2 [61], информационный критерий Акаике [12], статистика C_p Мэллоуса [46], байесовский информационный критерий BIC [54], информационный критерий Хеннана – Куинна ($2 \log \log n$) [35], сфокусированный информационный критерий FIC [63], RIC (risk inflation criterion) [32].

Для решения проблемы выбора статистической модели в диссертации воспользуемся FDR-мерой множественной проверки гипотез, преимущества которой будут приведены ниже.

Для начала рассмотрим задачу проверки статистических гипотез. Пусть имеется выборка X размера m , с помощью статистики T (функции от выборки) проверяется нулевая гипотеза H_0 против общей альтернативы H_1 . Для статистики T задано некоторое нулевое распределение, т.е. распределение для случая, когда нулевая гипотеза верна. Ошибки, возникающие при применении такого алгоритма, делятся на ошибки первого и второго рода (см. таблица 1).

Ошибка первого рода заключается в том, что отвергается верная нулевая гипотеза. Ситуация, при которой принимается нулевая гипотеза в то время, как она не верна, называется ошибкой второго рода. Ограничим сверху вероятность

Таблица 1 — Типы ошибок при проверке гипотез

	H_0 верна	H_0 неверна
H_0 принимается	H_0 верно принята	ошибка II рода
H_0 отвергается	ошибка I рода	H_0 верно отвергнута

ошибки первого рода на уровне α , т.е.

$$\mathbb{P}(\text{ошибка I рода}) = \mathbb{P}(T \geq t | H_0) \leq \alpha,$$

где t - критическое пороговое значение. Получив область, для которой выполнены ограничения на вероятность ошибки первого рода, можно минимизировать ошибку второго рода, таким образом максимизируя статистическую мощность.

Однако задача существенно усложняется в случае множественной проверки гипотез. Теперь имеется n различных выборок, каждой из которых соответствует своя нулевая гипотеза $\{H_{0_i}, i = 1, \dots, n\}$ и альтернатива $\{H_{1_i}, i = 1, \dots, n\}$. Гипотезы проверяются статистиками T_i с заданными нулевыми распределениями и вычисляются достигаемые уровни значимости $\{p_i, i = 1, \dots, n\}$. На основе получаемых значений p_i принимается решение об отвержении нулевой гипотезы $\{H_{0_i}\}$ для каждого i . Обозначим через M_0 множество индексов верных нулевых гипотез, а через R множество индексов отвергаемых гипотез. Тогда $V = |M_0 \cap R|$ — число ошибок первого рода. Задача заключается в минимизации числа ошибок первого рода за счет изменения параметра R .

Существует множество статистических процедур, предлагающих различные методы решения задачи множественной проверки гипотез, самым простым из которых является метод Бонферрони. Суть этого метода заключается в сравнении достигаемого уровня значимости каждой гипотезы со значением α/n . Однако при увеличении количества проверяемых гипотез мощность метода Бонферрони резко падает, что приводит к необходимости рассматривать

более сложные методы решения проблемы множественных сравнений. В частности, одной из первых была предложена мера, получившая название FWER (Family-Wise Error Rate) [59] и обобщающая ошибку первого рода. Мера FWER представляет собой вероятность совершения хотя бы одной ошибки первого рода, тем самым избавляя от необходимости контролировать ошибки первого рода для каждого теста. Формально FWER определяется следующим образом:

$$\text{FWER} = \mathbb{P}(V \geq 1) \leq \alpha.$$

Однако такое сильное ограничение на ошибки первого рода значительно увеличивает ошибку второго рода особенно в ситуации, когда n велико.

Поэтому в ситуациях, когда количество проверяемых гипотез большое и предпочтительнее допустить некоторое количество ошибок первого рода с целью увеличения статистической мощности критерия, зачастую используют меру FDR (False Discovery Rate), предложенную в работе [15]. FDR мера заключается в контроле ожидаемой доли ложных отклонений, т.е.

$$\text{FDR} = \mathbb{E} \left(\frac{V}{\max(R, 1)} \right).$$

Для контроля FDR-меры чаще всего используется метод Бенжамини–Хочберга, который позволяет ограничить сверху ожидаемую долю ложных отклонений параметром α :

$$\mathbb{E} \left(\frac{V}{\max(R, 1)} \right) \leq \alpha.$$

Для гипотез H_{0_i} уровни значимости α_i меняются линейно и определяются следующим образом:

$$\alpha_i = \frac{\alpha \cdot i}{n}, i = 1, \dots, n.$$

Затем строится вариационный ряд из достигаемых уровней значимости p_i :

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}.$$

По построенному вариационному ряду находится $k \in [1, n]$ — максимальный индекс такой, что для него выполнено условие

$$p_{(i)} \leq \alpha_i,$$

где $p_{(i)}$ — i -й член вариационного ряда из достигаемых уровней значимости. И отвергаются все гипотезы H_{0_1}, \dots, H_{0_k} .

Помимо рассмотренных выше методов решения задачи множественной проверки гипотез в работе [59] рассматривалось q — *value*, обеспечивающее контроль положительной доли ложных отклонений pFDR (positive False Discovery Rate). Контроль показателя FCR (false coverage rate) предлагает решение проблемы множественной проверки гипотез с точки зрения доверительных интервалов [14]. Работы [64; 65] посвящены методу НМР (harmonic mean p-value).

В диссертации будет использован метод FDR, т.к. этот метод хорошо адаптирован к работе с большими объемами данных. Также в работе [60] была доказана связь между FDR-мерой и байесовским подходом. Помимо этого в работе [11] было доказано, что при использовании FDR-меры можно получить асимптотически минимаксные оценки риска для различного вида разреженных данных.

1.2 Пороговая обработка

Рассмотрим задачу оценивания математического ожидания гауссова вектора:

$$X_i = \mu_i + \varepsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

где X_i — наблюдаемые зашумленные данные, ε_i — независимые нормально распределенные случайные величины с нулевым математическим ожиданием и известной дисперсией σ^2 , а $\mu = (\mu_1, \dots, \mu_n)$ — неизвестный вектор истинных данных, принадлежащий некоторому заданному классу “разреженности”. Для

каждого $i \in [1, n]$ рассмотрим нулевую гипотезу H_{0_i} , включающую в себя предположение о распределении X_i , а также о равенстве нулю μ_i .

Предположение о разреженности вектора является ключевым в диссертации и предполагает, что лишь небольшое количество компонент вектора значительно отличается от нуля. Такая постановка задачи возникает в различных областях знаний, где имеет место обработка и анализ сигналов, содержащих шум. При этом разреженное представление сигнала обеспечивается за счет предварительной обработки сигнала, в частности, применения дискретного вейвлет-преобразования.

Рассмотрение в диссертации именно разреженных данных обусловлено практической значимостью таких моделей. В частности, в работе [57] было доказано, что удаление шумовых компонент из сигнала изображений приводит к разреженному вектору данных, т.е. при вейвлет-разложении сигнала многие коэффициенты получаются небольшими по амплитуде и лишь небольшое количество коэффициентов существенно велики. Разреженность данных применяется в различных направлениях компьютерного зрения [17], включая распознавание лиц [52], сегментацию движения и данных [30], [49], моделирование фона [19], [20] и классификацию изображений [21]. Помимо задач компьютерного зрения разреженные данные встречаются в задачах обработки аудио сигнала [13], результатов электроэнцефалографии [55] и др.

В диссертации будут рассмотрены несколько определений “разреженности”. Пусть $\|\mu\|_0 = \#\{\mu_i : \mu_i \neq 0\}$ обозначает число компонент μ , отличных от нуля. Фиксируя η_n , определим класс

$$L_0(\eta_n) = \{\mu \in \mathbb{R}^n : \|\mu\|_0 \leq \eta_n n\}.$$

При малых значениях η_n лишь небольшое число компонент вектора из $L_0(\eta_n)$ отлично от нуля.

Другой возможный способ определения “разреженности” заключается в ограничении абсолютных значений компонент μ . Для этого рассмотрим упоря-

доченные по абсолютной величине значения

$$|\mu|_{(1)} \geq \dots \geq |\mu|_{(n)}$$

и при $0 < p < 2$ определим класс

$$L_p(\eta_n) = \{\mu \in \mathbb{R}^n : |\mu|_{(k)} \leq \eta_n n^{1/p} k^{-1/p} \text{ для всех } k = 1, \dots, n\}.$$

Также “разреженность” можно моделировать с помощью ℓ_p -нормы

$$\|\mu\|_p = \left(\sum_{i=1}^n |\mu_i|^p \right)^{1/p}.$$

В этом случае “разреженный” класс определяется как

$$M_p(\eta_n) = \{\mu \in \mathbb{R}^n : \sum_{i=1}^n |\mu_i|^p \leq \eta_n^p n\}.$$

Между этими классами существует взаимосвязь, отмеченная в работе [11]. Также отметим, что при $p \rightarrow 0$ справедливо $\|\mu\|_p^p \rightarrow \|\mu\|_0$. Помимо этого для рассматриваемых классов разреженности справедливо вложение

$$M_p(\eta_n) \subset L_p(\eta_n) \not\subset M_{p'}(\eta_n), \quad p' > p.$$

Для построения оценки вектора μ при наличии шума воспользуемся методом пороговой обработки. Рассмотрим жесткую (hard) пороговую обработку для каждой компоненты вектора:

$$\hat{\mu}_i = \rho_H(X_i, T) = \begin{cases} X_i & \text{при } |X_i| > T, \\ 0 & \text{при } |X_i| \leq T, \end{cases} \quad (1.2)$$

т. е. компонента вектора обнуляется, если ее абсолютное значение не превосходит критического порога T . При использовании FDR-метода пороговое значение T выбирается по следующему правилу: по абсолютным значениям величин исходного вектора строится вариационный ряд

$$|X|_{(1)} \geq \dots \geq |X|_{(n)},$$

и $|X|_{(k)}$ сравнивается с квантилями Гауссова распределения $t_k = \sigma z(\alpha/2 \cdot k/n)$. Пусть k_F — наибольший индекс k , для которого $|X|_{(k)} \geq t_k$, тогда выбирается порог $T^F = t_{k_F}$. Применение жесткой порогой обработки для компонент вектора μ в задаче (1.1) эквивалентно процедуре множественной проверки гипотез, описанной выше.

В сочетании с методами проверки гипотез также широко используется метод штрафов, при котором минимизируется невязка с добавлением штрафной функции [11; 29; 50]. В частном случае этот метод приводит к так называемой мягкой (soft) пороговой обработке, при которой оценки компонент вектора вычисляются по правилу

$$\hat{\mu}_i = \rho_S(X_i, T) = \begin{cases} X_i - T & \text{при } X_i > T, \\ X_i + T & \text{при } X_i < -T, \\ 0 & \text{при } |X_i| \leq T. \end{cases} \quad (1.3)$$

Данный подход в некоторых случаях оказывается более адекватным, чем (1.2), поскольку функция ρ_S в (1.3) непрерывна по X_i . Графики функций для жесткой и мягкой порогой обработки приведены на рисунке (1.1).

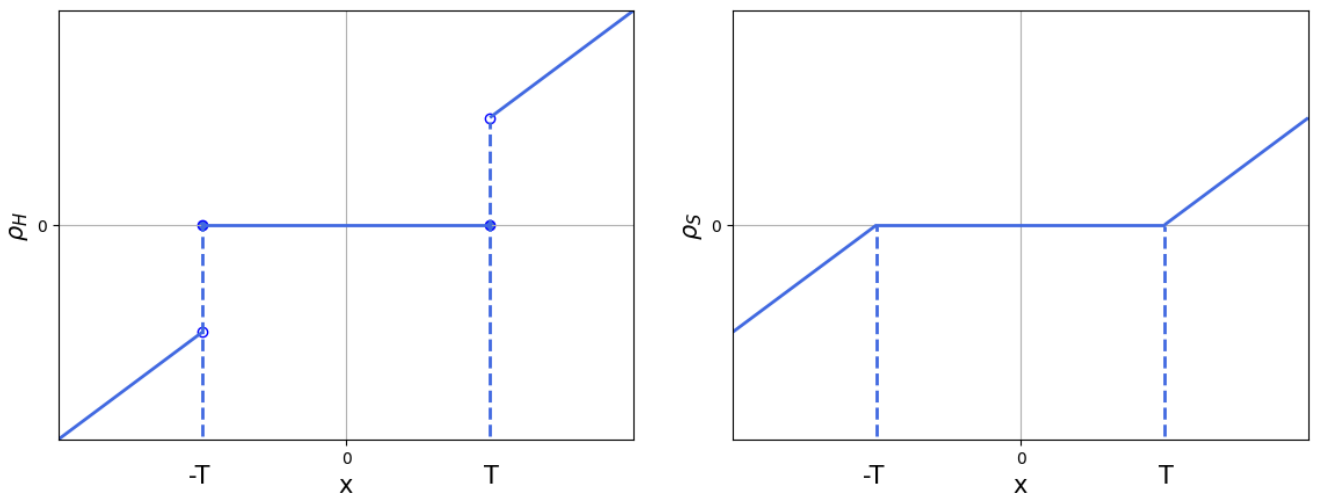


Рисунок 1.1 — Функции мягкой и жесткой пороговой обработки

В данной диссертации будут использоваться только введенные выше функции мягкой и жесткой пороговой обработки. Однако можно рассмотреть и другие виды пороговых функций ρ , в частности в работе [34] была предложена

так называемая полумягкая пороговая функция (semisoft) ρ_{SS} :

$$\rho_{SS}(X_i, T_1, T_2) = \begin{cases} X_i & \text{при } |X_i| > T_2, \\ \operatorname{sgn}(X_i) \frac{T_2(|X_i| - T_1)}{T_2 - T_1} & \text{при } T_1 < |X_i| \leq T_2, \\ 0 & \text{при } |X_i| \leq T_1. \end{cases}$$

В работе [33] была рассмотрена более гладкая функция, по своим свойствам схожая с полумягкой пороговой функцией:

$$\rho_G(X_i, T) = \begin{cases} X_i - \frac{T^2}{X_i} & \text{при } |X_i| > T, \\ 0 & \text{при } |X_i| \leq T. \end{cases}$$

Еще один вид пороговой функции был приведен в работе [43]:

$$\rho_{DG}(X_i, T_1, T_2) = \sum_{k=1}^K a_k \cdot X_i \exp\left(- (k-1) \frac{X_i^2}{2T^2}\right).$$

Как в случае мягкой пороговой обработки, так и в случае жесткой пороговой обработки основной задачей становится выбор стратегии для определения порогового значения T , при котором полученная модель будет в некотором смысле наилучшей. Зачастую такую задачу сводят к поиску порогового значения, минимизирующего среднеквадратичную погрешность.

1.3 Оценка среднеквадратичного риска при пороговой обработке

Среднеквадратичная погрешность (или риск) рассмотренных процедур определяется следующим образом:

$$R(T) = \sum_{i=1}^n \mathbb{E}(\hat{\mu}_i - \mu_i)^2. \quad (1.4)$$

Методы выбора порогового значения T , как правило, ориентированы на минимизацию риска (1.4) при условии принадлежности вектора μ заданному классу. “Идеальным” значением порога является

$$T^{min} : R(T^{min}) = \min_T R(T).$$

Заметим, что в выражении (1.4) присутствуют неизвестные величины μ_i и вычислить значения $R(T)$ и T^{min} на практике не представляется возможным. Поэтому рассмотрим оценку среднеквадратичного риска, используемую в книге [45], которая определяется выражением

$$\hat{R}(T) = \sum_{i=1}^n F[X_i, T], \quad (1.5)$$

где в случае жесткой пороговой обработки

$$F[X_i, T] = \begin{cases} X_i^2 - \sigma^2 & \text{при } |X_i| \leq T, \\ \sigma^2 & \text{при } |X_i| > T \end{cases} \quad (1.6)$$

и в случае мягкой пороговой обработки

$$F[X_i, T] = \begin{cases} X_i^2 - \sigma^2 & \text{при } |X_i| \leq T, \\ \sigma^2 + T^2 & \text{при } |X_i| > T. \end{cases} \quad (1.7)$$

Благодаря тому, что при вычислении оценки риска используются только наблюдаемые данные, можно получать значения оценки риска на практике. Следовательно, оценки риска позволяют сравнивать между собой модели, полученные при применении различных порогов, основываясь исключительно на наблюдаемых данных.

В работе [24] был предложен порог $T^U = \sigma\sqrt{2\log n}$, получивший название универсальный. Для универсального порога [22; 27; 62] при применении методов мягкой и жесткой пороговой обработки было доказано, что получаемые значения риска близки к минимальному значению риска $R(T^{min})$ и выполнено неравенство [45]

$$R(T^U) \leq (2\log n + 1) \left(\frac{\sigma^2}{n} + R(T^{min}) \right).$$

При этом данное значение порога удаляет практически весь шум:

$$\mathbb{P}(\max_{i \in [1, n]} |\varepsilon_i| > \sigma \sqrt{2 \log n}) \rightarrow 0 \text{ при } n \rightarrow \infty$$

и его можно использовать при самых слабых ограничениях на вектора наблюдений [37].

Также в работах [23; 31; 38] было доказано, что универсальный порог является в определенном смысле максимальным, таким образом значения порога $T > T^U$ можно исключить из рассмотрения.

Много важных результатов, связанных с адаптацией различных порогов к неизвестной разреженности входных данных, было приведено в работе [11]. Не ограничивая общности, приведем ниже теоремы 1.1 и 1.2 из [11] для случая $\sigma^2 = 1$.

Теорема 1.1 *Рассмотрим вектор μ из классов разреженности $\mu \in L_0(\eta_n)$, $\mu \in L_p(\eta_n)$ и $\mu \in M_p(\eta_n)$, $0 \leq p < r \leq 2$, где $\eta_n \in [n^{-1}(\log n)^5, n^{-\gamma}]$ для $L_0(\eta_n)$ или $\eta_n^p \in [n^{-1}(\log n)^5, n^{-\gamma}]$ для $L_p(\eta_n)$ и $M_p(\eta_n)$, $\gamma > 0$, $\hat{\mu}_F$ - оценка, полученная при применении FDR-порога T^F с управляющим параметром α_n . Тогда при $n \rightarrow \infty$ выполняются неравенства:*

$$\begin{aligned} \sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r &\leq \inf_{\hat{\mu}} \sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu} - \mu\|_r^r \cdot \left(1 + (r - p) \cdot \frac{\alpha_n}{1 - \alpha_n} + o(1)\right), \\ \sup_{\mu \in L_p(\eta_n)} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r &\leq \inf_{\hat{\mu}} \sup_{\mu \in L_p(\eta_n)} \mathbb{E} \|\hat{\mu} - \mu\|_r^r \cdot \left(1 + (r - p) \cdot \frac{\alpha_n}{1 - \alpha_n} + o(1)\right), \\ \sup_{\mu \in M_p(\eta_n)} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r &\leq \inf_{\hat{\mu}} \sup_{\mu \in M_p(\eta_n)} \mathbb{E} \|\hat{\mu} - \mu\|_r^r \cdot \left(1 + (r - p) \cdot \frac{\alpha_n}{1 - \alpha_n} + o(1)\right). \end{aligned}$$

Следовательно, если управляющий параметр FDR-порога $\alpha_n \rightarrow 0$ при росте n , то

$$\begin{aligned} \sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r &\sim \inf_{\hat{\mu}} \sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu} - \mu\|_r^r, \\ \sup_{\mu \in L_p(\eta_n)} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r &\sim \inf_{\hat{\mu}} \sup_{\mu \in L_p(\eta_n)} \mathbb{E} \|\hat{\mu} - \mu\|_r^r, \end{aligned}$$

и

$$\sup_{\mu \in M_p[\eta_n]} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r \sim \inf_{\hat{\mu}} \sup_{\mu \in M_p[\eta_n]} \mathbb{E} \|\hat{\mu} - \mu\|_r^r,$$

в том смысле, что отношение приближается к 1 при $n \rightarrow \infty$. Таким образом, теорема 1.1 устанавливает асимптотическую минимаксность FDR оценки и более того, минимаксность оценки FDR адаптивна к различным классам разреженности.

Следующая теорема демонстрирует адаптивную минимаксность FDR-порога.

Теорема 1.2 Пусть $X \sim \mathcal{N}_n(\mu, I)$, $\hat{\mu}_F$ - оценка, полученная при применении FDR-порога T^F . Предположим, что

$$\liminf_n \alpha_n \log n > 0, \quad \limsup_n \alpha_n < 1.$$

Рассмотрим пространства $L_0(\eta_n)$ и $M_p(\eta_n)$ для $0 < p < r \leq 2$. Предположим, что η_n (или соответственно η_n^p) лежит в интервале $[n^{-1} \log^5 n, n^{-\gamma}]$ для некоторого $\gamma > 0$. Тогда при $n \rightarrow \infty$

$$\sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r \leq \inf_{\hat{\mu}} \sup_{\mu \in L_0[\eta_n]} \mathbb{E} \|\hat{\mu} - \mu\|_r^r \cdot \left(1 + c_{rp} \cdot \frac{\alpha_n}{1 - \alpha_n} + o(1)\right),$$

$$\sup_{\mu \in M_p[\eta_n]} \mathbb{E} \|\hat{\mu}_F - \mu\|_r^r \leq \inf_{\hat{\mu}} \sup_{\mu \in M_p[\eta_n]} \mathbb{E} \|\hat{\mu} - \mu\|_r^r \cdot \left(1 + c_{rp} \cdot \frac{\alpha_n}{1 - \alpha_n} + o(1)\right),$$

где $c_{rp} = 2 \left(1 - \frac{p}{r}\right)$. Для случая пространства $L_p(\eta_n)$ утверждение теоремы верно при $c_{rp} = 2$.

Из результатов теоремы 1.2 следует, что FDR-оценка одновременно асимптотически минимаксна для различных классов разреженности: L_0, M_p, L_p , где $p < r \leq 2$ и широкого диапазона значений $\eta_n \rightarrow 0$ таких, что управляющий параметр α_n приближается к 0.

Результат теоремы 1.2 позволяет сравнивать между собой поведение FDR-оценки и ряда других оценок. В частности, в этой работе (с использованием результатов, полученных ранее в статье [25]) было продемонстрировано, что универсальный порог T^U плохо адаптируется к неизвестной разреженности данных. А именно, было показано, что если $\hat{\mu}_U$ - это оценка, полученная при мягкой

или жесткой пороговой обработки при использовании универсального порога $T^U = \sqrt{2 \log n}$ (при $\sigma = 1$), то

$$\sup_{\mu \in l_p[n_n]} R(T^U) \sim \left(\frac{2}{2-p} \right)^{1-p/2} \inf_{\mu \in L_0[n_n]} \sup R(T).$$

Из этого следует, что этот универсальный порог T^U не адаптируется при значениях $p < 2$.

Рассмотрим еще один способ построения порога. Для этого будем рассматривать Stein Unbiased Risk Estimator оценки (SURE-оценки) [58], т.е. несмещенные оценки риска. Очевидно, что значения риска $R(T)$ можно уменьшить, выбирая порог, адаптирующийся к зашумленным данным X . Порог, который минимизирует SURE-оценки, называется SURE-порогом. Рассмотрим риск $R(T)$, полученный при применении мягкой пороговой обработки. Вычислим оценку риска $\tilde{R}(T)$ на основе наблюдаемых данных X . Из представления (1.3) следует, что для значений X_i , таких что $|X_i| \leq T$ вклад компоненты вектора X_i в значение риска будет равно нулю, а вклад в значение оценки риска будет равен $X_i^2 - \sigma^2$. Итоговый риск, получаемый при применении мягкой пороговой обработки, был приведен в выражении (1.3), а оценка риска - в выражении (1.7). В работе [26] была доказана теорема о несмещенности оценки риска $\tilde{R}(T)$.

Теорема 1.3 *Для мягкой пороговой обработки оценка риска $\tilde{R}(T)$ является несмещенной оценкой риска*

$$\mathbb{E} \tilde{R}(T) = R(T).$$

Таким образом, $\tilde{R}(T)$ является SURE-оценкой.

Чтобы найти порог \tilde{T} , минимизирующего SURE-оценку, необходимо построить ряд в порядке убывания абсолютных значений компонент вектора X_i , т.е.

$$|X|_{(1)} \geq \dots \geq |X|_{(n)}.$$

Пусть l , такой что $|X|_{(l+1)} \leq T < |X|_{(l)}$. Тогда оценку риска $\tilde{R}(T)$ можно записать в следующем виде:

$$\begin{aligned} \tilde{R}(T) &= \sum_{i=1}^n F[X_i, T] = \sum_{i=1}^l F[X_i, T] + \sum_{i=l+1}^n F[X_i, T] = \\ &= \sum_{i=1}^l (\sigma^2 + T^2) + \sum_{i=l+1}^n ((X_{(l)})^2 - \sigma^2) = \sum_{i=l+1}^n (X_{(l)})^2 - (n-l) \cdot \sigma^2 + l \cdot (\sigma^2 + T^2). \end{aligned} \tag{1.8}$$

Очевидно, что значение порога $\tilde{T} = |X|_{(l+1)}$ будет минимизировать выражение SURE-оценку риска, т.к. выражение (1.8) возрастает по T для фиксированного l . Таким образом, чтобы найти итоговое значение порога \tilde{T} , необходимо сравнить значения оценки риска в n точках, т.е. значения выражения (1.8) при $T = |X|_{(1)}, \dots, |X|_{(n)}$ и выбрать в качестве \tilde{T} то значение, при котором достигается минимум оценки риска $\tilde{R}(T)$. Этот метод поиска порогового значения \tilde{T} был описан в работе [24].

Однако SURE-порог дает плохие результаты в случае, когда незашумленные данные имеют низкую энергию относительно энергии шума, т.е. $\|\mu\| \ll \mathbb{E} \|\varepsilon\| = n\sigma^2$. В таком случае значение порога \tilde{T} будет маленьким и после применения пороговой обработки останется много шумовых компонент в сигнале. Поэтому в таком случае имеет смысл воспользоваться значением универсального порога T^U . Таким образом, в случае низкочастотного полезного сигнала и больших значений шумовых компонент порог имеет смысл определять следующим образом

$$T = \begin{cases} \sigma\sqrt{2\log n} & \text{при } \|X\|^2 - n\sigma^2 \leq \varepsilon_n, \\ \tilde{T} & \text{при } \|X\|^2 - n\sigma^2 > \varepsilon_n, \end{cases}$$

где $\varepsilon_n = \sigma^2 \sqrt{n(\log n)^3}$.

Таким образом, показано, что риск, соответствующий оптимальному выбору порога \tilde{T} , достигается с точностью до члена эмпирических флуктуаций порядка $\sqrt{n(\log n)^3}$. Следовательно полученный порог \tilde{T} является адаптивным

в случае $L_p(\eta_n)$ и $M_p(\eta_n)$ при $p > 1$, но при этом, из результатов теоремы 1.3 видно, что полученный член флуктуаций слишком груб для обработки разреженных ситуаций при $p \leq 1$.

В случае жесткой пороговой обработки оценка среднеквадратичного риска получается смещенной, из-за этого нельзя гарантировать что порог, минимизирующий риск, будет являться оптимальным для жестких пороговых оценок.

Приведенные утверждения демонстрируют, что рассмотренный метод построения FDR-оценок в модели (1.1) обладает очень схожими свойствами с методом, основанным на минимизации по параметру T оценки (1.5) (см. [56]).

Еще одно важное свойство FDR-порога было доказано в работе [11], а именно теорема 1.4, приведенная ниже показывает, что при выполнении условий теоремы вероятность того, что FDR-порог окажется меньше значения T^1 достаточно мала.

Теорема 1.4 Пусть $\mu \in L_0(\eta_n)$ или $\mu \in L_p(\eta_n)$, $0 < p < 2$, где $\eta_n \in [n^{-1}(\log n)^5, n^{-\gamma}]$ для $L_0(\eta_n)$ или $\eta_n^p \in [n^{-1}(\log n)^5, n^{-\gamma}]$ для $L_p(\eta_n)$, $0 < \gamma < 1$. Тогда существует такая константа $c > 0$, что для FDR-порога T^F с управляющим параметром $\alpha_n \rightarrow 0$ при достаточно больших n выполняются неравенства:

$$\sup_{\mu \in L_0[\eta_n]} \mathbb{P}(T^F < T^1) \leq 2n \exp\{-c \alpha_n \kappa_n \gamma_n^2\},$$

$$\sup_{\mu \in L_p(\eta_n)} \mathbb{P}(T^F < T^1) \leq 2n \exp\{-c \alpha_n \kappa_n \gamma_n^2\},$$

где для $L_0(\eta_n)$

$$\gamma_n = \frac{1}{\log \log n}, \quad \kappa_n = \frac{n \eta_n}{1 - \alpha_n - \gamma_n}, \quad T^1 = \sigma(2 \log \eta_n^{-1})^{1/2}, \quad (1.9)$$

и для $L_p(\eta_n)$

$$\gamma_n = \frac{1}{\log \log n}, \quad \tau_\eta = \sigma(2 \log \eta_n^{-p})^{1/2}, \quad \kappa_n = \frac{n \eta_n^p \tau_\eta^{-p}}{1 - \alpha_n - \gamma_n}, \quad T^1 = \sigma(2 \log \eta_n^{-p})^{1/2}. \quad (1.10)$$

Приведенные выше результаты из ряда статей демонстрируют хорошие асимптотические свойства FDR-порога, что делает целесообразным его применение при работе с большими объемами данных.

1.4 Сильная состоятельность оценки риска

В этом и следующем разделе будут доказаны асимптотические свойства оценки риска, а именно сильная состоятельность и асимптотическая нормальность оценки риска.

Теорема 1.5 была доказана в работе [66].

Теорема 1.5 Пусть $\mu \in L_0[\eta_n]$, где $\eta_n \in [n^{-1}(\log n)^5, n^{-\gamma}]$ или $\mu \in L_p(\eta_n)$, $0 < p < 2$, где $\eta_n^p \in [n^{-1}(\log n)^5, n^{-\gamma}]$ соответственно, $0 < \gamma < 1$. Пусть T^F – FDR-порог с управляющим параметром $\alpha_n \rightarrow 0$ и $\frac{\alpha_n \kappa_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, где κ_n и γ_n определены в (1.9) и (1.10) соответственно. Тогда

$$\frac{\hat{R}(T^F) - R(T^{\min})}{n} \rightarrow 0 \text{ п.в.}$$

Доказательство. Справедлива оценка

$$\begin{aligned} \mathbb{P} \left(\frac{|\hat{R}(T^F) - R(T^{\min})|}{n} > \varepsilon \right) &\leq \\ &\leq \mathbb{P} \left(\frac{|\hat{R}(T^F) - R(T^F)|}{n} > \varepsilon/2 \right) + \mathbb{P} \left(\frac{|R(T^F) - R(T^{\min})|}{n} > \varepsilon/2 \right). \end{aligned} \quad (1.11)$$

Разобьем первое слагаемое в (1.11) и запишем

$$\begin{aligned} \mathbb{P} \left(\frac{|\hat{R}(T^F) - R(T^F)|}{n} > \varepsilon/2 \right) &+ \mathbb{P} \left(\frac{|R(T^F) - R(T^{\min})|}{n} > \varepsilon/2 \right) \leq \\ \mathbb{P} \left(\frac{|\hat{R}(T^F) - R(T^F)|}{n} > \varepsilon/2, T^F \in [T^1, T^U] \right) &+ \mathbb{P} (T^F \notin [T^1, T^U]) + \\ &+ \mathbb{P} \left(\frac{|R(T^F) - R(T^{\min})|}{n} > \varepsilon/2 \right). \end{aligned} \quad (1.12)$$

Далее рассмотрим каждое слагаемое из последнего выражения отдельно. Заметим, что первое слагаемое из (1.12) не превосходит

$$\mathbb{P} \left(\sup_{T \in [T^1, T^U]} \frac{|\hat{R}(T) - R(T)|}{n} > \varepsilon/2 \right).$$

Введем обозначение:

$$\frac{\hat{R}(T) - R(T)}{n} = \frac{1}{n} \sum_{i=1}^n Y_i(T) = Z(T),$$

где $Y_i(T)$ - независимые случайные величины с нулевым математическим ожиданием. Для $|Y_i(T)|$ можно показать, что:

$$\begin{aligned} |Y_i(T)| &\leq |1 - 2 \cdot \mathbf{1}(X_i^2 \leq T^2) + X_i^2 \wedge T^2 - \mathbb{E}\hat{\mu}_i - \mu_i|^2 \leq \\ &\leq |1 - 2 \cdot \mathbf{1}(X_i^2 \leq T^2)| + |X_i^2 \wedge T^2 - \mathbb{E}\hat{\mu}_i - \mu_i|^2 \leq 2 + T^2. \end{aligned}$$

Тогда для $|Z(T)|$ можем записать неравенство Хёффдинга [36] для фиксированного T :

$$\mathbb{P} \left(|Z(T)| \geq \varepsilon \right) \leq 2 \exp \left\{ -\frac{2\varepsilon^2 n}{(2 + T^2)^2} \right\}.$$

Возьмем $T < T'$ и обозначим $N(T, T') = \#\{i : T < |X_i| \leq T'\}$. В [2] были получены следующие оценки для приращений $\hat{R}(T)$ и $R(T)$:

$$\left| \frac{\hat{R}(T) - \hat{R}(T')}{n} \right| \leq \frac{2}{n} N(T, T') + ((T')^2 - T^2), \quad (1.13)$$

$$\left| \frac{R(T) - R(T')}{n} \right| \leq 4T^U \delta_n. \quad (1.14)$$

Из оценок (1.13) и (1.14) следует, что

$$\begin{aligned} |Z(T) - Z(T')| &\leq \frac{2}{n} N(T, T') + (T' - T)(T' + T) + 4T^U \delta_n \leq \\ &\leq \frac{2}{n} N(T, T') + 6T^U \delta_n. \end{aligned} \quad (1.15)$$

Для дальнейшего доказательства разобьем отрезок $[T^1, T^U]$ на равные части:

$$T_j = T^1 + j\delta_n \in [T^1, T^U], \quad j = 1, \dots, n-1,$$

$$\delta_n = \frac{T^U - T^1}{n}.$$

Тогда можно записать следующее выражение:

$$A_n = \left\{ \sup_{T \in [T^1, T^U]} |Z(T)| \geq 3\varepsilon \right\} \subset D_n \cup E_n,$$

где

$$D_n = \left\{ \sup_j |Z(T_j)| \geq \varepsilon \right\}, \quad E_n = \left\{ \sup_j \sup_{[T_j, T_j + \delta_n]} |Z(T) - Z(T_j)| \geq 2\varepsilon \right\}.$$

Для E_n справедлива следующая цепочка вложений:

$$\begin{aligned} & \left\{ \sup_j \sup_{[T_j, T_j + \delta_n]} |Z(T) - Z(T_j)| \geq 2\varepsilon \right\} \subset \\ & \subset \left\{ \sup_j \sup_{[T_j, T_j + \delta_n]} \frac{2}{n} N(T_j, T) + 6T^U \delta_n \geq 2\varepsilon \right\} \subset \\ & \subset \left\{ \sup_j \frac{2}{n} N(T_j, T_j + \delta_n) + 6T^U \delta_n \geq 2\varepsilon \right\} \subset \\ & \subset \left\{ \sup_j N(T_j, T_j + \delta_n) \geq n\varepsilon - 3nT^U \delta_n \right\}. \end{aligned}$$

Заметим, что $N(T, T') = \#\{i : T < |X_i| \leq T'\}$ есть не что иное, как сумма индикаторов того, что случайная величина X_i попала в заданный интервал $N(T, T') = \sum_{i=1}^n \mathbf{1}(T < |X_i| \leq T')$.

Для того, чтобы получить оценку для E_n , необходимо привести левую часть неравенства к виду $N(T_j, T_j + \delta_n) - \mathbb{E}N(T_j, T_j + \delta_n)$. Нетрудно показать, что

$$\mathbb{E}N(T_j, T_j + \delta_n) \leq \sqrt{\frac{2}{\pi}} \cdot n \cdot \delta_n,$$

следовательно:

$$\begin{aligned} & \left\{ \sup_j (N(T_j, T_j + \delta_n) - \mathbb{E}N(T_j, T_j + \delta_n)) \geq n\varepsilon - 3nT^U \delta_n - \sqrt{\frac{2}{\pi}} n \delta_n \right\} \subset \\ & \subset \left\{ \sup_j \frac{1}{n} (N(T_j, T_j + \delta_n) - \mathbb{E}N(T_j, T_j + \delta_n)) \geq \varepsilon - 3T^U \delta_n - \sqrt{\frac{2}{\pi}} \delta_n \right\} = E_n''. \end{aligned}$$

Пусть $\varepsilon - 3T^U \delta_n - \sqrt{\frac{2}{\pi}} \delta_n = \varepsilon_2$. Применяя неравенство Хёффдинга, получаем:

$$\mathbb{P}(E_j'') = \mathbb{P}\left(\frac{1}{n} \left| N(T_j, T_j + \delta_n) - \mathbb{E}N(T_j, T_j + \delta_n) \right| \geq \varepsilon_2\right) \leq 2e^{-2n\varepsilon_2^2}.$$

Следовательно, для события E_n'' получаем

$$\mathbb{P}(E_n'') \leq \sum_j \mathbb{P}(E_j'') \leq 2 \frac{T^U}{\delta_n} e^{-2n\varepsilon_2^2}. \quad (1.16)$$

Теперь рассмотрим вероятность события D_n . Очевидно, что для случайных величин $Y_i(T_j)$ справедливо неравенство $|Y_i(T_j)| \leq 2 + (T^U)^2$. Тогда, применяя неравенство Хёффдинга, получим следующую оценку для D_n :

$$\begin{aligned} \mathbb{P}(D_n) &= \mathbb{P}\left(\sup_j |Z_n(T_j)| \geq \varepsilon\right) \leq \sum_j \mathbb{P}(|Z_n(T_j)| \geq \varepsilon) \leq \\ &\leq 2 \frac{T^U}{\delta_n} \exp\left\{-\frac{2n\varepsilon^2}{(2 + (T^U)^2)^2}\right\}. \end{aligned} \quad (1.17)$$

Тогда из оценок (1.16) и (1.17) следует, что для вероятности A_n справедлива оценка

$$\mathbb{P}(A_n) \leq \mathbb{P}(D_n) + \mathbb{P}(E_n'') \leq 2 \frac{T^U}{\delta_n} \left(\exp\left\{-\frac{2n\varepsilon^2}{(2 + (T^U)^2)^2}\right\} + \exp\{-2n\varepsilon_2^2\} \right).$$

Не ограничивая общности будем считать $\sigma^2 = 1$. Тогда, подставив значение универсального порога T^U при $\sigma^2 = 1$, получим

$$\begin{aligned} \mathbb{P}(A_n) &\leq \mathbb{P}\left(\sup_{[T^1, T^U]} |Z(T)| \geq 3\varepsilon\right) \leq \\ &\leq \frac{2}{\delta_n} \sqrt{2 \log n} \cdot \exp\left\{-\frac{n\varepsilon^2}{(1 + \log n)^2}\right\} + \\ &\quad + \frac{2}{\delta_n} \sqrt{2 \log n} \cdot \exp\left\{-2n\left(\varepsilon - 3\delta_n \sqrt{2 \log n} - \sqrt{\frac{2}{\pi}} \delta_n\right)^2\right\} = r_n. \end{aligned}$$

Тогда ряд из r_n будет сходиться, т.е.

$$\sum_{n=1}^{\infty} r_n < \infty.$$

Для второго слагаемого из (1.12) по теореме 1.4 имеем:

$$\mathbb{P}\left(T^F \notin [T^1, T^U]\right) \leq 2n \exp\{-c\alpha_n k_n \gamma_n^2\} = m_n.$$

При условии, что $\frac{\alpha_n k_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, ряд из m_n сходится:

$$\sum_{n=1}^{\infty} m_n < \infty.$$

Рассмотрим третье слагаемое из (1.12):

$$\mathbb{P}\left(\frac{|R(T^F) - R(T^{min})|}{n} > \varepsilon/2\right).$$

По аналогии с первым слагаемым из (1.11) запишем:

$$\begin{aligned} \mathbb{P}\left(\frac{|R(T^F) - R(T^{min})|}{n} > \varepsilon/2\right) &\leq \\ &\leq \mathbb{P}\left(\frac{|R(T^F) - R(T^{min})|}{n} > \varepsilon/2, T^F \in [T^1, T^U]\right) + \mathbb{P}(T^F \notin [T^1, T^U]) \leq \\ &\leq \mathbb{P}\left(\sup_{T \in [T^1, T^U]} \frac{|R(T) - R(T^{min})|}{n} > \varepsilon/2\right) + \mathbb{P}(T^F \notin [T^1, T^U]). \end{aligned}$$

Для $\mathbb{P}(T^F \notin [T^1, T^U])$ снова применим теорему 1.4 и, как было отмечено выше, при выполнении условий теоремы ряд из m_n будет сходиться. В ходе доказательства состоятельности оценки риска при выборе FDR-порога в [2] было показано, что

$$\sup_{T \in [T^1, T^U]} \frac{|R(T) - R(T^{min})|}{n} \rightarrow 0 \text{ при } n \rightarrow \infty. \quad (1.18)$$

Заметим, что ряд из вероятностей сходится:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{|R(T^F) - R(T^{min})|}{n} > \varepsilon/2\right) < \infty,$$

в силу того, что выполнено (1.18), а следовательно, начиная с некоторого номера n ,

$$\mathbb{P}\left(\sup_{T \in [T^1, T^U]} \frac{|R(T) - R(T^{min})|}{n} > \varepsilon/2\right) = 0.$$

Теорема доказана. □

Таким образом, в теореме 1.5 доказана сильная состоятельность оценки риска $\hat{R}(T)$.

1.5 Асимптотическая нормальность оценки риска

Докажем утверждение об асимптотической нормальности оценки (1.5), которое, в частности, позволяет строить асимптотические доверительные интервалы для среднеквадратичного риска (1.4). При доказательстве мы будем использовать обозначение C для некоторых положительных констант, которые могут зависеть от параметров рассматриваемых классов и методов, но не зависят от n . При этом в разных соотношениях одной буквой могут обозначаться разные константы.

Теоремы 1.6 и 1.7, приведенные ниже, были доказаны в работе [70]. Сначала рассмотрим класс разреженности $L_0(\eta_n)$.

Теорема 1.6. Пусть $\mu \in L_0(\eta_n)$, $\eta_n \in [n^{-1}(\log n)^5, n^{-\gamma}]$, $1/2 < \gamma < 1$. Пусть T_F – FDR-порог с управляющим параметром $\alpha_n \rightarrow 0$ и $\frac{\alpha_n \kappa_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, где κ_n и γ_n определены в (1.9). Тогда

$$\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} \Rightarrow \mathcal{N}(0,1).$$

Доказательство. Докажем теорему для метода мягкой пороговой обработки. В случае жесткой пороговой обработки доказательство аналогично.

Обозначим

$$U(T) = \hat{R}(T) - \hat{R}(T^{min}) = \sum_{i=1}^n H_i(T, T^{min}),$$

где

$$H_i(T, T^{min}) = F[X_i, T] - F[X_i, T^{min}],$$

и запишем

$$\hat{R}(T^F) - R(T^{min}) + \hat{R}(T^{min}) - \hat{R}(T^{min}) = \hat{R}(T^{min}) - R(T^{min}) + U(T^F).$$

Покажем, что

$$\frac{\hat{R}(T^{min}) - R(T^{min})}{\sigma^2 \sqrt{2n}} \Rightarrow \mathcal{N}(0,1). \quad (1.19)$$

При мягкой пороговой обработке $\hat{R}(T^{min})$ является несмещенной оценкой $R(T^{min})$, а при жесткой пороговой обработке и выполнении условий теоремы смещение стремится к нулю при делении на \sqrt{n} [45]. Для дисперсии числителя имеем [5]

$$\lim_{n \rightarrow \infty} \frac{\mathbb{D} \sum_{i=1}^n (F[X_i, T^{min}] - \mathbb{E}F[X_i, T^{min}])}{\mathbb{D} \sum_{i=1}^n X_i^2} = 1. \quad (1.20)$$

Кроме того, поскольку X_i независимы, $\mathbb{D}X_i^2 = 2\sigma^4 + 4\sigma^2\mu_i^2$ и число не равных нулю μ_i не превосходит $\eta_n n$, получаем

$$\lim_{n \rightarrow \infty} \frac{\mathbb{D} \sum_{i=1}^n X_i^2}{2\sigma^4 n} = 1. \quad (1.21)$$

Наконец, выполнено условие Линдеберга: для любого $\varepsilon > 0$ при $n \rightarrow \infty$

$$\frac{1}{V_n^2} \sum_{i=1}^n \mathbb{E} \left[(F[X_i, T^{min}] - \mathbb{E}F[X_i, T^{min}])^2 \times \right. \\ \left. \times \mathbf{1}(|F[X_i, T^{min}] - \mathbb{E}F[X_i, T^{min}]| > \varepsilon V_n) \right] \rightarrow 0, \quad (1.22)$$

где $V_n^2 = \mathbb{D} \sum_{i=1}^n (F[X_i, T^{min}] - \mathbb{E}F[X_i, T^{min}])$. Действительно, в силу (1.20) и (1.21) и т. к. слагаемые в $\hat{R}(T^{min})$ по модулю ограничены величиной $(T^U)^2 + \sigma^2$, начиная с некоторого n все индикаторы в (1.22) обращаются в нуль.

Следовательно, выполнено (1.19) и для доказательства теоремы осталось показать, что

$$\frac{U(T^F)}{\sqrt{n}_{30}} \xrightarrow{\mathbb{P}} 0.$$

Повторяя рассуждения из [3; 4; 37] можно показать, что $T^{min} \geq T^1 - \alpha_n$, где $|\alpha_n| \leq C \frac{\log \log n}{\sqrt{\log n}}$. Для сокращения записи без ущерба для доказательства можно опустить α_n и считать, что $T^{min} \geq T^1$.

Для произвольного $\varepsilon > 0$

$$\begin{aligned} \mathbb{P} \left(\frac{|U(T^F)|}{\sqrt{n}} > \varepsilon \right) &\leq \mathbb{P}(T^F \leq T^1) + \mathbb{P} \left(\frac{\sup_{T \in [T^1, T^U]} |U(T)|}{\sqrt{n}} > \varepsilon \right) \leq \\ &\leq \mathbb{P}(T^F \leq T^1) + \mathbb{P} \left(\frac{\sup_{T \in [T^1, T^U]} |U(T) - \mathbb{E}U(T)| + \sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)|}{\sqrt{n}} > \varepsilon \right). \end{aligned}$$

Пусть $U(T) = S_1(T) + S_2(T)$, $T \in [T^1, T^U]$, где сумма $S_2(T)$ содержит слагаемые, для которых $\mu_i = 0$, а $S_1(T)$ – остальные слагаемые. В силу определения класса $L_0(\eta_n)$ число слагаемых в $S_1(T)$ не превосходит $n_1 \approx \eta_n n$. Кроме того, каждое слагаемое по модулю не превосходит величины $(T^U)^2 + \sigma^2$. Для удобства будем считать, что $S_1(T)$ содержит слагаемые с номерами от 1 до n_1 , т. е.

$$S_1(T) = \sum_{i=1}^{n_1} H_i(T, T^{min}), \quad S_2(T) = \sum_{i=n_1+1}^n H_i(T, T^{min}).$$

Далее

$$\mathbb{E}U(T) = \mathbb{E}S_1(T) + \mathbb{E}S_2(T)$$

и

$$\sup_{T \in [T^1, T^U]} |\mathbb{E}S_1(T)| \leq n_1 \cdot \left((T^U)^2 + \sigma^2 \right).$$

Учитывая определение класса $L_0(\eta_n)$ и вид T^1 , можно убедиться, что для слагаемых из S_2 при $T \in [T^1, T^U]$ справедлива оценка $|\mathbb{E}H_i(T, T^{min})| \leq C(\log n)^{1/2} n^{-\gamma}$. Таким образом, объединяя оценки для $\mathbb{E}S_1(T)$ и $\mathbb{E}S_2(T)$, получаем

$$\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)| \leq C \log n n^{1-\gamma}.$$

Следовательно при $\gamma > 1/2$ и $n \rightarrow \infty$

$$\frac{\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)|}{\sqrt{n}_{31}} \rightarrow 0. \quad (1.23)$$

Далее возьмем $T < T'$ и обозначим

$$N_1(T, T') = \#\{i : T < |X_i| \leq T', i \in [1, n_1]\},$$

т.е.

$$N_1(T, T') = \sum_{i=1}^{n_1} \mathbf{1}(T < |X_i| \leq T').$$

Пусть

$$Z_1(T) = S_1(T) - \mathbb{E}S_1(T),$$

тогда справедлива оценка [23]:

$$|Z_1(T) - Z_1(T')| \leq 4\sigma^2 N_1(T, T') + 2n_1((T')^2 - T^2) \quad \text{п. в.} \quad (1.24)$$

Разобьем отрезок $[T^1, T^U]$ на равные части:

$$T_j = T^1 + j\delta_{n_1} \in [T^1, T^U], \quad j = 1, \dots, n_1 - 1,$$

$$\delta_{n_1} = \frac{T^U - T^1}{n_1}.$$

Тогда

$$A_n = \left\{ \sup_{T \in [T^1, T^U]} |S_1(T) - \mathbb{E}S_1(T)| \geq 5\varepsilon\sqrt{n} \right\} \subset D_n \cup E_n,$$

где

$$D_n = \left\{ \sup_j |Z_1(T_j)| > \varepsilon\sqrt{n} \right\},$$

$$E_n = \left\{ \sup_j \sup_{T \in [T_j, T_j + \delta_{n_1}]} |Z_1(T) - Z_1(T_j)| \geq 4\varepsilon\sqrt{n} \right\}.$$

Применяя неравенство Хеффдинга [36] для D_n , получаем оценку

$$\begin{aligned} \mathbb{P}(D_n) &\leq \sum_j \mathbb{P}(|Z_1(T_j)| > \varepsilon\sqrt{n}) \leq 2n_1 \exp \left\{ -\frac{\varepsilon^2 n}{2((T^U)^2 + 2\sigma^2)^2 n_1} \right\} \leq \\ &\leq 2n^{1-\gamma} \exp \left\{ -\frac{C\varepsilon^2 n^\gamma}{\log n} \right\}. \end{aligned}$$

Далее, учитывая (1.24), получим

$$\begin{aligned}
E_n &\subset \left\{ \sup_j \sup_{T \in [T_j, T_j + \delta_{n_1}]} 4\sigma^2 N_1(T_j, T) + 4n_1 \cdot T^U \delta_{n_1} \geq 4\varepsilon\sqrt{n} \right\} \subset \\
&\subset \left\{ \sup_j \sigma^2 N_1(T_j, T_j + \delta_{n_1}) + n_1 T^U \delta_{n_1} \geq \varepsilon\sqrt{n} \right\} = \\
&= \left\{ \sup_j \sigma^2 N_1(T_j, T_j + \delta_{n_1}) \geq \varepsilon\sqrt{n} - n_1 T^U \delta_{n_1} \right\}.
\end{aligned}$$

Нетрудно показать, что $\mathbb{E}N_1(T_j, T_j + \delta_{n_1}) \leq C n_1 \delta_{n_1}$. Следовательно,

$$\begin{aligned}
&\left\{ \sup_j \sigma^2 N_1(T_j, T_j + \delta_{n_1}) \geq \varepsilon\sqrt{n} - n_1 T^U \delta_{n_1} \right\} \subset \\
&\subset \left\{ \sup_j \frac{1}{n_1} |N_1(T_j, T_j + \delta_{n_1}) - \mathbb{E}N_1(T_j, T_j + \delta_{n_1})| \geq \frac{\varepsilon\sqrt{n}}{\sigma^2 n_1} - \frac{T^U \delta_{n_1}}{\sigma^2} - \frac{C}{\sigma^2} \delta_{n_1} \right\} = \\
&= E'_n.
\end{aligned}$$

Применяя неравенство Хеффдинга для E'_j , получаем

$$\begin{aligned}
\mathbb{P}(E'_j) &= \mathbb{P}\left(\frac{1}{n_1} |N_1(T_j, T_j + \delta_{n_1}) - \mathbb{E}N_1(T_j, T_j + \delta_{n_1})| \geq \frac{\varepsilon\sqrt{n}}{\sigma^2 n_1} - \frac{T^U \delta_{n_1}}{\sigma^2} - \frac{C}{\sigma^2} \delta_{n_1} \right) \leq \\
&\leq 2 \exp \left\{ -2 \left(\frac{\varepsilon\sqrt{n}}{\sigma^2 n_1} - \frac{T^U \delta_{n_1}}{\sigma^2} - \frac{c}{\sigma^2} \delta_{n_1} \right)^2 n_1 \right\} \leq 2 \exp \{ -C\varepsilon^2 n^\gamma \}.
\end{aligned}$$

Тогда для вероятности события E'_n справедлива оценка:

$$\mathbb{P}(E'_n) \leq \sum_{j=1}^{n_1} \mathbb{P}(E'_j) \leq C n^{1-\gamma} \exp \{ -C\varepsilon^2 n^\gamma \}.$$

Таким образом, для произвольного $\varepsilon > 0$

$$\mathbb{P} \left(\sup_{T \in [T^1, T^U]} \frac{|S_1(T) - \mathbb{E}S_1(T)|}{\sqrt{n}} > \varepsilon \right) \rightarrow 0 \quad (1.25)$$

при $n \rightarrow \infty$.

Рассмотрим теперь сумму $S_2(T)$. При больших значениях n число слагаемых в этой сумме равно $n - n_1 \approx n$. Повторяя приведенные выше рассуждения, разобьем отрезок $[T^1, T^U]$ на равные части:

$$T_j = T^1 + j\delta_{n_1} \in [T^1, T^U], \quad j = 1, \dots, n-1,$$

$$\delta_n = \frac{T^U - T^1}{n}.$$

Обозначим

$$Z_2(T) = S_2(T) - \mathbb{E}S_2(T), \quad N_2(T, T') = \sum_{i=n_1+1}^n \mathbf{1}(T < |x_i| \leq T'),$$

тогда

$$A_n = \left\{ \sup_{T \in [T^1, T^U]} |S_2(T) - \mathbb{E}S_2(T)| \geq 5\varepsilon\sqrt{n} \right\} \subset D_n \cup E_n,$$

где

$$D_n = \left\{ \sup_j |Z_2(T_j)| > \varepsilon\sqrt{n} \right\},$$

$$E_n = \left\{ \sup_j \sup_{T \in [T_j, T_j + \delta_n]} |Z_2(T) - Z_2(T_j)| \geq 4\varepsilon\sqrt{n} \right\}.$$

Учитывая определение класса $L_0(\eta_n)$ и вид T^1 , можно показать, что для дисперсии слагаемых в S_2 (а значит, и Z_2) справедлива оценка

$$\mathbb{D}H_i(T, T^{\min}) \leq C \left(\log \frac{n}{(\log n)^5} \right)^{3/2} n^{-\gamma}.$$

Применяя неравенство Бернштейна [16] для D_n , получаем

$$\begin{aligned} \mathbb{P}(D_n) &\leq \sum_j \mathbb{P}(|Z_2(T_j)| > \varepsilon\sqrt{n}) \leq \\ &\leq 2n \exp \left\{ - \frac{C\varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{3/2} n^{1-\gamma} + 2((T^U)^2 + \sigma^2)\sqrt{n}} \right\}. \end{aligned}$$

Введем множество E'_n , такое что $E_n \subset E'_n$ и

$$E'_n = \left\{ \sup_j |N_2(T_j, T_j + \delta_n) - \mathbb{E}N_2(T_j, T_j + \delta_n)| \geq \frac{\varepsilon\sqrt{n}}{\sigma^2} - \frac{nT^U\delta_n}{\sigma^2} - \frac{Cn}{\sigma^2}\delta_n \right\}.$$

Заметим, что дисперсия слагаемых в $N_2(T_j, T_j + \delta_n)$ ограничена величиной $C \left(\log \frac{n}{(\log n)^5} \right)^{1/2} \cdot n^{-\gamma}$. Применяя неравенство Бернштейна, получаем

$$\begin{aligned} \mathbb{P}(E'_n) &= \mathbb{P} \left(\left| N_1(T_j, T_j + \delta_n) - \mathbb{E}N_1(T_j, T_j + \delta_n) \right| \geq \frac{\varepsilon\sqrt{n}}{\sigma^2} - \frac{nT^U\delta_n}{\sigma^2} - \frac{Cn}{\sigma^2}\delta_n \right) \leq \\ &\leq 2 \exp \left\{ - \frac{C\varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{1/2} n^{1-\gamma} + 2\sqrt{n}} \right\}. \end{aligned}$$

Следовательно,

$$\mathbb{P}(E'_n) \leq \sum_{j=1}^n \mathbb{P}(E''_j) \leq 2n \exp \left\{ -\frac{C\varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{1/2} n^{1-\gamma} + 2\sqrt{n}} \right\}.$$

Таким образом, для произвольного $\varepsilon > 0$

$$\mathbb{P} \left(\sup_{T \in [T^1, T^U]} \frac{|S_2(T) - \mathbb{E}S_2(T)|}{\sqrt{n}} > \varepsilon \right) \rightarrow 0 \quad (1.26)$$

при $n \rightarrow \infty$.

Объединяя (1.19), (1.23), (1.25) и (1.26), получаем утверждение теоремы. \square

Аналогичное утверждение справедливо для класса $L_p(\boldsymbol{\eta}_n)$.

Теорема 1.7 Пусть $\boldsymbol{\mu} \in L_p(\boldsymbol{\eta}_n)$, $0 < p < 2$, $\boldsymbol{\eta}_n^p \in [n^{-1}(\log n)^5, n^{-\gamma}]$, $1/2 < \gamma < 1$. Пусть T_F – FDR-порог с управляющим параметром $\alpha_n \rightarrow 0$ и $\frac{\alpha_n \kappa_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, где κ_n и γ_n определены в (1.10). Тогда

$$\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} \Rightarrow \mathcal{N}(0,1).$$

Доказательство. Основные этапы доказательства этой теоремы повторяют доказательство теоремы 1.6. Запишем

$$\hat{R}(T^F) - R(T^{min}) + \hat{R}(T^{min}) - \hat{R}(T^{min}) = \hat{R}(T^{min}) - R(T^{min}) + U(T^F).$$

Утверждение

$$\frac{\hat{R}(T^{min}) - R(T^{min})}{\sigma^2 \sqrt{2n}} \Rightarrow N(0,1)$$

доказывается полностью аналогично утверждению (1.19).

Пусть

$$U(T) = S_1(T) + S_2(T), \quad T \in [T^1, T^U],$$

где сумма $S_2(T)$ содержит слагаемые, для которых $|\boldsymbol{\mu}_i| \leq C/T^1$, а $S_1(T)$ – остальные слагаемые. В силу определения класса $L_p(\boldsymbol{\eta}_n)$ число слагаемых в $S_1(T)$ не превосходит $n_1 \approx C \boldsymbol{\eta}_n^p n (\log \boldsymbol{\eta}_n^{-p})^{p/2}$ и каждое слагаемое не превосходит

$(T^U)^2 + \sigma^2$. Учитывая вид T^1 , можно убедиться, что математические ожидания слагаемых в S_2 не превосходят $C(\log n)^{1/2}n^{-\gamma}$, а их дисперсии не превосходят $C\left(\log \frac{n}{(\log n)^5}\right)^{3/2}n^{-\gamma}$. Далее, рассуждая, как в теореме 1.6, убеждаемся, что

$$\frac{\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)|}{\sqrt{n}} \rightarrow 0$$

и

$$\mathbb{P}\left(\sup_{T \in [T^1, T^U]} \frac{|U(T) - \mathbb{E}U(T)|}{\sqrt{n}} > \varepsilon\right) \rightarrow 0$$

для произвольного $\varepsilon > 0$ при $n \rightarrow \infty$. Таким образом, поскольку $\mathbb{P}(T^F \leq T^1) \rightarrow 0$,

$$\frac{U(T^F)}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0,$$

при $n \rightarrow \infty$, что завершает доказательство теоремы. \square

Результаты теорем 1.6 и 1.7 позволяют строить асимптотические доверительные интервалы для оценки риска $R(T^{min})$ для классов разреженности $L_0(\eta_n)$ и $L_p(\eta_n)$. А именно, исходя из полученных результатов, что

$$\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} \Rightarrow \mathcal{N}(0,1),$$

можно записать следующее:

$$\mathbb{P}\left(-z_{\frac{1+\gamma}{2}} \leq \frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} \leq z_{\frac{1+\gamma}{2}}\right) = \mathbb{P}(T_1 \leq R(T^{min}) \leq T_2) \rightarrow \gamma.$$

где $z_{\frac{1+\gamma}{2}}$ - квантиль стандартного нормального распределения уровня $\frac{1+\gamma}{2}$ и

$$T_1 = \hat{R}(T^F) - \sigma^2 \sqrt{2n} \cdot z_{\frac{1+\gamma}{2}},$$

$$T_2 = \hat{R}(T^F) + \sigma^2 \sqrt{2n} \cdot z_{\frac{1+\gamma}{2}}.$$

Таким образом, границы доверительных интервалов минимального среднеквадратичного риска $R(T^{min})$ могут быть найдены через значения оценки риска при использовании FDR-порога.

1.6 Оценка скорости сходимости распределения оценки риска

В дальнейшем через $c, C, C_i, i \in \mathbb{N}$ будем обозначать некоторые положительные константы, которые могут быть разными в разных выражениях.

В работе [67] были доказаны следующие теоремы 1.8 и 1.9 для различных классов разреженности данных.

Теорема 1.8 Пусть $\mu \in L_0(\eta_n)$, где $\eta_n \in [n^{-1}(\log n)^5, n^{-\gamma}]$, $1/2 < \gamma < 1$. Пусть T_F — FDR-порог с управляющим параметром $\alpha_n \rightarrow 0$ и $\frac{\alpha_n \kappa_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, где κ_n и γ_n определены в (1.9). Тогда для $x \in \mathbb{R}$ выполнены следующие неравенства:

для $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$

$$\sup_x \left| \mathbb{P}\left(\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} < x\right) - \Phi(x) \right| \leq C \cdot n^{1/2-\gamma} \cdot \log n,$$

для $\gamma \in \left(\frac{3}{4}, 1\right)$

$$\sup_x \left| \mathbb{P}\left(\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} < x\right) - \Phi(x) \right| \leq C \cdot n^{-1/4} \cdot \log n,$$

где в обоих случаях C — это константы, вообще говоря, разные.

Доказательство. В статье [6] было приведено следующее неравенство Бернштейна для $Z_n = V_n + W_n$, любого $\varepsilon > 0$ и $x \in \mathbb{R}$:

$$|\mathbb{P}(V_n + W_n < x) - \Phi(x)| \leq \sup_x |\mathbb{P}(V_n < x) - \Phi(x)| + \frac{\varepsilon}{\sqrt{2\pi}} + \mathbb{P}(|W_n| \geq \varepsilon). \quad (1.27)$$

Обозначим

$$U(T) = \hat{R}(T) - \hat{R}(T^{min}) = \sum_{i=1}^n H_i(T, T^{min}),$$

где

$$H_i(T, T^{min}) = F[X_i, T] - F[X_i, T^{min}].$$

В данном случае возьмем в качестве Z_n следующее выражение

$$Z_n = \frac{\hat{R}(T^{min}) - R(T^{min}) + U(T^F)}{\sigma^2 \sqrt{2n}},$$

тогда

$$V_n = \frac{\hat{R}(T^{min}) - R(T^{min})}{\sigma^2 \sqrt{2n}} \quad \text{и} \quad W_n = \frac{U(T^F)}{\sigma^2 \sqrt{2n}}.$$

Далее без ограничения общности будем считать $\sigma^2 = 1$. Для первого слагаемого из (1.27), применяя неравенство Берри-Эссеена, получаем следующую оценку:

$$\sup_x |\mathbb{P}(V_n < x) - \Phi(x)| \leq \frac{C}{\sqrt{n}}. \quad (1.28)$$

В ходе доказательств теоремы 1.6 и теоремы 1.7 было получено неравенство для $U(T^F)$:

$$\begin{aligned} \mathbb{P}\left(\frac{|U_{T^F}|}{\sqrt{n}} \geq \varepsilon\right) &\leq \mathbb{P}(T^F \leq T^1) + \\ &+ \mathbb{P}\left(\frac{\sup_{T \in [T^1, T^U]} |U(T) - \mathbb{E}U(T)| + \sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)|}{\sqrt{n}} > \varepsilon\right). \end{aligned} \quad (1.29)$$

Для первого слагаемого из (1.29) вновь воспользуемся оценкой из теоремы 1.4:

$$\sup_{\mu \in L_0[\eta_n]} \mathbb{P}(T^F < T^1) \leq 2n \exp\{-c \alpha_n \kappa_n \gamma_n^2\}, \quad (1.30)$$

В ходе доказательства предыдущих теорем также были получены следующие оценки:

$$\begin{aligned} \mathbb{P}\left(\frac{\sup_{T \in [T^1, T^U]} |U(T) - \mathbb{E}U(T)|}{\sqrt{n}} > \varepsilon\right) &\leq 2n^{1-\gamma} \exp\left\{-\frac{C_3 \varepsilon^2 n^\gamma}{\log n}\right\} + \\ &+ C_4 n^{1-\gamma} \exp\{-C_5 \varepsilon^2 n^\gamma\} + \\ &+ 2n \exp\left\{-\frac{C_6 \varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5}\right)^{3/2} n^{1-\gamma} + 2((T^U)^2 + \sigma^2)\sqrt{n}}\right\} + \\ &+ 2n \exp\left\{-\frac{C_7 \varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5}\right)^{1/2} n^{1-\gamma} + 2\sqrt{n}}\right\}, \end{aligned} \quad (1.31)$$

$$\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)| \leq C_8 n^{1-\gamma} \log n, \quad (1.32)$$

для $\gamma > 1/2$

$$\frac{\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)|}{\sqrt{n}} \rightarrow 0$$

при $n \rightarrow \infty$.

Заметим, что в выражении (1.31) первые два слагаемых не влияют на порядок, поэтому порядок выражения (1.31):

$$2n \exp \left\{ -\frac{C_6 \varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{3/2} n^{1-\gamma} + 2((T^U)^2 + \sigma^2)\sqrt{n}} \right\} + \\ + 2n \exp \left\{ -\frac{C_7 \varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{1/2} n^{1-\gamma} + 2\sqrt{n}} \right\}. \quad (1.33)$$

Возьмем за $\varepsilon_1(n)$ такую функцию, что $\varepsilon_1(n)\sqrt{n}$ будет превосходить оценку из неравенства (1.32):

$$\varepsilon_1(n) = C \cdot \log n \cdot n^{1/2-\gamma}.$$

Рассмотрим первое слагаемое из (1.33), учитывая, что $\gamma > \frac{1}{2}$, получаем следующую цепочку:

$$2n \exp \left\{ -\frac{C_6 \varepsilon^2 n}{2 \left(\log \frac{n}{(\log n)^5} \right)^{3/2} n^{1-\gamma} + 2((T^U)^2 + \sigma^2)\sqrt{n}} \right\} \sim \\ \sim 2n \exp \left\{ -\frac{C \varepsilon^2 n}{\left(\log \frac{n}{(\log n)^5} \right)^{3/2} n^{1-\gamma} + (\sigma^2 \log n + \sigma^2)\sqrt{n}} \right\} \sim \\ \sim 2n \exp \left\{ -\frac{C \varepsilon^2 n^{1/2}}{\log n} \right\} \sim 2n \exp \left\{ -C n^{3/2-2\gamma} \cdot \log n \right\}.$$

Проделав аналогичные рассуждения для второго слагаемого из (1.33), получаем, что выражение (1.31) будет иметь порядок

$$n \exp \left\{ -C n^{3/2-2\gamma} \cdot \log n \right\} + \frac{n \exp \left\{ -C n^{3/2-2\gamma} \cdot (\log n)^2 \right\}}{39}.$$

В этом случае для второго и третьего слагаемых из неравенства (1.27) имеем

$$\begin{aligned} \frac{\varepsilon(n)}{\sqrt{2\pi}} + \mathbb{P}(|W_n| \geq \varepsilon(n)) &\leq \frac{C \log n \cdot n^{1/2-\gamma}}{\sqrt{2\pi}} + n \exp\{-c \alpha_n \kappa_n \gamma_n^2\} + \\ &+ n \exp\{-C n^{3/2-2\gamma} \cdot \log n\} + n \exp\{-C n^{3/2-2\gamma} \cdot (\log n)^2\}. \end{aligned} \quad (1.34)$$

Таким образом, порядок последнего выражения для $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ будет равен $n^{1/2-\gamma} \cdot \log n$.

Теперь рассмотрим случай, когда $\gamma \in \left(\frac{3}{4}, 1\right)$. Для этого возьмем

$$\varepsilon_2(n) = C \cdot n^{-1/4} \cdot \log n.$$

В этом случае порядок второго и третьего слагаемых из неравенства (1.27) будет равен

$$\begin{aligned} \frac{\varepsilon(n)}{\sqrt{2\pi}} + \mathbb{P}(|W_n| \geq \varepsilon(n)) &\leq \frac{C \log n \cdot n^{-1/4}}{\sqrt{2\pi}} + 2 \exp\{-c \alpha_n \kappa_n \gamma_n^2\} + \\ &+ n \exp\{-C \log n\} + n \exp\{-C (\log n)^2\}. \end{aligned} \quad (1.35)$$

Следовательно, для случая $\gamma \in \left(\frac{3}{4}, 1\right)$ порядок выражения (1.35) будет равен $n^{-1/4} \cdot \log n$.

Таким образом, порядок итоговой оценки скорости сходимости распределения оценки риска к нормальному закону при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ равен $C \cdot n^{1/2-\gamma} \cdot \log n$, а при $\gamma \in \left(\frac{3}{4}, 1\right)$ равен $C \cdot n^{-1/4} \cdot (\log n)$.

Теорема доказана. □

Далее рассмотрим теорему об оценке скорости сходимости распределения оценки риска к нормальному закону в случае класса разреженности $L_p(\boldsymbol{\eta}_n)$.

Теорема 1.9 Пусть $\boldsymbol{\mu} \in L_p(\boldsymbol{\eta}_n)$, $0 < p < 2$, где $\boldsymbol{\eta}_n^p \in [n^{-1}(\log n)^5, n^{-\gamma}]$, $1/2 < \gamma < 1$. Пусть T_F — FDR-порог с управляющим параметром $\alpha_n \rightarrow 0$ и $\frac{\alpha_n \kappa_n \gamma_n^2}{\log n} \rightarrow \infty$ при $n \rightarrow \infty$, где κ_n и γ_n определены в (1.10). Тогда для $x \in \mathbb{R}$ выполнены следующие неравенства:

для $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$

$$\sup_x \left| \mathbb{P}\left(\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} < x\right) - \Phi(x) \right| \leq C \cdot n^{1/2-\gamma} \cdot \log n^{1+p/2},$$

для $\gamma \in \left(\frac{3}{4}, 1\right)$

$$\sup_x \left| \mathbb{P}\left(\frac{\hat{R}(T^F) - R(T^{min})}{\sigma^2 \sqrt{2n}} < x\right) - \Phi(x) \right| \leq C \cdot n^{-1/4} \cdot \log n,$$

где в обоих случаях C - это константы, вообще говоря, разные.

Доказательство. Доказательство по большей части аналогично доказательству теоремы 1.8. Коротко приведем основные отличия.

Для первого слагаемого из (1.29) необходимо будет использовать оценку, приведенную в теореме 1.4, для случая разреженности $L_p(\eta_n)$:

$$\sup_{\mu \in L_p(\eta_n)} \mathbb{P}(T^F < T^1) \leq 2n \exp\{-c \alpha_n \kappa_n \gamma_n^2\}. \quad (1.36)$$

В силу определения класса $L_p(\eta_n)$ число слагаемых в $S_1(T)$ не превосходит $n_1 \approx C \eta_n^p n (\log \eta_n^p)^{p/2}$, что эквивалентно $n_1 \approx C \eta_n^p n (\log n)^{p/2}$. Тогда оценка (1.32) в случае класса разреженности L_p примет следующий вид:

$$\sup_{T \in [T^1, T^U]} |\mathbb{E}U(T)| \leq C \eta_n^p n (\log n)^{1+p/2} \leq C n^{1-\gamma} (\log n)^{1+p/2}. \quad (1.37)$$

Возьмем за $\varepsilon_1(n)$ такую функцию, что $\varepsilon_1(n) \sqrt{n}$ будет превосходить оценку из неравенства (1.37):

$$\varepsilon_1(n) = C \cdot (\log n)^{1+p/2} \cdot n^{1/2-\gamma}.$$

Тогда выражение (1.31) будет иметь порядок

$$2n \exp \left\{ -\frac{C n^{3/2-2\gamma} (\log n)^{p+2}}{\left(\log \frac{n}{(\log n)^5}\right)^{3/2} + ((TU)^2 + \sigma^2)} \right\} + 2n \exp \left\{ -\frac{C n^{3/2-2\gamma} (\log n)^{p+2}}{\left(\log \frac{n}{(\log n)^5}\right)^{1/2} + 1} \right\}.$$

В этом случае для второго и третьего слагаемых из неравенства (1.27) имеем

$$\begin{aligned} \frac{\varepsilon(n)}{\sqrt{2\pi}} + \mathbb{P}(|W_n| \geq \varepsilon(n)) &\leq \frac{C(\log n)^{1+p/2} \cdot n^{1/2-\gamma}}{\sqrt{2\pi}} + 2n \exp\{-c \alpha_n \kappa_n \gamma_n^2\} + \\ &+ 2n \exp\left\{-\frac{C n^{3/2-2\gamma} (\log n)^{p+2}}{\left(\log \frac{n}{(\log n)^5}\right)^{3/2} + ((TU)^2 + \sigma^2)}\right\} + 2n \exp\left\{-\frac{C n^{3/2-2\gamma} (\log n)^{p+2}}{\left(\log \frac{n}{(\log n)^5}\right)^{1/2} + 1}\right\}. \end{aligned} \quad (1.38)$$

Таким образом, порядок последнего выражения для $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ будет равен $(\log n)^{p+2} \cdot n^{1/2-\gamma}$.

Теперь рассмотрим случай, когда $\gamma \in \left(\frac{3}{4}, 1\right)$. Для этого возьмем

$$\varepsilon_2(n) = C \cdot n^{-1/4} \cdot (\log n)^{p+2}.$$

В этом случае порядок второго и третьего слагаемых из неравенства (1.27) будет равен $C \cdot n^{-1/4} \cdot (\log n)$.

Таким образом, порядок итоговой оценки при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ равен $C \cdot n^{1/2-\gamma} \cdot \log n$, а при $\gamma \in \left(\frac{3}{4}, 1\right)$ равен $C \cdot n^{-1/4} \cdot (\log n)^2$, что завершает доказательство теоремы. \square

Полученные оценки скорости сходимости могут быть использованы в асимптотических методах статистики, например, при построении доверительных интервалов необходимо знать меру близости исходного и предельного распределений.

Глава 2. Обращение линейных однородных операторов

2.1 Линейные однородные операторы

Рассмотрим постановку задачи, в которой необходимо оценить исходные данные $f \in L^2(\mathbb{R})$, но при этом для наблюдения доступны данные, подвергнутые действию линейного преобразователя K , помимо этого в наблюдаемых данных присутствует шум. Подобного рода задачи можно встретить при анализе результатов компьютерной томографии, физики плазмы и телекоммуникационного трафика. Модель можно записать формально в следующем виде:

$$X_i = (Kf)_i + \varepsilon_i, \quad i = 1, \dots, n,$$

где $f \in L^2(\mathbb{R})$ — функция исходных данных, ε_i — независимые нормально распределенные случайные величины с нулевым математическим ожиданием и дисперсией σ^2 , а K — линейное преобразование.

Для того, чтобы выразить исходные данные f , можно было бы применить обратный оператор K^{-1} , однако такой оператор может либо не существовать, либо не быть ограниченным, что приводит к так называемым некорректно поставленным статистическим задачам, для решения которых используются методы регуляризации в сочетании с различными методами разложения. Одним из видов разложений является применение аппарата вейвлет-анализа. В частности, рассмотрим метод вейвлет-вейвлет-разложения, который был предложен в работе [10].

В диссертации будет рассматриваться случай, когда оператор K является однородным (о причинах выбора именно таких операторов будет сказано далее). Оператор K называется однородным с показателем $\beta \geq 0$, если для

любого $a > 0$ выполнено

$$K[f(ax)] = a^{-\beta}(Kf)[ax]. \quad (2.1)$$

В случае, когда показатель $\beta \geq 0$ задача обращения оператора K будет являться некорректно поставленной статистической задачей. При $\beta < 0$ задача будет считаться корректной и такой случай не будет рассматриваться в рамках диссертации.

Можно показать, что из однородности линейного преобразования K с показателем β вытекает однородность преобразования K^{-1} с показателем $-\beta$. Обозначим $g = Kf$, тогда

$$\begin{aligned} K^{-1}[g(ax)] &= K^{-1}[(Kf)[ax]] = a^\beta K^{-1}[K[f(ax)]] = \\ &= a^\beta f(ax) = a^\beta (K^{-1}g)[ax]. \end{aligned}$$

Примером линейного однородного оператора является преобразование Гильберта [53]

$$H(u)(t) = \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} \frac{u(\tau)}{t - \tau} d\tau,$$

которое получило широкое применение при цифровой обработке сигналов, к примеру, для определения мгновенной частоты, обработки радиолокационных сигналов, получении огибающей и т.д. Также преобразование Гильберта тесно связано с понятием аналитического сигнала - представления сигнала в виде набора комплексных чисел, что позволяет разложить исходный сигнал на амплитудную и фазовую составляющие. Преобразование Гильберта и его обобщения применяются при анализе, модуляции и демодуляции цифровых сигналов. Помимо преобразования Гильберта также широко известны другие примеры линейных однородных операторов такие, как преобразование Абеля [8]

$$F(y) = 2\frac{1}{\pi} \text{p.v.} \int_y^\infty \frac{f(r)r}{\sqrt{r^2 - y^2}} dr,$$

и потенциал Рисса [51]

$$V_\alpha(x) = \int_{\mathbb{R}^d} \frac{d\mu(y)}{|x - y|^\alpha}, \quad \alpha > 0,$$

где μ - положительная борелевская мера с компактным носителем на евклидовом пространстве \mathbb{R}^d , $d \geq 2$, $|x - y|$ - расстояние между точками в пространстве \mathbb{R}^d .

Также примерами линейных однородных преобразований являются операторы интегрирования и дифференцирования.

2.2 Вейвлет-разложение и вейвлет-коэффициенты

Вейвлет-разложение подразумевает под собой представление функции в виде колебаний, локализованных по частоте и времени. Это выгодно отличает вейвлет-преобразования от преобразований Фурье [45], которые не дают информации о том, на какой частоте и в какой момент времени в исходных данных присутствовали колебания. Также популярность вейвлет-разложения обусловлена тем, что в отличие от рядов Фурье, вейвлет-анализ хорошо применим не только к пространственно однородным функциям, но и к функциям, имеющим локальные сингулярности.

Аппарат вейвлет-анализа базируется на понятии «материнской» вейвлет-функции $\psi(x) \in L^2(\mathbb{R})$, которая представляет собой затухающие на прямой колебания. Для того, чтобы ψ была вейвлет-функцией необходимо выполнение условия допустимости:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

Из этого равенства следует, что любая вейвлет-функция осциллирует вокруг оси x и имеет хотя бы один нулевой момент.

Вейвлет-функцию можно задавать таким образом, чтобы она удовлетворяла необходимым условиям, в частности была дифференцируема необходимое

количество раз или имела необходимое количество нулевых моментов M_0 , т.е.

$$\int_{-\infty}^{\infty} t^k \psi(t) dt = 0, \quad k \in [0, M_0 - 1].$$

Наиболее простым примером вейвлет-функции является вейвлет-функция Хаара, которая имеет только один нулевой момент:

$$\psi(x) = \begin{cases} 1 & \text{при } 0 \leq x < \frac{1}{2}, \\ -1 & \text{при } \frac{1}{2} \leq x < 1, \\ 0 & \text{иначе.} \end{cases} \quad (2.2)$$

Другим примером являются вейвлеты Добеши [18], которые представляют собой семейство вейвлет-функций с компактным носителем, т.е. вне конечного отрезка на прямой значение вейвлет-функции тождественно равно нулю. Параметром вейвлетов Добеши является число нулевых моментов, в частности, вейвлетом Добеши первого порядка является рассмотренный выше вейвлет Хаара. С ростом количества нулевых моментов растет гладкость вейвлет-функции Добеши, а также увеличивается носитель. Отметим, что вейвлет Хаара является единственной вейвлет-функцией Добеши, которая может быть записана в аналитическом виде. Этот факт не мешает широкому применению вейвлетов Добеши на практике, т.к. они могут быть вычислены с любой наперед заданной точностью.

В качестве еще одного примера рассмотрим вейвлет Мейера [47]. Преобразование Фурье вейвлет-функции Мейера $\psi(x)$ имеет вид

$$\hat{\psi}(x) = \begin{cases} \frac{1}{\sqrt{2}} \cdot e^{-\frac{iw}{2}} \cdot \hat{h}\left(\frac{w}{2} + \pi\right) & \text{при } \frac{2\pi}{3} \leq x < \frac{4\pi}{3}, \\ \frac{1}{\sqrt{2}} \cdot e^{-\frac{iw}{2}} \cdot \hat{h}\left(\frac{w}{4}\right) & \text{при } \frac{4\pi}{3} \leq x < \frac{8\pi}{3}, \\ 0 & \text{иначе,} \end{cases}$$

где

$$\hat{h}(w) = \begin{cases} \sqrt{2} & \text{при } -\frac{\pi}{3} \leq w \leq \frac{\pi}{3}, \\ 0 & \text{при } \frac{2\pi}{3} \leq |x| < \pi. \end{cases}$$

Для значений $\frac{\pi}{3} \leq |w| \leq \frac{2\pi}{3}$ функция $\hat{h}(w)$ может иметь любое значение такое, что выполнено условие

$$\left| \hat{h}(w) \right|^2 + \left| \hat{h}(w + \pi) \right|^2 = 2.$$

Доопределить функцию $\hat{h}(w)$ можно так, чтобы получить необходимое число непрерывных производных в точках $|w| = \frac{\pi}{3}$ и $|w| = \frac{2\pi}{3}$. Эта гладкость преобразования Фурье вейвлет-функции Мейера обеспечивает быстрое асимптотическое затухание во времени. График вейвлет-функции Мейера приведен на рисунке (2.5).

Еще одним примером вейвлет-функции является МНАТ-вейвлет (Mexican HAT), называемый также “мексиканской шляпой” из-за сходства её графика с сомбреро (см. рисунок (2.5)). Вейвлет-функция МНАТ получается путем двукратного дифференцирования функции Гаусса [1]:

$$\psi(x) = \frac{2}{\pi^{1/4}\sqrt{3}}(1 - x^2)e^{-\frac{x^2}{2}}.$$

Такая вейвлет-функция имеет два нулевых момента.

Так же, как и для преобразований Фурье, можно рассматривать интегральную форму (непрерывное вейвлет-преобразование) и разложение в ряд (двоичное вейвлет-преобразование). В диссертации будем рассматривать двоичное вейвлет-преобразование, которое зачастую в литературе называют дискретным. Пусть имеется исходная функция $f \in L^2(\mathbb{R})$, тогда ее вейвлет-разложение выглядит следующим образом:

$$f = \sum_{j,l \in \mathbb{Z}} \langle f, \psi_{j,l} \rangle \psi_{j,l}, \quad (2.3)$$

где $\psi_{j,l}(x) = 2^{j/2}\psi(2^j x - l)$ образуют ортонормированный базис в $L^2(\mathbb{R})$. При этом набор коэффициентов $\{\langle f, \psi_{j,l} \rangle\}$ называется двоичным вейвлет-преобразованием функции f и для любой функции $f \in L^2(\mathbb{R})$ можно записать представление

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{l=-\frac{\infty}{4}}^{\frac{\infty}{4}} \langle f, \psi_{j,l} \rangle \psi_{j,l}(x).$$

С вейвлет-функцией ψ тесно связана так называемая масштабирующая функция φ [45] такая, что

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1.$$

Используя масштабирующую функцию φ , можно записать следующее разложение функции f до уровня j_0 :

$$f(x) = \sum_{l=-\infty}^{\infty} \langle f, \varphi_{j_0, l} \rangle \varphi_{j_0, l} + \sum_{j \geq j_0} \sum_{l=-\infty}^{\infty} \langle f, \psi_{j, l} \rangle \psi_{j, l},$$

где $j_0 \in \mathbb{Z}$. Коэффициенты $\langle f, \varphi_{j_0, l} \rangle$ из последнего разложения называются коэффициентами аппроксимации, а $\langle f, \psi_{j, l} \rangle$ - коэффициентами деталей. Примеры некоторых вейвлетов и соответствующих им масштабирующих функций приведены на рисунках (2.1), (2.2), (2.3), (2.4). Стоит отметить, что для некоторых вейвлет-функций ψ масштабирующая функция может не существовать, такие примеры в частности приведены в работе [44].

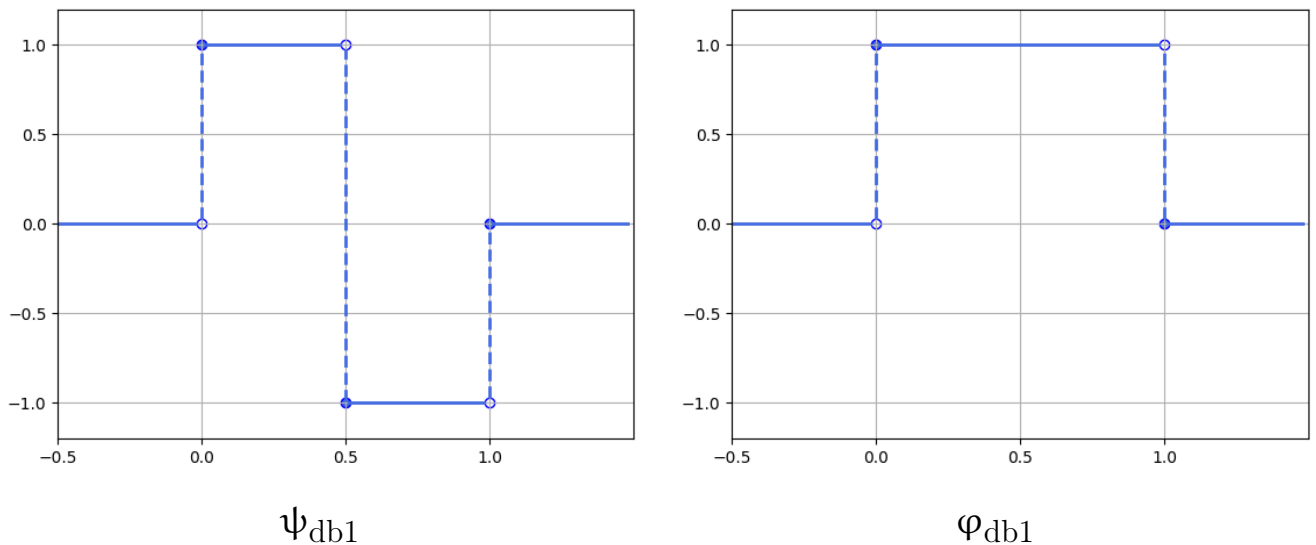


Рисунок 2.1 — Вейвлет-функция Добеши 1 порядка (вейвлет Хаара) ψ_{db1} и соответствующая ей масштабирующая функция φ_{db1}

Отметим, что для оконного преобразования Фурье частотно-временное окно имеет фиксированную ширину при анализе различных диапазонов частот. Такой подход создает определенные неудобства при рассмотрении сигнала, который имеет как участки с высокочастотными данными, так и участки с низкочастотными данными. Как известно, частота прямо пропорциональна числу

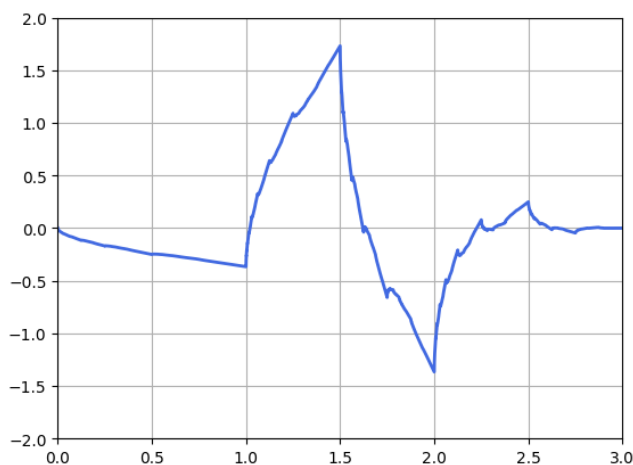
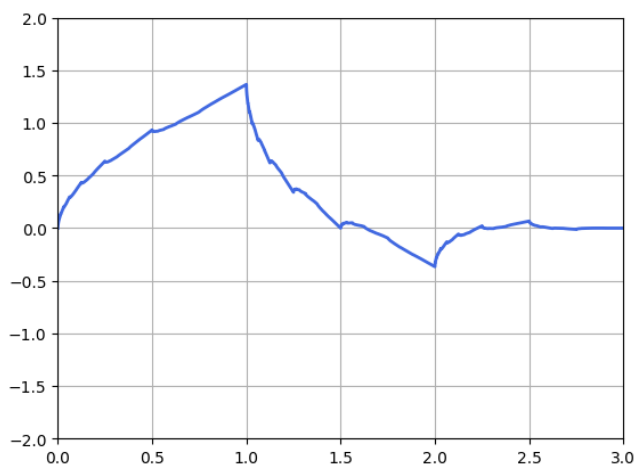

 ψ_{db2}

 φ_{db2}

Рисунок 2.2 — Вейвлет-функция Добеши 2 порядка ψ_{db2} и соответствующая ей масштабирующая функция φ_{db2}

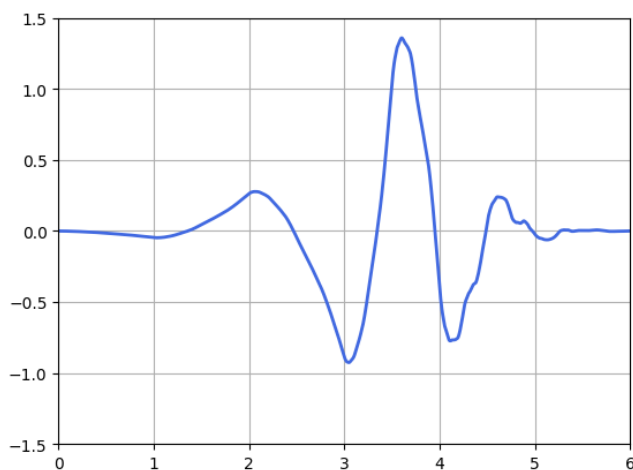
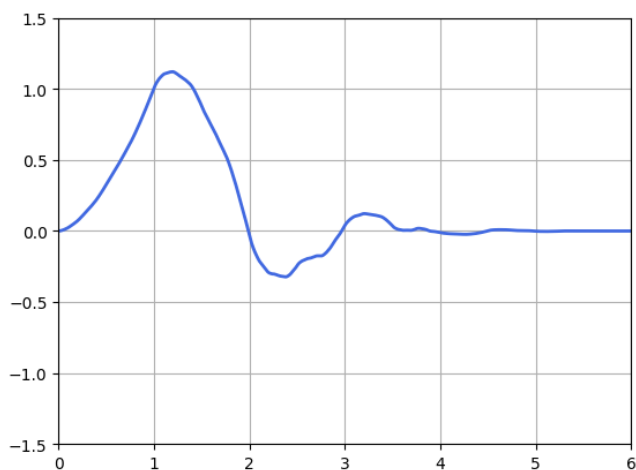

 ψ_{db4}

 φ_{db4}

Рисунок 2.3 — Вейвлет-функция Добеши 4 порядка ψ_{db4} и соответствующая ей масштабирующая функция φ_{db4}

периодов в единицу времени, а следовательно для анализа высокочастотных сигналов необходима более узкая ширина окна, а для низкочастотных - более широкая. В отличие от преобразования Фурье вейвлет-преобразование имеет изменяющееся частотно-временное окно, которое сужается при рассмотрении участка с высокочастотными составляющими сигнала и расширяется на низкочастотных областях. При рассмотрении дискретного вейвлет-преобразования (2.3) за положение частотно-временного окна отвечает параметр l , а за размер

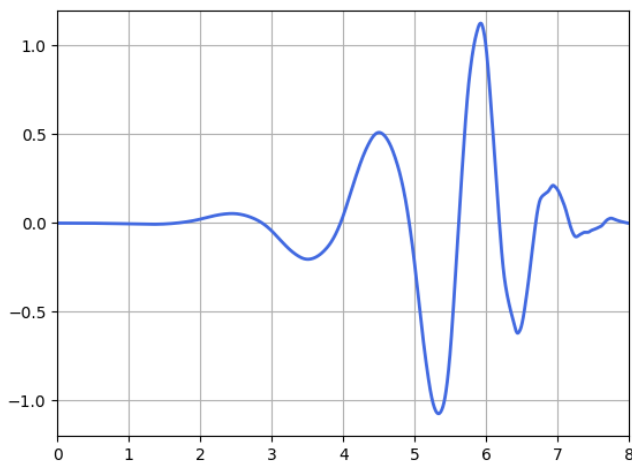
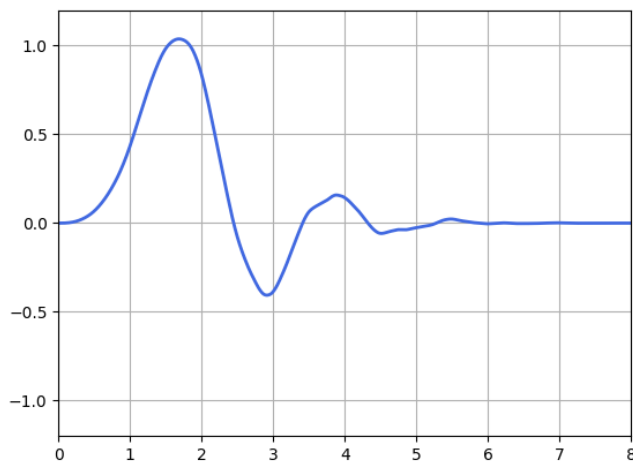

 ψ_{db6}

 φ_{db6}

Рисунок 2.4 — Вейвлет-функция Добеши 6 порядка ψ_{db6} и соответствующая ей масштабирующая функция φ_{db6}

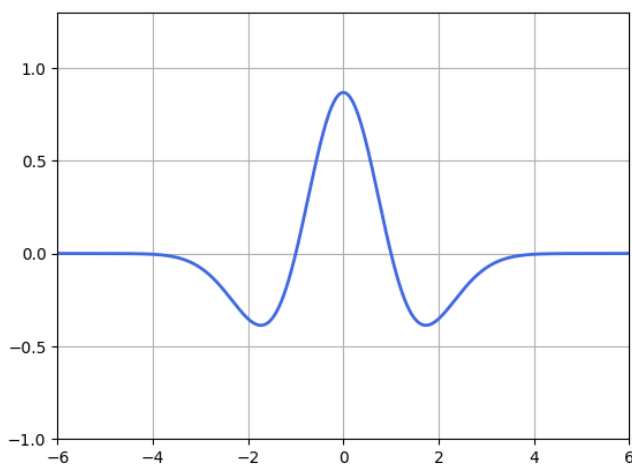
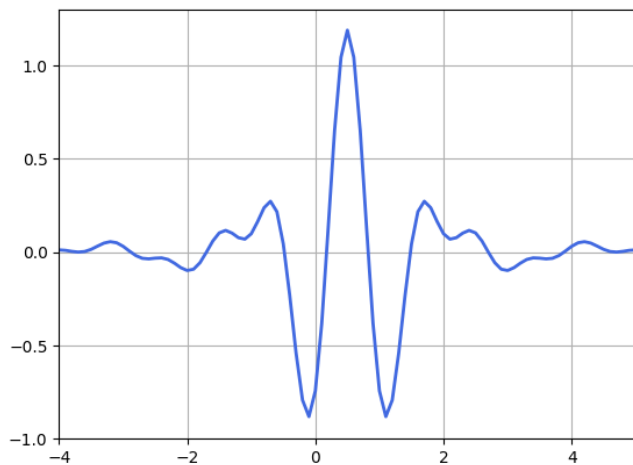

 ψ_{mhat}

 ψ_{meyer}

Рисунок 2.5 — Примеры вейвлетов ψ : вейвлет “мексиканская шляпа” и вейвлет Мейера

деталей этого окна - параметр j . Поэтому параметр j называют уровнем разложения или “масштабом”, а индекс l — “сдвигом”.

2.3 Вейвлет-разложение в случае обращения линейного однородного оператора

В методе вейвлет-вейвлет разложения результат линейного преобразования функции f раскладывается в ряд из сдвигов и растяжений некоторой вейвлет-функции ψ :

$$Kf = \sum_{j,l \in \mathbb{Z}} \langle Kf, \psi_{j,l} \rangle \psi_{j,l},$$

где $\psi_{j,l}(x) = 2^{j/2} \psi(2^j x - l)$, причем базис, по которому раскладывается преобразование Kf , получается ортонормированным. Этот важный факт позволяет сделать вывод о том, что зашумленные коэффициенты разложения по-прежнему будут являться независимыми случайными величинами, что существенно облегчает дальнейшие исследования. Для сравнения в методе вейвлет-вейвлет, который также используется для решения задач обращения линейных однородных операторов (см. работы [23; 41; 42]), разложения вейвлет-коэффициенты не будут являться независимыми

Потребуем, чтобы для оператора K существовали такие константы $\beta_{j,l}$, чтобы последовательность функций $u_{j,l} = K^{-1} \psi_{j,l} / \beta_{j,l}$ образовывала устойчивый базис, т.е. существовали такие константы $0 < A \leq B < \infty$, что

$$A \sum_{j,l} c_{j,l}^2 \leq \left\| \sum_{j,l} c_{j,l} u_{j,l} \right\|_{L^2}^2 \leq B \sum_{j,l} c_{j,l}^2. \quad (2.4)$$

Функции $u_{j,l}$, для которых выполнено $\langle f, u_{j,l} \rangle = \langle Kf, \psi_{j,l} \rangle$, называют «вейвглетами». Очевидно, что не для всех операторов K найдется такое семейство вейвлетов $u_{j,l}$, что будет выполнено условие (2.4). Поэтому в дальнейшем будем рассматривать только однородные преобразования K , для которых гарантируется выполнение (2.4) при соблюдении некоторых условий регулярности для материнской вейвлет-функции ψ [10].

Пусть $\beta_{j,l} = \|K^{-1}\psi_{j,l}\|_{L^2}$. В работе [10] было показано, что

$$\beta_{j,l} = 2^{\beta \cdot j} \beta_{0,0}.$$

Тогда разложение исходной функции f представимо в виде

$$F = \sum_{j,l \in \mathbb{Z}} 2^{\beta j} \beta_{0,0} \langle Kf, \psi_{j,l} \rangle u_{j,l}. \quad (2.5)$$

Заметим, что последовательность функций $\{u_{j,l}\}$ не образует ортонормированный базис, однако при выполнении некоторых условий гладкости для $K^{-1}\psi$ и $K^*\psi$, которые были приведены в [42], последовательность функций $\{u_{j,l}\}$ является устойчивым базисом.

Таким образом, в методе вейглет-вейвлет-разложения происходит разложение наблюдаемого преобразованного сигнала Kf по ортонормированному базису, затем применяется пороговая обработка к коэффициентам разложения и наконец исходные незашумленные данные f выражаются через обновленные коэффициенты и вейглеты.

Аппроксимация функции Kf записывается в виде суммы из сдвигов и растяжений вейвлет-функции ψ :

$$Kf \approx \sum_{j=0}^{J-1} \sum_{l=0}^{2^j-1} \langle Kf, \psi_{j,l} \rangle \psi_{j,l},$$

где $\psi_{j,l}(x) = 2^{j/2} \psi(2^j x - l)$.

Следовательно, аппроксимация исходной функции f представима в виде

$$F = \sum_{j=0}^{J-1} \sum_{l=0}^{2^j-1} 2^{\beta j} \beta_{0,0} \langle Kf, \psi_{j,l} \rangle u_{j,l}, \quad (2.6)$$

где $u_{j,l} = K^{-1}\psi_{j,l}/\beta_{j,l}$.

Как правило, на практике функция f задана дискретно и имеет конечный носитель. Поэтому перейдем к рассмотрению функции f и ее линейного преобразования Kf в точках i/N , где $N = 2^J$ и $i = 1, \dots, N$. Дискретное вейвлет-преобразование представляет собой умножение вектора значений функции

на ортогональную матрицу, порождаемую вейвлет-функцией ψ [45]. Обозначим через $\mu_{j,l}$ соответствующие коэффициенты, полученные после дискретного вейвлет-преобразования. Коэффициенты $\mu_{j,l}$ образуют вектор $\mu \in \mathbb{R}^d$, $d = 2^J$.

Для каждого уровня $j \in [0, J-1]$ рассмотрим упорядоченные по абсолютной величине значения $\mu_{j,(k)}$, $k = 1, \dots, 2^j$:

$$|\mu_{j,(1)}| \geq \dots \geq |\mu_{j,(2^j)}|.$$

Обозначим $\eta = (\eta_0, \eta_1, \dots, \eta_{J-1})$, где $\eta_j \in [0, 1]$, $j \in [0, J-1]$. Как и в главе 1 будем рассматривать векторы μ , которые принадлежат классу разреженности L_p для $0 < p < 2$:

$$L_p(\eta) = \{\mu \in \mathbb{R}^d : |\mu_{j,(k)}| \leq \eta_j \cdot 2^{j/p} \cdot k^{-1/p} \text{ для всех } k = \overline{1, 2^j}, j = \overline{0, J-1}\}.$$

2.4 Пороговая обработки и оценка среднеквадратичного риска в случае обращения линейных однородных операторов

Для пороговой обработки используются функции жесткой (1.2) и мягкой (1.3) пороговой обработки, определенные в главе 1. Однако в случае обращения линейного однородного оператора для каждого масштаба разложения возникает корректировка значения среднеквадратичного риска, поэтому необходимо вычислять значения порога на каждом уровне j . В таком случае аргументами для пороговых функций будут $X_{j,k}^W$ и T_j , для каждого $j \in [0, J-1]$, где $X_{j,k}^W$ — зашумленные вейвлет-коэффициенты из разложения (2.5) (см. работу [9]).

Пороговое значение T_j , вычисляемое для каждого уровня разложения j , в случае использования FDR-метода выбирается следующим образом: по исходной выборке строится вариационный ряд убывающих по абсолютному значению

величин

$$|X|_j^{(1)} \geq |X|_j^{(2)} \geq \dots \geq |X|_j^{(2^j)},$$

и $|X|_j^{(k)}$ сравнивается с квантилями Гауссова распределения $t_k = \sigma z(\alpha/2 \cdot k/2^j)$. Пусть k_F — наибольший индекс k , для которого $|X|_j^{(k)} \geq t_k$, тогда выбирается порог $T_j^F = t_{k_F}$.

Определение среднеквадратичного риска в случае дискретного вейвлет-разложения несколько отличается от определения, приводимого в главе 1, а именно:

$$R(\mathbf{T}) = \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} \beta_{j,k}^2 \mathbb{E} (\rho(X_{j,k}^W, T_j) - \mu_{j,k})^2, \quad (2.7)$$

где $\mathbf{T} = (T_0, T_1, \dots, T_{J-1})$, а ρ — функция используемой пороговой обработки.

Обозначим значение порога, при котором риск достигает минимума, \mathbf{T}^{min} :

$$\mathbf{T}^{min} : R(\mathbf{T}^{min}) = \min_{\mathbf{T}} R(\mathbf{T}),$$

где $\mathbf{T}^{min} = (T_0^{min}, T_1^{min}, \dots, T_{J-1}^{min})$. Как уже отмечалось в главе 1 в выражении $R(\mathbf{T})$ присутствуют неизвестные величины $\mu_{j,k}$, из-за чего вычислить значения $R(\mathbf{T})$ и \mathbf{T}^{min} на практике нельзя. Поэтому будем рассматривать оценку риска, которая вычисляется на основе только наблюдаемых данных и определяется следующим выражением:

$$\hat{R}(\mathbf{T}) = \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} 2^{2\beta_j} \beta_{0,0} F[X_{j,k}^W, T_j], \quad (2.8)$$

где в случае жесткой пороговой обработки

$$F[X_{j,k}^W, T_j] = ((X_{j,k}^W)^2 - \sigma^2) \cdot \mathbf{1}(|X_{j,k}^W| \leq T_j) + \sigma^2 \cdot \mathbf{1}(|X_{j,k}^W| > T_j)$$

и в случае мягкой пороговой обработки

$$F[X_{j,k}^W, T_j] = ((X_{j,k}^W)^2 - \sigma^2) \cdot \mathbf{1}(|X_{j,k}^W| \leq T_j) + (\sigma^2 + T_j^2) \cdot \mathbf{1}(|X_{j,k}^W| > T_j).$$

При применении методов пороговой обработки для решения обратных статистических задач обозначим универсальный порог $\mathbf{T}^U : T_j^U = \sigma \sqrt{2 \log 2^j}$. Как

уже было отмечено в главе 1, в силу того, что универсальный порог в определенном смысле близок к максимальному, всюду далее будем считать, что $T_j \leq T_j^U$, $j = 0, \dots, J - 1$.

В последующих параграфах главы 2 будут исследованы некоторые статистические свойства оценки среднеквадратичного риска при выборке FDR-порога по алгоритму Бенжамини-Хочберга. Аналогичные свойства для векторного случая были рассмотрены в главе 1, далее приведем доказательства этих асимптотических свойств оценки риска для случая обращения линейного однородного оператора.

2.5 Сильная состоятельность оценки риска

Как уже было отмечено в главе 1, в силу того, что на практике вычислить риск (2.7) и порог \mathbf{T}^{min} , минимизирующий этот риск, не представляется возможным из-за вхождения неизвестных компонент незашумленного сигнала, имеет смысл перейти к рассмотрению оценки риска. Оценка риска, определяемая выражением (2.8), в свою очередь не зависит от неизвестных компонент исходного сигнала, а значит может быть вычислена на практике. Более того, с помощью оценки риска можно сравнивать между собой различные значения порогов с целью нахождения в некотором смысле наилучшего порогового значения.

Для удобства дальнейшего изложения введем ряд обозначений:

$$\gamma_j = \frac{1}{\log \log(2^j)}, \quad \tau_j = \sigma (2 \log \eta_j^{-p})^{1/2}, \quad \kappa_j = \frac{2^j \eta_j^p \tau_j^{-p}}{1 - \alpha_j - \gamma_j} \quad \text{для } L_p(\eta_j). \quad (2.9)$$

Следующая теорема была доказана в статье [69].

Теорема 2.1. Пусть $\mu \in L_p(\boldsymbol{\eta})$, $\eta_j^p \in [2^{-j}(\log 2^j)^5, 2^{-j\gamma}]$, $0 < \gamma < 1$, $j \in [0, J-1]$. Пусть \mathbf{T}^F — вектор FDR-порогов с управляющими параметрами

$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{J-1})$ такими что, $\alpha_j \rightarrow 0$, $\frac{\alpha_j \kappa_j \gamma_j^2}{\log 2^j} \rightarrow \infty$ при $j \geq tJ$ для некоторого $0 < t < 1$ и $J \rightarrow \infty$, где κ_j и γ_j определены в (2.9). Тогда при $J \rightarrow \infty$

$$\frac{\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min})}{2^{J(1+2\beta)}} \rightarrow 0 \text{ п. в.}$$

Доказательство. Не ограничивая общности будем считать $\sigma^2 = 1$. Для удобства дальнейшего доказательства введем следующие обозначения:

$$R(\mathbf{T}) = \sum_{j=0}^{J-1} R_j(T_j), \quad \hat{R}(\mathbf{T}) = \sum_{j=0}^{J-1} \hat{R}_j(T_j)$$

Заметим, что

$$\left| \frac{\sum_{j=0}^{Jt-1} (\hat{R}_j(T_j^F) - R_j(T_j^{min}))}{2^{J(1+2\beta)}} \right| > \frac{\varepsilon}{J} \rightarrow 0 \text{ п.в. при } J \rightarrow \infty,$$

т.к. слагаемые в $\hat{R}_j(T_j) - R_j(T_j)$ ограничены по модулю величиной $2\beta_{0,0} \cdot 2^{2\beta j} \left((T_j^U)^2 + 1 \right)$ и количество слагаемых в сумме мало по сравнению с 2^J .

Следовательно далее будем рассматривать только часть суммы по j от Jt до $J - 1$. Справедлива следующая оценка

$$\begin{aligned} & \mathbb{P} \left(\left| \frac{\sum_{j=Jt}^{J-1} (\hat{R}_j(T_j^F) - R_j(T_j^{min}))}{2^{J(1+2\beta)}} \right| > \frac{\varepsilon}{J} \right) \leq \\ & \leq \mathbb{P} \left(\left| \frac{\sum_{j=Jt}^{J-1} (\hat{R}_j(T_j^F) - R_j(T_j^F))}{2^{J(1+2\beta)}} \right| > \frac{\varepsilon}{2} \right) + \mathbb{P} \left(\left| \frac{\sum_{j=Jt}^{J-1} (R_j(T_j^F) - R_j(T_j^{min}))}{2^{J(1+2\beta)}} \right| > \frac{\varepsilon}{2} \right). \end{aligned} \tag{2.10}$$

Введем обозначение:

$$\frac{\hat{R}_j(T_j) - R_j(T_j)}{2^{J(1+2\beta)}} = \frac{2^{2\beta j} \beta_{0,0}}{2^{J(1+2\beta)}} \sum_{k=1}^{2^j} Y_k(T_j) = 2^{2\beta j} \beta_{0,0} \cdot Z_j(T_j),$$

где $Y_k(T_j)$ - независимые случайные величины с нулевым математическим ожиданием. Для $|Y_k(T_j)|$ можно показать, что:

$$\begin{aligned} |Y_k(T_j)| &\leq \left| 1 - 2 \cdot \mathbf{1}((X_{j,k}^W)^2 \leq T_j^2) + (X_{j,k}^W)^2 \wedge T_j^2 - \mathbb{E}(\hat{\mu}_{j,k} - \mu_{j,k})^2 \right| \leq \\ &\leq \left| 1 - 2 \cdot \mathbf{1}((X_{j,k}^W)^2 \leq T_j^2) \right| + \left| (X_{j,k}^W)^2 \wedge T_j^2 - \mathbb{E}(\hat{\mu}_{j,k} - \mu_{j,k})^2 \right| \leq 2 + T_j^2. \end{aligned}$$

Рассмотрим первое слагаемое из (2.10) и запишем его с помощью введенных выше обозначений:

$$\begin{aligned} \mathbb{P} \left(\sum_{j=Jt}^{J-1} 2^{2\beta j} \beta_{0,0} \cdot Z_j(T_j^F) > \frac{\varepsilon}{2} \right) &\leq \\ &\leq \sum_{j=Jt}^{J-1} \mathbb{P} \left(2^{2\beta j} \beta_{0,0} \cdot Z_j(T_j^F) > \frac{\varepsilon}{2J}, T_j^F \in [\tau_j, T_j^U] \right) + \\ &\quad + \sum_{j=Jt}^{J-1} \mathbb{P} (T_j^F \notin [\tau_j, T_j^U]). \quad (2.11) \end{aligned}$$

Далее рассмотрим каждую сумму из последнего выражения отдельно. Заметим, что первое слагаемое из (2.11) не превосходит

$$\begin{aligned} \sum_{j=Jt}^{J-1} \mathbb{P} \left(2^{2\beta j} \beta_{0,0} \cdot Z_j(T_j^F) > \frac{\varepsilon}{2J}, T_j^F \in [\tau_j, T_j^U] \right) &\leq \\ &\leq \sum_{j=Jt}^{J-1} \mathbb{P} \left(\sup_{T_j \in [\tau_j, T_j^U]} 2^{2\beta j} \beta_{0,0} \cdot Z_j(T_j) > \frac{\varepsilon}{2J} \right). \end{aligned}$$

Тогда для $|Z_j(T_j)|$ можем записать неравенство Хёффдинга [36] для фиксированного T_j :

$$\begin{aligned} \mathbb{P} \left(|2^{2\beta j} \cdot Z_j(T_j)| \geq \frac{\varepsilon}{J} \right) &= \mathbb{P} \left(\left| \sum_{k=1}^{2^j} Y_k(T_j) \right| \geq \frac{2^{J(1+2\beta)} \cdot \varepsilon}{J \cdot 2^{2\beta j}} \right) \leq \\ &\leq 2 \exp \left\{ -\frac{2\varepsilon^2 \cdot 2^{2J(1+2\beta)}}{J^2 \cdot 2^{4\beta j} \cdot 2^j \cdot (2 + T_j^2)^2} \right\} = 2 \exp \left\{ -\frac{2^{2J(1+2\beta)-4\beta j-j+1} \cdot \varepsilon^2}{J^2 \cdot (2 + T_j^2)^2} \right\}. \end{aligned}$$

Возьмем $T < T'$ и обозначим $N_j(T, T') = \#\{k : T < |X_{j,k}^W| \leq T'\}$. Для дальнейшего доказательства разобьем отрезок $[\tau_j, T_j^U]$ на равные части:

$$T_j^l = \tau_j + l\delta_{2^j} \in [\tau_j, T_j^U], \quad l = 1, \dots, 2^j - 1,$$

$$\delta_{2^j} = \frac{T_j^U - \tau_j}{2^j}.$$

Заметим, что $N_j(T, T')$ есть не что иное, как сумма индикаторов того, что случайная величина $X_{j,k}^W$ попала в заданный интервал

$N_j(T, T') = \sum_{k=1}^{2^j} \mathbf{1}(T < |X_{j,k}^W| \leq T')$. Приведем следующие оценки для приращений $\hat{R}_j(T)$ и $R_j(T)$:

$$\begin{aligned} \left| \frac{\hat{R}_j(T_j) - \hat{R}_j(T'_j)}{2^{J(1+2\beta)}} \right| &\leq \left| \frac{2}{2^{J(1+2\beta)}} \sum_{k=1}^{2^j} \mathbf{1} \left(T_j^2 < (X_{j,k}^W)^2 \leq (T'_j)^2 \right) \right| + \\ &+ \left| \frac{1}{2^{J(1+2\beta)}} \sum_{k=1}^{2^j} \left((X_{j,k}^W)^2 \wedge T_j^2 - (X_{j,k}^W)^2 \wedge (T'_j)^2 \right) \right| \leq \\ &\leq \frac{2}{2^{J(1+2\beta)}} N_j(T, T') + \frac{2^j}{2^{J(1+2\beta)}} \left((T'_j)^2 - T_j^2 \right), \quad (2.12) \end{aligned}$$

$$\left| \frac{R_j(T_j) - R_j(T'_j)}{2^{J(1+2\beta)}} \right| \leq \frac{4 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j}. \quad (2.13)$$

Из оценок (2.12) и (2.13) следует, что

$$\begin{aligned} |Z_j(T_j) - Z_j(T'_j)| &\leq \\ &\leq \frac{2}{2^{J(1+2\beta)}} N_j(T_j, T'_j) + \frac{2^j}{2^{J(1+2\beta)}} (T'_j - T_j)(T'_j + T_j) + \frac{4 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j} \leq \\ &\leq \frac{2}{2^{J(1+2\beta)}} N_j(T_j, T'_j) + \frac{6 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j}, \quad (2.14) \end{aligned}$$

таким образом

$$2^{2\beta j} \cdot |Z_j(T_j) - Z_j(T'_j)| \leq 2^{2\beta j} \cdot \left(\frac{2}{2^{J(1+2\beta)}} N_j(T_j, T'_j) + \frac{6 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j} \right).$$

Тогда можно записать следующее выражение:

$$A_j = \left\{ \sup_{T \in [\tau_j, T_j^U]} 2^{2\beta j} \cdot |Z_j(T)| \geq \frac{3\varepsilon}{J} \right\} \subset D_j \cup E_j,$$

где

$$D_j = \left\{ \sup_l |Z_j(T_j^l)| \geq \frac{\varepsilon}{J \cdot 2^{2\beta j}} \right\},$$

$$E_j = \left\{ \sup_l \sup_{T \in [T_j^l, T_j^l + \delta_{2^j})} |Z_j(T) - Z_j(T_j^l)| \geq \frac{2\varepsilon}{J \cdot 2^{2\beta j}} \right\}.$$

Для E_j справедлива следующая цепочка вложений:

$$\begin{aligned} & \left\{ \sup_l \sup_{T \in [T_j^l, T_j^l + \delta_{2^j})} |Z_j(T) - Z_j(T_j^l)| \geq \frac{2\varepsilon}{J \cdot 2^{2\beta j}} \right\} \subset \\ & \subset \left\{ \sup_j \sup_{T \in [T_j^l, T_j^l + \delta_{2^j})} \frac{2}{2^{J(1+2\beta)}} \cdot N_j(T_j^l, T) + \frac{6 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j} \geq \frac{2\varepsilon}{J \cdot 2^{2\beta j}} \right\} \subset \\ & \subset \left\{ \sup_l \frac{2}{2^{J(1+2\beta)}} \cdot N_j(T_j^l, T_j^l + \delta_{2^j}) + \frac{6 \cdot 2^j}{2^{J(1+2\beta)}} \cdot T_j^U \delta_{2^j} \geq \frac{2\varepsilon}{J \cdot 2^{2\beta j}} \right\} \subset \\ & \subset \left\{ \sup_l N_j(T_j^l, T_j^l + \delta_{2^j}) \geq \frac{2^{J(1+2\beta)-2\beta j} \cdot \varepsilon}{J} - 3 \cdot 2^j T_j^U \delta_{2^j} \right\}. \end{aligned}$$

Для того, чтобы получить оценку для E_j , необходимо привести левую часть неравенства к виду $N_j(T_j^l, T_j^l + \delta_{2^j}) - \mathbb{E}N_j(T_j^l, T_j^l + \delta_{2^j})$. Нетрудно показать, что

$$\mathbb{E}N_j(T_j^l, T_j^l + \delta_{2^j}) \leq \sqrt{\frac{2}{\pi}} \cdot 2^j \cdot \delta_{2^j},$$

следовательно $E_j \subset E'_j$, где $q_j = 2^{2\beta(J-j)+J}$ и

$$\begin{aligned} & \left\{ \sup_l (N_j(T_j^l, T_j^l + \delta_{2^j}) - \mathbb{E}N_j(T_j^l, T_j^l + \delta_{2^j})) \geq \frac{q_j \varepsilon}{J} - 3 \cdot 2^j T_j^U \delta_{2^j} - C \cdot 2^j \cdot \delta_{2^j} \right\} \subset \\ & \subset \left\{ \sup_l \frac{1}{2^j} (N_j(T_j^l, T_j^l + \delta_{2^j}) - \mathbb{E}N_j(T_j^l, T_j^l + \delta_{2^j})) \geq \frac{q_j \varepsilon}{J \cdot 2^j} - 3T_j^U \delta_{2^j} - C\delta_{2^j} \right\} = E''_l. \end{aligned}$$

Пусть

$$\frac{2^{J(1+2\beta)-j(2\beta+1)} \varepsilon}{J} - 3T_j^U \delta_{2^j} - C\delta_{2^j} = \varepsilon_2.$$

Применяя неравенство Хёффдинга, получаем:

$$\mathbb{P}(E''_l) = \mathbb{P}\left(\frac{1}{2^j} \left| N_j(T_j^l, T_j^l + \delta_{2^j}) - \mathbb{E}N_j(T_j^l, T_j^l + \delta_{2^j}) \right| \geq \varepsilon_2\right) \leq 2e^{-2 \cdot 2^j \varepsilon_2^2}.$$

Следовательно, для события E''_j получаем

$$\mathbb{P}(E''_j) \leq \sum_l \mathbb{P}(E''_l) \leq 2 \frac{T_j^U}{\delta_{2^j}} e^{-2^{j+1} \varepsilon_2^2}. \quad (2.15)$$

Теперь рассмотрим вероятность события D_j . Очевидно, что для случайных величин $Y_k(T_j)$ справедливо неравенство $|Y_k(T_j)| \leq 2 + (T_j^U)^2$. Тогда, применяя неравенство Хёффдинга, получим следующую оценку для D_j :

$$\begin{aligned} \mathbb{P}(D_j) &= \mathbb{P}\left(\sup_j |Z_j(T_j)| \geq \frac{\varepsilon}{J \cdot 2^{2\beta j}}\right) \leq \sum_j \mathbb{P}\left(|Z_j(T_j)| \geq \frac{\varepsilon}{J \cdot 2^{2\beta j}}\right) \leq \\ &\leq 2 \frac{T_j^U}{\delta_{2^j}} \exp\left\{-\frac{2^{2J(1+2\beta)-4\beta j-j+1} \cdot \varepsilon^2}{J^2 \cdot (2 + (T_j^U)^2)^2}\right\} \leq 2 \frac{T_j^U}{\delta_{2^j}} \exp\left\{-\frac{2^{J+1} \cdot \varepsilon^2}{J^2 \cdot (2 + (T_j^U)^2)^2}\right\}. \end{aligned} \quad (2.16)$$

Таким образом, из оценок (2.15) и (2.16) следует, что для вероятности A_j справедлива оценка

$$\mathbb{P}(A_j) \leq \mathbb{P}(D_j) + \mathbb{P}(E_j'') \leq 2 \frac{T_j^U}{\delta_{2^j}} \left(\exp\left\{-\frac{2^{J+1} \varepsilon^2}{J^2 \cdot (2 + (T_j^U)^2)^2}\right\} + \exp\{-2^{j+1} \varepsilon_2^2\} \right).$$

Подставив значение универсального порога T_j^U при $\sigma^2 = 1$, получим

$$\begin{aligned} \mathbb{P}(A_j) &\leq \mathbb{P}\left(\sup_{[\tau_j, T_j^U]} |Z_j(T_j)| \geq \frac{3\varepsilon}{J \cdot 2^{2\beta j}}\right) \leq \\ &\leq \frac{2}{\delta_{2^j}} \sqrt{2 \log 2^j} \cdot \exp\left\{-\frac{2^{J-1} \varepsilon^2}{J^2 (1 + \log 2^j)^2}\right\} + \\ &+ \frac{2}{\delta_{2^j}} \sqrt{2 \log 2^j} \cdot \exp\left\{-2^{j+1} \left(\frac{2^{J(2\beta+1)-j(2\beta+1)} \varepsilon}{J} - 3\sqrt{2 \log 2^j} \delta_{2^j} - C\delta_{2^j}\right)^2\right\} = r_j. \end{aligned}$$

Следовательно будет сходиться ряд

$$\sum_{J=1}^{\infty} \sum_{j=tJ}^{J-1} r_j < \infty.$$

Для второго слагаемого из (2.11) по теореме 1.4 имеем:

$$\mathbb{P}(T_j^F \notin [\tau_j, T_j^U]) \leq 2^{j+1} \exp\{-c\alpha_j k_j \gamma_j^2\} = m_j.$$

При условии, что $\frac{\alpha_j k_j \gamma_j^2}{\log 2^j} \rightarrow \infty$ при $J \rightarrow \infty$, сходится ряд:

$$\sum_{J=1}^{\infty} \sum_{j=tJ}^{J-1} m_j < \infty.$$

Рассмотрим второе слагаемое из (2.10):

$$\mathbb{P} \left(\frac{\left| \sum_{j=Jt}^{J-1} (R_j(T_j^F) - R_j(T_j^{min})) \right|}{2^{J(1+2\beta)}} > \frac{\varepsilon}{2} \right).$$

По аналогии с первым слагаемым запишем:

$$\begin{aligned} & \mathbb{P} \left(\frac{\left| \sum_{j=Jt}^{J-1} (R_j(T_j^F) - R_j(T_j^{min})) \right|}{2^{J(1+2\beta)}} > \frac{\varepsilon}{2} \right) \leq \\ & \leq \sum_{j=Jt}^{J-1} \mathbb{P} \left(\sup_{T_j \in [\tau_j, T_j^U]} \frac{|R_j(T_j) - R_j(T_j^{min})|}{2^{J(1+2\beta)}} > \frac{\varepsilon}{2J}, T_j^F \in [\tau_j, T_j^U] \right) + \\ & \quad + \sum_{j=Jt}^{J-1} \mathbb{P} (T_j^F \notin [\tau_j, T_j^U]). \end{aligned}$$

Для $\mathbb{P} (T_j^F \notin [\tau_j, T_j^U])$ снова применим теорему 1.4 и, как было отмечено выше, при выполнении условий теоремы ряд из m_j будет сходиться. Повторив рассуждения из доказательства теоремы про состоятельность оценки риска при выборе FDR-порога из статьи [2], можно показать, что для $j \geq tJ$

$$\sup_{T_j \in [\tau_j, T_j^U]} \frac{|R_j(T_j) - R_j(T_j^{min})|}{2^{J(1+2\beta)}} \rightarrow 0 \text{ при } J \rightarrow \infty. \quad (2.17)$$

Заметим, что ряд из вероятностей сходится

$$\sum_{J=1}^{\infty} \sum_{j=tJ}^{J-1} \mathbb{P} \left(\frac{|R_j(T_j^F) - R_j(T_j^{min})|}{2^{J(1+2\beta)}} > \frac{\varepsilon}{2J} \right) < \infty,$$

в силу того, что выполнено (2.17), а следовательно, начиная с некоторого номера J ,

$$\mathbb{P} \left(\sup_{T_j \in [\tau_j, T_j^U]} \frac{|R_j(T_j^F) - R_j(T_j^{min})|}{2^{J(1+2\beta)}} > \frac{\varepsilon}{2J} \right) = 0,$$

что и завершает доказательство теоремы. □

Из теоремы 2.1 следует, что для рассматриваемой оценки среднеквадратичного риска $\hat{R}(\mathbf{T}^F)$ при использовании FDR-порога имеет место сходимость почти наверное к минимальному риску $R(\mathbf{T}^{min})$.

2.6 Асимптотическая нормальность оценки риска

Далее рассмотрим асимптотическую нормальность оценки риска $\hat{R}(\mathbf{T}^F)$ (это свойство было доказано в статье [69]).

Теорема 2.2. Пусть $\mu \in L_p(\boldsymbol{\eta})$, $\eta_j^p \in [2^{-j}(\log 2^j)^5, 2^{-j\gamma}]$, $1/2 < \gamma < 1$. Пусть \mathbf{T}^F — вектор FDR-порогов с управляющими параметрами $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{J-1})$ такими что, $\alpha_j \rightarrow 0$, $\frac{\alpha_j \kappa_j \gamma_j^2}{\log 2^j} \rightarrow \infty$ при $j \geq tJ$ для некоторого $0 < t < 1/2$ и $J \rightarrow \infty$, где κ_j и γ_j определены в (2.9). Тогда

$$\frac{\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \Rightarrow \mathcal{N}(0,1).$$

Доказательство. Докажем теорему для метода мягкой пороговой обработки. В случае жесткой пороговой обработки доказательство аналогично. Обозначим

$$W(\mathbf{T}) = \hat{R}(\mathbf{T}) - \hat{R}(\mathbf{T}^{min}) = \sum_{j=0}^{J-1} 2^{2\beta j} \beta_{0,0} \cdot U_j(T_j),$$

где

$$U_j(T_j) = \sum_{k=1}^{2^j} (F[X_{j,k}^W, T_j] - F[X_{j,k}^W, T_j^{min}])$$

и запишем

$$\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min}) + \hat{R}(\mathbf{T}^{min}) - \hat{R}(\mathbf{T}^{min}) = \hat{R}(\mathbf{T}^{min}) - R(\mathbf{T}^{min}) + W(\mathbf{T}^F).$$

Покажем, что

$$\frac{\hat{R}(\mathbf{T}^{min}) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \Rightarrow \mathcal{N}(0,1). \quad (2.18)$$

При мягкой пороговой обработке $\hat{R}(\mathbf{T}^{min})$ является несмещенной оценкой $R(\mathbf{T}^{min})$, а при жесткой пороговой обработке и выполнении условий теоремы смещение стремится к нулю при делении на $\sqrt{2^{J+4\beta J}}$ [45]. Для дисперсии числителя имеем [5]

$$\lim_{j \rightarrow \infty} \frac{\mathbb{D} \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} 2^{2\beta j} \beta_{0,0} \cdot (F[X_{j,k}^W, T_j^{min}] - \mathbb{E}F[X_{j,k}^W, T_j^{min}])}{\mathbb{D} \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} 2^{2\beta j} \beta_{0,0} \cdot (X_{j,k}^W)^2} = 1. \quad (2.19)$$

Кроме того, поскольку $X_{j,k}^W$ являются независимыми случайными величинами при применении метода вейглет-вейвлет разложения, $\mathbb{D}(X_{j,k}^W)^2 = 2\sigma^4 + 4\sigma^2\mu_{j,k}^2$ и в силу определения класса $L_p(\boldsymbol{\eta})$ получаем

$$\lim_{J \rightarrow \infty} \frac{\mathbb{D} \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} 2^{2\beta j} \beta_{0,0} \cdot (X_{j,k}^W)^2}{\sigma^4(2^{J+4\beta J} \beta_{0,0}^4)} = 1. \quad (2.20)$$

Наконец, выполнено условие Линдеберга: для любого $\varepsilon > 0$ при $J \rightarrow \infty$

$$\begin{aligned} & \frac{1}{V_j^2} \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} \mathbb{E} [2^{2\beta j} \beta_{0,0} \cdot (F[X_{j,k}^W, T_j^{min}] - \mathbb{E}F[X_{j,k}^W, T_j^{min}])^2 \times \\ & \times \mathbf{1}(2^{2\beta j} \beta_{0,0} \cdot |F[X_{j,k}^W, T_j^{min}] - \mathbb{E}F[X_{j,k}^W, T_j^{min}]| > \varepsilon V_j)] \rightarrow 0, \end{aligned} \quad (2.21)$$

где

$$V_j^2 = \mathbb{D} \sum_{j=0}^{J-1} \sum_{k=1}^{2^j} 2^{2\beta j} \beta_{0,0} \cdot (F[X_{j,k}^W, T_j^{min}] - \mathbb{E}F[X_{j,k}^W, T_j^{min}]).$$

Поскольку слагаемые $\hat{R}(T_j^{min})$ ограничены по модулю величиной $(T_j^U)^2 + \sigma^2$

$$|F[X_{j,k}^W, T_j^{min}] - \mathbb{E}F[X_{j,k}^W, T_j^{min}]| \leq 2((T_j^U)^2 + \sigma^2) \quad (2.22)$$

и в силу (2.19), (2.20) и (2.22) начиная с некоторого j все индикаторы в (2.21) обращаются в нуль. Таким образом, справедливо (2.18).

Осталось показать, что

$$\frac{W(\mathbf{T}^F)}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \xrightarrow{\mathbb{P}} 0. \quad (2.23)$$

Разобьем $W(\mathbf{T}^F)$ на две компоненты:

$$W(\mathbf{T}^F) = \sum_{j=0}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) = \sum_{j=0}^{Jt} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) + \sum_{j=Jt}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \quad (2.24)$$

для некоторого $t \in (0, 1/2)$.

Рассмотрим вторую сумму от Jt до $J-1$ (для сокращения записи будем считать, что Jt – целое число). Для любого $\varepsilon > 0$

$$\begin{aligned} \mathbb{P} \left(\frac{\left| \sum_{j=Jt}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \right|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \varepsilon \right) &\leq \sum_{j=Jt}^{J-1} \mathbb{P} \left(\frac{|2^{2\beta j} \beta_{0,0} U_j(T_j^F)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right) \leq \\ &\leq \sum_{j=Jt}^{J-1} \left(\mathbb{P}(T_j^F \leq \tau_j) + \mathbb{P} \left(\frac{2^{2\beta j} \beta_{0,0} \cdot \sup_{T_j \in [\tau_j, T_j^U]} |U_j(T_j)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right) \right) \leq \\ &\leq \sum_{j=Jt}^{J-1} \mathbb{P}(T_j^F \leq \tau_j) + \\ &+ \sum_{j=Jt}^{J-1} \mathbb{P} \left(\frac{2^{2\beta j} \beta_{0,0} \left(\sup_{T_j \in [\tau_j, T_j^U]} |U_j(T_j) - \mathbb{E}U_j(T_j)| + \sup_{T_j \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T_j)| \right)}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right). \end{aligned} \quad (2.25)$$

Пусть $U_j(T_j) = S'_j(T_j) + S''_j(T_j)$, $T_j \in [\tau_j, T_j^U]$, где $S''_j(T_j)$ состоит из слагаемых, для которых $|\mu_{j,(k)}| \leq C/\tau_j$, а $S'_j(T_j)$ содержит все остальные.

По определению класса $L_p(\eta_j)$ число слагаемых в $S'_j(T_j)$ не превосходит

$$C \cdot 2^j T_j^p \eta_j^p = \{T_j^p \sim (\log 2^j)^{\frac{p}{2}}\} = C \cdot 2^j (\log 2^j)^{\frac{p}{2}} \eta_j^p \equiv m_j.$$

Так как каждое слагаемое из $S'_j(T_j)$ не превосходит по модулю $2((T_j^U)^2 + \sigma^2)$,

$$|S'_j(T_j)| \leq C \cdot 2^j (\log 2^j)^{1+\frac{p}{2}} \cdot \eta_j^p.$$

Следовательно,

$$\frac{2^{2\beta j} \beta_{0,0} \cdot \sup_{T_j \in [\tau_j, T_j^U]} |S'_j(T_j) - \mathbb{E}S'_j(T_j)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq \frac{C \eta_j^p \cdot 2^{j(2\beta+1)} (\log 2^j)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq \frac{C \cdot 2^{j(2\beta-\gamma+1)} (\log 2^J)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}}. \quad (2.26)$$

Далее, в силу определения T_j^{min} ,

$$|\mathbb{E}S''_j(T_j)| \leq C \cdot |\mathbb{E}S'_j(T_j)|.$$

Таким образом,

$$\sup_{T \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T)| \leq C \cdot 2^j (\log 2^j)^{1+\frac{p}{2}} \cdot \eta_j^p$$

и

$$\frac{2^{2\beta j} \beta_{0,0} \cdot \sup_{T_j \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T_j)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq \frac{C \eta_j^p \cdot 2^{j(2\beta+1)} (\log 2^j)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq \frac{C \cdot 2^{j(2\beta-\gamma+1)} (\log 2^J)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}}. \quad (2.27)$$

При $\gamma > 1/2$ полученные выражения (2.26) и (2.27) начиная с некоторого J меньше, чем ε/J , при всех $j \in [Jt, J-1]$.

Рассмотрим теперь сумму $S''_j(T)$. При больших значениях 2^j число слагаемых в этой сумме равно $2^j - m_j \approx 2^j$. Для удобства будем считать, что $S''_j(T)$ содержит слагаемые с номерами от $m_j + 1$ до 2^j . Повторяя рассуждения из [3; 4; 37], можно показать, что

$$T_j^{min} \geq \tau_j - \alpha_j,$$

где

$$|\alpha_j| \leq C \frac{\log \log 2^j}{\sqrt{\log 2^j}}.$$

Причем без ущерба для доказательства можно опустить α_j и считать, что $T_j^{min} \geq \tau_j$. Разобьем отрезок $[\tau_j, T_j^U]$ на одинаковые части:

$$T_j^l = \tau_j + l\delta_{2^j} \in [\tau_j, T_j^U], \quad l = 1, \dots, 2^j - 1,$$

$$\delta_{2^j} = \frac{T_j^U - \tau_j}{2^j}.$$

Обозначим

$$Z_j''(T) = S_j''(T) - \mathbb{E}S_j''(T), \quad N_j''(T, T') = \sum_{i=m_j+1}^{2^j} \mathbf{1}(T < |X_{j,k}^W| \leq T').$$

Тогда

$$A_j = \left\{ \frac{2^{2\beta j} \cdot \sup_{T \in [\tau_j, T_j^U]} |S_j''(T) - \mathbb{E}S_j''(T)|}{2^{J(2\beta+1/2)}} \geq 5 \cdot \frac{\varepsilon}{J} \right\} \subset D_j \cup E_j,$$

где

$$D_j = \left\{ \sup_l |Z_j''(T_j^l)| > 2^{2\beta(J-j)+J/2} \cdot \frac{\varepsilon}{J} \right\},$$

$$E_j = \left\{ \sup_l \sup_{T \in [T_j^l, T_j^l + \delta_{2^j}]} |Z_j''(T) - Z_j''(T_j^l)| \geq 4 \cdot 2^{2\beta(J-j)+J/2} \cdot \frac{\varepsilon}{J} \right\}.$$

Обозначим $q_j = 2^{2\beta(J-j)+J/2}$. Учитывая определение класса $L_p(\boldsymbol{\eta})$ и вид τ_j , можно показать, что дисперсии слагаемых в $S_j''(T)$ (а значит, и $Z_j''(T)$) не превосходят $C \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{3/2} \cdot 2^{-j\gamma}$. Применяя неравенство Бернштейна [16] для D_j , получим

$$\mathbb{P}(D_j) \leq \sum_l \mathbb{P}(|Z_j''(T_j^l)| > q_j \varepsilon / J) \leq$$

$$\leq 2^{j+1} \exp \left\{ - \frac{C J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{3/2} + 2^{1+j/2} ((T_j^U)^2 + \sigma^2)} \right\}.$$

Далее, повторяя рассуждения из [70], получаем, что $E_j \subset E'_j$, где

$$E'_j = \left\{ \sup_l |N_j''(T_j^l, T_j^l + \delta_{2^j}) - \mathbb{E}N_j''(T_j^l, T_j^l + \delta_{2^j})| \geq \frac{q_j \varepsilon}{\sigma^2 J} - \frac{2^j T_j^U \delta_{2^j}}{\sigma^2} - \frac{2^j C \delta_{2^j}}{\sigma^2} \right\}.$$

Дисперсия слагаемых в $N_j''(T_j^l, T_j^l + \delta_{2j})$ ограничена величиной $C \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} \cdot 2^{-j\gamma}$. Снова применяя неравенство Бернштейна, получаем

$$\begin{aligned} \mathbb{P}(E_l'') &= \mathbb{P} \left(\left| N_j''(T_j^l, T_j^l + \delta_{2j}) - \mathbb{E} N_j''(T_j^l, T_j^l + \delta_{2j}) \right| \geq \frac{q_j \varepsilon}{\sigma^2 J} - \frac{2^j T_j^U \delta_{2j}}{\sigma^2} - \frac{2^j C \delta_{2j}}{\sigma^2} \right) \leq \\ &\leq 2 \exp \left\{ - \frac{C J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} + 2^{1+j/2}} \right\}. \end{aligned}$$

Следовательно,

$$\mathbb{P}(E_j') \leq \sum_l \mathbb{P}(E_l'') \leq 2^{j+1} \exp \left\{ - \frac{C J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} + 2^{1+j/2}} \right\}.$$

Таким образом, для произвольного $\varepsilon > 0$

$$\sum_{j=Jt}^{J-1} \mathbb{P} \left(\sup_{T \in [\tau_j, T_j^U]} \frac{2^{2\beta j} \cdot |S_j''(T) - \mathbb{E} S_j''(T)|}{2^{J(2\beta+1/2)}} > \frac{\varepsilon}{J} \right) \rightarrow 0$$

при $J \rightarrow \infty$.

Для первого слагаемого из (2.24) в силу (2.22) имеем

$$\left| \sum_{j=0}^{Jt} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \right| \leq C \cdot 2^{(2\beta+1)tJ} \cdot \log 2^J.$$

Следовательно, при $J \rightarrow \infty$

$$\frac{\left| \sum_{j=0}^{Jt} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \right|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \rightarrow 0 \text{ п. в.} \quad (2.28)$$

Для оценки $\mathbb{P}(T_j^F < \tau_j)$ воспользуемся результатом, полученным в теореме 1.4. Имеем

$$\sum_{j=Jt}^{J-1} \sup_{\mu \in L_p(\eta)} \mathbb{P}(T_j^F < \tau_j) \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp\{-c \alpha_j \kappa_j \gamma_j^2\}.$$

Объединяя полученные выше оценки, для выражения (2.25) будем иметь

$$\begin{aligned}
& \sum_{j=Jt}^{J-1} \mathbb{P}(T_j^F \leq \tau_j) + \\
& + \sum_{j=Jt}^{J-1} \mathbb{P} \left(\frac{2^{2\beta j} \beta_{0,0} \left(\sup_{T_j \in [\tau_j, T_j^U]} |U_j(T_j) - \mathbb{E}U_j(T_j)| + \sup_{T_j \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T_j)| \right)}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right) \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp\{-c \alpha_j \kappa_j \gamma_j^2\} + \\
& + \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} + 2^{1+j/2}} \right\} + \\
& \quad + \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{3/2} + 2^{1+j/2} ((T_j^U)^2 + \sigma^2)} \right\}.
\end{aligned}$$

При $J \rightarrow \infty$ и выполнении условий теоремы полученное выражение стремится к нулю и, учитывая (2.28), получаем, что справедливо (2.23).

Объединяя, получаем утверждение теоремы. \square

Аналогично замечаниям, приведенным в главе 1, результат теоремы 2.2 позволяет строить доверительные интервалы для минимального среднеквадратичного риска $R(\mathbf{T}^{min})$. Далее приведем их вид.

$$\mathbb{P} \left(-z_{\frac{1+\gamma}{2}} \leq \frac{\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq z_{\frac{1+\gamma}{2}} \right) = \mathbb{P}(T_1 \leq R(\mathbf{T}^{min}) \leq T_2) \rightarrow \gamma.$$

где $z_{\frac{1+\gamma}{2}}$ - квантиль стандартного нормального распределения уровня $\frac{1+\gamma}{2}$,

$$T_1 = \hat{R}(\mathbf{T}^F) - \sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4} \cdot z_{\frac{1+\gamma}{2}},$$

$$T_2 = \hat{R}(\mathbf{T}^F) + \sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4} \cdot z_{\frac{1+\gamma}{2}}.$$

2.7 Оценка скорости сходимости

В следующей теореме, доказанной в работе [68], найдем оценку скорости сходимости к нормальному закону оценки риска для случая обращения линейных однородных операторов.

Теорема 2.3 Пусть $\mu \in L_p(\boldsymbol{\eta})$, $\eta_j^p \in [2^{-j}(\log 2^j)^5, 2^{-j\gamma}]$, $1/2 < \gamma < 1$. Пусть \mathbf{T}^F — вектор FDR-порогов с управляющими параметрами $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{J-1})$ такими что, $\alpha_j \rightarrow 0$, $\frac{\alpha_j \kappa_j \gamma_j^2}{\log 2^j} \rightarrow \infty$ при $j \geq tJ$ для некоторого $0 < t < \frac{1-\gamma}{2\beta+1}$ при $J \rightarrow \infty$, где κ_j и γ_j определены в (2.9). Тогда для $x \in \mathbb{R}$ выполнено неравенство:

$$\sup_x \left| \mathbb{P} \left(\frac{\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}} < x \right) - \Phi(x) \right| \leq C \cdot J^{2+p/2} \cdot 2^{-J(\gamma-1/2)} \text{ для } \gamma \in \left(\frac{1}{2}, \frac{3}{4} \right],$$

$$\sup_x \left| \mathbb{P} \left(\frac{\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}} < x \right) - \Phi(x) \right| \leq C \cdot J^2 \cdot 2^{-J/4} \text{ для } \gamma \in \left(\frac{3}{4}, 1 \right),$$

где в обоих случаях C — это константы, вообще говоря, разные.

Доказательство. Введем обозначение

$$W(\mathbf{T}) = \hat{R}(\mathbf{T}) - \hat{R}(\mathbf{T}^{min}) = \sum_{j=0}^{J-1} 2^{2\beta j} \beta_{0,0} \cdot U_j(T_j),$$

где

$$U_j(T_j) = \sum_{k=1}^{2^j} (F[X_{j,k}^W, T_j] - F[X_{j,k}^W, T_j^{min}]),$$

и запишем

$$\hat{R}(\mathbf{T}^F) - R(\mathbf{T}^{min}) + \hat{R}(\mathbf{T}^{min}) - \hat{R}(\mathbf{T}^{min}) = \hat{R}(\mathbf{T}^{min}) - R(\mathbf{T}^{min}) + W(\mathbf{T}^F).$$

В статье [6] было приведено следующее неравенство Бернштейна для $V_J + Z_J$, любого $\varepsilon > 0$ и $x \in \mathbb{R}$:

$$|\mathbb{P}(V_J + Z_J < x) - \Phi(x)| \leq \sup_x |\mathbb{P}(V_J < x) - \Phi(x)| + \frac{\varepsilon}{\sqrt{2\pi}} + \mathbb{P}(|Z_J| \geq \varepsilon). \quad (2.29)$$

В данном случае возьмем

$$V_J = \frac{\hat{R}(\mathbf{T}^{min}) - R(\mathbf{T}^{min})}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}}, \quad Z_J = \frac{W(\mathbf{T}^F)}{\sigma^2 \sqrt{2^{J+4\beta J} \beta_{0,0}^4}}.$$

Далее без ограничения общности возьмем $\sigma^2 = 1$.

Разобьем $W(\mathbf{T}^F)$ на две компоненты (для сокращения записи будем считать, что Jt — целое число):

$$W(\mathbf{T}^F) = \sum_{j=0}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) = \sum_{j=0}^{Jt} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) + \sum_{j=Jt}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \quad (2.30)$$

для некоторого $t \in (0,1)$.

Обозначим $q_j = 2^{2\beta(J-j)+J/2}$. В ходе доказательства теоремы 2.2 для сумм из (2.30) были получены следующие неравенства:

$$\left| \sum_{j=0}^{Jt} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \right| \leq C \cdot 2^{(2\beta+1)tJ} \cdot \log 2^J, \quad (2.31)$$

$$\begin{aligned} & \mathbb{P} \left(\frac{\left| \sum_{j=Jt}^{J-1} 2^{2\beta j} \beta_{0,0} U_j(T_j^F) \right|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \varepsilon \right) \leq \sum_{j=Jt}^{J-1} \mathbb{P} \left(\frac{|2^{2\beta j} \beta_{0,0} U_j(T_j^F)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right) \leq \\ & \leq \sum_{j=Jt}^{J-1} \mathbb{P} (T_j^F \leq \tau_j) + \\ & + \sum_{j=Jt}^{J-1} \mathbb{P} \left(\frac{2^{2\beta j} \beta_{0,0} \left(\sup_{T_j \in [\tau_j, T_j^U]} |U_j(T_j) - \mathbb{E}U_j(T_j)| + \sup_{T_j \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T_j)| \right)}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} > \frac{\varepsilon}{J} \right) \leq \\ & \leq \sum_{j=Jt}^{J-1} \left(2^{j+1} \exp\{-c \alpha_j \kappa_j \gamma_j^2\} + 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} + 2^{1+j/2}} \right\} + \right. \\ & \left. + 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \varepsilon^2 q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{3/2} + 2^{1+j/2} q_j \left((T_j^U)^2 + \sigma^2 \right)} \right\} \right). \quad (2.32) \end{aligned}$$

Заметим, что последнее выражение из (2.32) содержит под знаком суммы три слагаемых.

Для второго слагаемого из (2.32), учитывая определение q_j и тот факт, что $\gamma > 1/2$, имеем

$$\begin{aligned}
& \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{1/2} + 2^{1+j/2}} \right\} \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot q_j^2}{2^{1+j/2}} \right\} \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{4\beta J - 4\beta j + J - j/2 - 1} \right\} \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2} \right\} \leq \\
& \leq C_2 \cdot 2^J \exp \left\{ -C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2} \right\}.
\end{aligned}$$

Для третьего слагаемого из (2.32) имеем

$$\begin{aligned}
& \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot q_j^2}{2^{1+j(1-\gamma)} \left(\log \frac{2^j}{(\log 2^j)^5} \right)^{3/2} + 2^{1+j/2} ((T_j^U)^2 + \sigma^2)} \right\} \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{4\beta J - 4\beta j + J - j/2 - 1}}{((T_j^U)^2 + \sigma^2)} \right\} \leq \\
& \leq \sum_{j=Jt}^{J-1} 2^{j+1} \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2}}{(T_j^U)^2 + \sigma^2} \right\} \leq \\
& \leq C_2 \cdot 2^J \exp \left\{ -\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2}}{(T_j^U)^2 + \sigma^2} \right\}
\end{aligned}$$

для любого $\varepsilon > 0$.

Вернемся к рассмотрению выражения (2.29). Заметим, что первый член из (2.30) не является главным и соответственно не будет влиять на порядок оценки скорости. Для второго и третьего слагаемого из (2.29) с учетом оценок,

полученных выше для выражения (2.32), имеем

$$\begin{aligned} \frac{\varepsilon}{\sqrt{2\pi}} + \mathbb{P}(|Z_J| \geq \varepsilon) &\leq \frac{\varepsilon(J)}{\sqrt{2\pi}} + \sum_{j=Jt}^{J-1} 2^{j+1} \exp\{-c \alpha_j \kappa_j \gamma_j^2\} + \\ &+ C_2 \cdot 2^J \exp\left\{-C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2}\right\} + C_2 \cdot 2^J \exp\left\{-\frac{C \cdot J^{-2} \cdot \varepsilon^2 \cdot 2^{J/2}}{(T_j^U)^2 + \sigma^2}\right\}. \end{aligned} \quad (2.33)$$

В ходе доказательства теоремы 2.2 было получено следующее неравенство:

$$\frac{2^{2\beta j} \beta_{0,0} \cdot \sup_{T_j \in [\tau_j, T_j^U]} |\mathbb{E}U_j(T_j)|}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} \leq \frac{C \cdot 2^{j(2\beta-\gamma+1)} (\log 2^J)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}}, \quad (2.34)$$

которое при $\gamma > 1/2$, начиная с некоторого J , меньше чем ε/J при всех $j \in [Jt, J-1]$.

Тогда возьмем $\varepsilon_1(J)$ так, чтобы

$$\frac{\varepsilon_1(J)}{J} > \frac{C \cdot 2^{j(2\beta-\gamma+1)} (\log 2^J)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}},$$

т.е.

$$\varepsilon_1(J) = \frac{C \cdot J \cdot 2^{j(2\beta-\gamma+1)} (\log 2^J)^{1+p/2}}{\sqrt{2^{J+4\beta J} \beta_{0,0}^4}} = C_3 \cdot J^{2+p/2} \cdot 2^{-J\gamma+J/2}. \quad (2.35)$$

Из (2.34) и (2.31) вытекает ограничение на параметр t

$$t < \frac{1-\gamma}{2\beta+1}. \quad (2.36)$$

Тогда порядок слагаемых в (2.33) будет следующим (второе слагаемое убывает быстрее остальных, поэтому не будет влиять на порядок):

$$\frac{C_4 \cdot J^{2+p/2}}{2^{J(\gamma-1/2)}} + C_5 \cdot 2^J \exp\left\{-2^{J(3/2-2\gamma)} \cdot J^{2+p}\right\} + C_6 \cdot 2^J \exp\left\{-2^{J(3/2-2\gamma)} \cdot J^{2+p}\right\}. \quad (2.37)$$

Тогда при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ в (2.37) порядок будет $J^{2+p/2} \cdot 2^{-J(\gamma-1/2)}$.

Для первого слагаемого из (2.29), используя неравенство Берри–Эссеена и действуя аналогично рассуждениям из статьи [7], получаем следующую оценку:

$$\sup_x |\mathbb{P}(V_J < x) - \Phi(x)| \leq C_3 \cdot J^{2+p/2} \cdot 2^{-J\gamma+J/2}. \quad (2.38)$$

Таким образом, при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ получаем скорость сходимости оценки риска порядка

$$J^{2+p/2} \cdot 2^{-J(\gamma-1/2)}.$$

Рассмотрим случай $\gamma \in \left(\frac{3}{4}, 1\right)$. Пусть $\varepsilon_2(J) = C_7 \cdot 2^{-J/4} J^2$. Тогда порядок слагаемых в (2.33) будет $C_8 \cdot J^2 \cdot 2^{-\frac{J}{4}}$.

Следовательно, скорость сходимости оценки риска в случае $\gamma \in \left(\frac{3}{4}, 1\right)$ будет иметь порядок

$$J^2 \cdot 2^{-\frac{J}{4}}.$$

Таким образом, утверждение теоремы доказано. □

Оценка скорости сходимости, полученная в теореме 2.3, имеет порядок при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ равный $J^{2+p/2} \cdot 2^{-J(\gamma-1/2)}$, а при $\gamma \in \left(\frac{3}{4}, 1\right)$ — $J^2 \cdot 2^{-\frac{J}{4}}$. Вспомним, что во второй главе диссертации $2^J = N$, следовательно порядок при $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right]$ равный $(\log N)^{2+p/2} \cdot N^{1/2-\gamma}$, а при $\gamma \in \left(\frac{3}{4}, 1\right)$ — $(\log N)^2 \cdot N^{-\frac{1}{4}}$.

Как отмечалось ранее в главе 1, оценки скорости сходимости из теоремы 2.3 могут быть использованы для построения доверительных интервалов, а также, например, для уточнения уровня доверия в доверительных интервалах.

Заключение

1. Обзор проведенного исследования. Тематика диссертации относится к области математической статистики. В работе рассмотрены задачи множественной проверки гипотез при применении FDR-порога в случае рассмотрения векторов большой размерности (глава 1), а также в случае обращения линейного однородного оператора (глава 2). Также в работе исследуются асимптотические свойства оценки среднеквадратичного риска при пороговой обработке. Основные результаты диссертации состоят в следующем:

- **Доказана** сильная состоятельность оценки риска при применении FDR-порога в случае классов разреженности исходных данных $L_0(\boldsymbol{\eta}_n)$ и $L_p(\boldsymbol{\eta}_n)$, а также сильная состоятельность оценки риска в задаче обращения линейных однородных операторов для класса разреженности $L_p(\boldsymbol{\eta})$.
- **Доказана** асимптотическая нормальность оценки риска при применении FDR-порога в случае классов разреженности исходных данных $L_0(\boldsymbol{\eta}_n)$ и $L_p(\boldsymbol{\eta}_n)$, а также асимптотическая нормальность оценки риска в задаче обращения линейных однородных операторов для класса разреженности $L_p(\boldsymbol{\eta})$. Помимо этого приведены виды асимптотических доверительных интервалов для минимального риска в обоих случаях рассматриваемых постановок задачи.
- **Получены** оценки скорости сходимости распределения оценки риска к нормальному закону при применении FDR-порога в случае рассмотрения векторов большой размерности для классов разреженности $L_0(\boldsymbol{\eta}_n)$ и $L_p(\boldsymbol{\eta}_n)$, а также в случае рассмотрения задачи обращения линейного однородного оператора для $L_p(\boldsymbol{\eta})$ класса разреженности исходных данных.

2. Рекомендации и перспективы по дальнейшей разработке темы диссертации.

- Исследовать возможность уточнения оценок скорости сходимости распределения оценки риска к нормальному закону при применении FDR-порога в случае рассмотрения векторов большой размерности для классов разреженности $L_0(\boldsymbol{\eta}_n)$ и $L_p(\boldsymbol{\eta}_n)$, а также в случае рассмотрения задачи обращения линейного однородного оператора для $L_p(\boldsymbol{\eta})$ класса разреженности исходных данных.
- Исследовать асимптотические свойства оценок среднеквадратичного риска при применении FDR-порога в случае класса разреженности исходных данных $M_p(\boldsymbol{\eta}_n)$, а также асимптотические свойства оценок риска в задаче обращения линейных однородных операторов для класса разреженности $M_p(\boldsymbol{\eta})$.
- Получить оценки скорости сходимости распределения оценки риска к нормальному закону при применении FDR-порога в случае рассмотрения векторов большой размерности для класса разреженности $M_p(\boldsymbol{\eta}_n)$, а также для задачи обращения линейного однородного оператора для $M_p(\boldsymbol{\eta})$ класса разреженности исходных данных.

Список литературы

1. *Добеши И.* Десять лекций по вейвлетам. — Ижевск : НИЦ «Регулярная и хаотическая динамика», 2001. — 464 с.
2. *Засна А. Ю., Шестаков О. В.* Состоятельность оценки риска при множественной проверке гипотез с FDR-порогом // Вестник Тверского государственного университета. Серия: Прикладная математика. — 2017. — № 1. — С. 5—16.
3. *Кудрявцев А. А., Шестаков О. В.* Асимптотически оптимальная пороговая обработка вейвлет-коэффициентов в моделях с негауссовым распределением шума // Доклады Академии наук. — 2016. — Т. 471, № 1. — С. 11—15.
4. *Кудрявцев А. А., Шестаков О. В.* Асимптотическое поведение порога, минимизирующего усредненную вероятность ошибки вычисления вейвлет-коэффициентов // Доклады Академии наук. — 2016. — Т. 468, № 5. — С. 487—491.
5. *Маркин А.* Предельное распределение оценки риска при пороговой обработке вейвлет-коэффициентов // Информатика и её применения. — 2009. — Т. 3, № 4. — С. 57—63.
6. *Сунклодас Й.* Аппроксимация распределений сумм слабо зависимых случайных величин нормальным распределением // Теория вероятностей—6. Предельные теоремы теории вероятностей : Итоги науки и техн. Сер. Современ. пробл. мат. Фундам. направления. — 1991. — Т. 81. — С. 140—199.
7. *Шестаков О. В.* Аппроксимация распределения оценки риска пороговой обработки вейвлет-коэффициентов нормальным распределением при использовании выборочной дисперсии // Информатика и её применения. — 2010. — Т. 4, № 4. — С. 73—81.
8. *Abel N. H.* Resolution d'un Probleme // Journal fur die reine und angewandte Mathematik. — 1826. — Vol. 1. — P. 153—157.

9. *Abramovich F., Benjamini Y.* Adaptive thresholding of wavelet coefficients // Computational statistics and data analysis. — 1996. — Vol. 22, no. 4. — P. 351—361.
10. *Abramovich F., Silverman B. W.* Wavelet decomposition approaches to statistical inverse problems // Biometrika. — 1998. — Vol. 85, no. 1. — P. 115—129.
11. Adapting to unknown sparsity by controlling the false discovery rate / F. Abramovich [et al.] // The Annals of Statistics. — 2006. — Vol. 34, no. 2. — P. 584—653.
12. *Akaike H.* Information theory and an extension of the maximum likelihood principle // Second International Symposium on Information Theory. — Akademiai Kiado, Budapest, 1973. — P. 267—281.
13. *Benedetto J. J., A. T.* A wavelet auditory model and data compression // Applied and Computational Harmonic Analysis. — 1993. — Vol. 1. — P. 3—28.
14. *Benjamini Y., Yekutieli D.* False discovery rate-adjusted multiple confidence intervals for selected parameters // Journal of the American Statistical Association. — 2005. — Vol. 100. — P. 71—81.
15. *Benjamini Y., Hochberg Y.* Controlling the false discovery rate: a practical and powerful approach to multiple testing // Journal of the royal statistical society series b-methodological. — 1995. — Vol. 57. — P. 289—300.
16. *Bennett G.* Probability inequalities for the sum of independent random variables // Journal of the American Statistical Association. — 1962. — Vol. 57. — P. 33—45.
17. *Chang S. G., Yu B., M. V.* Adaptive wavelet thresholding for image denoising and compression // IEEE transactions on image processing. — 2000. — Vol. 9, no. 9. — P. 1532—1546.
18. *Cohen A., Daubechies I., Vial P.* Wavelets on the interval and fast wavelet transforms // Applied and computational harmonic analysis. — 1993. — Vol. 1, no. 1. — P. 54—81.

19. Compressive Sensing for Background Subtraction / V. Cevher [et al.] // European Conference on Computer Vision : Lecture Notes in Computer Science. — 2008. — Vol. 5303. — P. 155—168.
20. *Dikmen M., Huang T. S.* Robust estimation of foreground in surveillance videos by sparse error estimation // 19th International Conference on Pattern Recognition. — 2008. — P. 1—4.
21. Discriminative learned dictionaries for local image analysis / J. Mairal [et al.] // IEEE Conference on Computer Vision and Pattern Recognition. — 2008. — P. 1—8.
22. *Donoho D.* De-noising by soft-thresholding // IEEE transactions on information theory. — 1995. — Vol. 41, no. 3. — P. 613—627.
23. *Donoho D.* Nonlinear Solution of Linear Inverse Problems by Wavelet–Vaguelette Decomposition // Applied and Computational Harmonic Analysis. — 1995. — Vol. 2. — P. 101—126.
24. *Donoho D., Johnstone I. M.* Ideal spatial adaptation via wavelet shrinkage // Biometrika. — 1994. — Vol. 81, no. 3. — P. 425—455.
25. *Donoho D., Johnstone I. M.* Minimax risk over lp-balls for lq-error // Probability Theory and Related Fields. — 1994. — Vol. 99. — P. 277—303.
26. *Donoho D., Johnstone I. M.* Adapting to unknown smoothness via wavelet shrinkage // J. Amer. Statist. Assoc. — 1995. — Vol. 90. — P. 1200—1224.
27. *Donoho D., Johnstone I. M.* Neo-classical minimax problems, thresholding and adaptive function estimation // Bernoulli. — 1996. — Vol. 2, no. 1. — P. 39—62.
28. *Donoho D., Johnstone I. M.* Minimax estimation via wavelet shrinkage // The Annals of Statistics. — 1998. — Vol. 26, no. 3. — P. 879—921.
29. *Donoho D. L., Jin J.* Asymptotic minimaxity of false discovery rate thresholding for sparse exponential data // The Annals of Statistics. — 2006. — Vol. 34. — P. 2980—3018.

30. *Elhamifar E., Vidal R.* Sparse Subspace Clustering: Algorithm, Theory, and Applications // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2013. — Vol. 35. — P. 2765—2781.
31. Exact Risk Analysis of Wavelet Regression / J. S. Marron [et al.] // Journal of Computational and Graphical Statistics. — 1998. — Vol. 7. — P. 278—309.
32. *Foster D. P., George E. I.* The risk inflation criterion for multiple regression // The Annals of Statistics. — 1994. — Vol. 22. — P. 1947—1975.
33. *Gao H. Y.* Wavelet shrinkage denoising using the non-negative garrote // Journal of computational and graphical statistics. — 1998. — Vol. 7, no. 4. — P. 469—488.
34. *Gao H. Y., Bruce A. G.* Waveshrink with firm shrinkage // Statistica Sinica. — 1997. — Vol. 7, no. 4. — P. 855—874.
35. *Hannan E. J., Quinn B. G.* The determination of the order of an autoregression // Journal of the royal statistical society series b-methodological. — 1979. — Vol. 41. — P. 190—195.
36. *Hoeffding W.* Probability inequalities for sums of bounded random variables // J. American Statistical Association. — 1963. — Vol. 58, no. 301. — P. 13—30.
37. *Jansen M.* Noise Reduction by Wavelet Thresholding. — New York: Springer : Lecture Notes in Statistics, 2001. — 161 p.
38. *Johnstone I. M.* Wavelet shrinkage for correlated data and inverse problems: adaptivity results // Statistica Sinica. — 1999. — Vol. 9. — P. 51—83.
39. *Kalifa J., Laine A., Esser P. D.* A wavelet shrinkage approach to tomographic image reconstruction // Journal of the American Statistical Association. — 1996. — Vol. 91, no. 435. — P. 1079—1090.
40. *Kalifa J., Laine A., Esser P. D.* Tomographic reconstruction with non-linear diagonal estimators // Proceedings of SPIE, the International Society for Optical Engineering. — 2000. — Vol. 4119. — P. 576—586.

41. *Kalifa J., Mallat S.* Thresholding estimators for linear inverse problems and deconvolutions // The Annals of Statistics. — 2003. — Vol. 31, no. 1. — P. 58—109.
42. *Lee N.* Wavelet-vaguelette decompositions and homogeneous equations: PhD dissertation. — Purdue University, 1997.
43. *Luisier F., Blu T., Unser M.* A new SURE approach to Image denoising: interscale orthonormal wavelet thresholding // IEEE transactions on image processing. — 2007. — Vol. 38, no. 5. — P. 1323—1342.
44. *Mallat S.* Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$ // Transactions of the American Mathematical Society. — 1989. — Vol. 315, no. 1. — P. 69—87.
45. *Mallat S.* A Wavelet Tour of Signal Processing. — N. Y. : Academic Press, 1999. — 857 p.
46. *Mallows C. L.* Some Comments on C_p // Technometrics. — 1973. — Vol. 15, no. 4. — P. 661—675.
47. *Meyer Y.* Wavelets and Operators. — New York : Cambridge University Press, 1992. — 223 p. — (Cambridge Studies in Advanced Mathematics).
48. *Miller R.* Simultaneous Statistical Inference. — 2nd. — New York : Springer Verlag, 1981. — 272 p.
49. Motion segmentation via robust subspace separation in the presence of outlying, incomplete, or corrupted trajectories / S. R. Rao [et al.] // 2008 IEEE Conference on Computer Vision and Pattern Recognition. — 2008. — P. 1—8.
50. *Neuvial P., Roquain É.* On false discovery rate thresholding for classification under sparsity // The Annals of Statistics. — 2012. — Vol. 40. — P. 2572—2600.
51. *Riesz M.* Integrales de Riemann-Liouville et potentiels // Acta sci. math. Szeged. — 1938. — Vol. 9. — P. 1—42.

52. Robust Face Recognition via Sparse Representation / J. Wright [et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2009. — Vol. 31. — P. 210—227.
53. *Schwartz L.* Theorie des distributions. — Paris : Hermann, 1950. — 427 p.
54. *Schwarz.* Estimating the dimension of a model // The Annals of Statistics. — 1978. — Vol. 6. — P. 461—464.
55. *Selesnick I. W.* Sparse signal representations using the tunable Q-factor wavelet transform // Proceedings of SPIE : Wavelets and Sparsity XIV. — 2011. — Vol. 8138. — P. 1—15.
56. *Shestakov O.* Asymptotic normality of adaptive wavelet thresholding risk estimation // Doklady Mathematics. — 2012. — Vol. 86, no. 1. — P. 556—558.
57. *Simoncelli E. P.* Bayesian Denoising of Visual Images in the Wavelet Domain // Bayesian Inference in Wavelet-Based Models. — 1999. — P. 291—308.
58. *Stein C.* Estimation of the mean of a multivariate normal distribution // The Annals of Statistics. — 1981. — Vol. 9, no. 6. — P. 1135—1151.
59. *Storey J. D.* A direct approach to false discovery rates // Journal of the Royal Statistical Society: Series B (Statistical Methodology). — 2002. — Vol. 64, no. 3. — P. 479—498.
60. *Storey J. D.* The positive false discovery rate: a Bayesian interpretation and the q-value // The Annals of Statistics. — 2003. — Vol. 31. — P. 2013—2035.
61. *Theil H.* Economic Forecasts And Policy. — 2nd. — Amsterdam : North-Holland, 1961. — 567 p.
62. Wavelet shrinkage: asymptopia? / D. Donoho [et al.] // Journal of the Royal Statistical Society, series B. — 1995. — Vol. 57. — P. 301—369.
63. *Wie C.* On Predictive Least Squares Principles // The Annals of Statistics. — 1992. — Vol. 20. — P. 1—42.

64. *Wilson D. J.* Reply to Held: When is a harmonic mean p-value a Bayes factor? // Proceedings of the National Academy of Sciences. — 2019. — Vol. 116. — P. 5857—5858.
65. *Wilson D. J.* The harmonic mean p-value for combining dependent tests // Proceedings of the National Academy of Sciences of the United States of America. — 2019. — Vol. 116. — P. 1195—1200.

Публикации автора по теме диссертации

В изданиях из списка Web of Science, Scopus, RSCI

66. *Палионная С. И.* Сильная состоятельность оценки риска при множественной проверке гипотез с FDR-порогом // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. — 2020. — № 4. — С. 34—39.
67. *Палионная С. И.* Скорость сходимости оценки риска к нормальному закону в задаче множественной проверки гипотез с использованием FDR-порога // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. — 2021. — № 3. — С. 31—36.
68. *Палионная С. И.* Скорость сходимости распределения оценки риска к нормальному закону с использованием FDR-метода множественной проверки гипотез при обращении линейных однородных операторов // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. — 2022. — № 3. — С. 49—55.
69. *Палионная С. И., Шестаков О. В.* Использование FDR-метода множественной проверки гипотез при обращении линейных однородных операторов // Информатика и ее применения. — 2022. — Т. 16, № 2. — С. 44—51.

70. *Palionnaya S. I., Shestakov O. V.* Asymptotic Properties of MSE Estimate for the False Discovery Rate Controlling Procedures in Multiple Hypothesis Testing // *Mathematics*. — 2020. — Vol. 8, no. 11. — P. 1913.

В сборниках трудов конференций

71. *Палионная С. И., Шестаков О. В.* Свойства оценки риска при множественной проверке гипотез с использованием FDR-метода // *Тихоновские чтения*. — 2020. — С. 31—31.
72. *Палионная С. И., Шестаков О. В.* Асимптотическое поведение оценки риска FDR-метода в задаче множественной проверки гипотез // *Ломоносовские чтения-2021*. — 2021. — С. 120—121.
73. *Палионная С. И., Шестаков О. В.* Свойства оценок риска в задачах обращения линейных операторов при использовании FDR-метода множественной проверки гипотез // *Ломоносовские чтения-2022*. — 2022. — С. 171—172.
74. *Palionnaya S. I., Shestakov O. V.* Asymptotic behavior of a risk estimate for the FDR-method in the problem of multiple hypothesis testing // *XXXVI International Seminar on Stability Problems for Stochastic Models*. — 2021. — Available at: <http://isspsm2021.krc.karelia.ru/ru/conf/7/presentation>.