ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 519.24

 М.В. Бурнашев

О МИНИМАКСНОМ ОБНАРУЖЕНИИ ГАУССОВСКИХ СТОХАСТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С НЕТОЧНО ИЗВЕСТНЫМИ СРЕДНИМИ И КОВАРИАЦИОННЫМИ МАТРИЦАМИ[1]

Рассматривается задача обнаружения (проверки) гауссовских стохастических последовательностей (сигналов) с неточно известными средними и ковариационными матрицами. Альтернативой являются независимые одинаково распределенные гауссовские случайные величины с нулевыми средними и единичными дисперсиями. При заданной вероятности "*ложной тревоги*" (вероятности ошибки 1-го рода) качество минимаксного обнаружения определяется наилучшей экспонентой "*вероятности пропуска*" (вероятности ошибки 2-го рода) при растущем интервале наблюдений. Исследуется максимальное множество средних и ковариационных матриц (сложная гипотеза), такое что его минимаксную проверку можно заменить проверкой одной конкретной пары среднего и ковариационной матрицы (простая гипотеза) без ухудшения экспоненты обнаружения. В статье полностью описывается это максимальное множество.

Ключевые слова: минимаксная проверка гипотез, вероятность ошибки 1-го рода, вероятность ошибки 2-го рода, экспонента вероятности ошибки, лемма Стейна.

DOI: `10.31857/S0555292322030068`, **EDN:** EANVWL

§ 1. Введение и основные результаты

1.1. Постановка задачи. Одна из традиционных задач проверки простых гипотез $\mathcal{H}_0$ и $\mathcal{H}_1$ относительно гауссовского сигнального вектора $\boldsymbol{\eta}_n$ на фоне гауссовского шума $\boldsymbol{\xi}_n$ (т.е. задачи обнаружения сигнала на фоне шума) по наблюдениям $\boldsymbol{y}_n^T = \boldsymbol{y}'_n = (y_1, \ldots, y_n) \in \mathbb{R}^n$ имеет вид

$$\begin{aligned} &\mathcal{H}_0 \colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n), \\ &\mathcal{H}_1 \colon \boldsymbol{y}_n = \boldsymbol{\eta}_n, \quad \boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n), \end{aligned} \tag{1}$$

где выборка $\boldsymbol{\xi}_n^T = (\xi_1, \ldots, \xi_n)$ представляет собой "шум" и состоит из независимых одинаково распределенных гауссовских случайных величин с нулевыми средними и дисперсиями 1, а $\boldsymbol{I}_n$ – единичная ковариационная матрица. Стохастический "сигнал" $\boldsymbol{\eta}_n$ является гауссовским случайным вектором с известным средним $\boldsymbol{a}_n$ и известной ковариационной матрицей $\boldsymbol{M}_n$.

Однако на практике среднее $\boldsymbol{a}_n$ и матрица $\boldsymbol{M}_n$ обычно не известны в точности, и поэтому в реальности модель (1) принимает вид

$$\begin{aligned} &\mathcal{H}_0 \colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n), \\ &\mathcal{H}_1 \colon \boldsymbol{y}_n = \boldsymbol{\eta}_n, \quad \boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n), \quad \boldsymbol{a}_n \in \mathcal{A}_n, \quad \boldsymbol{M}_n \in \mathcal{M}_n, \end{aligned} \tag{2}$$

[1] Исследование выполнено при частичной финансовой поддержке Российского фонда фундаментальных исследований (номер проекта 19-01-00364).

где $\mathcal{A}_n$ – заданное множество возможных средних $\boldsymbol{a}_n$, а $\mathcal{M}_n$ – заданное множество возможных ковариационных матриц $\boldsymbol{M}_n$ (возможно, зависящее от $\boldsymbol{a}_n$). Будем также обозначать для удобства

$$\boldsymbol{B}_n = (\boldsymbol{b}_n, \boldsymbol{V}_n), \quad \boldsymbol{b}_n \in \mathcal{A}_n, \quad \boldsymbol{V}_n \in \mathcal{M}_n, \quad \mathcal{F}_n = \{\boldsymbol{B}_n\} = (\mathcal{A}_n, \mathcal{M}_n).$$

Далее для модели (2) рассматривается задача минимаксной проверки [1–3] простой гипотезы $\mathcal{H}_0$ против сложной альтернативы $\mathcal{H}_1$ по наблюдениям $\boldsymbol{y}_n^T = \boldsymbol{y}_n' = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Если для принятия решения в пользу $\mathcal{H}_0$ выбирается область $\mathcal{D} \in \mathbb{R}^n$, такая что

$$\boldsymbol{y}_n \in \mathcal{D} \;\Rightarrow\; \mathcal{H}_0, \quad \boldsymbol{y}_n \notin \mathcal{D} \;\Rightarrow\; \mathcal{H}_1, \tag{3}$$

то тогда вероятности ошибки 1-го рода (“*ложной тревоги*”) $\alpha(\mathcal{D})$ и 2-го рода (“*вероятность пропуска*”) $\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n)$ определяются, соответственно, формулами

$$\alpha(\mathcal{D}) = \mathbf{P}(\boldsymbol{y}_n \notin \mathcal{D} \,|\, \mathcal{H}_0) \tag{4}$$

и

$$\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n) = \sup_{\boldsymbol{a}_n \in \mathcal{A}_n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \mathbf{P}(\boldsymbol{y}_n \in \mathcal{D} \,|\, \boldsymbol{M}_n, \boldsymbol{a}_n). \tag{5}$$

Нас интересует минимально возможная вероятность $\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n)$ ошибки 2-го рода (см. (4) и (5)) при заданной вероятности ошибки 1-го рода α, $0 < \alpha < 1$:

$$\beta(\alpha, \mathcal{A}_n, \mathcal{M}_n) = \inf_{\mathcal{D}\colon \alpha(\mathcal{D}) \leqslant \alpha} \beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n), \tag{6}$$

и соответствующая оптимальная область принятия решения $\mathcal{D}(\alpha)$ из (3).

В статье рассматривается случай, когда величина α фиксирована (или медленно убывает при $n \to \infty$). Этот случай иногда называют задачей Неймана – Пирсона минимаксного различения гипотез. В этом случае ошибки 1-го и 2-го рода влекут очень разные потери для статистика, и он в основном хочет минимизировать вероятность ошибки 2-го рода $\beta = \mathbf{P}\{\mathcal{H}_0 \,|\, \mathcal{H}_1\}$. Этот случай очень популярен в разных приложениях (см., например, работу [4] и библиографию в ней).

Для заданных среднего $\boldsymbol{a}_n$, матрицы $\boldsymbol{M}_n$ и величины α через $\beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим минимально возможную вероятность ошибки 2-го рода (см. (6)). Соответствующая оптимальная область принятия решения $\mathcal{D}(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ описывается леммой Неймана – Пирсона [1, 2]. Ясно, что

$$\sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \beta(\alpha, \boldsymbol{a}_n, \mathcal{M}_n). \tag{7}$$

Также для фиксированного α и заданных множеств $\mathcal{A}_n, \mathcal{M}_n$ обозначим через $\beta(\alpha, \mathcal{A}_n, \mathcal{M}_n)$ минимально возможную вероятность ошибки 2-го рода (см. (6)). Тогда аналогично (7) имеем

$$\sup_{\boldsymbol{a}_n \in \mathcal{A}_n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n). \tag{8}$$

Во многих практических случаях величина $\beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ убывает экспоненциально по $n \to \infty$. Поэтому естественно (во всяком случае, проще и продуктивнее) исследовать соответствующие экспоненты $n^{-1} \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ и $n^{-1} \ln \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n)$ при $n \to \infty$ (некоторые результаты относительно равенства в (7) имеются в [5]).

В данной статье исследуются множества $\mathcal{F}_n = (\mathcal{A}_n, \mathcal{M}_n)$, для которых в соотношении (8) выполняется следующее асимптотическое равенство:

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n). \tag{9}$$

Мотивация исследования минимаксной проверки гипотез (обнаружения сигналов) подробно описана в [1–4]. Если для заданных множеств средних $\mathcal{A}_n$ и матриц $\mathcal{M}_n$ выполняется соотношение (9), то тогда можно заменить (без асимптотических потерь) все множество $\mathcal{F}_n$ одной конкретной парой $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Напомним, что оптимальный тест для конкретной пары $(\boldsymbol{a}_n, \boldsymbol{M}_n)$ описывается леммой Неймана–Пирсона и сводится к простому тесту отношения правдоподобия (LR-тесту). В противном случае (без соотношения (9)) оптимальным минимаксным тестом является значительно более сложный байесовский тест относительно *наименее благоприятного* априорного распределения на множестве $\mathcal{F}_n$. Поэтому естественно исследовать, когда заданное множество $\mathcal{F}_n$ можно заменить конкретной парой $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$. Однако технически удобнее рассмотреть эквивалентную задачу: для заданной пары $\boldsymbol{F}_n$ найти максимальное множество пар $\mathcal{F}_n(\boldsymbol{F}_n)$, которое может быть заменено парой $\boldsymbol{F}_n$. Эта задача в основном и рассматривается в настоящей статье.

Замечание 1. Модели (1) и (2) можно свести к эквивалентным моделям с диагональной матрицей $\boldsymbol{M}_n$. Действительно, так как $\boldsymbol{M}_n$ – ковариационная матрица (т.е. симметричная и положительно определенная), то существуют ортогональная матрица $\boldsymbol{T}_n$ и диагональная матрица $\boldsymbol{\Lambda}_n$, такие что $\boldsymbol{M}_n = \boldsymbol{T}_n \boldsymbol{\Lambda}_n \boldsymbol{T}_n'$ (см. [6, §§ 4.7–4.9; 7, теорема 4.1.5]). При этом диагональная матрица $\boldsymbol{\Lambda}_n = \boldsymbol{T}_n' \boldsymbol{M}_n \boldsymbol{T}_n$ состоит из собственных значений $\{\lambda_i\}$ матрицы $\boldsymbol{M}_n$. Отметим также, что для любой ортогональной матрицы $\boldsymbol{T}_n$ вектор $\boldsymbol{T}_n' \boldsymbol{\xi}_n$ имеет то же самое распределение, что и сам вектор $\boldsymbol{\xi}_n$ (для простой гипотезы $\mathcal{H}_0$ в (2)). Поэтому, умножая обе части (2) на $\boldsymbol{T}_n'$, можно свести модель (2) к эквивалентной модели с диагональной матрицей $\boldsymbol{M}_n$.

Определение 1. Для фиксированного α и заданной последовательности пар $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим через $\mathcal{F}_0(\boldsymbol{F}_n)$ последовательность наибольших множеств пар, таких что равенство (9) принимает вид

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{F}_0(\boldsymbol{F}_n)) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{F}_n). \tag{10}$$

Ясно, что $\boldsymbol{F}_n \in \mathcal{F}_0(\boldsymbol{F}_n)$.

Другими словами, для заданной вероятности ошибки 1-го рода α последовательность $\mathcal{F}_0(\boldsymbol{F}_n)$ представляет собой наибольшее множество пар, которое может быть заменено (без асимптотических потерь для $\beta(\mathcal{F}_0(\boldsymbol{F}_n)))$ одной парой $\boldsymbol{F}_n$. Далее (теорема 1) описывается наибольшее множество $\mathcal{F}_0(\boldsymbol{F}_n)$, удовлетворяющее (10). Это обобщает аналогичный результат из [8], где рассматривался случай $\boldsymbol{a}_n = \boldsymbol{0}_n$. Это также усиливает аналогичный результат из [4], где для множества $\mathcal{F}_0(\boldsymbol{0}_n, \boldsymbol{M}_n)$ были получены некоторые оценки снизу.

Удобно сначала исследовать максимальные множества $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$, аналогичные $\mathcal{F}_0(\boldsymbol{F}_n)$, но которые возникают, если используется LR-детектор (см. определение 2). Будет показано, что $\mathcal{F}_0(\boldsymbol{F}_n) = \mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$, т.е. LR-детектор является асимптотически оптимальным.

В моделях (1) и (2) обозначим через $\mathbf{P}_{\boldsymbol{I}_n}$ распределение величины $\boldsymbol{y}_n = \boldsymbol{\xi}_n$, где $\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$. Аналогично обозначим через $\mathbf{Q}_{\boldsymbol{F}_n}$, $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$, распределение величины $\boldsymbol{y}_n = \boldsymbol{\eta}_n$, где $\boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Обозначим также через $p_{\boldsymbol{I}_n}(\boldsymbol{y}_n)$ и $p_{\boldsymbol{F}_n}(\boldsymbol{y}_n)$, $\boldsymbol{y}_n \in \mathbb{R}^n$, соответствующие плотности распределения вероятностей. Для $(n \times n)$-матрицы $\boldsymbol{M}_n$ будем обозначать $|\boldsymbol{M}_n| = \det \boldsymbol{M}_n$. Заметим, что если $|\boldsymbol{M}_n| \neq 0$, то

$$\begin{aligned} \ln p_{\boldsymbol{I}_n}(\boldsymbol{y}_n) &= -\frac{1}{2}[n\ln(2\pi) + (\boldsymbol{y}_n, \boldsymbol{y}_n)], \quad \boldsymbol{y}_n \in \mathbb{R}^n, \\ \ln p_{\boldsymbol{F}_n}(\boldsymbol{y}_n) &= -\frac{1}{2}\left[n\ln(2\pi) + \ln|\boldsymbol{M}_n| + (\boldsymbol{y}_n - \boldsymbol{a}_n, \boldsymbol{M}_n^{-1}(\boldsymbol{y}_n - \boldsymbol{a}_n))\right]. \end{aligned} \tag{11}$$

При $|\boldsymbol{M}_n| \neq 0$ введем также логарифм отношения правдоподобия (см. (11))

$$\begin{aligned} r_{\boldsymbol{F}_n}(\boldsymbol{y}_n) &= \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{F}_n}}(\boldsymbol{y}) = \\ &= \frac{1}{2}\Big[\ln|\boldsymbol{M}_n| + (\boldsymbol{y}, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{y}) - 2(\boldsymbol{y}, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\Big]. \end{aligned} \tag{12}$$

Рассмотрим сначала LR-детекторы. Введем соответствующие области принятия решения $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha)$ в пользу гипотезы $\mathcal{H}_0$ (т.е. в пользу матрицы $\boldsymbol{I}_n$), когда проверяются простые гипотезы $\boldsymbol{I}_n$ и $\boldsymbol{F}_n$:

$$\mathcal{D}_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha) = \{\boldsymbol{y}_n \in \mathbb{R}^n : r_{\boldsymbol{F}_n}(\boldsymbol{y}_n) \geqslant \gamma\}, \tag{13}$$

где γ такое, что (см. (12))

$$\begin{aligned} \alpha &= \mathbf{P}_{\boldsymbol{I}_n}\{\mathcal{D}^c_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha)\} = \mathbf{P}_{\boldsymbol{I}_n}\{r_{\boldsymbol{F}_n}(\boldsymbol{\xi}_n) \leqslant \gamma\} = \mathbf{P}_{\boldsymbol{I}_n}\Big\{[\ln|\boldsymbol{M}_n| + \\ &+ (\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{\xi}_n) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)] \leqslant 2\gamma\Big\}. \end{aligned} \tag{14}$$

Определение 2. Для фиксированного α и заданной последовательности пар $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим через $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$ последовательность максимальных множеств пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$, таких что

$$\lim_{n\to\infty} \frac{1}{n} \ln \sup_{(\boldsymbol{b}_n, \boldsymbol{V}_n) \in \mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)} \beta(\boldsymbol{b}_n, \boldsymbol{V}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{a}_n, \boldsymbol{M}_n), \tag{15}$$

при условии, что используются решающие области $\mathcal{D}_{\mathrm{LR}}(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$.

Ниже в теореме 2 описывается множество $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ для модели (2).

Нам понадобится также следующее определение [9].

Определение 3. Для вероятностных мер $\mathbf{P}$ и $\mathbf{Q}$ на измеримом пространстве $(\mathcal{X}, \mathcal{B})$ введем функцию (расстояние (или дивергенция) Кульбака – Лейблера для мер $\mathbf{P}$ и $\mathbf{Q}$)

$$D(\mathbf{P} \,\|\, \mathbf{Q}) = \mathbf{E}_{\mathbf{P}} \ln \frac{d\,\mathbf{P}}{d\,\mathbf{Q}}(\boldsymbol{x}) \geqslant 0, \tag{16}$$

где математическое ожидание берется по мере $\mathbf{P}$.

С помощью формул (11) и (16) имеем

$$\begin{aligned} D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n, \boldsymbol{M}_n}) &= \mathbf{E}_{\boldsymbol{\xi}_n} \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \\ &= \frac{1}{2}\Big[\ln|\boldsymbol{M}_n| + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + \mathbf{E}_{\boldsymbol{\xi}_n}(\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{\xi}_n)\Big] = \\ &= \frac{1}{2}\left[\sum_{i=1}^n (\ln\lambda_i + \frac{1}{\lambda_i} - 1) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\right], \end{aligned} \tag{17}$$

где $\{\lambda_1, \ldots, \lambda_n\}$ – собственные значения (все положительные) ковариационной матрицы $\boldsymbol{M}_n$, а $\boldsymbol{a}_n = (a_1, \ldots, a_n)$.

1.2. Предположения. В модели (2) обозначим через $\lambda_1(\boldsymbol{M}_n), \ldots, \lambda_n(\boldsymbol{M}_n)$ собственные значения (все положительные) ковариационной матрицы $\boldsymbol{M}_n$. Предположим, что выполняются следующие предположения:

I. Для всех ковариационных матриц $\boldsymbol{M}_n \in \mathcal{M}_n(\boldsymbol{M}_n)$ существует предел (см. (17))

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(\ln \lambda_i(\boldsymbol{M}_n) + \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right) \tag{18}$$

(заметим, что $\ln z + 1/z - 1 \geqslant 0$, $z > 0$).

II. Для какого-нибудь $\delta > 0$ имеем

$$\lim_{n\to\infty} \frac{1}{n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \sum_{i=1}^{n} \left| \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right|^{1+\delta} < \infty. \tag{19}$$

1.3. Основные результаты. Сделаем сначала важное пояснение.

Замечание 2. При описании максимальных множеств $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ имеется следующая техническая трудность. Соотношение (9) имеет асимптотический (при $n \to \infty$) характер. Поэтому и максимальные множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ можно описать только асимптотически (при $n \to \infty$). Для этого удобнее всего описать наиболее простую последовательность множеств, которые в пределе дают максимальные множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

В статье для $(n \times n)$-матрицы $\boldsymbol{A}_n$ будем обозначать $|\boldsymbol{A}_n| = \det \boldsymbol{A}_n$. Через $(\boldsymbol{x}, \boldsymbol{y})$ будем обозначать скалярное произведение векторов $\boldsymbol{x}, \boldsymbol{y}$. Будем писать $\boldsymbol{A}_n > \boldsymbol{0}$, если матрица $\boldsymbol{A}_n$ положительно определена.

Пусть $\mathcal{C}_n$ – множество всех ковариационных (т.е. симметричных и положительно определенных) $(n \times n)$-матриц в $\mathbb{R}^n$. Для любых $\boldsymbol{M}_n, \boldsymbol{V}_n \in \mathcal{C}_n$, и $\boldsymbol{a}_n, \boldsymbol{b}_n \in \mathbb{R}^n$ определим функцию

$$f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) = \frac{|\boldsymbol{M}_n| e^{-K}}{|\boldsymbol{V}_n||\boldsymbol{B}_n|}, \tag{20}$$

где

$$\begin{aligned} &\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}, \quad \boldsymbol{d} = \boldsymbol{B}_n^{-1}(\boldsymbol{V}_n^{-1}\boldsymbol{b}_n - \boldsymbol{M}_n^{-1}\boldsymbol{a}_n), \\ &K = (\boldsymbol{b}_n, \boldsymbol{V}_n^{-1}\boldsymbol{b}_n) - (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - (\boldsymbol{d}, \boldsymbol{B}_n\boldsymbol{d}). \end{aligned} \tag{21}$$

Для последовательности пар $(\boldsymbol{a}_n, \boldsymbol{M}_n)$ введем следующую последовательность множеств пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$:

$$\begin{aligned} \mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n) &= \left\{ (\boldsymbol{b}_n, \boldsymbol{V}_n) : \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)} \right\} = \\ &= \left\{ (\boldsymbol{b}_n, \boldsymbol{V}_n) : \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1} > \boldsymbol{0},\ f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant e^{o(n)} \right\}, \end{aligned} \tag{22}$$

где функция $f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ определена в (20).

Следующая теорема является главным результатом статьи и описывает множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ из (10) и (15) соответственно.

Т е о р е м а 1. *Если выполняются предположения* (18), (19), *то при* $n \to \infty$

$$\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n), \tag{23}$$

где равенства понимаются в смысле замечания 2.

Замечание 3. Ясно, что $(\boldsymbol{a}_n, \boldsymbol{M}_n) \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Также множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ являются выпуклыми по $\boldsymbol{b}_n, \boldsymbol{V}_n$. Действительно, известно [6, § 8.5, теорема 4; 7, теорема 7.6.7], что функция $f(\boldsymbol{A}_n) = \ln|\boldsymbol{A}_n|$ является строго вогнутой на выпуклом множестве всех положительно определенных симметричных матриц в $\mathbb{R}^n$. Поэтому множество $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ является выпуклым, т.е. любые матрицы $\boldsymbol{V}_n^{(1)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\boldsymbol{V}_n^{(2)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ удовлетворяют условию

$$a\boldsymbol{V}_n^{(1)} + (1-a)\boldsymbol{V}_n^{(2)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n) \quad \text{для любого } 0 \leqslant a \leqslant 1.$$

В определенном смысле $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ – это множество $\mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n)$, расширенное на "тонкий слой" с шириной порядка $o(n)$. Другими словами, $\mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n)$ можно рассматривать как "ядро" множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

Приведем также следующее упрощающее следствие из теоремы 1. Без ограничения общности можно считать, что матрица $\boldsymbol{M}_n$ является диагональной (см. замечание 1) с собственными значениями $\{\lambda_i\}$ (все положительные). Также ограничимся в (23) только диагональными матрицами $\boldsymbol{V}_n$ с положительными собственными значениями $\{\nu_i\}$. Матрица $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ является диагональной с собственными значениями $\{\mu_i\}$:

$$\mu_i = 1 + \frac{1}{\nu_i} - \frac{1}{\lambda_i}, \quad i = 1, \ldots, n. \tag{24}$$

Тогда для $\boldsymbol{a}_n = (a_{1,n}, \ldots, a_{n,n})$, $\boldsymbol{b}_n = (b_{1,n}, \ldots, b_{n,n})$ из (21) имеем

$$K = \sum_{i=1}^{n} \left[\frac{b_{i,n}^2}{\nu_i} - \frac{a_{i,n}^2}{\lambda_i} - \frac{1}{\mu_i}\Big(\frac{b_{i,n}}{\nu_i} - \frac{a_{i,n}}{\lambda_i}\Big)^2\right]. \tag{25}$$

Введем выпуклое множество $\mathcal{C}_{\mathrm{diag},n}$ диагональных положительно определенных матриц $\boldsymbol{V}_n$:

$$\mathcal{C}_{\mathrm{diag},n} = \{\boldsymbol{V}_n \in \mathcal{C}_n : \ \boldsymbol{V}_n > \boldsymbol{0} \ \text{ и } \ \boldsymbol{V}_n \text{ – диагональная матрица}\}.$$

Если $\boldsymbol{M}_n, \boldsymbol{V}_n \in \mathcal{C}_{\mathrm{diag},n}$, то функция $f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ из (20) принимает вид

$$f^{(0)}_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) = e^{-K} \prod_{i=1}^{n} \Big(\frac{\lambda_i}{\nu_i \mu_i}\Big), \tag{26}$$

где $\{\mu_i\}$ определены в (24), а K определено в (25). Предполагается также, что $\mu_i > 0$, $i = 1, \ldots, n$.

Для последовательности пар $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, $\boldsymbol{M}_n \in \mathcal{C}_{\mathrm{diag},n}$, введем следующее множество пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$, $\boldsymbol{V}_n \in \mathcal{C}_{\mathrm{diag},n}$:

$$\begin{aligned} &\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \\ &\quad = \Big\{(\boldsymbol{b}_n, \boldsymbol{V}_n) : \ 1 + 1/\nu_i - 1/\lambda_i > 0, \ i = 1, \ldots, n, \ \ln f^{(0)}_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant o(n)\Big\}, \end{aligned} \tag{27}$$

где функция $f^{(0)}_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ определена в (26).

Тогда справедлива следующая "внутренняя граница" для $\mathcal{M}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

Т е о р е м а 2. *Если выполняются предположения* (18), (19), *то множество* $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ *содержит множество* $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$:

$$\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n) \subseteq \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n), \quad \boldsymbol{M}_n \in \mathcal{C}_{\mathrm{diag},n}, \tag{28}$$

где множество $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ *определено в* (27).

Множество $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ выпукло по $\boldsymbol{V}_n$ (см. замечание 3).

Далее в § 2 приводится вспомогательная теорема 3. В § 3 доказывается теорема 1, а в § 4 в качестве примеров рассматриваются некоторые частные случаи задачи.

§ 2. Вспомогательная теорема

В моделях (1), (2) рассмотрим сначала проверку простых гипотез: пара $(\mathbf{0}_n, \boldsymbol{I}_n)$ против пары $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Обозначим

$$D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) = D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n, \boldsymbol{M}_n}).$$

Следующая теорема является основным вспомогательным результатом, используемым в настоящей статье. Ее доказательство следует доказательству теоремы 3 в [8]. Более общий результат содержится в [10].

Т е о р е м а 3. *Для минимально возможного* $\beta(\alpha)$, $0 < \alpha < 1$, *справедливы границы*

$$\ln \beta(\alpha) \geqslant -\frac{D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) + h(\alpha)}{1-\alpha}, \quad h(\alpha) = -\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha), \tag{29}$$

и

$$\ln \beta(\alpha) \leqslant -D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) + \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n), \tag{30}$$

где $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ *определяется соотношением*

$$\mathbf{P}_{\boldsymbol{I}_n}\left\{ \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \mu_0 \right\} = \alpha. \tag{31}$$

Отметим, что обе границы (29) и (30) являются чисто аналитическими соотношениями, без каких-либо предельных операций. Нижняя граница (29) и верхняя граница (30) близки друг к другу, если величина $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ значительно меньше, чем $D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n)$ (которая обычно имеет порядок n).

Следующий результат дает для величины $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ из (31) оценку сверху порядка $n^{1/p}$, $p > 1$. Ее доказательство (см. Приложение) следует доказательству леммы 1 в [8].

Л е м м а 1. *Для* $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ *из* (30) *справедлива оценка сверху* (*см.* (19))

$$\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \left(\frac{24}{\alpha} \sum_{i=1}^{n} \left| \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right|^p \right)^{1/p} + 3\left\| \boldsymbol{M}_n^{-1} \boldsymbol{a}_n \right\| \sqrt{\ln(1/\alpha)}. \tag{32}$$

§ 3. Доказательство теоремы 1

Так как $\mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n) \subseteq \mathcal{F}_n(\boldsymbol{a}_n, \boldsymbol{M}_n)$, то для доказательства теоремы 1 достаточно получить “внутреннюю границу” для $\mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$, а затем получить аналогичную “внешнюю границу” для $\mathcal{F}_n(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

3.1. “Внутренняя граница” для $\boldsymbol{\mathcal{F}_n^{\mathrm{LR}}(a_n, M_n)}$. Оценим сначала сверху величину $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$. Для этого в модели (2) рассмотрим проверку простой гипотезы $(\mathbf{0}_n, \boldsymbol{I}_n)$ против простой альтернативы $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, когда $\boldsymbol{a}_n$ известна. Используем оптимальный LR-тест с решающей областью $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$ (см. (13), (14)), где $\mu_0 = \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) > 0$ определено в (31). Рассмотрим какую-либо другую пару $(\boldsymbol{b}_n, \boldsymbol{V}_n)$ и оценим вероятность ошибки 2-го рода $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$ при условии, что используется решающая область $\mathcal{A}_{\mu_0}$. Тогда

$$\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n) = \int\limits_{\mathcal{A}_{\mu_0}} p_{\boldsymbol{b}_n,\boldsymbol{V}_n}(\boldsymbol{x})d\boldsymbol{x} = \int\limits_{\mathcal{A}_{\mu_0}} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\frac{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}{p_{\boldsymbol{I}_n}}(\boldsymbol{x})p_{\boldsymbol{I}_n}(\boldsymbol{x})d\boldsymbol{x} =$$

$$= e^{-D(\boldsymbol{I}_n \| \boldsymbol{a}_n,\boldsymbol{M}_n)+\mu_1} \int\limits_{\mathcal{A}_{\mu_0}} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})p_{\boldsymbol{I}_n}(\boldsymbol{x})d\boldsymbol{x} \leqslant$$

$$\leqslant \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)e^{\mu_0}\, \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}), \tag{33}$$

где $0 \leqslant \mu_1 \leqslant \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$. В силу предположения (19) и оценки (32) имеем

$$\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) = O(n^{1/(1+\delta)}) = o(n), \quad n \to \infty. \tag{34}$$

Поэтому если

$$\sup_{(\boldsymbol{b}_n,\boldsymbol{V}_n)\in\mathcal{F}(\boldsymbol{a}_n,\boldsymbol{M}_n)} \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \quad n \to \infty, \tag{35}$$

то в силу (33)–(35) при $n \to \infty$

$$\sup_{(\boldsymbol{b}_n,\boldsymbol{V}_n)\in\mathcal{F}(\boldsymbol{a}_n,\boldsymbol{M}_n)} \ln \beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) + o(n). \tag{36}$$

3.2. “Внешняя граница” для $\boldsymbol{\mathcal{F}_n(a_n, M_n)}$. Получим теперь аналогичную оценку снизу для $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$. Рассмотрим сначала проверку простой гипотезы $(\mathbf{0}_n, \boldsymbol{I}_n)$ против простой альтернативы $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Используем для этого оптимальный LR-тест с решающей областью $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$ (см. (13), (14)). Тогда, обозначая $p = p_{\boldsymbol{I}_n}$ и $q = p_{\boldsymbol{a}_n,\boldsymbol{M}_n}$, для вероятностей ошибок имеем

$$\alpha = \mathbf{P}_{\boldsymbol{I}_n}(\mathcal{A}_{\mu_0}), \quad \beta_{\boldsymbol{a}_n,\boldsymbol{M}_n} = \int\limits_{\mathcal{A}_{\mu_0}} q(\boldsymbol{x})\, d\boldsymbol{x} = \beta(\alpha). \tag{37}$$

Рассмотрим какую-либо другую пару $(\boldsymbol{b}_n, \boldsymbol{V}_n)$. Пусть $\mathcal{D} \in \mathbb{R}^n$ – какая-то решающая область в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$, а $\beta_{\boldsymbol{b}_n,\boldsymbol{V}_n}(\mathcal{D})$ и $\alpha = \alpha(\mathcal{D})$ – соответствующие вероятности ошибок. Тогда, обозначая $q_1 = p_{\boldsymbol{b}_n,\boldsymbol{V}_n}$, для вероятности ошибки 2-го рода $\beta_{\boldsymbol{b}_n,\boldsymbol{V}_n}(\mathcal{D})$ (см. (37)) должно выполняться

$$\beta_{\boldsymbol{b}_n,\boldsymbol{V}_n}(\mathcal{D}) = \int\limits_{\mathcal{D}} q_1(\boldsymbol{x})\, d\boldsymbol{x} \leqslant \beta(\alpha)e^{o(n)}, \quad \alpha = \mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D}^c). \tag{38}$$

Для некоторого δ, $0 \leqslant \delta \leqslant 1$, рассмотрим также плотность вероятности

$$q_\delta(\boldsymbol{x}) = (1-\delta)q(\boldsymbol{x}) + \delta q_1(\boldsymbol{x}) \tag{39}$$

и соответствующую величину β_δ для нее:

$$\beta_\delta = \int\limits_{\mathcal{D}} q_\delta(\boldsymbol{x})\, d\boldsymbol{x} = (1-\delta)\beta(\alpha) + \delta\beta_{\boldsymbol{b}_n,\boldsymbol{V}_n}. \tag{40}$$

В силу (38) и (40) имеем

$$\beta_\delta \leqslant \beta(\alpha)\big(1 - \delta + \delta e^{o(n)}\big). \tag{41}$$

Отметим, что плотность вероятности $q_\delta(\boldsymbol{x})$ соответствует байесовской постановке задачи, когда альтернативная гипотеза $\mathcal{H}_1$ с вероятностью $1-\delta$ совпадает с $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, а с вероятностью δ – с $(\boldsymbol{b}_n, \boldsymbol{V}_n)$. Величина β_δ является соответствующей вероятностью ошибки 2-го рода.

Оценим снизу величину β_δ. Сначала имеем

$$\ln\frac{\beta_\delta}{1-\alpha} = \ln\left[\frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}\right] \geqslant \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} =$$
$$= -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))}{1-\alpha} - \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}^c} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}. \tag{42}$$

Для последнего члена в правой части (42) имеем

$$\int\limits_{\mathcal{D}^c} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} \leqslant \alpha\ln\left[\frac{1}{\alpha}\int\limits_{\mathcal{D}^c} q_\delta(\boldsymbol{x})\,d\boldsymbol{x}\right] = \alpha\ln\frac{1-\beta_\delta}{\alpha} \leqslant \alpha\ln\frac{1}{\alpha}.$$

Поэтому получаем

$$\ln\beta_\delta \geqslant -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) + h(\alpha)}{1-\alpha}. \tag{43}$$

Рассмотрим величину $D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))$ в правой части (43). Обозначая

$$r(\boldsymbol{x}) = \frac{q_1(\boldsymbol{x})}{q(\boldsymbol{x})}, \tag{44}$$

в силу (39) и (44) имеем

$$\frac{q_\delta(\boldsymbol{x})}{q(\boldsymbol{x})} = 1-\delta+\delta\frac{q_1(\boldsymbol{x})}{q(\boldsymbol{x})} = 1-\delta+\delta r(\boldsymbol{x}).$$

Поэтому

$$D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) = -\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} = D(p(\boldsymbol{x})\,\|\,q(\boldsymbol{x})) + g(\delta), \tag{45}$$

где

$$g(\delta) = -\int\limits_{\mathbb{R}^n} p(\mathbf{x})\ln[1-\delta+\delta r(\boldsymbol{x})]\,d\boldsymbol{x}. \tag{46}$$

Поэтому в силу (41), (45) и (46) необходимо иметь

$$g(\delta) \geqslant -\ln(1-\delta+\delta e^{o(n)}) \quad \text{для всех } 0<\delta\leqslant 1. \tag{47}$$

Заметим, что так как $\ln\mathbf{E}\,\xi \geqslant \mathbf{E}\ln\xi$, то из (46) имеем

$$g(\delta) \leqslant \ln\int\limits_{\mathbb{R}^n}\frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \quad \text{для всех } 0<\delta\leqslant 1.$$

Поэтому для того чтобы выполнялось (47), необходимо иметь

$$\int\limits_{\mathbb{R}^n}\frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \geqslant \frac{1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{48}$$

Так как $\int p(\boldsymbol{x})\,d\boldsymbol{x}=1$, то соотношение (48) эквивалентно условию

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})(r(\boldsymbol{x})-1)}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{e^{o(n)}-1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{49}$$

Заметим, что

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{1}{1-\delta}, \quad 0<\delta\leqslant 1.$$

Поэтому для того чтобы выполнялось (49), необходимо по меньшей мере иметь

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})r(\boldsymbol{x})}{1+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{e^{o(n)}}{(1-\delta)(1-\delta+\delta e^{o(n)})}. \tag{50}$$

Полагая $\delta\downarrow 0$, получаем из (50) необходимое условие

$$\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})r(\boldsymbol{x})\,d\boldsymbol{x} = \mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \tag{51}$$

что дает “внешнюю границу” для $\mathcal{F}_n(\boldsymbol{a}_n,\boldsymbol{M}_n)$ (см. (23)).

Заметим, что “внутренняя граница” (35), (36) для $\mathcal{F}_n(\boldsymbol{a}_n,\boldsymbol{M}_n)$ совпадает с (51). Поэтому для завершения доказательства теоремы 1 остается выразить условие (51) через матрицы $\boldsymbol{M}_n,\boldsymbol{V}_n$ и средние $\boldsymbol{a}_n,\boldsymbol{b}_n$. Для этого используем следующий результат.

Л е м м а 2. *Если* $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}>\boldsymbol{0}$, *то справедлива формула* (*см.* (20)–(22))

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2}e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}|\boldsymbol{B}_n|^{1/2}} = f^{1/2}_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n,\boldsymbol{V}_n), \tag{52}$$

где функция $f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n,\boldsymbol{V}_n)$ *определена в* (20).

Если матрица $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}$ *не является положительно определенной, то*

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \infty. \tag{53}$$

Д о к а з а т е л ь с т в о. Обозначая

$$\zeta = \big(\boldsymbol{\xi}_n-\boldsymbol{b}_n,\boldsymbol{V}_n^{-1}(\boldsymbol{\xi}_n-\boldsymbol{b}_n)\big)-\big(\boldsymbol{\xi}_n-\boldsymbol{a}_n,\boldsymbol{M}_n^{-1}(\boldsymbol{\xi}_n-\boldsymbol{a}_n)\big),$$

с помощью (11) получаем

$$\begin{aligned}\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) &= \mathbf{E}_{\boldsymbol{\xi}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{V}_n|^{1/2}}\,\mathbf{E}_{\boldsymbol{\xi}_n}\,e^{-\zeta/2} = \\ &= \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}}\int\limits_{\mathbb{R}^n} e^{-\frac12\left[(\boldsymbol{x},\boldsymbol{x})+(\boldsymbol{x}-\boldsymbol{b}_n,\boldsymbol{V}_n^{-1}(\boldsymbol{x}-\boldsymbol{b}_n))-(\boldsymbol{x}-\boldsymbol{a}_n,\boldsymbol{M}_n^{-1}(\boldsymbol{x}-\boldsymbol{a}_n))\right]}\,d\boldsymbol{x}.\end{aligned} \tag{54}$$

Заметим, что (см. (54))

$$(\boldsymbol{x},\boldsymbol{x})+(\boldsymbol{x}-\boldsymbol{b},\boldsymbol{V}^{-1}(\boldsymbol{x}-\boldsymbol{b}))-(\boldsymbol{x}-\boldsymbol{a},\boldsymbol{M}^{-1}(\boldsymbol{x}-\boldsymbol{a})) = (\boldsymbol{x}-\boldsymbol{d},\boldsymbol{B}(\boldsymbol{x}-\boldsymbol{d}))+K,$$

где (см. также (21))

$$\boldsymbol{B} = \boldsymbol{I} + \boldsymbol{V}^{-1} - \boldsymbol{M}^{-1}, \quad \boldsymbol{d} = \boldsymbol{B}^{-1}(\boldsymbol{V}^{-1}\boldsymbol{b} - \boldsymbol{M}^{-1}\boldsymbol{a}),$$
$$K = (\boldsymbol{b}, \boldsymbol{V}^{-1}\boldsymbol{b}) - (\boldsymbol{a}, \boldsymbol{M}^{-1}\boldsymbol{a}) - (\boldsymbol{d}, \boldsymbol{B}\boldsymbol{d}).$$

Поэтому (54) можно продолжить следующим образом:

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2}e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-((\boldsymbol{x}-\boldsymbol{d}),\boldsymbol{B}_n(\boldsymbol{x}-\boldsymbol{d}))/2}\, d\boldsymbol{x} =$$
$$= \frac{|\boldsymbol{M}_n|^{1/2}e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x},\boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x}. \tag{55}$$

Рассмотрим интеграл в правой части (55). Если $\boldsymbol{B}_n > \mathbf{0}$, то [6, § 6.9, теорема 3]

$$\int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x},\boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x} = \frac{(2\pi)^{n/2}}{|\boldsymbol{B}_n|^{1/2}}. \tag{56}$$

В противном случае

$$\int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x},\boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x} = \infty. \tag{57}$$

Предположим сначала, что $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1} > \mathbf{0}$, т.е. матрица $\boldsymbol{B}_n$ положительно определена. Тогда из (55), (56) получаем

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2}e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}|\boldsymbol{B}_n|^{1/2}}. \tag{58}$$

Если же матрица $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ не является положительно определенной, то в силу (57)

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \infty, \tag{59}$$

и поэтому условие (51) не может быть выполнено. Из (58), (59) следует лемма 2. ▲

Продолжим доказательство теоремы. Определим $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ как максимальное множество, удовлетворяющее условию

$$f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant e^{o(n)}, \quad n \to \infty. \tag{60}$$

Это множество совпадает с определением (22). Поэтому из (35), (51), (52) и (60) следует теорема 1.

§ 4. Примеры. Частные случаи

4.1. Известные среднее $\boldsymbol{a}_n$ и ковариационная матрица $\boldsymbol{M}_n$. Рассмотрим сначала простейший случай известных среднего $\boldsymbol{a}_n = (a_1, \ldots, a_n)$ и матрицы $\boldsymbol{M}_n$ и применим теорему 3. Это позволит оценить скорость сходимости в теореме 1. Без ограничения общности в модели (2) можно считать ковариационную матрицу $\boldsymbol{M}_n$ диагональной с положительными собственными значениями $\lambda_1, \ldots, \lambda_n$ (см. замечание 1).

Тогда (см. (17))

$$D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) = D(\boldsymbol{\xi}_n \,\|\, \boldsymbol{a}_n + \boldsymbol{\eta}_n) = \frac{1}{2}\left[\sum_{i=1}^{n}\left(\ln\lambda_i + \frac{1}{\lambda_i} - 1 + \frac{a_i^2}{\lambda_i}\right)\right]. \tag{61}$$

В силу (29), (30) для $D = D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n})$ получаем

$$-\frac{D+1}{1-\alpha} \leqslant \ln\beta(\alpha) \leqslant -D + \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n), \tag{62}$$

где $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ оценено в (32).

Для того чтобы оценить $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ проще, чем (32), предположим дополнительно, что выполняется следующее условие:

III. Существует $C > 0$, такое что

$$\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 \leqslant C^2 D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}). \tag{63}$$

Тогда, используя неравенство Чебышева, имеем

$$\begin{aligned}\alpha_\mu &= \mathbf{P}_{\boldsymbol{I}_n}\left\{D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}_n) \geqslant \mu\right\} \leqslant \\ &\leqslant \mu^{-2}\,\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 \leqslant C^2\mu^{-2} D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}).\end{aligned} \tag{64}$$

Для того чтобы правая часть (64) не превышала α, достаточно положить

$$\mu = C\sqrt{\frac{D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n})}{\alpha}},$$

и тогда (62) принимает вид

$$-\frac{D+1}{1-\alpha} \leqslant \ln\beta(\alpha) \leqslant -D + C\sqrt{\frac{D}{\alpha}},$$

который оценивает скорость сходимости в (62).

Заметим также, что аналогично (74), (75) нетрудно получить

$$\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 = \frac{1}{2}\sum_{i=1}^{n}\left[\left(1 - \frac{1}{\lambda_i}\right)^2 + 2\frac{a_i^2}{\lambda_i^2}\right]. \tag{65}$$

Поэтому условие **III** эквивалентно неравенству (см. (61) и (65))

$$\sum_{i=1}^{n}\left[\left(1 - \frac{1}{\lambda_i}\right)^2 + 2\frac{a_i^2}{\lambda_i^2}\right] \leqslant C^2\left[\sum_{i=1}^{n}\left(\ln\lambda_i + \frac{1}{\lambda_i} - 1 + \frac{a_i^2}{\lambda_i}\right)\right].$$

Замечание 4. Предположение (63) выполняется, например, в естественном "регулярном" случае, когда элементы $\boldsymbol{a}_{n+1}$, $\boldsymbol{M}_{n+1}$ являются "продолжениями" элементов $\boldsymbol{a}_n$, $\boldsymbol{M}_n$.

4.2. Неизвестное среднее $\boldsymbol{a_n}$ и известная ковариационная матрица $\boldsymbol{M_n}$. Рассмотрим случай модели (2), в котором известна ковариационная матрица $\boldsymbol{M}_n$, но не известно среднее $\boldsymbol{a}_n$. Без ограничения общности можно считать ковариационную матрицу $\boldsymbol{M}_n$ диагональной с положительными собственными значениями $\lambda_1, \ldots, \lambda_n$ (см. замечание 1). Тогда функция $f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{M}_n)$ из (20) принимает вид

$$f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{M}_n) = e^{-K},$$

где при $\boldsymbol{a}_n = (a_{1,n}, \ldots, a_{n,n})$ и $\boldsymbol{b}_n = (b_{1,n}, \ldots, b_{n,n})$ имеем

$$\begin{aligned} K &= (\boldsymbol{b}_n, \boldsymbol{M}_n^{-1}\boldsymbol{b}_n) - (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - \big(\boldsymbol{M}_n^{-1}(\boldsymbol{b}_n - \boldsymbol{a}_n), \boldsymbol{M}_n^{-1}(\boldsymbol{b}_n - \boldsymbol{a}_n)\big) = \\ &= \sum_{i=1}^{n}\left[\frac{b_{i,n}^2 - a_{i,n}^2}{\lambda_i} - \frac{(b_{i,n} - a_{i,n})^2}{\lambda_i^2}\right]. \end{aligned} \tag{66}$$

Соответствующее максимальное множество $\mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{M}_n) = \{\boldsymbol{b}_n\}$ в этом случае принимает вид (см. (22))

$$\mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{M}_n) = \{\boldsymbol{b}_n :\ K \geqslant o(n)\}, \tag{67}$$

где функция $K = K(\boldsymbol{a}_n, \boldsymbol{M}_n, \boldsymbol{b}_n)$ определена в (66).

Отметим, что в случае $\boldsymbol{M}_n = \boldsymbol{I}_n$ (т.е. когда гипотезы различаются только сдвигом $\boldsymbol{a}_n$) формулы (66), (67) принимают особенно простой вид:

$$K = 2(\boldsymbol{a}_n, \boldsymbol{b}_n - \boldsymbol{a}_n), \quad \mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{I}_n) = \{\boldsymbol{b}_n :\ (\boldsymbol{a}_n, \boldsymbol{b}_n - \boldsymbol{a}_n) \geqslant o(n)\}. \tag{68}$$

Эти результаты следуют также из работ [11, 12] (где эта задача рассматривалась в гильбертовом и банаховом пространствах).

4.3. Известное среднее $\boldsymbol{a_n}$ и неизвестная ковариационная матрица $\boldsymbol{M_n}$. Ограничимся случаем $\boldsymbol{a}_n = \boldsymbol{0}_n$. Тогда функция $f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ из (20) при $\boldsymbol{a}_n = \boldsymbol{b}_n = \boldsymbol{0}_n$ принимает вид

$$f_{\boldsymbol{0}_n,\boldsymbol{M}_n}(\boldsymbol{0}_n, \boldsymbol{V}_n) = \frac{|\boldsymbol{M}_n|}{|\boldsymbol{V}_n| \cdot \left|\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}\right|}. \tag{69}$$

Соответствующее максимальное множество $\mathcal{F}_1(\boldsymbol{0}_n, \boldsymbol{M}_n) = \{\boldsymbol{V}_n\}$ в этом случае принимает вид (см. (22))

$$\mathcal{F}_1(\boldsymbol{0}_n, \boldsymbol{M}_n) = \{\boldsymbol{V}_n :\ f_{\boldsymbol{0}_n,\boldsymbol{M}_n}(\boldsymbol{0}_n, \boldsymbol{V}_n) \leqslant e^{o(n)}\}. \tag{70}$$

Формулы (69), (70) совпадают с соответствующими результатами в [8, теорема 1].

ПРИЛОЖЕНИЕ: ДОКАЗАТЕЛЬСТВО ЛЕММЫ 1

Пусть $\boldsymbol{\xi}_n$ – гауссовский случайный вектор с распределением $\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, а $\boldsymbol{A}_n$ – симметричная $(n \times n)$-матрица с собственными значениями $\{a_i\}$. Рассмотрим квадратичную форму $(\boldsymbol{\xi}_n, \boldsymbol{A}_n\boldsymbol{\xi}_n)$. Существует ортогональная матрица $\boldsymbol{T}_n$, такая что $\boldsymbol{T}_n'\boldsymbol{A}_n\boldsymbol{T}_n = \boldsymbol{B}_n$, где $\boldsymbol{B}_n$ – диагональная матрица с диагональными элементами $\{a_i\}$ [6, § 4.7]. Так как $\boldsymbol{T}_n\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, то квадратичные формы $(\boldsymbol{\xi}_n, \boldsymbol{A}_n\boldsymbol{\xi}_n)$ и $(\boldsymbol{\xi}_n, \boldsymbol{B}_n\boldsymbol{\xi}_n)$ имеют одинаковые распределения. Поэтому из формулы (12) имеем

$$\ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{y}_n) \stackrel{d}{=} \frac{1}{2}\big[\ln|\boldsymbol{M}_n| + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + \eta_n\big], \tag{71}$$

где

$$\eta_n = (\boldsymbol{y}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I})\boldsymbol{y}_n) - 2(\boldsymbol{y}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n). \tag{72}$$

Введем величину (см. (31))

$$\alpha_\mu = \mathbf{P}_{\boldsymbol{I}_n}\left\{\ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}_n) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \mu\right\}. \tag{73}$$

Тогда с помощью (71), (72) и (17) для α_μ из (73) имеем

$$\begin{aligned}\alpha_\mu &\leqslant \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|(\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I})\boldsymbol{\xi}_n) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - \sum_{i=1}^n\Bigl(\frac{1}{\lambda_i} - 1\Bigr)\right| > 2\mu\right\} = \\ &= \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\sum_{i=1}^n\Bigl(\frac{1}{\lambda_i} - 1\Bigr)(\xi_i^2 - 1) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\right| > 2\mu\right\} \leqslant P_1 + P_2,\end{aligned} \tag{74}$$

где

$$\begin{aligned}P_1 &= \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\sum_{i=1}^n\Bigl(\frac{1}{\lambda_i} - 1\Bigr)(\xi_i^2 - 1)\right| > \mu\right\}, \\ P_2 &= \mathbf{P}_{\boldsymbol{\xi}_n}\bigl\{\bigl|(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\bigr| > \mu/2\bigr\}.\end{aligned} \tag{75}$$

Для того чтобы оценить величину P_1 из (75), используем следующий результат [13, п. III.5.15]: пусть $\zeta_1, \ldots, \zeta_n$ – независимые случайные величины с $\mathbf{E}\,\zeta_i = 0$, $i = 1, \ldots, n$. Тогда для любого $1 \leqslant p \leqslant 2$

$$\mathbf{E}\left|\sum_{i=1}^n \zeta_i\right|^p \leqslant 2\sum_{i=1}^n \mathbf{E}\,|\zeta_i|^p. \tag{76}$$

Поэтому, используя для P_1 неравенство Чебышева и (76), получаем

$$\begin{aligned}P_1 &\leqslant \mu^{-p}\,\mathbf{E}\left|\sum_{i=1}^n\Bigl(\frac{1}{\lambda_i} - 1\Bigr)(\xi_i^2 - 1)\right|^p \leqslant 2\mu^{-p}\sum_{i=1}^n\left|\frac{1}{\lambda_i} - 1\right|^p \mathbf{E}\,|\xi_i^2 - 1|^p \leqslant \\ &\leqslant 2\mu^{-p}\sum_{i=1}^n\left|\frac{1}{\lambda_i} - 1\right|^p\bigl(\mathbf{E}\,|\xi^2 - 1|^2\bigr)^{p/2} \leqslant 2\mu^{-p}6^{p/2}\sum_{i=1}^n\left|\frac{1}{\lambda_i} - 1\right|^p \leqslant \\ &\leqslant 12\mu^{-p}\sum_{i=1}^n\left|\frac{1}{\lambda_i} - 1\right|^p.\end{aligned} \tag{77}$$

Для того чтобы оценить величину P_2 из (74), (75), заметим, что

$$(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) \sim \mathcal{N}(0, \|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|),$$

и тогда

$$(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) \stackrel{d}{=} \|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|\xi.$$

Поэтому, используя стандартную оценку

$$\mathbf{P}(|\xi| \geqslant z) \leqslant e^{-z^2/2}, \quad z \geqslant 0,$$

получаем ($\xi_i \sim \mathcal{N}(0,1)$)

$$P_2 = \mathbf{P}_{\boldsymbol{\xi}_n}\bigl\{|(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)| > \mu/2\bigr\} \leqslant e^{-\mu^2/(8\|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|^2)}. \tag{78}$$

Для выполнения условия $\alpha_\mu \leqslant \alpha$ выберем μ так, что $\max\{P_1, P_2\} \leqslant \alpha/2$. В силу (77) и (78) для этого достаточно положить μ удовлетворяющим оценке (32).

СПИСОК ЛИТЕРАТУРЫ

1. *Вальд А.* Статистические решающие функции // Позиционные игры. М.: Наука, 1967. С. 300–522.
2. *Леман Э.* Проверка статистических гипотез. М.: Наука, 1979.
3. *Poor H.V.* An Introduction to Signal Detection and Estimation. New York: Springer, 1994.
4. *Zhang W., Poor H.V.* On Minimax Robust Detection of Stationary Gaussian Signals in White Gaussian Noise // IEEE Trans. Inform. Theory. 2011. V. 57. № 6. P. 3915–3924. https://doi.org/10.1109/TIT.2011.2136210
5. *Бурнашев М.В.* Об обнаружении гауссовских стохастических последовательностей // Пробл. передачи информ. 2017. Т. 53. № 4. С. 49–68. http://mi.mathnet.ru/ppi2252
6. *Беллман Р.* Введение в теорию матриц. М.: Наука, 1976.
7. *Хорн Р., Джонсон Ч.* Матричный анализ. М.: Мир, 1989.
8. *Бурнашев М.В.* О минимаксном обнаружении гауссовских стохастических последовательностей и гауссовских стационарных сигналов // Пробл. передачи информ. 2021. Т. 57. № 3. С. 55–72. https://doi.org/10.31857/S0555292321030049
9. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967.
10. *Burnashev M.V.* On Stein's Lemma in Hypotheses Testing in General Non-Asymptotic Case // Stat. Inference Stoch. Process. 2022. Online First article. https://doi.org/10.1007/s11203-022-09278-4
11. *Бурнашев М.В.* О минимаксном обнаружении неточно известного сигнала на фоне белого гауссовского шума // Теория вероятн. и ее примен. 1979. Т. 24. № 1. С. 106–118. http://mi.mathnet.ru/tvp957
12. *Бурнашев М.В.* О различении гипотез для гауссовских мер и одна геометрическая характеризация гауссовского распределения // Матем. заметки. 1982. Т. 32. № 4. С. 549–556. http://mi.mathnet.ru/mz6021
13. *Петров В.В.* Суммы независимых случайных величин. М.: Наука, 1972.

Бурнашев Марат Валиевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
burn@iitp.ru

Поступила в редакцию
28.03.2022
После доработки
18.08.2022
Принята к публикации
19.08.2022