Автоматика и телемеханика, № 10, 2019

© 2019 г. А. КАРПАНТЬЕ, д-р (alexandra.carpentier@ovgu.de)
(Магдебургский университет, Германия),
О. КОЛЬЕ, д-р (olivier.collier@u-paris10.fr)
(Modal'X, Университет Париж-Нантер и CREST, Франция),
Л. КОМЕНЖ, д-р (comminges@ceremade.dauphine.fr)
(CEREMADE, Университет Париж-Дофин и CREST, Франция),
А.Б. ЦЫБАКОВ, д-р физ.-мат. наук (alexandre.tsybakov@ensae.fr)
(CREST, ENSAE, Париж, Франция),
Ю. ВАНГ (yuhaow@mit.edu)
(LIDS-IDSS, MIT, Кембридж, Массачусетс, США)

МИНИМАКСНАЯ ТОЧНОСТЬ ПРОВЕРКИ ГИПОТЕЗ В РАЗРЕЖЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ[1]

Рассматривается задача проверки гипотезы о том, что параметр модели линейной регрессии равен нулю против s-разреженной альтернативы, отделенной от нуля в ℓ_2-расстоянии. Показывается, что в модели гауссовской линейной регрессии с $p < n$, где p — размерность параметра, а n — размер выборки, неасимптотическая минимаксная скорость тестирования имеет вид $\sqrt{(s/n)\log(1+\sqrt{p}/s)}$. Также показывается, что это минимаксная скорость оценивания ℓ_2-нормы параметра регрессии.

Ключевые слова: линейная регрессия, разреженность, обнаружение сигнала.

DOI: 10.1134/S0005231019100040

1. Введение

Настоящая работа посвящена задаче проверки гипотез о параметре модели линейной регрессии при разреженных альтернативах. Эта задача имеет различные приложения в генетике, передаче и обнаружении сигналов, сжатии данных. Подробное описание этих приложений можно найти, например, в [6]. Важную роль играет поиск оптимальных методов проверки таких гипотез, причем один из естественных подходов состоит в том, чтобы понятие оптимальности определять в минимаксном смысле. Минимаксная постановка задачи проверки гипотез при разреженных альтернативах была впервые изучена в [1, 2], где рассматривалась модель гауссовской последовательности. В этих работах применялся асимптотический подход в предположении, что

[1] Работа А. Карпантье поддержана грантами Emmy Noether grant MuSyAD CA 1488/1-1, GK 2297 MathCoRe on Mathematical Complexity Reduction 314838170, GRK 2297 MathCoRe, SFB 1294 Data Assimilation "The seamless integration of data and models", Project A03, финансируемыми фондом Deutsche Foschungsgemeinschaft (DFG) и французско-немецким университетом в рамках программы докторского колледжа Потсдам-Тулуза CDFA 01-18. Работа О. Колье выполнена в рамках проекта Labex MME-DII (ANR11-LBX-0023-01). Работа А.Б. Цыбакова поддержана институтом GENES и грантами IPANEMA (ANR-13-BSH1-0004-02) и Labex Ecodec (ANR-11-LABEX-0047).

индекс разреженности зависит от размерности степенным образом. Неасимптотическая постановка для модели гауссовской последовательности была рассмотрена в [3], где границы для минимаксной скорости тестирования были установлены с точностью до логарифмического множителя. Наконец, точная неасимптотическая минимаксная скорость тестирования для модели гауссовской последовательности была получена в [4]. В настоящей работе результаты [4] обобщаются на линейную регрессионную модель с гауссовским шумом. Заметим, что задача минимаксного тестирования для линейной регрессии при разреженных альтернативах уже изучалась в [5–7]. А именно, в [5, 6] рассматривалась асимптотическая постановка при дополнительных предположениях о параметрах задачи, а в [7] были установлены неасимптотические границы с точностью до логарифмического множителя, аналогичные [3]. Цель данной работы состоит в том, чтобы найти неасимптотическую минимаксную скорость тестирования в модели гауссовской линейной регрессии без каких-либо специфических предположений о параметрах задачи. Здесь дается ответ на этот вопрос в случае, когда $p < n$, где p — размерность, а n — размер выборки.

Рассмотрим модель

$$Y = X\theta + \sigma\xi, \tag{1}$$

где $\sigma > 0$, $\xi \in \mathbb{R}^n$ — векторный гауссовский белый шум, т.е. $\xi \sim \mathcal{N}(0, I_n)$, I_n — единичная матрица размера $n \times n$, X — матрица размера $n \times p$ со случайными элементами и $\theta \in \mathbb{R}^p$ — неизвестный параметр. Далее везде предполагается, что X не зависит от ξ.

Ниже будут использованы следующие обозначения. Для $u = (u_1, \dots, u_p) \in \in \mathbb{R}^p$ через $\|\cdot\|_2$ обозначим ℓ_2-норму, т.е.

$$\|u\|_2^2 = \sum_{i=1}^{p} u_i^2,$$

и пусть $\|\cdot\|_0$ обозначает ℓ_0-полунорму, т.е.

$$\|u\|_0 = \sum_{i=1}^{p} \mathbf{1}_{\{u_i \neq 0\}},$$

где $\mathbf{1}_{\{.\}}$ — индикаторная функция. Обозначим через $\langle u, v\rangle = u^{\mathrm{T}} v$ скалярное произведение векторов $u \in \mathbb{R}^p$ и $v \in \mathbb{R}^p$, а через $\lambda_{\min}(M)$ и $\mathrm{tr}[M]$ минимальное собственное значение и след матрицы $M \in \mathbb{R}^{p\times p}$. Для целого $s \in \{1, \dots, p\}$ рассмотрим множество $B_0(s)$ всех s-разреженных векторов в пространстве $\mathbb{R}^p$:

$$B_0(s) := \{u \in \mathbb{R}^p : \|u\|_0 \leqslant s\}.$$

Рассмотрим следующую задачу проверки гипотез по наблюдениям (X, Y):

$$H_0 : \theta = 0 \quad \text{против альтернативы} \quad H_1 : \theta \in \Theta(s, \tau), \tag{2}$$

где

$$\Theta(s,\tau)=\{\theta\in B_0(s):\|\theta\|_2\geqslant\tau\}$$

для некоторых $s\in\{1,\dots,p\}$ и $\tau>0$. Пусть $\Delta=\Delta(X,Y)$ — статистика со значениями в $\{0,1\}$. Определим риск теста, основанного на Δ, как сумму ошибки первого рода и максимальной ошибки второго рода:

$$\mathbf{P}_0(\Delta=1)+\sup_{\theta\in\Theta(s,\tau)}\mathbf{P}_\theta(\Delta=0),$$

где $\mathbf{P}_\theta$ обозначает совместное распределение пары (X,Y), удовлетворяющей (1). Наименьшее возможное значение этого риска равно минимаксному риску

$$\mathcal{R}_{s,\tau}:=\inf_{\Delta}\left\{\mathbf{P}_0(\Delta=1)+\sup_{\theta\in\Theta(s,\tau)}\mathbf{P}_\theta(\Delta=0)\right\},$$

где нижняя грань inf берется по всем $\{0,1\}$-значным статистикам Δ. Определим минимаксную скорость тестирования *для класса* $B_0(s)$ *относительно* ℓ_2*-расстояния* как значение $\lambda>0$, для которого выполняются следующие два свойства:

(i) (*верхняя граница*) для любого $\varepsilon\in(0,1)$ существует $A_\varepsilon>0$, не зависящее от p,n,s,σ, такое что для всех $A>A_\varepsilon$

$$\mathcal{R}_{s,A\lambda}\leqslant\varepsilon, \tag{3}$$

(ii) (*нижняя граница*) для любого $\varepsilon\in(0,1)$ существует $a_\varepsilon>0$, не зависящее от p,n,s,σ, такое что для всех $0<A<a_\varepsilon$

$$\mathcal{R}_{s,A\lambda}\geqslant 1-\varepsilon. \tag{4}$$

Заметим, что определенная таким образом скорость λ неасимптотическая, и поэтому данное определение отличается от классического асимптотического определения минимаксной скорости тестирования, которое можно найти, например, в [8]. В [4] показано, что когда X — единичная матрица и $p=n$ (что соответствует модели гауссовской последовательности), неасимптотическая минимаксная скорость тестирования на классе $B_0(s)$ по отношению к ℓ_2-расстоянию имеет следующий вид:

$$\lambda=\begin{cases}\sigma\sqrt{s\log(1+p/s^2)}, & \text{если}\quad s<\sqrt{p},\\ \sigma p^{1/4}, & \text{если}\quad s\geqslant\sqrt{p}.\end{cases} \tag{5}$$

Для модели регрессии со случайным X, удовлетворяющей некоторым сильным предположениям, асимптотическая минимаксная скорость тестирования при n,p и s стремящимся к ∞ так, что $s=p^a$ для некоторого $0<a<1$, изучена в [5]. В частности, в [5] показано, что для этой конфигурации параметров и при условии, что матрица X имеет независимые одинаково распределенные (н.о.р.) стандартные нормальные компоненты, асимптотическая скорость имеет вид

$$\lambda = \sigma \min\left(\sqrt{\frac{s\log(p)}{n}}, n^{-1/4}, \frac{p^{1/4}}{\sqrt{n}}\right). \tag{6}$$

Аналогичный результат для несколько иначе определенной альтернативы H_1 получен в [6].

Ниже будет показано, что без каких-либо специфических ограничений на параметры n, p и s неасимптотическая нижняя граница для минимаксной скорости тестирования имеет вид

$$\lambda = \sigma \min\left(\sqrt{\frac{s\log(2+p/s^2)}{n}}, n^{-1/4}, \frac{p^{1/4}}{\sqrt{n}}\right), \tag{7}$$

какова бы ни была матрица X с изотропным распределением и независимыми субгауссовскими элементами (определения изотропности и субгауссовости распределения даны в разделе 3). Кроме того, будет показано, что верхняя граница совпадает по порядку с нижней в случае, когда X — матрица с н.о.р. стандартными гауссовскими элементами и $p < n$. Заметим, что при $p < n$ выражение (7) принимает вид

$$\lambda = \sigma \min\left(\sqrt{\frac{s\log(2+p/s^2)}{n}}, \frac{p^{1/4}}{\sqrt{n}}\right). \tag{8}$$

Также полезно отметить, что, поскольку для $s \leqslant \sqrt{p}$ функция $s \mapsto s\log(2+ +p/s^2)$ возрастает и удовлетворяет неравенству $\log(2+p/s^2) \leqslant 2\log(1+p/s^2)$, скорость (8) можно эквивалентным образом (с точностью до числового постоянного множителя) записать в виде

$$\lambda = \begin{cases} \sigma\sqrt{\dfrac{s\log(1+p/s^2)}{n}}, & \text{если} \quad s < \sqrt{p}, \\ \sigma\dfrac{p^{1/4}}{\sqrt{n}}, & \text{если} \quad s \geqslant \sqrt{p}. \end{cases} \tag{9}$$

Это выражение аналогично (5). Наконец, заметим, что скорость может быть записана в следующей более компактной форме:

$$\sigma \min\left(\sqrt{\frac{s\log(2+p/s^2)}{n}}, \frac{p^{1/4}}{\sqrt{n}}\right) \asymp \sigma\sqrt{\frac{s\log(1+\sqrt{p}/s)}{n}}, \tag{10}$$

где $\asymp$ обозначает эквивалентность с точностью до числового постоянного множителя.

2. Верхние границы для минимаксных скоростей

В этом разделе предполагается, что $p < n$ и X — матрица с н.о.р. стандартными гауссовскими элементами, и устанавливается верхняя граница (9) для минимаксной скорости тестирования. Для этого используется связь между проверкой гипотез и оценкой функционалов. Сначала будет введена оценка $\hat{Q}$ квадратичного функционала $\|\theta\|_2^2$ и установлена верхняя граница для его риска. Затем из этого результата будет выведена верхняя граница для риска оценки $\hat{N}$ нормы $\|\theta\|_2$, определяемой следующим образом:

$$\hat{N} = \sqrt{\max(\hat{Q}, 0)}.$$

Наконец, используя $\hat{N}$ как статистику критерия, получим верхнюю оценку минимаксной скорости тестирования.

Введем обозначение

$$\alpha_s = \mathbf{E}\left(Z^2 | Z^2 > 2\log(1 + p/s^2)\right),$$

где Z — стандартная нормальная величина, и положим

$$y_i = \left\{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y\right\}_i,$$

где $\{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y\}_i$ — i-я компонента оценки наименьших квадратов $(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y$. Заметим, что обратная матрица $(X^{\mathrm{T}}X)^{-1}$ существует почти наверное, так как в этом разделе предполагается, что X является матрицей с н.о.р. стандартными гауссовскими элементами и $p < n$, так что X почти наверное имеет полный ранг. Рассмотрим следующую оценку квадратичного функционала $\|\theta\|_2^2$:

$$\hat{Q} := \begin{cases} \displaystyle\sum_{i=1}^{p} y_i^2 - \sigma^2 \mathrm{tr}[(X^{\mathrm{T}}X)^{-1}], & \text{если} \quad s \geqslant \sqrt{p}, \\ \displaystyle\sum_{i=1}^{p} \left[y_i^2 - \sigma^2 (X^{\mathrm{T}}X)^{-1}_{ii}\alpha_s\right] \mathbf{1}_{\{y_i^2 > 2\sigma^2 (X^{\mathrm{T}}X)^{-1}_{ii}\log(1+p/s^2)\}}, & \text{если} \quad s < \sqrt{p}. \end{cases}$$

Здесь и далее $(X^{\mathrm{T}}X)^{-1}_{ij}$ обозначает (i,j)-й элемент матрицы $(X^{\mathrm{T}}X)^{-1}$.

Для любых целых чисел n, p, s таких, что $s \leqslant p$, положим

$$\psi(s,p) = \begin{cases} \dfrac{s\log(1+p/s^2)}{n}, & \text{если} \quad s < \sqrt{p}, \\ \dfrac{p^{1/2}}{n}, & \text{если} \quad s \geqslant \sqrt{p}. \end{cases}$$

Теорема 1. *Пусть n, p, s — целые числа такие, что $s \leqslant p$, $n \geqslant 9$ и $p \leqslant \min(\gamma n, n-8)$ для некоторой константы $0 < \gamma < 1$. Пусть $r > 0$, $\sigma > 0$. Предположим, что все элементы матрицы X — н.о.р. стандартные гауссовские случайные величины. Тогда существует константа $c > 0$, зависящая только от γ, такая что*

$$\sup_{\theta: \|\theta\|_0 \leqslant s, \|\theta\|_2 \leqslant r} \mathbf{E}_\theta \left[(\hat{Q} - \|\theta\|_2^2)^2\right] \leqslant c\left(\sigma^2 \frac{r^2}{n} + \sigma^4\psi^2(s,p)\right).$$

Доказательство теоремы 1 дано в Приложении.

Рассуждая точно так же, как в доказательстве теоремы 8 в [4], из теоремы 1 выводим следующую верхнюю оценку квадратичного риска оценки $\hat{N}$.

Теорема 2. *Пусть выполнены условия теоремы 1. Тогда существует константа $c' > 0$, зависящая только от γ, такая что*

$$\sup_{\theta \in B_0(s)} \mathbf{E}_\theta \left[(\hat{N} - \|\theta\|_2)^2 \right] \leqslant c' \sigma^2 \psi(s,p).$$

Из теоремы 2 следует, что для теста $\Delta^* = \mathbf{1}_{\{\hat{N} > A\lambda/2\}}$, где $\lambda = \sigma\sqrt{\psi(s,p)}$ (т.е. то же значение λ, что и в (9)), справедливы неравенства

$$\mathbf{P}_0(\Delta^* = 1) + \sup_{\theta \in \Theta(s, A\lambda)} \mathbf{P}_\theta(\Delta^* = 0) \leqslant$$
$$\leqslant \mathbf{P}_0(\hat{N} > A\lambda/2) + \sup_{\theta \in B_0(s)} \mathbf{P}_\theta(\hat{N} - \|\theta\|_2 \leqslant -A\lambda/2) \leqslant$$
$$\leqslant 2 \sup_{\theta \in B_0(s)} \frac{\mathbf{E}_\theta[(\hat{N} - \|\theta\|_2)^2]}{(A/2)^2\lambda^2} \leqslant C_* A^{-2},$$

где $C_* > 0$ — некоторая константа. Используя это замечание и выбирая $A_\varepsilon = (C_*/\varepsilon)^{1/2}$, получаем верхнюю оценку (i), определенную в предыдущем разделе. Этот вывод сформулирован в следующей теореме.

Теорема 3. *Пусть выполнены условия теоремы 1 и пусть λ определяется через* (9). *Тогда для любого $\varepsilon \in (0,1)$ существует $A_\varepsilon > 0$, зависящее только от ε и γ, такое что*

$$\mathcal{R}_{s, A\lambda} \leqslant \varepsilon$$

для всех $A > A_\varepsilon$.

3. Нижние границы для минимаксных скоростей

В этом разделе предполагается, что распределение матрицы X изотропно и имеет независимые σ_X-субгауссовские строки для некоторого $\sigma_X > 0$. Изотропия X означает, что $E_X(X^{\mathrm{T}}X/n) = I_p$, где E_X обозначает математическое ожидание по распределению P_X матрицы X.

Определение 1. *Пусть $b > 0$. Вещественнозначная случайная величина ζ называется b-субгауссовской, если*

$$\mathbf{E}\exp(t\zeta) \leqslant \exp(b^2t^2/2), \quad \forall t \in \mathbb{R}.$$

Случайный вектор η со значениями в $\mathbb{R}^d$ называется b-субгауссовским, если все скалярные произведения $\langle \eta, v \rangle$ с векторами $v \in \mathbb{R}^d$, такими что $\|v\|_2 = 1$, являются b-субгауссовскими случайными величинами.

Следующая теорема о нижней границе является неасимптотической и она справедлива без ограничений на параметры n, p, s за исключением неизбежного условия $s \leqslant p$.

Теорема 4. Пусть $\varepsilon \in (0,1)$, $\sigma > 0$ и пусть целые n, p, s таковы, что $s \leqslant p$. Предположим, что распределение матрицы X изотропно и X имеет независимые σ_X-субгауссовские строки при некотором $\sigma_X > 0$. Тогда существует $a_\varepsilon > 0$, зависящее только от ε, и σ_X, такое что для

$$\tau = A\sigma \min\left(\sqrt{\frac{s\log(2+p/s^2)}{n}}, n^{-1/4}, \frac{p^{1/4}}{\sqrt{n}}\right) \tag{11}$$

с произвольным A, удовлетворяющим $0 < A < a_\varepsilon$, имеем

$$\mathcal{R}_{s,\tau} \geqslant 1 - \varepsilon.$$

Доказательство теоремы 4 дано в Приложении.

Следующий результат является непосредственным следствием теорем 3 и 4.

Следствие 1. Пусть выполнены условия теоремы 1. Тогда минимаксная скорость тестирования на классе $B_0(s)$ по отношению к ℓ_2-расстоянию задается через (8).

Кроме того, из теоремы 4 получаем следующую нижнюю границу минимаксного риска оценки ℓ_2-нормы $\|\theta\|_2$.

Теорема 5. Пусть выполнены условия теоремы 4 и пусть λ определено в (7). Тогда существует константа $c_* > 0$, зависящая только от σ_X, такая что

$$\inf_{\hat{T}} \sup_{\theta \in B_0(s)} \mathbf{E}_\theta \left[(\hat{T} - \|\theta\|_2)^2\right] \geqslant c_* \lambda^2,$$

где $\inf_{\hat{T}}$ обозначает инфимум по всем оценкам.

Результат теоремы 5 вытекает из теоремы 4, если заметить, что для τ в (11) и λ в (7) имеем $\tau = A\lambda$ и для любой оценки $\hat{T}$ справедливы неравенства

$$\begin{aligned}\sup_{\theta \in B_0(s)} \mathbf{E}_\theta \left[(\hat{T} - \|\theta\|_2)^2\right] &\geqslant \frac{1}{2}\left[\mathbf{E}_0[\hat{T}^2] + \sup_{\theta \in \Theta(s,\tau)} \mathbf{E}_\theta \left[(\hat{T} - \|\theta\|_2)^2\right]\right] \geqslant \\ &\geqslant \frac{\tau^2}{8}\left[\mathbf{P}_0(\hat{T} > \tau/2) + \sup_{\theta \in \Theta(s,\tau)} \mathbf{P}_\theta(\hat{T} \leqslant \tau/2)\right] \geqslant \\ &\geqslant \frac{(A\lambda)^2}{8}\mathcal{R}_{s,\tau}.\end{aligned}$$

Следствие 2. Пусть выполнены условия теоремы 1 и пусть λ определено в (8). Тогда минимаксная скорость оценки нормы $\|\theta\|_2$ относительно среднеквадратичного риска для класса $B_0(s)$ равна λ^2, т.е.

$$c_* \lambda^2 \leqslant \inf_{\hat{T}} \sup_{\theta \in B_0(s)} \mathbf{E}_\theta \left[(\hat{T} - \|\theta\|_2)^2\right] \leqslant c' \lambda^2,$$

где $c_* > 0$ — числовая константа, а $c' > 0$ — константа, зависящая только от γ.

Это утверждение является непосредственным следствием теорем 2 и 5.

Замечание 1. Из доказательств, приведенных в Приложении, видно, что результаты этого раздела остаются в силе, если заменить везде ℓ_0-шар $B_0(s)$ на ℓ_0-сферу $\bar{B}_0(s) = \{u \in \mathbb{R}^p : \|u\|_0 = s\}$.

ПРИЛОЖЕНИЕ

П.1. Вспомогательные леммы для доказательства теоремы 1

Приведем два основных вспомогательных результата, используемых в доказательстве теоремы 1. Первый из них позволяет оценить математическое ожидание от отрицательных степеней минимального собственного значения эмпирической матрицы ковариаций. Второй используется при оценке ошибки идентификации ненулевых компонент в разреженной постановке. Для этого нужны точные границы для корреляций между урезанными вариантами двух центрированных коррелированных случайных величин с распределением χ_1^2. Напомним сначала два общих факта, которые будут использоваться для получения первого результата.

В дальнейшем будем обозначать одним и тем же символом C различные положительные константы, встречающиеся в выкладках.

Лемма 1 ([9], см. также [10]). *Пусть матрица X удовлетворяет условиям теоремы 1 и пусть $\lambda_{\min}(\hat{\Sigma})$ — наименьшее собственное значение выборочной ковариационной матрицы $\hat{\Sigma} = \frac{1}{n} X^{\mathrm{T}} X$. Тогда для любого $t > 0$ с вероятностью не менее $1 - 2\exp(-t^2/2)$ имеем*

$$1 - \sqrt{\frac{p}{n}} - \frac{t}{\sqrt{n}} \leqslant \sqrt{\lambda_{\min}(\hat{\Sigma})} \leqslant 1 + \sqrt{\frac{p}{n}} + \frac{t}{\sqrt{n}}.$$

Лемма 2 ([11, лемма A4], см. также [12, лемма 4.14]). *Пусть $1 \leqslant p \leqslant n$. Пусть R_i — i-й столбец матрицы $X \in \mathbb{R}^{n \times p}$ и R_{-i} — линейная оболочка векторов $\{R_j : j \neq i\}$. Если X — матрица полного ранга, то*

$$(X^{\mathrm{T}} X)^{-1}_{ii} = \operatorname{dist}(R_i, R_{-i})^{-2},$$

где $\operatorname{dist}(R_i, R_{-i})$ — евклидово расстояние вектора R_i до пространства R_{-i}.

Лемма 3. Пусть $n \geqslant 9$ и $p \leqslant \min(\gamma n, n - 8)$ для некоторой постоянной γ такой, что $0 < \gamma < 1$. Предположим, что все компоненты матрицы $X \in \mathbb{R}^{n \times p}$ — н.о.р. стандартные гауссовские случайные величины. Тогда существует константа $c > 0$, зависящая только от γ, такая что

$$\mathbf{E}\left[\lambda_{\min}^{-2}(\hat{\Sigma})\right] \leqslant c. \tag{П.1}$$

Доказательство. Положим $\beta = \sqrt{\gamma}$. Из неравенства $p \leqslant \gamma n$ и леммы 1 имеем

$$\mathbf{P}\left(\sqrt{\lambda_{\min}(\hat{\Sigma})} < 1 - \beta - \frac{t}{\sqrt{n}}\right) \leqslant \mathbf{P}\left(\sqrt{\lambda_{\min}(\hat{\Sigma})} < 1 - \sqrt{\frac{p}{n}} - \frac{t}{\sqrt{n}}\right) \leqslant$$
$$\leqslant 2\exp(-t^2/2).$$

Полагая здесь $t = \sqrt{n}(1-\beta)/2$, приходим к неравенству

$$\mathbf{P}\left(\lambda_{\min}(\hat{\Sigma}) < \left(\frac{1-\beta}{2}\right)^2\right) \leqslant 2\exp\left(-\frac{n(1-\beta)^2}{8}\right).$$

Используя это неравенство, получим

$$\mathbf{E}\left[\lambda_{\min}^{-2}(\hat{\Sigma})\right] \leqslant \left(\frac{1-\beta}{2}\right)^{-4} + \sqrt{\mathbf{E}\left[\lambda_{\min}^{-4}(\hat{\Sigma})\right]}\sqrt{2}\exp\left(-\frac{n(1-\beta)^2}{16}\right). \tag{П.2}$$

Оценим теперь сверху математическое ожидание $\mathbf{E}[\lambda_{\min}^{-4}(\hat{\Sigma})]$. Ясно, что

$$\lambda_{\min}^{-1}(\hat{\Sigma}) \leqslant \operatorname{tr}\left[\hat{\Sigma}^{-1}\right]. \tag{П.3}$$

Из леммы 2 следует, что почти наверное

$$\left(\operatorname{tr}[\hat{\Sigma}^{-1}]\right)^4 = n^4\left[\sum_{i=1}^{p}\operatorname{dist}(R_i, R_{-i})^{-2}\right]^4 \leqslant n^4 p^3 \sum_{i=1}^{p}\operatorname{dist}(R_i, R_{-i})^{-8}.$$

Поскольку случайные величины $\operatorname{dist}(R_i, R_{-i})$ одинаково распределены и $p \leqslant$ $\leqslant n$, имеем

$$\mathbf{E}\left[\left(\operatorname{tr}[\hat{\Sigma}^{-1}]\right)^4\right] \leqslant n^8\mathbf{E}\left[\operatorname{dist}(R_1, R_{-1})^{-8}\right]. \tag{П.4}$$

Остается только оценить сверху $\mathbf{E}[\operatorname{dist}(R_1, R_{-1})^{-8}]$. Если $\mathcal{S}$ — подпространство $\mathbb{R}^n$ размерности $p-1$, то случайная величина $\operatorname{dist}(R_1, \mathcal{S})^2$ имеет χ^2-распределение χ^2_{n-p+1} с $n-p+1$ степенями свободы. Следовательно, так как R_{-1} — линейная оболочка случайных векторов, не зависящих от R_1, и пространство R_{-1} — почти наверное $(p-1)$-мерное, то

$$\begin{aligned}\mathbf{E}[\operatorname{dist}(R_1, R_{-1})^{-8}] &= \mathbf{E}\left[\frac{1}{(\chi^2_{n-p+1})^4}\right] = \\ &= \frac{1}{(n-p-1)(n-p-3)(n-p-5)(n-p-7)} \leqslant \frac{1}{105}.\end{aligned} \tag{П.5}$$

Объединяя (П.2)–(П.5), получим

$$\mathbf{E}\left[\lambda_{\min}^{-2}(\hat{\Sigma})\right] \leqslant \left(\frac{1-\beta}{2}\right)^{-4} + \frac{n^8}{105}\sqrt{2}\exp\left(-\frac{n(1-\beta)^2}{16}\right),$$

что доказывает лемму.

Теперь перейдем ко второму вопросу данного раздела, т.е. к оценке сверху корреляций. Здесь понадобится следующая лемма о хвостах стандартного нормального распределения.

Лемма 4. Для $\eta \sim \mathcal{N}(0,1)$ и всякого $x > 0$ имеем

$$
\text{(П.6)} \qquad \begin{aligned} \frac{4}{\sqrt{2\pi}(x+\sqrt{x^2+4})}\exp(-x^2/2) \leqslant \mathbf{P}(|\eta| > x) \leqslant \\ \leqslant \frac{4}{\sqrt{2\pi}(x+\sqrt{x^2+2})}\exp(-x^2/2), \end{aligned}
$$

$$
\text{(П.7)} \qquad \mathbf{E}[\eta^2 \mathbf{1}_{\{|\eta|>x\}}] \leqslant \sqrt{\frac{2}{\pi}}\left(x+\frac{2}{x}\right)\exp(-x^2/2),
$$

$$
\text{(П.8)} \qquad \mathbf{E}[\eta^4 \mathbf{1}_{\{|\eta|>x\}}] \leqslant \sqrt{\frac{2}{\pi}}\left(x^3+3x+\frac{1}{x}\right)\exp(-x^2/2).
$$

Более того, если $x \geqslant 1$, то

$$
\text{(П.9)} \qquad x^2 < \mathbf{E}[\eta^2 \mid |\eta| > x] \leqslant 5x^2.
$$

Неравенства (П.6)–(П.8) даны, например, в [4, лемма 4], а (П.9) легко вытекает из (П.6) и (П.7).

Лемма 5. Пусть (η, ζ) — гауссовский вектор со средним 0 и матрицей ковариаций $\Gamma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $0 < \rho < 1$. Положим $\alpha = \mathbf{E}[\eta^2 \mid |\eta| > x]$. Тогда существует числовая постоянная $C > 0$ такая, что

$$
\mathbf{E}\left[(\eta^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\eta|>x\}}\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant C\rho^2 x^4 \exp(-x^2/2)
$$

для всякого $x \geqslant 1$.

Доказательство. Из (П.9) получаем, что $\alpha \leqslant 5x^2$. Таким образом, используя (П.8) и условие, что $x \geqslant 1$, находим

$$
\text{(П.10)} \qquad \begin{aligned} \mathbf{E}\left[(\zeta^2-\alpha)^2\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant \mathbf{E}\left[(\zeta^4+\alpha^2)\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant \\ \leqslant 26\mathbf{E}\left[\zeta^4\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant Cx^3\exp\left(-x^2/2\right). \end{aligned}
$$

Следовательно,

$$
\begin{gathered} \mathbf{E}\left[(\eta^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\eta|>x\}}\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant \\ \leqslant \mathbf{E}\left[(\eta^2-\alpha)^2\mathbf{1}_{\{|\eta|>x\}}\right] + \mathbf{E}\left[(\zeta^2-\alpha)^2\mathbf{1}_{\{|\zeta|>x\}}\right] \leqslant Cx^3\exp\left(-x^2/2\right). \end{gathered}
$$

Это доказывает лемму для $\rho \geqslant 1/\sqrt{5}$.

Рассмотрим теперь случай, когда $0 < \rho < 1/\sqrt{5}$. Заметим, что, поскольку $\alpha \leqslant 5x^2$, для $0 < \rho < 1/\sqrt{5}$ также выполняется неравенство

$$
\rho < \frac{x}{\sqrt{\alpha}}.
$$

В силу симметричности распределения (η, ζ) имеем

$$\text{(П.11)}\qquad \begin{aligned} &\mathbf{E}\Big[(\eta^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\eta|>x\}}\mathbf{1}_{\{|\zeta|>x\}}\Big] = \\ &= 2\mathbf{E}\Big[(\eta^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\eta|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big]. \end{aligned}$$

Используем теперь тот факт, что $(\eta, \zeta) \overset{d}{=} (\rho\zeta + \sqrt{1-\rho^2}Z, \zeta)$, где $\overset{d}{=}$ означает равенство по распределению и Z — стандартная гауссовская случайная величина, не зависящая от ζ. Таким образом,

$$\text{(П.12)}\qquad \begin{aligned} \mathbf{E}\Big[(\eta^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\eta|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] &= \rho^2\mathbf{E}\Big[(\zeta^2-\alpha)^2\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] + \\ &+ 2\rho\sqrt{1-\rho^2}\mathbf{E}\Big[\zeta Z(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] + \\ &+ (1-\rho^2)\mathbf{E}\Big[(Z^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big]. \end{aligned}$$

Оценим сверху каждое из трех слагаемых в правой части (П.12). Для первого слагаемого, используя (П.10), получаем оценку

$$\text{(П.13)}\qquad \begin{aligned} &\rho^2\mathbf{E}\Big[(\zeta^2-\alpha)^2\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] \leqslant \\ &\leqslant 26\rho^2\mathbf{E}\Big[\zeta^4\mathbf{1}_{\{\zeta>x\}}\Big] \leqslant C\rho^2x^3\exp\left(-\frac{x^2}{2}\right). \end{aligned}$$

Чтобы оценить второе слагаемое, запишем сначала

$$\text{(П.14)}\qquad \mathbf{E}\Big[\zeta Z(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] = \mathbf{E}\Big[\zeta(\zeta^2-\alpha)g(\zeta)\mathbf{1}_{\{\zeta>x\}}\Big],$$

где $g(\zeta) := \mathbf{E}[Z\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}} \mid \zeta]$. Непосредственно проверяется, что

$$g(\zeta) = \exp\left(-\frac{(x-\rho\zeta)^2}{2(1-\rho^2)}\right) - \exp\left(-\frac{(x+\rho\zeta)^2}{2(1-\rho^2)}\right).$$

Таким образом, $g(\zeta)$ положительна при $\zeta > x$. Следовательно, имеем

$$\text{(П.15)}\qquad \mathbf{E}\Big[\zeta(\zeta^2-\alpha)g(\zeta)\mathbf{1}_{\{\zeta>x\}}\Big] \leqslant \mathbf{E}\Big[\zeta^3g(\zeta)\mathbf{1}_{\{\zeta>x\}}\Big].$$

Кроме того,

$$\text{(П.16)}\qquad \begin{aligned} g(\zeta) &= \exp\left(-\frac{(x-\rho\zeta)^2}{2(1-\rho^2)}\right)\left(1-\exp\left(-\frac{2x\rho\zeta}{1-\rho^2}\right)\right) \leqslant \\ &\leqslant 1-\exp\left(-\frac{2x\rho\zeta}{1-\rho^2}\right) \leqslant \frac{2x\rho\zeta}{1-\rho^2}. \end{aligned}$$

Объединяя (П.14)–(П.16) с (П.8) и учитывая, что $\rho \leqslant \frac{1}{2}$, получаем

$$2\rho\sqrt{1-\rho^2}\mathbf{E}\Big[\zeta Z(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] \leqslant C\rho^2x^4\exp\left(-\frac{x^2}{2}\right). \tag{П.17}$$

Рассмотрим теперь третье слагаемое правой части (П.12). Докажем, что

$$\mathbf{E}\Big[(Z^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] \leqslant 0. \tag{П.18}$$

Имеем

$$\mathbf{E}\Big[(Z^2-\alpha)(\zeta^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}}\mathbf{1}_{\{\zeta>x\}}\Big] = \mathbf{E}\Big[(\zeta^2-\alpha)f(\zeta)\mathbf{1}_{\{\zeta>x\}}\Big],$$

где

$$f(\zeta) := \mathbf{E}\Big[(Z^2-\alpha)\mathbf{1}_{\{|\rho\zeta+\sqrt{1-\rho^2}Z|>x\}} \mid \zeta\Big] =$$
$$= \int\limits_{\frac{x-\rho\zeta}{\sqrt{1-\rho^2}}}^{\infty} (z^2-\alpha)\exp\left(-\frac{z^2}{2}\right)dz + \int\limits_{-\infty}^{-\frac{x+\rho\zeta}{\sqrt{1-\rho^2}}} (z^2-\alpha)\exp\left(-\frac{z^2}{2}\right)dz.$$

Заметим, что $x < \sqrt{\alpha}$ в силу (П.9). Чтобы доказать (П.18), достаточно показать, что

$$\forall\,\zeta \in [x,\sqrt{\alpha}], \qquad f(\zeta) \geqslant f(\sqrt{\alpha}) \tag{П.19}$$

и

$$\forall\,\zeta \in [\sqrt{\alpha},\infty), \qquad f(\zeta) \leqslant f(\sqrt{\alpha}). \tag{П.20}$$

Действительно, предположим, что (П.19) и (П.20) выполнены. Тогда имеем

$$\mathbf{E}\Big[(\zeta^2-\alpha)f(\zeta)\mathbf{1}_{\{x<\zeta\leqslant\sqrt{\alpha}\}}\Big] \leqslant \mathbf{E}\Big[(\zeta^2-\alpha)f(\sqrt{\alpha})\mathbf{1}_{\{x<\zeta\leqslant\sqrt{\alpha}\}}\Big] =$$
$$= -\mathbf{E}\Big[(\zeta^2-\alpha)f(\sqrt{\alpha})\mathbf{1}_{\{\zeta>\sqrt{\alpha}\}}\Big] \leqslant -\mathbf{E}\Big[(\zeta^2-\alpha)f(\zeta)\mathbf{1}_{\{\zeta>\sqrt{\alpha}\}}\Big],$$

где равенство следует из того, что

$$\mathbf{E}\Big[(\zeta^2-\alpha)\mathbf{1}_{\{\zeta>x\}}\Big] = \frac{1}{2}\mathbf{E}\Big[(\zeta^2-\alpha)\mathbf{1}_{\{|\zeta|>x\}}\Big] = 0$$

ввиду симметричности нормального распределения и определения α. Таким образом, чтобы закончить доказательство леммы, остается показать (П.19) и (П.20). Сначала установим (П.19), для чего достаточно показать, что $f'(\zeta) < 0$ при $\zeta \in [x,\sqrt{\alpha}]$. Поскольку $0 < \rho < x/\sqrt{\alpha}$ и $x < \sqrt{\alpha}$, имеем

$$\frac{(x-\rho y)^2}{1-\rho^2} < \alpha \quad \text{при всех} \quad y \in [x,\sqrt{\alpha}]. \tag{П.21}$$

Используя (П.21), для всех $\zeta \in [x, \sqrt{\alpha}]$ будем иметь

$$f'(\zeta) = \frac{\rho}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{x+\rho\zeta}{\sqrt{1-\rho^2}}\right)^2\right)\left(\left(\frac{(x-\rho\zeta)^2}{1-\rho^2} - \alpha\right)\exp\left(\frac{2\rho x\zeta}{1-\rho^2}\right) - \left(\frac{(x+\rho\zeta)^2}{1-\rho^2} - \alpha\right)\right) \leqslant$$

$$\leqslant \frac{\rho}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{x+\rho\zeta}{\sqrt{1-\rho^2}}\right)^2\right)\left(\left(\frac{(x-\rho\zeta)^2}{1-\rho^2} - \alpha\right) - \left(\frac{(x+\rho\zeta)^2}{1-\rho^2} - \alpha\right)\right) =$$

$$= -\frac{\rho}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{x+\rho\zeta}{\sqrt{1-\rho^2}}\right)^2\right)\frac{4x\rho\zeta}{1-\rho^2} < 0.$$

Это влечет (П.19). Наконец, докажем (П.20). Для этого достаточно установить следующие три факта:
(i) f' непрерывна и $f'(\sqrt{\alpha}) < 0$;
(ii) уравнение $f'(y) = 0$ имеет не более одного решения в интервале $[\sqrt{\alpha}, +\infty)$;
(iii) $f(\infty) = \lim_{y\to\infty} f(y) \leqslant f(\sqrt{\alpha})$.
Свойство (i) уже доказано выше. Чтобы доказать (ii), прежде всего заметим, что решение уравнения $\frac{d}{dy}f(y) = 0$ также является решением уравнения $h(y) = 0$, где

$$h(y) := \left(\frac{(x-\rho y)^2}{1-\rho^2} - \alpha\right)\left(\exp\left(\frac{2\rho xy}{1-\rho^2}\right) - 1\right) - \frac{4\rho xy}{1-\rho^2}.$$

Пусть теперь y_1 и y_2 — решения квадратного уравнения $\frac{(x-\rho y)^2}{1-\rho^2} = \alpha$:

$$y_1 = \frac{x - \sqrt{\alpha(1-\rho^2)}}{\rho} \quad \text{и} \quad y_2 = \frac{x + \sqrt{\alpha(1-\rho^2)}}{\rho}.$$

В силу (П.21) имеем $y_1 < \sqrt{\alpha} < y_2$. Следовательно, $h(y) < 0$ на интервале $[\sqrt{\alpha}, y_2]$. Кроме того, на интервале $(y_2, +\infty)$ функция h строго выпукла и $h(y_2) < 0$. Отсюда следует, что h принимает нулевое значение только в одной точке интервала $(y_2, +\infty)$. Следовательно, (ii) доказано.

Остается показать, что $f(\sqrt{\alpha}) \geqslant f(\infty) = \int_{-\infty}^{\infty}(z^2-\alpha)\exp(-z^2/2)dz$. Переписывая $f(\sqrt{\alpha})$ в виде

$$f(\sqrt{\alpha}) = f(\infty) - \int\limits_{-\frac{x+\rho\sqrt{\alpha}}{\sqrt{1-\rho^2}}}^{\frac{x-\rho\sqrt{\alpha}}{\sqrt{1-\rho^2}}} (z^2-\alpha)\exp\left(-\frac{z^2}{2}\right)dz,$$

получаем, что неравенство $f(\infty) \leqslant f(\sqrt{\alpha})$ вытекает из (П.21). Это доказывает (iii) и, значит, (П.20). Доказательство (П.18) завершено. Лемма теперь следует из (П.11)–(П.13), (П.17) и (П.18).

П.2. Доказательство теоремы 1

Рассмотрим отдельно случаи $s \geqslant \sqrt{p}$ и $s < \sqrt{p}$.

Случай $s \geqslant \sqrt{p}$. Из (1) получаем, что почти наверное

$$(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y = \theta + \tilde{\epsilon},$$

где

$$\tilde{\epsilon} = \sigma(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\xi.$$

Учитывая это, имеем

$$\begin{aligned}\mathbf{E}_\theta\left[\left(\hat{Q} - \|\theta\|_2^2\right)^2\right] &= \mathbf{E}_\theta\left(2\theta^{\mathrm{T}}\tilde{\epsilon} + \|\tilde{\epsilon}\|_2^2 - \sigma^2\mathrm{tr}\left[(X^{\mathrm{T}}X)^{-1}\right]\right)^2 \leqslant \\ &\leqslant 8\,\mathbf{E}_\theta\left(\theta^{\mathrm{T}}\tilde{\epsilon}\right)^2 + 2\,\mathbf{E}_\theta\left(\|\tilde{\epsilon}\|_2^2 - \sigma^2\mathrm{tr}\left[(X^{\mathrm{T}}X)^{-1}\right]\right)^2.\end{aligned} \tag{П.22}$$

Заметим, что при фиксированном X случайный вектор $\tilde{\epsilon}$ имеет нормальное распределение со средним 0 и матрицей ковариаций $\sigma^2(X^{\mathrm{T}}X)^{-1}$. Следовательно, при фиксированном X случайная величина $\theta^{\mathrm{T}}\tilde{\epsilon}$ имеет нормальное распределение со средним 0 и дисперсией $\sigma^2\theta^{\mathrm{T}}(X^{\mathrm{T}}X)^{-1}\theta$. Таким образом, $\mathbf{E}_\theta\left(\theta^{\mathrm{T}}\tilde{\epsilon}\right)^2 \leqslant \sigma^2 r^2\mathbf{E}\left[\lambda_{\min}^{-1}(X^{\mathrm{T}}X)\right]$. Значит, применяя лемму 3, имеем

$$\mathbf{E}_\theta\left(\theta^{\mathrm{T}}\tilde{\epsilon}\right)^2 \leqslant C\sigma^2\frac{r^2}{n}, \tag{П.23}$$

где C — постоянная, зависящая только от γ. Рассмотрим теперь второй член в правой части (П.22). Обозначим через λ_i, $i = 1, \ldots, p$, собственные значения матрицы $(X^{\mathrm{T}}X)^{-1}$ и через u_i, $i = 1, \ldots, p$, соответствующие ортонормальные собственные векторы. Положим $v_i = \sqrt{\lambda_i}u_i^{\mathrm{T}}X^{\mathrm{T}}\xi$. Имеем

$$\mathbf{E}_\theta\left(\|\tilde{\epsilon}\|_2^2 - \sigma^2\mathrm{tr}\left[(X^{\mathrm{T}}X)^{-1}\right]\right)^2 = \sigma^4\mathbf{E}\left(\sum_{i=1}^{p}\lambda_i[v_i^2 - 1]\right)^2.$$

При фиксированном X случайные величины $v_1, \ldots, v_p$ являются н.о.р. стандартными гауссовскими. Используя этот факт и лемму 3, получаем

$$\begin{aligned}&\mathbf{E}_\theta\left(\|\tilde{\epsilon}\|_2^2 - \sigma^2\mathrm{tr}\left[(X^{\mathrm{T}}X)^{-1}\right]\right)^2 = \\ &= 2\sigma^4\mathbf{E}\left(\sum_{i=1}^{p}\lambda_i^2\right) \leqslant 2p\sigma^4\mathbf{E}\left[\lambda_{\min}^{-2}\left(X^{\mathrm{T}}X\right)\right] \leqslant C\frac{\sigma^4 p}{n^2},\end{aligned} \tag{П.24}$$

где C — постоянная, зависящая только от γ, Объединяя (П.22)–(П.24), получаем результат теоремы при $s \geqslant \sqrt{p}$.

Случай $s < \sqrt{p}$. Положим $S = \{i : \theta_i \neq 0\}$. Имеем

$$\begin{aligned}\mathbf{E}_\theta\left(\hat{Q} - \|\theta\|_2^2\right)^2 &\leqslant 3\mathbf{E}_\theta\left(\sum_{i\in S}(y_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s - \theta_i^2)\right)^2 + \\ &+3\mathbf{E}_\theta\left(\sum_{i\in S}\left[y_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s\right]\mathbf{1}_{\{y_i^2 \leqslant 2\sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\log(1+p/s^2)\}}\right)^2 + \\ &+3\mathbf{E}_\theta\left(\sum_{i\notin S}\left[\tilde{\epsilon}_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s\right]\mathbf{1}_{\{y_i^2 > 2\sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\log(1+p/s^2)\}}\right)^2,\end{aligned} \tag{П.25}$$

где $\tilde{\epsilon}_i$ обозначает i-ю компоненту $\tilde{\epsilon}$. Оценим теперь сверху каждый из трех членов в правой части (П.25). Для оценки первого члена заметим, что

$$\begin{aligned}&\mathbf{E}_\theta\left(\sum_{i\in S}(y_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s - \theta_i^2)\right)^2 \leqslant \\ &\leqslant 8\mathbf{E}_\theta\left(\sum_{i\in S}\theta_i\tilde{\epsilon}_i\right)^2 + 2\mathbf{E}_\theta\left(\sum_{i\in S}(\tilde{\epsilon}_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s)\right)^2.\end{aligned} \tag{П.26}$$

Второе слагаемое в правой части (П.26) удовлетворяет неравенству

$$\begin{aligned}&\mathbf{E}_\theta\left(\sum_{i\in S}(\tilde{\epsilon}_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s)\right)^2 \leqslant \\ &\leqslant 2\sigma^4(\alpha_s^2+3)\mathbf{E}\sum_{i\in S}\sum_{j\in S}(X^{\mathrm{T}}X)_{ii}^{-1}(X^{\mathrm{T}}X)_{jj}^{-1} \leqslant \\ &\leqslant 2\sigma^4(\alpha_s^2+3)s^2\mathbf{E}\left[\lambda_{\min}^{-2}(X^{\mathrm{T}}X)\right].\end{aligned} \tag{П.27}$$

Из (П.9) получаем

$$\alpha_s \leqslant 10\log(1+p/s^2). \tag{П.28}$$

Используя (П.26)–(П.28) вместе с леммой 3 и (П.23), находим, что

$$\mathbf{E}_\theta\left(\sum_{i\in S}(y_i^2 - \sigma^2(X^{\mathrm{T}}X)_{ii}^{-1}\alpha_s - \theta_i^2)\right)^2 \leqslant C\sigma^4 s^2\log^2(1+p/s^2)/n^2, \tag{П.29}$$

где константа C зависит только от γ.

Кроме того, нетрудно видеть, что второй член в правой части (П.25), меньше, с точностью до числового постоянного множителя, чем

$$\mathbf{E}\sigma^4\sum_{i\in S}\sum_{j\in S}(X^{\mathrm{T}}X)_{ii}^{-1}(X^{\mathrm{T}}X)_{jj}^{-1}(\alpha_s^2 + 4\log^2(1+p/s^2)).$$

Действуя так же, как в (П.27), и применяя лемму 3 и (П.28), получаем

$$\text{(П.30)}\qquad \mathbf{E}_\theta \left(\sum_{i \in S} \left[y_i^2 - \sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \alpha_s \right] \mathbf{1}_{\{y_i^2 \leqslant 2\sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \log(1+p/s^2)\}} \right)^2 \leqslant \\ \leqslant C\sigma^4 s^2 \log^2(1+p/s^2)/n^2,$$

где константа C зависит только от γ. Для третьего члена правой части (П.25) имеем

$$\text{(П.31)}\qquad \mathbf{E}_\theta \left(\sum_{i \notin S} \left[\tilde{\epsilon}_i^2 - \sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \alpha_s \right] \mathbf{1}_{\{y_i^2 > 2\sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \log(1+p/s^2)\}} \right)^2 = \\ = \sigma^4 \sum_{i \notin S} \sum_{j \notin S} \mathbf{E} \left((X^{\mathrm{T}} X)^{-1}_{ii} (X^{\mathrm{T}} X)^{-1}_{jj} (\tilde{\xi}_i^2 - \alpha_s)(\tilde{\xi}_j^2 - \alpha_s) \mathbf{1}_{\{|\tilde{\xi}_i| > x\}} \mathbf{1}_{\{|\tilde{\xi}_j| > x\}} \right),$$

где

$$x = \sqrt{2\log(1+p/s^2)}, \quad \tilde{\xi}_i = \frac{\tilde{\epsilon}_i}{\sqrt{\sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii}}}.$$

Заметим, что $\mathbf{E}(\tilde{\xi}_i^2 | X) = \mathbf{E}(\tilde{\xi}_j^2 | X) = 1$ и при фиксированном X вектор $(\tilde{\xi}_i, \tilde{\xi}_j) \in \in \mathbb{R}^2$ имеет гауссовское распределение с нулевым средним и ковариацией

$$\rho_{ij} = \frac{(X^{\mathrm{T}} X)^{-1}_{ij}}{\sqrt{(X^{\mathrm{T}} X)^{-1}_{ii}} \sqrt{(X^{\mathrm{T}} X)^{-1}_{jj}}}.$$

Используя лемму 5, получим, что существуют числовые константы C такие, что

$$\sum_{i \notin S} \sum_{j \notin S} \mathbf{E} \left((X^{\mathrm{T}} X)^{-1}_{ii} (X^{\mathrm{T}} X)^{-1}_{jj} (\tilde{\xi}_i^2 - \alpha_s)(\tilde{\xi}_j^2 - \alpha_s) \mathbf{1}_{\{|\tilde{\xi}_i| > x\}} \mathbf{1}_{\{|\tilde{\xi}_j| > x\}} \right) = \\ = \sum_{i \notin S} \sum_{j \notin S} \mathbf{E} \left((X^{\mathrm{T}} X)^{-1}_{ii} (X^{\mathrm{T}} X)^{-1}_{jj} \mathbf{E} \left[(\tilde{\xi}_i^2 - \alpha_s)(\tilde{\xi}_j^2 - \alpha_s) \mathbf{1}_{\{|\tilde{\xi}_i| > x\}} \mathbf{1}_{\{|\tilde{\xi}_j| > x\}} \mid X \right] \right) \leqslant \\ \leqslant C \sum_{i,j=1}^{p} \mathbf{E} \left[(X^{\mathrm{T}} X)^{-1}_{ii} (X^{\mathrm{T}} X)^{-1}_{jj} \rho_{ij}^2 \right] x^4 \exp(-x^2/2) = \\ = C\mathbf{E} \left[\|(X^{\mathrm{T}} X)^{-1}\|_F^2 \right] x^4 \exp(-x^2/2) \leqslant \\ \leqslant C\mathbf{E} \left[\|(X^{\mathrm{T}} X)^{-1}\|_F^2 \right] \frac{s^2}{p} \log^2(1+p/s^2) \leqslant \\ \leqslant C\mathbf{E} \left[\lambda_{\min}^{-2}(X^{\mathrm{T}} X) \right] s^2 \log^2(1+p/s^2),$$

где $\|(X^{\mathrm{T}} X)^{-1}\|_F$ — норма Фробениуса матрицы $(X^{\mathrm{T}} X)^{-1}$. Наконец, лемма 3, (П.31) и последняя выкладка влекут, что существует постоянная C, зависящая только от γ, такая что

$$
\text{(П.32)} \qquad \mathbf{E}_\theta \left(\sum_{i \notin S} \left[\tilde{\epsilon}_i^2 - \sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \alpha_s \right] \mathbf{1}_{\{y_i^2 > 2\sigma^2 (X^{\mathrm{T}} X)^{-1}_{ii} \log(1+p/s^2)\}} \right)^2 \leqslant \\ \leqslant C \frac{\sigma^4 s^2 \log^2(1+p/s^2)}{n^2}.
$$

Доказательство завершается соединением неравенств (П.25), (П.29), (П.30) и (П.32).

П.3. Вспомогательные леммы для доказательства теоремы 4

Напомним сначала некоторые общие факты о нижних границах для рисков в задаче проверки гипотез. Пусть Θ — измеримое множество, не обязательно множество $\Theta(s,\tau)$, и пусть μ — вероятностная мера на Θ. Рассмотрим произвольное семейство вероятностных мер $\mathbf{P}_\theta$, индексированное параметром $\theta \in \Theta$. Обозначим через $\mathbb{P}_\mu$ смесь вероятностных мер

$$
\mathbb{P}_\mu = \int\limits_\Theta \mathbf{P}_\theta \, \mu(d\theta).
$$

Определим χ^2-расхождение между двумя вероятностными мерами P' и P как

$$
\chi^2(P', P) = \int (dP'/dP)^2 dP - 1,
$$

если $P' \ll P$, и положим $\chi^2(P', P) = +\infty$ в противном случае. Следующая лемма играет ключевую роль в методе Ле Кама доказательства нижних границ (см., например, [4, лемма 3]).

Лемма 6. *Пусть μ — вероятностная мера на Θ и пусть $\{\mathbf{P}_\theta : \theta \in \Theta\}$ — семейство вероятностных мер на $\mathcal{X}$, индексированных параметром $\theta \in \Theta$. Тогда для любой вероятностной меры Q на $\mathcal{X}$ имеем*

$$
\inf_\Delta \left\{ Q(\Delta = 1) + \sup_{\theta \in \Theta} \mathbf{P}_\theta(\Delta = 0) \right\} \geqslant 1 - \sqrt{\chi^2(\mathbb{P}_\mu, Q)},
$$

где нижняя грань inf *берется по всем статистикам Δ со значениями в $\{0,1\}$.*

Применяя лемму 6 с $Q = \mathbf{P}_0$, видим, что достаточно выбрать подходящую меру μ и ограничить $\chi^2(\mathbb{P}_\mu, \mathbf{P}_0)$ сверху величиной, строго меньшей 1, чтобы получить искомую нижнюю границу для $\mathcal{R}_{s,\tau}$. Следующая лемма полезна при оценке $\chi^2(\mathbb{P}_\mu, \mathbf{P}_0)$.

Лемма 7. *Пусть μ — вероятностная мера на Θ и пусть $\{\mathbf{P}_\theta : \theta \in \Theta\}$ — семейство вероятностных мер на $\mathcal{X}$, индексированных параметром $\theta \in \Theta$.*

Пусть Q — вероятностная мера на $\mathcal{X}$ такая, что $\mathbf{P}_\theta \ll Q$ для всех $\theta \in \Theta$. Тогда

$$\chi^2(\mathbb{P}_\mu, Q) = \mathbb{E}_{(\theta,\theta')\sim\mu^2}\left(\int \frac{d\mathbf{P}_\theta d\mathbf{P}_{\theta'}}{dQ}\right) - 1.$$

Здесь $\mathbb{E}_{(\theta,\theta')\sim\mu^2}$ обозначает математическое ожидание относительно распределения пары (θ,θ'), где θ и θ' — независимые случайные величины, каждая из которых распределена в соответствии с μ.

Доказательство. Достаточно заметить, что

$$\chi^2(\mathbb{P}_\mu, Q) = \int \frac{(d\mathbb{P}_\mu)^2}{dQ} - 1,$$

тогда как

$$\int \frac{(d\mathbb{P}_\mu)^2}{dQ} = \int \frac{\int_\Theta d\mathbf{P}_\theta \mu(d\theta) \int_\Theta d\mathbf{P}_{\theta'}\mu(d\theta')}{dQ} = \int\limits_\Theta \int\limits_\Theta \mu(d\theta)\mu(d\theta') \int \frac{d\mathbf{P}_\theta d\mathbf{P}_{\theta'}}{dQ}.$$

Найдем теперь выражение для χ^2-расхождения в лемме 7, когда $\mathbf{P}_\theta$ — вероятностное распределение, порожденное моделью (1) и $Q = \mathbf{P}_0$.

Лемма 8. Пусть $\mathbf{P}_\theta$ — распределение (X, Y), удовлетворяющее (1). Тогда

$$\chi^2(\mathbb{P}_\mu, \mathbf{P}_0) = \mathbb{E}_{(\theta,\theta')\sim\mu^2} E_X \exp(\langle X\theta, X\theta'\rangle/\sigma^2) - 1.$$

Доказательство. Применяем лемму 7 и замечаем, что для любых $(\theta,\theta') \in \Theta \times \Theta$ имеют место равенства

$$\int \frac{d\mathbf{P}_\theta d\mathbf{P}_{\theta'}}{d\mathbf{P}_0} =$$

$$= \frac{1}{(2\pi\sigma)^{n/2}} E_X \int\limits_{\mathbb{R}^n} \exp\left(-\frac{1}{2\sigma^2}(\|y - X\theta\|_2^2 + \|y - X\theta'\|_2^2 - \|y\|_2^2)\right) dy =$$

$$= \frac{1}{(2\pi\sigma)^{n/2}} E_X \int\limits_{\mathbb{R}^n} \exp\left(-\frac{1}{2\sigma^2}(\|y\|_2^2 - 2\langle y, X(\theta+\theta')\rangle + \|X(\theta+\theta')\|_2^2 - 2\langle X\theta, X\theta'\rangle)\right) dy =$$

$$= E_X\left(\frac{\exp(\langle X\theta, X\theta'\rangle/\sigma^2)}{(2\pi\sigma)^{n/2}} \int\limits_{\mathbb{R}^n} \exp\left(-\frac{1}{2\sigma^2}\|y - X(\theta+\theta')\|_2^2\right) dy\right) =$$

$$= E_X \exp(\langle X\theta, X\theta'\rangle/\sigma^2).$$

Лемма 9. Пусть $a \in \mathbb{R}$ — постоянная и пусть W — случайная величина. Тогда

$$\mathbf{E}\exp(W) \leqslant \exp(a)\left(1 + \int\limits_0^\infty e^t p(t) dt\right),$$

где $p(t) = \mathbf{P}(|W - a| \geqslant t)$.

Доказательство. Имеем

$$\begin{aligned}\mathbf{E}\exp(W) \leqslant \exp(a)\mathbf{E}\exp(|W-a|) = \\ = \exp(a)\int_0^\infty \mathbf{P}\left(\exp(|W-a|) \geqslant x\right) dx = \\ = \exp(a)\left[1 + \int_1^\infty \mathbf{P}\left(\exp(|W-a|) \geqslant x\right) dx\right] = \\ = \exp(a)\left[1 + \int_0^\infty e^t p(t) dt\right].\end{aligned}$$

Лемма 10. Предположим, что матрица X имеет изотропное распределение с независимыми σ_X-субгауссовскими строками при некотором $\sigma_X > 0$. Тогда для всех $x > 0$ и всех $\theta, \theta' \in \mathbb{R}^p$ имеем

$$P_X\left(|\langle X\theta, X\theta'\rangle - n\langle\theta, \theta'\rangle| \geqslant \|\theta\|_2\|\theta'\|_2 x\right) \leqslant 6\exp(-C_1\min(x, x^2/n)),$$

где константа $C_1 > 0$ зависит только от σ_X.

Доказательство. Используя однородность, достаточно рассмотреть случай, когда $\|\theta\|_2 = \|\theta'\|_2 = 1$. Это и будет предполагаться далее в доказательстве. Тогда имеем

$$\langle X\theta, X\theta'\rangle = \frac{\|X\theta\|_2^2 + \|X\theta'\|_2^2 - \|X(\theta-\theta')\|_2^2}{2}, \qquad \langle\theta, \theta'\rangle = \frac{2 - \|\theta-\theta'\|_2^2}{2},$$

что приводит к неравенству

$$\begin{aligned}\left|\frac{1}{n}\langle X\theta, X\theta'\rangle - \langle\theta,\theta'\rangle\right| \leqslant \frac{1}{2}\left(\left|\frac{1}{n}\|X\theta\|_2^2 - 1\right| + \left|\frac{1}{n}\|X\theta'\|_2^2 - 1\right| + \right. \\ \left. + \left|\frac{1}{n}\|X(\theta-\theta')\|_2^2 - \|\theta-\theta'\|_2^2\right|\right).\end{aligned} \tag{П.33}$$

Третье слагаемое в правой части (П.33) приводится перенормировкой к тому же виду, что и первые два слагаемых. Учитывая это, для доказательства леммы достаточно показать, что

$$P_X\left(\left|\frac{1}{n}\|X\theta\|_2^2 - 1\right| \geqslant v\right) \leqslant 2\exp(-C_1'\min(v, v^2)n), \quad \forall\ v > 0, \|\theta\|_2 = 1, \tag{П.34}$$

где константа $C_1' > 0$ зависит только от σ_X.

Обозначим i-ю строку матрицы X через $\mathbf{x}_i$. Тогда

$$\frac{1}{n}\|X\theta\|_2^2 - 1 = \frac{1}{n}\sum_{i=1}^n (Z_i^2 - 1),$$

где $Z_i = \mathbf{x}_i^{\mathrm{T}}\theta$ — независимые σ_X-субгауссовские случайные величины, такие что $\mathbf{E}(Z_i^2) = 1$ при $i = 1, \ldots, n$. Следовательно, $Z_i^2 - 1$, $i = 1, \ldots, n$, — независимые субэкспоненциальные случайные величины с нулевыми средними, и (П.34) немедленно вытекает из неравенства Бернштейна для субэкспоненциальных случайных величин (ср., например, [10, следствие 5.17]).

Лемма 11. Предположим, что матрица X имеет изотропное распределение с независимыми σ_X-субгауссовскими строками при некотором $\sigma_X > 0$. Тогда существует число $u_0 > 0$, зависящее только от σ_X, такое что для всех θ, θ' с $\|\theta\|_2, \|\theta'\|_2 \leqslant un^{-1/4}$ и всех $u \in (0, u_0)$ имеем

$$E_X \exp(\langle X\theta, X\theta'\rangle) \leqslant \exp(n\langle\theta, \theta'\rangle)(1 + C_0u^2),$$

где константа $C_0 > 0$ зависит только от σ_X.

Доказательство. По лемме 10 для всех $x > 0$ с P_X-вероятностью не менее $1 - 6e^{-C_1 \min(x, x^2/n)}$ имеем

$$\left|\langle X\theta, X\theta'\rangle - n\langle\theta, \theta'\rangle\right| \leqslant \|\theta\|_2\|\theta'\|_2 x \leqslant u^2 n^{-1/2} x.$$

Следовательно, для всех $t > 0$ с P_X-вероятностью не менее

$$1 - 6e^{-C_1 \min(\sqrt{n}t/u^2, t^2/u^4)}$$

имеем

$$\left|\langle X\theta, X\theta'\rangle - n\langle\theta, \theta'\rangle\right| \leqslant t.$$

Отсюда и из леммы 9 вытекает, что для всех $u \leqslant u_0 := (C_1/2)^{1/2}$ выполнены неравенства

$$\begin{aligned}
&E_X \exp(\langle X\theta, X\theta'\rangle) \leqslant \exp(n\langle\theta, \theta'\rangle)\left(1 + 6\int_0^\infty e^{t - C_1 \min\left(\sqrt{n}t/u^2, t^2/u^4\right)} dt\right) \leqslant \\
&\leqslant \exp(n\langle\theta, \theta'\rangle)\left(1 + 6\int_0^\infty e^{t(1 - C_1\sqrt{n}/u^2)} dt + 6\int_0^\infty e^{t - C_1 t^2/u^4} dt\right) \leqslant \\
&\leqslant \exp(n\langle\theta, \theta'\rangle)\left(1 + 6\int_0^\infty e^{-C_1\sqrt{n}t/(2u^2)} dt + 6\int_0^\infty e^{-t(C_1 t/u^4 - 1)} dt\right) \leqslant \\
&\leqslant \exp(n\langle\theta, \theta'\rangle)\left(1 + \frac{12u^2}{C_1\sqrt{n}} + \frac{12u^4}{C_1} e^{2u^4/C_1} + 6\int_{2u^4/C_1}^\infty e^{-t^2 C_1/(2u^4)} dt\right) \leqslant \\
&\leqslant \exp(n\langle\theta, \theta'\rangle)\left(1 + C_0u^2\right),
\end{aligned}$$

где использован тот факт, что $C_1\sqrt{n}/u^2 > 2$ и константа $C_0 > 0$ зависит только от C_1 и, следовательно, только от σ_X.

П.4. Доказательство теоремы 4

Пусть s — целое число, такое что $1 \leqslant s \leqslant p$, и пусть $\tau > 0$. Обозначим через μ_τ равномерное распределение на множестве всех векторов в $\mathbb{R}^p$, имеющих в точности s ненулевых коэффициентов, каждый из которых равен $\tau/\sqrt{s}$. Заметим, что носитель меры μ_τ содержится в $\Theta(s, \tau)$.

Возьмем теперь $\tau = \tau(s, n, p)$, определенное в (11), и положим $\mu = \mu_\tau$. В силу лемм 6–8, чтобы доказать теорему 4, достаточно показать, что

$$\mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} E_X \exp(\langle X\theta, X\theta'\rangle/\sigma^2) \leqslant 1 + o_A(1), \tag{П.35}$$

где $o_A(1)$ стремится к нулю при $A \to 0$.

Прежде чем доказывать (П.35), введем некоторые упрощения. Во-первых, заметим, что для τ, определенного в (11), левая часть (П.35) не зависит от σ. Таким образом, в дальнейшем будем полагать $\sigma = 1$ без потери общности. Далее заметим, что достаточно доказать теорему для случая $s \leqslant \sqrt{p}$. Действительно, для $s > \sqrt{p}$ можно использовать включения $\Theta(s, \tau(s, n, p)) \supseteq \supseteq \Theta(s', \tau(s, n, p)) \supseteq \Theta(s', \tau(s', n, p))$, где s' — наибольшее целое число, меньшее или равное $\sqrt{p}$. Поскольку

$$\tau(s', n, p) \asymp \min\left(\frac{p^{1/4}}{\sqrt{n}}, n^{-1/4}\right),$$

а скорость (11) также имеет этот порядок для $s > \sqrt{p}$, достаточно доказать нижнюю оценку для $s \leqslant s'$ и, следовательно, для $s \leqslant \sqrt{p}$. Принимая во внимание эти упрощения, в дальнейшем будем считать без ограничения общности, что $s \leqslant \sqrt{p}$, $\sigma = 1$ и

$$\tau := A \min\left(\sqrt{\frac{s\log(1 + p/s^2)}{n}}, n^{-1/4}\right). \tag{П.36}$$

Теперь докажем (П.35) при этих предположениях. По лемме 11 для всех $0 < A < u_0$ имеем

$$\mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} E_X \exp(\langle X\theta, X\theta'\rangle) \leqslant \mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} \exp(n\langle\theta,\theta'\rangle)(1 + C_0A^2). \tag{П.37}$$

Предположим, что $A < 1$. Рассуждая точно так же, как в доказательстве леммы 1 в [4], находим

$$\begin{aligned} \mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} \exp(n\langle\theta,\theta'\rangle) &= \mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} \exp\left(n\tau^2 s^{-1}\sum_{j=1}^p \mathbf{1}_{\{\theta_j\neq 0\}}\mathbf{1}_{\{\theta'_j\neq 0\}}\right) \leqslant \\ &\leqslant \left(1 - \frac{s}{p} + \frac{s}{p}\exp(n\tau^2 s^{-1})\right)^s \leqslant \\ &\leqslant \left(1 - \frac{s}{p} + \frac{s}{p}\left(1 + \frac{p}{s^2}\right)^{A^2}\right)^s \leqslant \\ &\leqslant \left(1 + \frac{A^2}{s}\right)^s \leqslant \exp(A^2), \end{aligned} \tag{П.38}$$

где было использовано неравенство $(1+x)^{A^2} - 1 \leqslant A^2 x$, выполненное при $0 < A < 1$ и $x > 0$. Из (П.37) и (П.38) получим, что для всех $0 < A < \min(1, u_0)$ имеет место неравенство

$$\mathbb{E}_{(\theta,\theta')\sim\mu_\tau^2} E_X \exp(\langle X\theta, X\theta'\rangle) \leqslant \exp(A^2)(1 + C_0 A^2)$$

при некоторых $u_0 > 0$ и $C_0 > 0$, зависящих только от σ_X. Это завершает доказательство теоремы.

СПИСОК ЛИТЕРАТУРЫ

1. *Ingster Y.I.* Some Problems of Hypothesis Testing Leading to Infinitely Divisible Distributions // Math. Methods Statist. 1997. No. 6. P. 47–49.
2. *Donoho D.L., Jin J.* Higher Criticism for Detecting Sparse Heterogeneous Mixtures // Ann. Statist. 2004. No. 32. P. 962–994.
3. *Baraud Y.* Non Asymptotic Minimax Rates of Testing in Signal Detection // Bernoulli. 2002. No. 8. P. 577–606.
4. *Collier O., Comminges L., Tsybakov A.B.* Minimax Estimation of Linear and Quadratic Functionals Under Sparsity Constraints // Ann. Statist. 2017. No. 45. P. 923–958.
5. *Ingster Y.I., Tsybakov A.B., Verzelen N.* Detection Boundary in Sparse Regression // Electron. J. Stat. 2010. No. 4. P. 1476–1526.
6. *Arias-Castro E., Candes E., Plan Y.* Global Testing Under Sparse Alternatives: ANOVA, Multiple Comparisons and the Higher Criticism // Ann. Statist. 2011. No. 39. P. 2533–2556.
7. *Verzelen N.* Minimax Risks for Sparse Regressions: Ultra-High Dimensional Phenomenons // Electron. J. Stat. 2012. No. 6. P. 38–90.
8. *Ingster Y.I., Suslina I.A.* Nonparametric Goodness-of-Fit Testing under Gaussian Models. N.Y.: Springer, 2003.
9. *Davidson K.R., Szarek S.J.* Local Operator Theory, Random Matrices and Banach Spaces. Handbook of the geometry of Banach spaces / Eds.: W.B. Johnson, J. Lindenstrauss. 2001. V. 1. P. 317–366
10. *Vershynin R.* Introduction to the Non-Asymptotic Analysis of Random Matrices. *Compressed sensing*. Cambridge: Cambridge Univ. Press. 2012. P. 210–268.
11. *Tao T., Vu V.* Random Matrices: Universality of Esds and the Circular Law // Ann. Probab. 2010. V. 38. No. 5. P. 2023–2065.
12. *Bordenave C., Chafaï D.* Around the Circular Law // Probab. Surveys. 2012. No. 9. P. 1–89.

Статья представлена к публикации членом редколлегии А.В. Назиным.

Поступила в редакцию 19.07.2018

После доработки 03.10.2018

Принята к публикации 08.11.2018