ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 57 **2021** **Вып. 3**

УДК 621.391 : 519.23

 М.В. Бурнашев

О МИНИМАКСНОМ ОБНАРУЖЕНИИ ГАУССОВСКИХ СТОХАСТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И ГАУССОВСКИХ СТАЦИОНАРНЫХ СИГНАЛОВ[1]

Рассматривается задача обнаружения гауссовских стохастических последовательностей (сигналов) с неизвестными ковариационными матрицами на фоне белого гауссовского шума. 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. Целью является нахождение максимального множества ковариационных матриц (сложная гипотеза), такого что его минимаксную проверку можно заменить проверкой одной конкретной ковариационной матрицы (простая гипотеза) без ухудшения экспоненты обнаружения. В статье полностью описывается это максимальное множество ковариационных матриц. Рассматриваются также некоторые следствия, касающиеся минимаксного обнаружения гауссовских стохастических сигналов в белом гауссовском шуме и обнаружения гауссовских стационарных сигналов.

Ключевые слова: минимаксная проверка гипотез, вероятности ошибки, экспонента вероятности ошибки, лемма Стейна.

DOI: 10.31857/S0555292321030049

§ 1. Введение

1.1. Введение и определения. Рассматривается задача минимаксной проверки простой гипотезы $\mathcal{H}_0$ против сложной альтернативы $\mathcal{H}_1$ [1–3] по наблюдениям $\boldsymbol{y}_n^T = \boldsymbol{y}'_n = (y_1, \ldots, y_n) \in \mathbb{R}^n$:

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

где выборка $\boldsymbol{\xi}_n^T = (\xi_1, \ldots, \xi_n)$ представляет “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 нулевым средним и неизвестной ковариационной матрицей $\boldsymbol{M}_n$. $\mathcal{M}_n$ – заданное множество возможных ковариационных матриц $\boldsymbol{M}_n$.

Без ограничения общности можно предполагать матрицу $\boldsymbol{M}_n$ положительно определенной, т.е. $|\boldsymbol{M}_n| = \det \boldsymbol{M}_n > 0$. Действительно, если $|\boldsymbol{M}_n| = 0$, то тогда меры $\mathbf{P}_{\boldsymbol{\xi}_n}$ и $\mathbf{P}_{\boldsymbol{\eta}_n}$ ортогональны, и поэтому гипотезы $\mathcal{H}_0$ и $\mathcal{H}_1$ можно различить без ошибок.

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

Для принятия решения при различении гипотез $\mathcal{H}_0$ и $\mathcal{H}_1$ выбирается область $\mathcal{D} \in \mathbb{R}^n$, такая что

$$\boldsymbol{y}_n \in \mathcal{D} \Rightarrow \mathcal{H}_0, \quad \boldsymbol{y}_n \notin \mathcal{D} \Rightarrow \mathcal{H}_1. \tag{2}$$

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

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

и

$$\beta(\mathcal{D}, \mathcal{M}_n) = \mathbf{P}(\boldsymbol{y}_n \in \mathcal{D} \,|\, \mathcal{H}_1) = \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \mathbf{P}(\boldsymbol{y}_n \in \mathcal{D} \,|\, \boldsymbol{M}_n).$$

При заданной вероятности ошибки 1-го рода α, $0 < \alpha < 1$, мы исследуем минимально возможную вероятность ошибки 2-го рода

$$\beta(\alpha, \mathcal{M}_n) = \inf_{\mathcal{D}:\alpha(\mathcal{D}) \leqslant \alpha} \beta(\mathcal{D}, \mathcal{M}_n) \tag{3}$$

и соответствующую оптимальную область принятия решения $\mathcal{D}(\alpha)$ (см. (2)).

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 ошибку 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 заданной матрицы $\boldsymbol{M}_n$ и фиксированного α обозначим через $\beta(\boldsymbol{M}_n) = \beta(\alpha, \boldsymbol{M}_n)$ минимально возможную вероятность ошибки 2-го рода. Аналогично для заданного множества $\mathcal{M}_n = \{\boldsymbol{M}_n\}$ и фиксированного α обозначим через $\beta(\mathcal{M}_n) = \beta(\alpha, \mathcal{M}_n)$ минимально возможную минимаксную вероятность ошибки 2-го рода (см. (3)). Ясно, что тогда

$$\sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\boldsymbol{M}_n) \leqslant \beta(\mathcal{M}_n), \tag{4}$$

или, эквивалентным образом,

$$\sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \inf_{\mathcal{D}} \beta(\mathcal{D}, \boldsymbol{M}_n) \leqslant \inf_{\mathcal{D}} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\mathcal{D}, \mathcal{M}_n).$$

Во многих практических случаях величина $\beta(\boldsymbol{M}_n)$ убывает экспоненциально по $n \to \infty$. Поэтому естественно (во всяком случае, проще и продуктивнее) исследовать соответствующие экспоненты $n^{-1} \ln \beta(\boldsymbol{M}_n)$ и $n^{-1} \ln \beta(\mathcal{M}_n)$ при $n \to \infty$ (некоторые результаты относительно равенства в (4) имеются в [5]).

В статье для фиксированного α и заданной последовательности матриц $\boldsymbol{M}_n$ исследуется последовательность множеств $\mathcal{M}_n$, таких что $\boldsymbol{M}_n \in \mathcal{M}_n$, $n = 1, 2, \ldots$, и выполняется следующее равенство:

$$\lim_{n \to \infty} \frac{1}{n} \ln \beta(\boldsymbol{M}_n) = \lim_{n \to \infty} \frac{1}{n} \ln \beta(\mathcal{M}_n(\boldsymbol{M}_n)). \tag{5}$$

Другими словами, при заданной вероятности ошибки 1-го рода α множество $\mathcal{M}_n$ есть множество ковариационных матриц, которое можно заменить одной матрицей $\boldsymbol{M}_n$ без потери для экспоненты вероятности ошибки 2-го рода. Далее будет описано максимальное из таких множеств $\mathcal{M}_n(\boldsymbol{M}_n)$ для модели (1), а также некоторые оценки для него.

Мотивация рассмотрения задачи минимаксной проверки гипотез (обнаружения сигналов) подробно описана в [1–3]. Если для заданного множества матриц $\mathcal{M}_n$ выполняется соотношение (5), то тогда можно заменить (без асимптотических потерь) все множество $\mathcal{M}_n$ одной конкретной матрицей $\boldsymbol{M}_n$. Напомним, что оптимальный тест для конкретной матрицы $\boldsymbol{M}_n$ описывается леммой Неймана – Пирсона и сводится к простому LR-тесту (likelihood ratio test detector). В противном случае (без соотношения (5)) оптимальным минимаксным тестом является значительно более сложный байессовский тест относительно *наименее благоприятного* априорного распределения на множестве $\mathcal{M}_n$. Поэтому естественно исследовать, когда заданное множество матриц $\mathcal{M}_n$ можно заменить конкретной матрицей $\boldsymbol{M}_n$. Однако технически удобнее рассмотреть эквивалентную задачу: для заданной матрицы $\boldsymbol{M}_n$ найти максимальное множество матриц $\mathcal{M}_n(\boldsymbol{M}_n)$, которое может быть заменено матрицей $\boldsymbol{M}_n$. Эта задача в основном и рассматривается в настоящей статье.

Определение 1. Для фиксированного α и заданной последовательности матриц $\boldsymbol{M}_n$, $n = 1, 2, \ldots$, обозначим через $\mathcal{M}_n(\boldsymbol{M}_n)$ последовательность максимальных множеств, таких что $\boldsymbol{M}_n \in \mathcal{M}_n(\boldsymbol{M}_n)$, и выполняется следующее равенство:

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{M}_n(\boldsymbol{M}_n)) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{M}_n).$$

Обозначим в модели (1) через $\mathbf{P}_{\boldsymbol{I}_n}$ распределение величины $\boldsymbol{y}_n = \boldsymbol{\xi}_n$, где $\boldsymbol{\xi}_n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n)$. Аналогично обозначим через $\mathbf{Q}_{\boldsymbol{M}_n}$ распределение величины $\boldsymbol{y}_n = \boldsymbol{\eta}_n$, где $\boldsymbol{\eta}_n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{M}_n)$. Обозначим также через $p_{\boldsymbol{I}_n}(\boldsymbol{y}_n)$ и $p_{\boldsymbol{M}_n}(\boldsymbol{y}_n)$, $\boldsymbol{y}_n \in \mathbb{R}^n$, соответствующие плотности распределения вероятностей.

Заметим, что если $|\boldsymbol{V}_n| \neq 0$ и $|\boldsymbol{M}_n| \neq 0$, то тогда

$$\ln \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{y}_n) = \frac{1}{2}\left[\ln \frac{|\boldsymbol{M}_n|}{|\boldsymbol{V}_n|} + \bigl(\boldsymbol{y}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{V}_n^{-1})\boldsymbol{y}_n\bigr)\right]. \tag{6}$$

Введем также аналогичные $\mathcal{M}_n(\boldsymbol{M}_n)$ максимальные множества $\mathcal{M}_n^{LR}(\boldsymbol{M}_n)$, возникающие, если используются LR-тесты (см. определение 2 ниже). Будет показано, что при естественных предположениях $\mathcal{M}_n(\boldsymbol{M}_n) = \mathcal{M}_n^{LR}(\boldsymbol{M}_n)$, т.е. LR-тесты являются асимптотически оптимальными.

Для этого введем логарифм отношения правдоподобия (см. (6))

$$f_{\boldsymbol{M}_n}(\boldsymbol{y}_n) = \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{y}_n) = \frac{1}{2}\Bigl[\ln |\boldsymbol{M}_n| + \bigl(\boldsymbol{y}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{y}_n\bigr)\Bigr]. \tag{7}$$

Область принятия решения $\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha)$ в пользу $\boldsymbol{I}_n$ при проверке матриц $\boldsymbol{I}_n$ и $\boldsymbol{M}_n$ имеет вид

$$\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha) = \{\boldsymbol{y}_n \in \mathbb{R}^n : \ f_{\boldsymbol{M}_n}(\boldsymbol{y}_n) \geqslant \gamma\}, \tag{8}$$

где γ – такое, что

$$\begin{aligned} \alpha &= \mathbf{P}_{\boldsymbol{I}_n}\{\mathcal{D}_{LR}^c(\boldsymbol{M}_n, \alpha)\} = \mathbf{P}_{\boldsymbol{I}_n}\{f_{\boldsymbol{M}_n}(\boldsymbol{y}_n) \leqslant \gamma\} = \\ &= \mathbf{P}_{\boldsymbol{I}_n}\bigl\{\bigl[(\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{\xi}_n) + \ln |\boldsymbol{M}_n|\bigr] \leqslant 2\gamma\bigr\}. \end{aligned} \tag{9}$$

Предположим, что в модели (1) для проверки матриц $\boldsymbol{I}_n$ и $\boldsymbol{M}_n$ (т.е. простых гипотез) используется оптимальный детектор (т.е. LR-тест) с областью принятия решения $\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha)$ (см. (8), (9)) в пользу $\boldsymbol{I}_n$. Для каких матриц $\boldsymbol{V}_n$ вместо $\boldsymbol{M}_n$ область принятия решения $\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha)$ не ухудшит вероятность ошибки 2-го рода $\beta(\alpha, \boldsymbol{M}_n)$? Для ответа на этот вопрос введем следующее

Определение 2. Для фиксированного α и заданной последовательности матриц $\boldsymbol{M}_n$, $n = 1, 2, \ldots$, обозначим через $\mathcal{M}_n^{LR}(\boldsymbol{M}_n)$ последовательность максимальных множеств матриц $\boldsymbol{V}_n$, таких что выполняется следующее равенство

$$\lim_{n\to\infty}\frac{1}{n}\ln\sup_{\boldsymbol{V}_n\in\mathcal{M}_n^{LR}(\boldsymbol{M}_n)}\beta(\boldsymbol{V}_n)=\lim_{n\to\infty}\frac{1}{n}\ln\beta(\boldsymbol{M}_n)$$

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

1.2. Расстояние Кульбака – Лейблера. Для случайных элементов $\boldsymbol{x}$ и $\boldsymbol{y}$ на измеримом пространстве $(\Omega,\mathcal{B})$ с вероятностными распределениями, соответственно, $\mathbf{P}_{\boldsymbol{x}}$ и $\mathbf{Q}_{\boldsymbol{y}}$ введем функцию (*расстояние* или *дивергенция Кульбака – Лейблера* для мер $\mathbf{P}_{\boldsymbol{x}}$ и $\mathbf{Q}_{\boldsymbol{y}}$)

$$D(\mathbf{P}_{\boldsymbol{x}}\,\|\,\mathbf{Q}_{\boldsymbol{y}})=\mathbf{E}_{\mathbf{P}_{\boldsymbol{x}}}\ln\frac{d\mathbf{P}_{\boldsymbol{x}}}{d\mathbf{Q}_{\boldsymbol{y}}}(\boldsymbol{u}). \tag{10}$$

В частности, если $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ и $\boldsymbol{x}\sim\mathcal{N}(\mathbf{0},\boldsymbol{V}_n)$, $\boldsymbol{y}\sim\mathcal{N}(\mathbf{0},\boldsymbol{M}_n)$, то тогда [6, гл. 9.1] (через tr обозначен след матрицы)

$$D(\mathbf{P}_{\boldsymbol{x}}\,\|\,\mathbf{Q}_{\boldsymbol{y}})=\frac{1}{2}\ln\frac{|\boldsymbol{M}_n|}{|\boldsymbol{V}_n|}+\frac{1}{2}\operatorname{tr}\bigl(\boldsymbol{V}_n\boldsymbol{M}_n^{-1}\bigr)-\frac{n}{2}.$$

Если $\boldsymbol{V}_n=\boldsymbol{I}_n$, то

$$D(\boldsymbol{I}_n\,\|\,\boldsymbol{M}_n)=D(\mathbf{P}_{\boldsymbol{x}}\,\|\,\mathbf{Q}_{\boldsymbol{y}})=\frac{1}{2}\sum_{i=1}^{n}\Bigl(\ln\lambda_i+\frac{1}{\lambda_i}-1\Bigr), \tag{11}$$

где $\lambda_1,\ldots,\lambda_n$ – собственные значения матрицы $\boldsymbol{M}_n$ (собственные значения матрицы $\boldsymbol{M}_n^{-1}$ суть $\lambda_1^{-1},\ldots,\lambda_n^{-1}$).

Расстояние Кульбака – Лейблера играет важную роль в проверке гипотез. Предположим, например, что в модели (1) для проверки гипотез $\boldsymbol{I}_n$ и $\boldsymbol{M}_n$ (т.е. простых гипотез) используется оптимальный тест, которым является LR-тест с областью принятия решения $\mathcal{D}_{LR}(\alpha,\boldsymbol{M}_n)$ (см. (8), (9)) в пользу $\boldsymbol{I}_n$. Тогда при некоторых естественных предположениях справедлива следующая формула:

$$\lim_{\alpha\to 0}\lim_{n\to\infty}\frac{1}{n}\ln\beta(\alpha)=-\lim_{n\to\infty}\frac{1}{n}D(\boldsymbol{I}_n\,\|\,\boldsymbol{M}_n). \tag{12}$$

Формула (12) называется леммой Стейна [7, 8]. В случае независимых одинаково распределенных случайных величин ее доказательство можно найти в [6, теорема 3.3; 9, теорема 12.8.1]. Естественно ожидать, что формула (12) справедлива не только при проверке простых гипотез, но и в более общих случаях проверки сложных гипотез. Некоторые частные аналоги формулы (12) уже появились для гауссовских стационарных [4] и пуассоновских [10] случайных процессов. В статье будут получены аналоги формулы (12) для моделей (1) и (68) (см. ниже).

§ 2. Предположения и основные результаты

2.1. Предположения. Пусть $\mathcal{C}_n$ – выпуклое множество всех ковариационных (т.е. симметричных и положительно определенных) $(n\times n)$-матриц в $\mathbb{R}^n$. Рассмотрим последовательность множеств $\mathcal{M}_i\in\mathcal{C}_i$ ковариационных матриц $\boldsymbol{M}_i\in\mathcal{M}_i$, $i=1,2,\ldots$, в “схеме серий”, т.е. множество $\mathcal{M}_{i+1}$ не обязательно является “продолжением” множества $\mathcal{M}_i$. Обозначим через $\lambda_1(\boldsymbol{M}_n),\ldots,\lambda_n(\boldsymbol{M}_n)$ собственные значения (все они положительны) матрицы $\boldsymbol{M}_n$. Далее нам понадобятся следующие предположения:

I. Для последовательности множеств $\mathcal{M}_n$ существует положительный предел (см. (11))

$$\lim_{n\to\infty}\frac{1}{n}\inf_{\boldsymbol{M}_n\in\mathcal{M}_n}\sum_{i=1}^{n}\left(\ln\lambda_i(\boldsymbol{M}_n)+\frac{1}{\lambda_i(\boldsymbol{M}_n)}-1\right)>0 \tag{13}$$

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

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

$$\lim_{n\to\infty}\frac{1}{n}\sup_{\boldsymbol{M}_n\in\mathcal{M}_n}\sum_{i=1}^{n}\lambda_i^{-1-\delta}(\boldsymbol{M}_n)<\infty. \tag{14}$$

2.2. Основные результаты. Для $(n\times n)$-матрицы $\boldsymbol{A}_n$ обозначим $|\boldsymbol{A}_n|=\det\boldsymbol{A}_n$. Пусть $\boldsymbol{A}_n>0$ означает, что матрица $\boldsymbol{A}_n$ положительно определена.

Для матриц $\boldsymbol{M}_n,\boldsymbol{V}_n\in\mathcal{C}_n$, таких что $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}>0$, определим функцию

$$f(\boldsymbol{M}_n,\boldsymbol{V}_n)=\frac{|\boldsymbol{I}_n+\boldsymbol{V}_n(\boldsymbol{I}_n-\boldsymbol{M}_n^{-1})|}{|\boldsymbol{M}_n|}=\frac{|\boldsymbol{M}_n+(\boldsymbol{V}_n-\boldsymbol{M}_n)(\boldsymbol{I}_n-\boldsymbol{M}_n^{-1})|}{|\boldsymbol{M}_n|}, \tag{15}$$

где использовалось полезное тождество

$$\boldsymbol{I}_n+\boldsymbol{V}_n(\boldsymbol{I}_n-\boldsymbol{M}_n^{-1})=\boldsymbol{M}_n+(\boldsymbol{V}_n-\boldsymbol{M}_n)(\boldsymbol{I}_n-\boldsymbol{M}_n^{-1}). \tag{16}$$

Основной результат статьи описывает множества $\mathcal{M}_n(\boldsymbol{M}_n)$ и $\mathcal{M}_n^{LR}(\boldsymbol{M}_n)$.

Т е о р е м а 1. *Если выполнены предположения* (13), (14), *то при* $n\to\infty$

$$\begin{aligned}&\mathcal{M}_n(\boldsymbol{M}_n)=\mathcal{M}_n^{LR}(\boldsymbol{M}_n)=\\&=\left\{\boldsymbol{V}_n\in\mathcal{C}_n:\ \lim_{n\to\infty}\sup_{\boldsymbol{V}_n\in\mathcal{M}_n(\boldsymbol{M}_n)}\frac{1}{n}\ln\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x})\leqslant 0\right\}=\\&=\left\{\boldsymbol{V}_n\in\mathcal{C}_n:\ \boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}>0,\ \lim_{n\to\infty}\inf_{\boldsymbol{V}_n\in\mathcal{M}_n(\boldsymbol{M}_n)}\frac{1}{n}\ln f(\boldsymbol{M}_n,\boldsymbol{V}_n)\geqslant 0\right\},\end{aligned} \tag{17}$$

где функция $f(\boldsymbol{M}_n,\boldsymbol{V}_n)$ *определена в* (15).

Ясно, что множества $\mathcal{M}_n(\boldsymbol{M}_n)$и $\mathcal{M}_n^{LR}(\boldsymbol{M}_n)$ являются выпуклыми.

Замечание 1. Известно [11, гл. 8.5, теорема 4], [12, теорема 7.6.7], что функция $f(\boldsymbol{A}_n)=\ln|\boldsymbol{A}_n|$ строго вогнута на выпуклом множестве $\mathcal{C}_n$ симметричных положительно определенных матриц в $\mathbb{R}^n$. Отсюда также следует выпуклость множества $\mathcal{M}_n(\boldsymbol{M}_n)$, т.е. если $\boldsymbol{V}_n^{(1)}\in\mathcal{M}_n(\boldsymbol{M}_n)$ и $\boldsymbol{V}_n^{(2)}\in\mathcal{M}_n(\boldsymbol{M}_n)$, то $a\boldsymbol{V}_n^{(1)}+(1-a)\boldsymbol{V}_n^{(2)}\in$ $\in\mathcal{M}_n(\boldsymbol{M}_n)$ для любого $0\leqslant a\leqslant 1$.

Упростим теорему 1, ограничиваясь в (17) только матрицами $\boldsymbol{V}_n$, коммутирующими с $\boldsymbol{M}_n$. Для ковариационной матрицы $\boldsymbol{M}_n\in\mathcal{C}_n$ введем выпуклое множество $\mathcal{C}_{\boldsymbol{M}_n}$ ковариационных матриц $\boldsymbol{V}_n$, коммутирующих с $\boldsymbol{M}_n$:

$$\mathcal{C}_{\boldsymbol{M}_n}=\{\boldsymbol{V}_n\in\mathcal{C}_n:\ \boldsymbol{M}_n\boldsymbol{V}_n=\boldsymbol{V}_n\boldsymbol{M}_n\}. \tag{18}$$

Обозначим через $\{\lambda_i\}$ собственные значения матрицы $\boldsymbol{M}_n$, а через $\{\nu_i\}$ – собственные значения матрицы $\boldsymbol{V}_n$. Тогда функция $f(\boldsymbol{M}_n,\boldsymbol{V}_n)$ из (15), (16) принимает вид

$$f(\boldsymbol{M}_n,\boldsymbol{V}_n)=\prod_{i=1}^{n}\frac{[\lambda_i+\nu_i(\lambda_i-1)]}{\lambda_i^2}=\prod_{i=1}^{n}\left[1+\frac{(\nu_i-\lambda_i)(\lambda_i-1)}{\lambda_i^2}\right] \tag{19}$$

при условии $\lambda_i+\nu_i(\lambda_i-1)>0$, $i=1,\ldots,n$.

Введем следующее подмножество множества $\mathcal{C}_{\boldsymbol{M}_n}$ при $n \to \infty$:

$$\mathcal{V}_n^{(0)}(\boldsymbol{M}_n) = \Big\{ \boldsymbol{V}_n : \inf_{\boldsymbol{V}_n \in \mathcal{V}_n^{(0)}(\boldsymbol{M}_n)} \ln f(\boldsymbol{M}_n, \boldsymbol{V}_n) \geqslant o(n) \Big\}, \tag{20}$$

где функция $f(\boldsymbol{M}_n, \boldsymbol{V}_n)$ определена в (19). Тогда справедлива следующая "внутренняя граница" $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$ для $\mathcal{M}_n(\boldsymbol{M}_n)$.

Теорема 2. *Если выполнены предположения* (13), (14), *то множество* $\mathcal{M}_n(\boldsymbol{M}_n)$ *содержит множество* $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$:

$$\mathcal{V}_n^{(0)}(\boldsymbol{M}_n) \subseteq \mathcal{M}_n(\boldsymbol{M}_n). \tag{21}$$

Множество $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$ является выпуклым по $\boldsymbol{V}_n$ (см. замечание 1).

Замечание 2. Наряду с предельной формой неравенств в (17) далее некоторые подобные условия записываются также в допредельном виде с использованием остаточных членов типа $o(1)$, $o(n)$ и т.п. Таким образом автор указывает, что эти остаточные члены можно оценить, что дает оценки скорости сходимости в соответствующих формулах.

Замечание 3. Заменяя $o(n)$ величиной 0 в правой части формулы (20), рассмотрим множество

$$\mathcal{V}_n^{(1)}(\boldsymbol{M}_n) = \Big\{ \boldsymbol{V}_n : \inf_{\boldsymbol{V}_n \in \mathcal{V}_n^{(1)}(\boldsymbol{M}_n)} \ln f(\boldsymbol{M}_n, \boldsymbol{V}_n) \geqslant 0 \Big\}. \tag{22}$$

Ясно, что $\mathcal{V}_n^{(1)}(\boldsymbol{M}_n) \in \mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$. В определенном смысле $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$ – это множество $\mathcal{V}_n^{(1)}(\boldsymbol{M}_n)$, расширенное на "тонкий слой" с шириной порядка $o(n)$. Другими словами, $\mathcal{V}_n^{(1)}(\boldsymbol{M}_n)$ из (22) можно рассматривать как "ядро" множества $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$.

Эта статья была написана под влиянием работы [4], где рассматривалась аналогичная задача для стохастических стационарных сигналов. В данной статье рассматривается более общий случай гауссовских случайных векторов с неизвестной ковариационной матрицей. При этом оказывается возможным провести, по существу, неасимптотическое исследование задачи. Стохастические стационарные сигналы являются тогда определенным предельным частным случаем задачи. Ранее подобный подход использовался в случае пуассоновских процессов [10]. Некоторые другие случаи с дополнительными ограничениями на вероятности ошибки α, β рассматривались в [13, 14].

2.3. Обратная задача. Следуя работе [4], рассмотрим также обратную задачу: когда проверку заданного множества $\mathcal{M}_n$ можно заменить проверкой некоторой матрицы $\boldsymbol{M}_n^{(0)}$?

Достаточное условие для этого следует из теоремы 1. Пусть задана последовательность множеств $\{\mathcal{M}_1, \mathcal{M}_2, \ldots\}$, $\mathcal{M}_i \subset \mathcal{C}_i$, $i = 1, 2, \ldots$, ковариационных матриц $\boldsymbol{M}_i \in \mathcal{M}_i$. Обозначим через $\{\boldsymbol{M}_i^{(0)}\}$ последовательность ковариационных матриц $\boldsymbol{M}_i^{(0)}$ (если таковая существует), удовлетворяющая следующему аналогу формулы (5):

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{M}_n^{(0)}) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{M}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta\big(\{\mathcal{M}_n, \boldsymbol{M}_n^{(0)}\}\big), \tag{23}$$

где $\{\mathcal{M}_n, \boldsymbol{M}_n^{(0)}\} = \mathcal{M}_n \cup \boldsymbol{M}_n^{(0)}$ – "расширение" множества $\mathcal{M}_n$ на матрицу $\boldsymbol{M}_n^{(0)}$. Мы не требуем, чтобы $\boldsymbol{M}_i^{(0)} \in \mathcal{M}_i$. Когда существует последовательность $\{\boldsymbol{M}_i^{(0)}\}$, удовлетворяющая (23)?

Из теоремы 1 получаем следующее достаточное условие для выполнения равенства (23).

Предложение 1. *Пусть* $\{\mathcal{M}_i\}$ *– какая-либо последовательность множеств, удовлетворяющая предположениям* (13), (14). *Если для последовательности матриц* $\{\boldsymbol{M}_i^{(0)}\}$ *выполняются следующие условия*:

1) $\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - (\boldsymbol{M}_n^{(0)})^{-1} > 0$ *для всех* $\boldsymbol{V}_n \in \mathcal{M}_n$;

2) *имеет место неравенство*

$$\inf_{\boldsymbol{V}_n \in \mathcal{M}_n} \ln f(\boldsymbol{M}_n^{(0)}, \boldsymbol{V}_n) \geqslant o(n), \quad n \to \infty, \tag{24}$$

то последовательность $\{\boldsymbol{M}_i^{(0)}\}$ *с LR-тестами удовлетворяет условию* (23).

Замечание 4. Если последовательность $\{\boldsymbol{M}_i^{(0)}\}$ удовлетворяет условию (24) для последовательности множеств $\{\mathcal{M}_i\}$, то она также удовлетворяет этому условию для последовательности множеств $\{\operatorname{conv}\mathcal{M}_i\}$, где $\operatorname{conv}\mathcal{M}_i$ – минимальное выпуклое множество матриц, содержащее множество $\mathcal{M}_i$. Ясно, что $\mathcal{M}_i \subseteq \operatorname{conv}\mathcal{M}_i$. Множество $\operatorname{conv}\mathcal{M}_i$ иногда называют "выпуклой оболочкой" множества $\mathcal{M}_i$.

Предложение 1 обобщает аналогичный результат в [4, теорема 1] (см. предложение 4 в п. 5.2).

Далее в § 3 приводится важная для нас вспомогательная теорема 3. В § 4 содержатся доказательства теорем 1 и 2, а также некоторые связанные с ними результаты. В § 5 рассматривается аналогичная (1) модель "сигнал + шум", а также случай стационарных стохастических сигналов. Все результаты § 5 являются следствиями теоремы 2. Некоторые примеры приводятся в § 6.

§ 3. Вспомогательная теорема 3

3.1. Простые гипотезы и теорема 3. Для исследования модели (1) рассмотрим сначала более простую модель

$$\begin{aligned} &\mathcal{H}_0\colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n), \\ &\mathcal{H}_1\colon \boldsymbol{y}_n = \boldsymbol{\eta}_n, \quad \boldsymbol{\eta}_n \sim \mathcal{N}(\mathbf{0}, \boldsymbol{M}_n), \end{aligned} \tag{25}$$

соответствующую проверке простых гипотез: $\boldsymbol{I}_n$ против $\boldsymbol{M}_n$. Анализ модели (1) далее будет развитием результатов для модели (25).

Заметим, что модель (25) можно свести к эквивалентной модели с диагональной матрицей $\boldsymbol{M}_n$. Действительно, так как $\boldsymbol{M}_n$ – ковариационная матрица (т.е. симметричная и положительно определенная), то существует ортогональная матрица $\boldsymbol{T}_n$ и диагональная матрица $\boldsymbol{\Lambda}_n$, такие что $\boldsymbol{M}_n = \boldsymbol{T}_n\boldsymbol{\Lambda}_n\boldsymbol{T}_n'$ [11, гл. 4.7-4.9; 12, 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 (25)). Поэтому, умножая обе части (25) на $\boldsymbol{T}_n'$, мы можем свести модель (25) к эквивалентной модели с диагональной матрицей $\boldsymbol{M}_n$.

Обозначим $D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) = D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{M}_n})$ (см. (10)). Пусть $\beta(\alpha, \boldsymbol{M}_n)$ – минимально возможная вероятность ошибки 2-го рода при заданной вероятности ошибки 1-го рода α. Следующая теорема является основным вспомогательным результатом статьи.

Т е о р е м а 3. *Величина $\beta(\alpha, \boldsymbol{M}_n)$, $0<\alpha<1$, удовлетворяет следующим неравенствам:*

$$\ln \beta(\alpha, \boldsymbol{M}_n) \geqslant -\frac{D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)+h(\alpha)}{1-\alpha}, \quad h(\alpha)=-\alpha \ln \alpha-(1-\alpha) \ln (1-\alpha), \tag{26}$$

и

$$\ln \beta(\alpha, \boldsymbol{M}_n) \leqslant -D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)+\mu_0(\alpha, \boldsymbol{M}_n), \tag{27}$$

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

$$\mathbf{P}_{\boldsymbol{I}_n}\Bigl\{\ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)-\mu_0\Bigr\}=\alpha. \tag{28}$$

Отметим, что обе границы (26) и (27) являются чисто аналитическими соотношениями, без каких-либо предельных операций. Нижняя граница (26) и верхняя граница (27) близки друг к другу, если величина $\mu_0(\alpha, \boldsymbol{M}_n)$ значительно меньше, чем $D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)$ (которая обычно имеет порядок n).

Следующий результат дает оценку сверху для величины $\mu_0(\alpha, \boldsymbol{M}_n)$ порядка $n^{1/p}$, $p>1$ (см. доказательство в § 4).

Л е м м а 1. *Пусть для какого-либо $1<p\leqslant 2$ выполнено условие*

$$\sup_n \frac{1}{n} \sum_{i=1}^n \lambda_i^{-p}(\boldsymbol{M}_n) \leqslant C_p<\infty. \tag{29}$$

Тогда для $\mu_0(\alpha, \boldsymbol{M}_n)$ из (27) справедлива оценка сверху

$$\mu_0(\alpha, \boldsymbol{M}_n) \leqslant\left[\frac{2(C_p+1) n}{\alpha}\right]^{1/p}. \tag{30}$$

В частности, если условие (29) выполняется при $p=2$, то тогда (30) дает для $\mu_0(\alpha, \boldsymbol{M}_n)$ оценку сверху порядка $\sqrt{n/\alpha}$. Далее лемма 1 будет использоваться совместно с предположением (14).

3.2. Доказательство теоремы 3. Докажем сначала оценку снизу (26). Пусть $\mathcal{D} \in \mathbb{R}^n$ – область принятия решения в пользу $\boldsymbol{I}_n$, и пусть $\beta=\beta(\mathcal{D})$, $\alpha=\alpha(\mathcal{D})$ – соответствующие вероятности ошибок. Обозначая $p=p_{\boldsymbol{I}_n}$ и $q=p_{\boldsymbol{M}_n}$, имеем

$$\beta=\mathbf{Q}_{\boldsymbol{M}_n}(\mathcal{D})=\int_{\mathcal{D}} p(\boldsymbol{x}) \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x}, \quad \alpha=\mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D}^c), \tag{31}$$

где $\mathcal{D}^c=\mathbb{R}^n \setminus \mathcal{D}$. Так как $\mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D})=1-\alpha$, то рассматривая $\mathbf{P}_{\boldsymbol{I}_n}/(1-\alpha)$ как вероятностное распределение на $\mathcal{D}$ и используя неравенство $\ln \mathbf{E}\,\xi \geqslant \mathbf{E} \ln \xi$, имеем

$$\begin{aligned}
\ln \frac{\beta}{1-\alpha} &= \ln \left[\frac{1}{(1-\alpha)} \int_{\mathcal{D}} p(\boldsymbol{x}) \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x}\right] \geqslant \frac{1}{(1-\alpha)} \int_{\mathcal{D}} p(\boldsymbol{x}) \ln \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x}= \\
&= -\frac{D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)}{1-\alpha}-\frac{1}{(1-\alpha)} \int_{\mathcal{D}^c} p(\boldsymbol{x}) \ln \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x}.
\end{aligned} \tag{32}$$

Так как $\mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D}^c) = \alpha$, то аналогично неравенству в (32) последний член в правой части (32) дает

$$\int\limits_{\mathcal{D}^c} p(\boldsymbol{x}) \ln \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x} \leqslant \alpha \ln \left[\frac{1}{\alpha} \int\limits_{\mathcal{D}^c} q(\boldsymbol{x})\, d\boldsymbol{x} \right] = \alpha \ln \frac{1-\beta}{\alpha} \leqslant \alpha \ln \frac{1}{\alpha}. \tag{33}$$

Поэтому из (32) и (33) получаем

$$\ln \frac{\beta}{1-\alpha} \geqslant -\frac{D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n)}{1-\alpha} - \frac{\alpha}{(1-\alpha)} \ln \frac{1}{\alpha},$$

откуда следует оценка (26).

Для того чтобы доказать оценку сверху (27), выберем величину $\mu > 0$ и определим область принятия решения в пользу $\boldsymbol{I}_n$ как

$$\mathcal{A}_\mu = \left\{ \boldsymbol{x} \in \mathcal{X} : \ \ln \frac{p}{q}(\boldsymbol{x}) \geqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) - \mu \right\}. \tag{34}$$

Обозначим через α_μ и β_μ вероятности ошибки, соответственно, 1-го и 2-го рода для области принятия решения $\mathcal{A}_\mu$. Тогда в силу (34)

$$\beta_\mu = \int\limits_{\mathcal{A}_\mu} p(\boldsymbol{x}) \frac{q}{p}(\boldsymbol{x})\, d\boldsymbol{x} = e^{-D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) + \mu_1}, \tag{35}$$

где $0 \leqslant \mu_1 \leqslant \mu$. Также имеем

$$\alpha_\mu = \mathbf{P}_{\boldsymbol{I}_n} \left\{ \ln \frac{p}{q}(\boldsymbol{x}) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) - \mu \right\} = \mathbf{P}(\eta \geqslant \mu), \tag{36}$$

где

$$\eta = D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) - \ln \frac{p}{q}(\boldsymbol{x}), \quad \mathbf{E}_{\boldsymbol{I}_n} \eta = 0.$$

Наилучшим является выбрать μ так, что $\alpha_\mu = \alpha$. Поэтому определим $\mu = \mu_0(\alpha, \boldsymbol{M}_n)$ формулой (28). Тогда

$$\alpha_{\mu_0(\alpha, \boldsymbol{M}_n)} = \alpha, \tag{37}$$

и в силу (35)

$$\ln \beta_{\mu_0(\alpha, \boldsymbol{M}_n)} \leqslant -D(\boldsymbol{I}_n \,\|\, \boldsymbol{M}_n) + \mu_0(\alpha, \boldsymbol{M}_n),$$

что доказывает оценку сверху (27). ▲

§ 4. Доказательства теорем 1 и 2

Так как $\mathcal{M}_n^{LR}(\boldsymbol{M}_n) \subseteq \mathcal{M}_n(\boldsymbol{M}_n)$, то для доказательства теоремы 1 достаточно получить "внутреннюю границу" для $\mathcal{M}_n^{LR}(\boldsymbol{M}_n)$ и аналогичную "внешнюю границу" для $\mathcal{M}_n(\boldsymbol{M}_n)$.

4.1. "Внутренняя граница" для $\boldsymbol{\mathcal{M}_n^{LR}(M_n)}$. Рассмотрим проверку простой гипотезы $\boldsymbol{I}_n$ против сложной альтернативы $\mathcal{M}_n$. Используем для этого оптимальный LR-тест для матрицы $\boldsymbol{M}_n \in \mathcal{M}_n$ с областью принятия решения $\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $\boldsymbol{I}_n$ (см. (8), (9) и (34)), где $\mu_0 = \mu_0(\alpha, \boldsymbol{M}_n) > 0$ определено в (28). Рассмотрим какую-либо другую матрицу $\boldsymbol{V}_n \in \mathcal{M}_n$ и оценим вероятность ошибки 2-го рода $\beta(\alpha, \boldsymbol{V}_n)$ при условии, что используется область принятия решения $\mathcal{A}_{\mu_0}$. Тогда с помощью (34), (35) имеем

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

$$= e^{-D(\boldsymbol{I}_n \| \boldsymbol{M}_n)+\mu_2} \int\limits_{\mathcal{A}_{\mu_0}} \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) p_{\boldsymbol{I}_n}(\boldsymbol{x})\, d\boldsymbol{x} \leqslant \beta(\alpha, \boldsymbol{M}_n) e^{\mu_2 - \mu_1} \, \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}), \tag{38}$$

где $0 \leqslant \max\{\mu_1, \mu_2\} \leqslant \mu_0(\alpha, \boldsymbol{M}_n)$. В силу предположения (14) и (30) имеем также

$$\mu_0(\alpha, \boldsymbol{M}_n) = O(n^{1/(1+\delta)}) = o(n), \quad n \to \infty. \tag{39}$$

Поэтому если

$$\sup_{\boldsymbol{V}_n \in \mathcal{M}_n^{LR}(\boldsymbol{M}_n)} \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \quad n \to \infty, \tag{40}$$

то из (38)–(40) при $n \to \infty$

$$\sup_{\boldsymbol{V}_n \in \mathcal{M}_n^{LR}(\boldsymbol{M}_n)} \ln \beta(\alpha, \boldsymbol{V}_n) \leqslant \ln \beta(\alpha, \boldsymbol{M}_n) + o(n). \tag{41}$$

4.2. “Внешняя граница” для $\boldsymbol{\mathcal{M}_n(M_n)}$. Пусть $\mathcal{D}_{LR}(\boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $\boldsymbol{I}_n$, и пусть $\beta_{\boldsymbol{M}_n} = \beta_{\boldsymbol{M}_n}(\mathcal{D})$, $\alpha = \alpha(\mathcal{D})$ – соответствующие вероятности ошибок. Тогда, обозначая $p = p_{\boldsymbol{I}_n}$ и $q = p_{\boldsymbol{M}_n}$, аналогично (31) имеем

$$\beta_{\boldsymbol{M}_n} = \mathbf{Q}_{\boldsymbol{M}_n}(\mathcal{D}) = \int\limits_{\mathcal{D}} q(\boldsymbol{x})\, d\boldsymbol{x}, \quad \alpha = \mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D}^c). \tag{42}$$

Рассмотрим какую-либо другую матрицу $\boldsymbol{V}_n \in \mathcal{M}_n(\boldsymbol{M}_n)$. Обозначая $q_1 = p_{\boldsymbol{V}_n}$, для вероятности ошибки 2-го рода $\beta_{\boldsymbol{V}_n} = \beta_{\boldsymbol{V}_n}(\mathcal{D})$ необходимо иметь

$$\beta_{\boldsymbol{V}_n} = \mathbf{Q}_{\boldsymbol{V}_n}(\mathcal{D}) = \int\limits_{\mathcal{D}} q_1(\boldsymbol{x})\, d\boldsymbol{x} \leqslant \beta_{\boldsymbol{M}_n} e^{o(n)}. \tag{43}$$

Для какого-нибудь δ, $0 \leqslant \delta \leqslant 1$, рассмотрим также функцию плотности распределения вероятностей

$$q_\delta(\boldsymbol{x}) = (1-\delta) q(\boldsymbol{x}) + \delta q_1(\boldsymbol{x}) \tag{44}$$

и связанную с ней величину

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

В силу (42), (43) имеем

$$\beta_\delta \leqslant \beta_{\boldsymbol{M}_n}(1 - \delta + \delta e^{o(n)}). \tag{45}$$

Отметим, что плотность распределения вероятностей $q_\delta(\boldsymbol{x})$ соответствует байессовской постановке задачи, в которой альтернативная гипотеза $\mathcal{H}_1$ с вероятностью $1-\delta$ совпадает с $\boldsymbol{M}_n$, а с вероятностью δ – с $\boldsymbol{V}_n$. Соответственно, величина β_δ есть вероятность ошибки 2-го рода.

Аналогично (32), (33) оценим снизу величину β_δ. Сперва аналогично (32) имеем

$$\ln\frac{\beta_\delta}{1-\alpha} = \ln\left[\frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}\right] \geqslant \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} =$$
$$= -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))}{1-\alpha} - \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}^c} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}. \tag{46}$$

Для последнего члена в правой части (46) аналогично (33) 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 аналогично (26) получаем

$$\ln\beta_\delta \geqslant -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) + h(\alpha)}{1-\alpha}. \tag{47}$$

Рассмотрим величину $D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))$ из правой части (47). Обозначая

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

в силу (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

$$D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) = -\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} = D(p(\boldsymbol{x})\,\|\,q(\boldsymbol{x})) + g(\delta), \tag{48}$$

где

$$g(\delta) = -\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})\ln[1-\delta+\delta r(\boldsymbol{x})]\,d\boldsymbol{x}. \tag{49}$$

В силу (45) и (48), (49) для любой функции $o(n)$ необходимо иметь

$$g(\delta) \geqslant -\ln(1-\delta+\delta e^{o(n)}) \quad \text{для всех } 0<\delta\leqslant 1. \tag{50}$$

Заметим, что в силу неравенства $\ln\mathbf{E}\,\xi \geqslant \mathbf{E}\ln\xi$ из (49) 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 (50), необходимо иметь

$$\int\limits_{\mathbb{R}^n}\frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}\,d\boldsymbol{x} \geqslant \frac{1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{51}$$

Так как $\int p(\boldsymbol{x})\,d\boldsymbol{x}=1$, то соотношение (51) эквивалентно условию

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})(r(\boldsymbol{x})-1)}{1-\delta+\delta r(\boldsymbol{x})}\,d\boldsymbol{x} \leqslant \frac{e^{o(n)}-1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{52}$$

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 (52), необходимо иметь

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})r(\boldsymbol{x})}{1+\delta r(\boldsymbol{x})}\,d\boldsymbol{x} \leqslant \frac{e^{o(n)}}{(1-\delta)(1-\delta+\delta e^{o(n)})} \quad \text{для всех } 0<\delta\leqslant 1. \tag{53}$$

Полагая $\delta\downarrow 0$, из (53) получаем необходимое условие

$$\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})r(\boldsymbol{x})\,d\boldsymbol{x} = \mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \tag{54}$$

которое совпадает с (40) для матрицы $\boldsymbol{M}_n$. В случае множества $\mathcal{M}_n(\boldsymbol{M}_n)$ условие (54) должно выполняться для всех $\boldsymbol{V}_n\in\mathcal{M}_n(\boldsymbol{M}_n)$, т.е. необходимо иметь

$$\sup_{\boldsymbol{V}_n\in\mathcal{M}_n(\boldsymbol{M}_n)} \mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \tag{55}$$

откуда следует “внешняя граница” для $\mathcal{M}_n(\boldsymbol{M}_n)$ (см. (17)).

Нам остается выразить условие (54) в терминах матриц $\boldsymbol{M}_n$ и $\boldsymbol{V}_n$.

Лемма 2. *Если $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}>0$, то справедлива формула* (*см.* (15))

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{I}_n+\boldsymbol{V}_n(\boldsymbol{I}_n-\boldsymbol{M}_n^{-1})|^{1/2}} = f^{-1/2}(\boldsymbol{M}_n,\boldsymbol{V}_n). \tag{56}$$

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

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) = \infty. \tag{57}$$

Доказательство. В силу (6)

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

Заметим, что если $\boldsymbol{I}_n+\boldsymbol{A}_n>0$, то для $\boldsymbol{\xi}_n\sim\mathcal{N}(\boldsymbol{0}_n,\boldsymbol{I}_n)$ из [11, гл. 6.9, теорема 3] имеем

$$\mathbf{E}_{\boldsymbol{\xi}_n} e^{-(\boldsymbol{\xi}_n,\boldsymbol{A}_n\boldsymbol{\xi}_n)/2} = \frac{1}{|\boldsymbol{I}_n+\boldsymbol{A}_n|^{1/2}}. \tag{59}$$

В противном случае

$$\mathbf{E}_{\boldsymbol{\xi}_n} e^{-(\boldsymbol{\xi}_n,\boldsymbol{A}_n\boldsymbol{\xi}_n)/2} = \infty. \tag{60}$$

Предположим сначала, что $\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1} > 0$. Тогда из (58) и (59) имеем

$$\mathbf{E}_{\boldsymbol{\xi}_n} \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{V}_n|^{1/2}|\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}|^{1/2}} = \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{I}_n + \boldsymbol{V}_n(\boldsymbol{I}_n - \boldsymbol{M}_n^{-1})|^{1/2}}. \tag{61}$$

Если же матрица $\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ не является положительно определенной, то в силу (58) и (60)

$$\mathbf{E}_{\boldsymbol{\xi}_n} \frac{p_{\boldsymbol{V}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \infty. \tag{62}$$

Из (58), (61) и (62) следует лемма 1. ▲

Поэтому в силу (56) условие (55) эквивалентно соотношению

$$\inf_{\boldsymbol{V}_n \in \mathcal{M}_n(\boldsymbol{M}_n)} \frac{|\boldsymbol{I}_n + \boldsymbol{V}_n(\boldsymbol{I}_n - \boldsymbol{M}_n^{-1})|}{|\boldsymbol{M}_n|} \geqslant e^{o(n)} \tag{63}$$

при условии, что матрица $\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ положительно определена.

Если же матрица $\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ не является положительно определенной, то в силу (57) и (62)

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

и поэтому условие (55) не может быть выполнено.

Определим $\mathcal{M}_n(\boldsymbol{M}_n)$ как максимальное множество, удовлетворяющее условию (63). Это множество совпадает с определением (15)–(17). Поэтому из (41), (55) и (56) следует теорема 1. ▲

4.3. Доказательство теоремы 2. Левую часть формулы (63) можно выразить через собственные значения матриц $\boldsymbol{M}_n$ и $\boldsymbol{V}_n$ следующим образом. Для ковариационной матрицы $\boldsymbol{M}_n$ с собственными значениями $\{\lambda_i\}$ рассмотрим ковариационные матрицы $\boldsymbol{V}_n$, коммутирующие с $\boldsymbol{M}_n$, т.е. $\boldsymbol{M}_n\boldsymbol{V}_n = \boldsymbol{V}_n\boldsymbol{M}_n$. Тогда каждая пара $\boldsymbol{M}_n, \boldsymbol{V}_n$ имеет одно и то же множество собственных векторов $\{\boldsymbol{x}_i\}$ [11, гл. 4.11, теорема 5]. Обозначим через $\{\nu_i\}$ собственные значения матрицы $\boldsymbol{V}_n$. Тогда матрица $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n(\boldsymbol{I}_n - \boldsymbol{M}_n^{-1})$ имеет собственные значения

$$1 + \nu_i - \nu_i/\lambda_i, \quad i = 1, \ldots, n.$$

Поэтому функция $f(\boldsymbol{M}_n, \boldsymbol{V}_n)$ из (15) принимает вид (19), откуда следует результат теоремы 2. ▲

4.4. 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 $\{a_i\}$ [11, гл. 4.7]. Так как $\boldsymbol{T}\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, то квадратичные формы $(\boldsymbol{\xi}_n, \boldsymbol{A}\boldsymbol{\xi}_n)$ и $(\boldsymbol{\xi}_n, \boldsymbol{B}\boldsymbol{\xi}_n)$ имеют одинаковые распределения. Поэтому из формул (7) и (11) имеем

$$\ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{y}_n) \stackrel{d}{=} \frac{1}{2}\bigl[\ln |\boldsymbol{M}_n| + \zeta_n\bigr], \tag{64}$$

где

$$\zeta_n = \bigl(\boldsymbol{y}_n, [\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n]\boldsymbol{y}_n\bigr). \tag{65}$$

Используем следующий результат [15, гл. III.5.15]: пусть $\zeta_1,\dots,\zeta_n$ – независимые случайные величины с $\mathbf{E}\,\zeta_i = 0$, $i=1,\dots,n$. Тогда для любого $1\leqslant p\leqslant 2$

$$\mathbf{E}\Big|\sum_{i=1}^{n}\zeta_i\Big|^p \leqslant 2\sum_{i=1}^{n}\mathbf{E}\,|\zeta_i|^p \tag{66}$$

при условии, что правая часть (66) конечна.

Для оценки величины α_μ из (36) используем неравенство Чебышева и формулы (11) и (64)–(66). Тогда при $1\leqslant p\leqslant 2$ получаем

$$\begin{aligned}
\alpha_\mu &= \mathbf{P}_{\boldsymbol{I}_n}\Big\{\ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant D(\boldsymbol{I}_n\,\|\,\boldsymbol{M}_n)-\mu\Big\} \leqslant \\
&\leqslant \mathbf{P}_{\boldsymbol{I}_n}\Big\{\Big|\ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{M}_n}}(\boldsymbol{x}) - D(\boldsymbol{I}_n\,\|\,\boldsymbol{M}_n)\Big| > \mu\Big\} = \\
&= \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\zeta_n - \sum_{i=1}^{n}\Big(\frac{1}{\lambda_i}-1\Big)\right| > 2\mu\right\} = \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\sum_{i=1}^{n}\Big(\frac{1}{\lambda_i}-1\Big)(\xi_i^2-1)\right| > 2\mu\right\} \leqslant \\
&\leqslant (2\mu)^{-p}\,\mathbf{E}\left|\sum_{i=1}^{n}\Big(\frac{1}{\lambda_i}-1\Big)(\xi_i^2-1)\right|^p \leqslant 2(2\mu)^{-p}\sum_{i=1}^{n}\mathbf{E}\left|\Big(\frac{1}{\lambda_i}-1\Big)(\xi_i^2-1)\right|^p \leqslant \\
&\leqslant 2\mu^{-p}\sum_{i=1}^{n}\left|\frac{1}{\lambda_i}-1\right|^p \leqslant 2(C_p+1)\mu^{-p}n,
\end{aligned} \tag{67}$$

так как $\mathbf{E}\,|\xi_i^2-1|^p \leqslant [\mathbf{E}(\xi_i^2-1)^2]^{p/2} = 2^{p/2}$, $0<p\leqslant 2$. Формула (30) следует из (67) и условия $\alpha_\mu\leqslant\alpha$. ▲

§ 5. Модель "сигнал + шум". Стационарные стохастические сигналы

Наряду с моделью (1) рассмотрим также следующую популярную модель "сигнал + шум":

$$\begin{aligned}
&\mathcal{H}_0\colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n\sim\mathcal{N}(\mathbf{0},\boldsymbol{I}_n),\\
&\mathcal{H}_1\colon \boldsymbol{y}_n = \boldsymbol{\xi}_n+\boldsymbol{s}_n, \quad \boldsymbol{s}_n\sim\mathcal{N}(\mathbf{0},\boldsymbol{S}_n), \quad \boldsymbol{S}_n\in\mathcal{S}_n,
\end{aligned} \tag{68}$$

где $\boldsymbol{s}_n$ – не зависящий от $\boldsymbol{\xi}_n$ гауссовский случайный вектор ("стохастический сигнал") с $\boldsymbol{s}_n\sim\mathcal{N}(\mathbf{0},\boldsymbol{S}_n)$, а $\mathcal{S}_n$ – заданное множество ковариационных матриц $\boldsymbol{S}_n$.

5.1. Модель (68). Так как модель (68) является частным случаем модели (1), то к ней можно применить теоремы 1 и 2, в которых следует положить $\boldsymbol{M}_n = \boldsymbol{S}_n+\boldsymbol{I}_n$. Обозначим через $\{\mu_i(\boldsymbol{S}_n)\}$) собственные значения (все они положительны) ковариационной матрицы $\boldsymbol{S}_n$. Тогда $\mu_i(\boldsymbol{S}_n) = \lambda_i(\boldsymbol{M}_n)-1$, $i=1,\dots,n$. Так как $\lambda_i(\boldsymbol{M}_n)>1$, $i=1,\dots,n$, то предположение (14) выполняется автоматически. Предположение (13) принимает следующий вид:

III. Для последовательности множеств $\mathcal{S}_n$ существует положительный предел

$$\lim_{n\to\infty}\frac{1}{n}\inf_{\boldsymbol{S}_n\in\mathcal{S}_n}\sum_{i=1}^{n}\left[\ln(\mu_i(\boldsymbol{S}_n)+1)+\frac{1}{\mu_i(\boldsymbol{S}_n)+1}-1\right] > 0. \tag{69}$$

Вместо функции $f(\boldsymbol{M}_n,\boldsymbol{V}_n)$ из (15) введем ее аналог $t(\boldsymbol{S}_n,\boldsymbol{V}_n)$. Для любых $\boldsymbol{S}_n,\boldsymbol{V}_n\in\mathcal{C}_n$, таких что $\boldsymbol{I}_n+(\boldsymbol{I}_n+\boldsymbol{V}_n)^{-1}-(\boldsymbol{I}_n+\boldsymbol{S}_n)^{-1}>0$, определим функцию

$$t(\boldsymbol{S}_n,\boldsymbol{V}_n) = \frac{|\boldsymbol{I}_n+\boldsymbol{S}_n+(\boldsymbol{V}_n-\boldsymbol{S}_n)\boldsymbol{S}_n(\boldsymbol{I}_n+\boldsymbol{S}_n)^{-1}|}{|\boldsymbol{I}_n+\boldsymbol{S}_n|}. \tag{70}$$

В определении (70) использовалась простая формула

$$\boldsymbol{I}_n - (\boldsymbol{I}_n + \boldsymbol{S}_n)^{-1} = \boldsymbol{S}_n(\boldsymbol{I}_n + \boldsymbol{S}_n)^{-1}.$$

В качестве прямого следствия из теоремы 1 получаем

Предложение 2. *Если для модели* (68) *выполнено предположение* (69), *то наибольшее множество* $\mathcal{S}_n(\boldsymbol{S}_n)$, *удовлетворяющее асимптотическому равенству*

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{S}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{S}_n), \tag{71}$$

имеет вид $(n \to \infty)$

$$\mathcal{S}_n(\boldsymbol{S}_n) = \Big\{\boldsymbol{V}_n : \ \boldsymbol{I}_n + (\boldsymbol{I}_n + \boldsymbol{V}_n)^{-1} - (\boldsymbol{I}_n + \boldsymbol{S}_n)^{-1} > 0,$$
$$\inf_{\boldsymbol{V}_n \in \mathcal{S}_n(\boldsymbol{S}_n)} \ln t(\boldsymbol{S}_n, \boldsymbol{V}_n) \geqslant o(n)\Big\}. \tag{72}$$

Ясно, что множество $\mathcal{S}_n(\boldsymbol{S}_n)$ является выпуклым по $\boldsymbol{V}_n$.

Упростим предложение 2 с помощью теоремы 2 следующим образом. Для матрицы $\boldsymbol{S}_n$ с собственными значениями $\{\mu_i\}$ рассмотрим в (72) только матрицы $\boldsymbol{V}_n$, коммутирующие с $\boldsymbol{S}_n$. Обозначим через $\{\nu_i\}$ собственные значения матрицы $\boldsymbol{V}_n$. Тогда аналогично (19) получаем

$$t(\boldsymbol{S}_n, \boldsymbol{V}_n) = \prod_{i=1}^{n} \left[1 + \frac{(\nu_i - \mu_i)\mu_i}{(1+\mu_i)^2}\right]. \tag{73}$$

Аналогично (18) для матрицы $\boldsymbol{S}_n$ введем выпуклое множество $\mathcal{C}_{\boldsymbol{S}_n}$ ковариационных матриц $\boldsymbol{V}_n$, коммутирующих с $\boldsymbol{S}_n$:

$$\mathcal{C}_{\boldsymbol{S}_n} = \{\boldsymbol{V}_n : \ \boldsymbol{S}_n \boldsymbol{V}_n = \boldsymbol{V}_n \boldsymbol{S}_n\}. \tag{74}$$

Введем также следующее подмножество этого множества (см. (73)):

$$\mathcal{V}_n^{(2)}(\boldsymbol{S}_n) = \Big\{\boldsymbol{V}_n \in \mathcal{C}_{\boldsymbol{S}_n} : \inf_{\boldsymbol{V}_n \in \mathcal{V}_n^{(2)}(\boldsymbol{S}_n)} \ln t(\boldsymbol{S}_n, \boldsymbol{V}_n) \geqslant o(n)\Big\}. \tag{75}$$

Множество $\mathcal{V}_n^{(2)}(\boldsymbol{S}_n)$ является выпуклым по $\boldsymbol{V}_n$, так как функция $\ln z$ выпукла по $z > 0$. Тогда аналогично теореме 2 получаем следующую "внутреннюю границу" $\mathcal{V}_n^{(2)}(\boldsymbol{S}_n)$ для $\mathcal{S}_n(\boldsymbol{S}_n)$.

Предложение 3. *Если для модели* (68) *выполнено предположение* (69), *то для наибольшего множества* $\mathcal{S}_n(\boldsymbol{S}_n)$, *удовлетворяющего формуле* (71), *имеем*

$$\mathcal{V}_n^{(2)}(\boldsymbol{S}_n) \subseteq \mathcal{S}_n(\boldsymbol{S}_n),$$

где множество $\mathcal{V}_n^{(2)}(\boldsymbol{S}_n)$ *определено в* (74), (75).

5.2. Стационарные стохастические сигналы. Рассмотрим модель (68), где $\boldsymbol{s}_n \sim \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{S}_n)$ – стационарный в широком смысле гауссовский случайный процесс с нулевым средним и спектральной плотностью $f_{\boldsymbol{S}_n}(\omega)$, $\omega \in [-\pi, \pi]$. Эта модель рассматривалась в работе [4]. Покажем, что основной результат [4, теорема 1] следует из предложения 3. Для этого рассмотрим какое-либо множество $\mathcal{K}_n$ ковариационных матриц $\boldsymbol{K}_n$, содержащее $\boldsymbol{S}_n$. Ограничимся только случаем, когда ковариационные матрицы $\boldsymbol{K}_n$ со спектральными плотностями $f_{\boldsymbol{K}_n}(\omega)$ коммутируют с $\boldsymbol{S}_n$. Тогда можно применить предложение 3, выраженное через спектральные плотности.

С помощью теоремы Сегё [16, гл. 5.2, теорема; 17, теорема 4.1] заменим предположение (69) его аналогом:

IV. Для последовательности множеств $\mathcal{K}_n$ существует положительный предел

$$\lim_{n\to\infty} \inf_{\boldsymbol{K}_n \in \mathcal{K}_n} \int\limits_{-\pi}^{\pi} \left[\ln(f_{\boldsymbol{K}_n}(\omega)+1) + \frac{1}{f_{\boldsymbol{K}_n}(\omega)+1} - 1 \right] d\omega > 0. \tag{76}$$

Для заданной спектральной плотности $f_{\boldsymbol{S}_n}(\omega)$ обозначим через $\mathcal{F}_n(f_{\boldsymbol{S}_n})$ наибольшее множество спектральных плотностей, удовлетворяющее следующему аналогу равенства (5):

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(f_{\boldsymbol{S}_n}) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{F}_n(f_{\boldsymbol{S}_n})). \tag{77}$$

Другими словами, при заданной вероятности ошибки 1-го рода α множество $\mathcal{F}_n(f_{\boldsymbol{S}_n})$ есть максимальное множество спектральных плотностей, которое можно заменить одной спектральной плотностью $f_{\boldsymbol{S}_n}$ (без потери для экспоненты вероятности ошибки 2-го рода $\beta(\mathcal{F}_n(f_{\boldsymbol{S}_n}))$).

Для того чтобы описать “внутреннюю границу” для множества $\mathcal{F}_n(f_{\boldsymbol{S}_n})$, введем следующий функционал:

$$f(\boldsymbol{S}_n, \boldsymbol{K}_n) = \int\limits_{-\pi}^{\pi} \ln\left[1 + \frac{f_{\boldsymbol{S}_n}(\omega)[f_{\boldsymbol{K}_n}(\omega) - f_{\boldsymbol{S}_n}(\omega)]}{(1+f_{\boldsymbol{S}_n}(\omega))^2}\right] d\omega. \tag{78}$$

Функционал $f(\boldsymbol{S}_n, \boldsymbol{K}_n)$ является аналогом функционала $\ln t(\boldsymbol{S}_n, \boldsymbol{V}_n)$ из (73), который следует в силу теоремы Сегё. Введем также множество (при $n \to \infty$)

$$\mathcal{F}_n^{(1)}(f_{\boldsymbol{S}_n}) = \Big\{ f_{\boldsymbol{K}_n}(\omega) : \inf_{f_{\boldsymbol{K}_n}(\omega) \in \mathcal{F}_n^{(1)}(f_{\boldsymbol{S}_n})} f(\boldsymbol{S}_n, \boldsymbol{K}_n) \geqslant o(1) \Big\}. \tag{79}$$

Множество $\mathcal{F}_n^{(1)}(f_{\boldsymbol{S}_n})$ является аналогом множества $\mathcal{V}_n^{(2)}(\boldsymbol{S}_n)$ из (75).

Как прямое следствие предложения 3 получаем

П р е д л о ж е н и е 4. *Если для модели* (68) *выполнено предположение* (76), *то для наибольшего множества* $\mathcal{F}_n(f_{\boldsymbol{S}_n})$, *для которого выполняется формула* (77), *имеем*

$$\mathcal{F}_n^{(1)}(\boldsymbol{S}_n) \subseteq \mathcal{F}_n(f_{\boldsymbol{S}_n}), \tag{80}$$

где множество $\mathcal{F}_n^{(1)}(\boldsymbol{S}_n)$ *определено в* (78), (79).

Множество $\mathcal{F}_n^{(1)}(\boldsymbol{S}_n)$ является выпуклым по $f_{\boldsymbol{K}_n}(\omega)$, так как функция $\ln z$ выпукла по $z > 0$. Другими словами, если $f_{\boldsymbol{K}_n^{(0)}}(\omega), f_{\boldsymbol{K}_n^{(1)}}(\omega) \in \mathcal{F}_n^{(1)}(\boldsymbol{S}_n)$, то тогда

$$c f_{\boldsymbol{K}_n^{(0)}}(\omega) + (1-c) f_{\boldsymbol{K}_n^{(1)}}(\omega) \in \mathcal{F}_n^{(1)}(\boldsymbol{S}_n) \quad \text{для любого } 0 \leqslant c \leqslant 1.$$

Замечание 5. Множество $\mathcal{F}_n^{(1)}(\boldsymbol{S}_n)$ из (78), (79) (при $n \to \infty$) использовалось в [4, теорема 1].

§ 6. Примеры

П р и м е р 1. В некоторых симметричных случаях множества $\mathcal{M}_n(\boldsymbol{M}_n)$ из (17) и $\mathcal{V}_n^{(0)}(\boldsymbol{M}_n)$ из (20) могут совпадать. Например, пусть в модели (1) $\boldsymbol{M}_n = c\boldsymbol{I}_n$, где

$c > 0$. Тогда для любой $\boldsymbol{V}_n$ матрицы $\boldsymbol{M}_n$ и $\boldsymbol{V}_n$ коммутируют, т.е. $\boldsymbol{M}_n\boldsymbol{V}_n = \boldsymbol{V}_n\boldsymbol{M}_n$. В этом случае теоремы 1 и 2 дают

$$\mathcal{M}_n(\boldsymbol{M}_n) = \mathcal{V}_n^{(0)}(\boldsymbol{M}_n).$$

Пример 2. Пусть в модели (68) сигнальный процесс $\{s_i\}$ имеет инвариантную во времени структуру

$$\begin{aligned} &s_{i+1} = as_i + \sqrt{1-a^2}\,u_i, \quad i = 1,\ldots,n,\\ &s_1 \sim \mathcal{N}(0,1), \quad u_i \sim \mathcal{N}(0,1), \end{aligned} \tag{81}$$

где a – известная величина (параметр скорости обновления), такая что $0 \leqslant a < 1$. Предположим, что шумовой процесс $\{u_i\}$ не зависит от зашумленных наблюдений $\{\xi_i\}$, а начальное состояние s_1 не зависит от $\{u_i\}$ для всех i. Сигнальная последовательность $\{s_i\}$ есть процесс авторегрессии первого порядка $AR(1)$ и является стационарным процессом. Этот пример является главным в [4, 18]. Основной результат для этого примера также следует из предложения 3. Действительно, спектр наблюдаемого процесса (81) при гипотезах $\mathcal{H}_0$ и $\mathcal{H}_1$ описывается формулами, соответственно,

$$f^{(0)}(\omega) = 1, \quad f^{(1)}(\omega) = 1 + f_{\boldsymbol{S}_n}(\omega), \quad \omega \in [-\pi,\pi],$$

где спектр сигнала дается пуассоновским ядром

$$f_{\boldsymbol{S}_n}(\omega) = \frac{1-a^2}{1-2a\cos\omega + a^2}.$$

Тогда множество $\mathcal{F}_n^{(1)}(\boldsymbol{S}_n)$ описывается формулами (78), (79). Аналогичные (но более громоздкие результаты) можно получить и для процессов авторегрессии $AR(k)$ порядка $k > 1$.

Автор благодарит рецензента за конструктивные критические замечания, улучшившие статью.

СПИСОК ЛИТЕРАТУРЫ

1. *Wald A.* Statistical Decision Functions. New York: Wiley, 1950.
2. *Леман Э.* Проверка статистических гипотез. М.: Наука, 1979.
3. *Poor H.V.* An Introduction to Signal Detection and Estimation. New York: Springer-Verlag, 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. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967.
7. *Stein C.* Information and Comparison of Experiments. Unpublished.
8. *Chernoff H.* Large-Sample Theory: Parametric Case // Ann. Math. Statist. 1956. V. 27. № 1. P. 1–22. https://doi.org/10.1214/aoms/1177728347
9. *Cover T.M., Thomas J.A.* Elements of Information Theory. New York: Wiley, 1991.
10. *Burnashev M.V.* On Neyman-Pearson Minimax Detection of Poisson Process Intensity // Stat. Inference Stoch. Process. 2021. V. 21. № 1. P. 211–221. https://doi.org/10.1007/s11203-020-09230-4
11. *Беллман Р.* Введение в теорию матриц. М.: Наука, 1976.
12. *Horn R.A., Johnson C.R.* Matrix Analysis. Cambridge: Cabmridge Univ. Press, 1985.

13. *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

14. *Бурнашев М.В.* Две теоремы сравнения распределений гауссовских квадратичных форм // Пробл. передачи информ. 2017. Т. 53. № 3. С. 3–15. http://mi.mathnet.ru/ppi2238

15. *Петров В.В.* Суммы независимых случайных величин. М.: Наука, 1972.

16. *Гренандер У., Сегё Г.* Теплицевы формы и их приложения. М.: Изд-во иностр. лит., 1961.

17. *Gray R.M.* On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices // IEEE Trans. Inform. Theory. 1972. V. 18. № 6. P. 725–730. https://doi.org/10.1109/TIT.1972.1054924

18. *Sung Y., Tong L., Poor H.V.* Neyman–Pearson Detection of Gauss–Markov Signals in Noise: Closed-Form Error Exponent and Properties // IEEE Trans. Inform. Theory. 2006. V. 52. № 4. P. 1354–1365. https://doi.org/10.1109/TIT.2006.871599

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

Поступила в редакцию
15.04.2021
После доработки
16.06.2021
Принята к публикации
29.06.2021