

© 2025 г. Я. БЕКРИ, д-р философии (yannisbekri@hotmail.com)
(Университет Гренобль-Альпы, Сен-Мартен-д'Эр, Франция),
А. НЕМИРОВСКИЙ, д-р физ.-мат. наук (nemirovs@isye.gatech.edu)
(Технологический институт Джорджии, Атланта, США),
А. ЮДИЦКИЙ, канд. техн. наук (anatoli.juditsky@univ-grenoble-alpes.fr)
(Университет Гренобль-Альпы, Сен-Мартен-д'Эр, Франция)

О НАДЕЖНОМ ВОССТАНОВЛЕНИИ СИГНАЛОВ ПО НЕПРЯМЫМ НАБЛЮДЕНИЯМ¹

Рассматривается следующая *линейная обратная задача с неопределенностью*. При имеющемся наблюдении $\omega = Ax_* + \zeta$, где $A \in \mathbb{R}^{m \times p}$ и $\zeta \in \mathbb{R}^m$ – шум наблюдения, требуется восстановить неизвестный сигнал x_* , который принадлежит заданному выпуклому множеству $\mathcal{X} \subset \mathbb{R}^n$. В отличие от стандартной постановки такой задачи предполагается, что модельный шум ζ является “искаженным”, т.е. содержит детерминированную, плотную или сингулярную, компоненту. В частности, ζ допускает разложение $\zeta = N\nu_* + \xi$, где ξ – случайный шум, а $N\nu_*$ – “адверсарияльное загрязнение” с известным множеством $\mathcal{N} \subset \mathbb{R}^n$, т.е. $\nu_* \in \mathcal{N}$ и $N \in \mathbb{R}^{m \times n}$. Рассматриваются две постановки, в которых $\mathcal{N}$ является либо выпуклым ограниченным множеством, либо множеством разреженных векторов (с не более чем s ненулевыми элементами). Исследуются свойства “устойчивых” по отношению к загрязнению *полиэдральных оценок* – специального класса нелинейных оценок, представленных в [1, 2]. Показывается, как оценки такого рода с гарантированным качеством могут быть построены в ситуации, когда множество сигналов является *элитопом* (по существу, симметричным выпуклым множеством, ограниченным квадратичными поверхностями), с помощью эффективных процедур выпуклой оптимизации.

Ключевые слова: робастное оценивание, линейные обратные задачи с загрязненными наблюдениями, оценивание сигнала в сингулярном шуме.

DOI: 10.31857/S0005231025080025, EDN: UTATKM

1. Постановка задачи и цели исследования

1.1. Введение

Проблема “робастного оценивания” находится в центре внимания классической статистической литературы с момента появления термина в 1950-х гг. Невозможно дать исчерпывающий обзор существующих публикаций по робастному оцениванию, и авторы не пытаются сделать это ниже; чтобы получить базовое представление о “классической” литературе, читатель может

¹ Работа выполнена при финансовой поддержке Мультидисциплинарного института искусственного интеллекта (МИАИ), @ Гренобль-Альпы (грант ANR-19-P3IA-0003).

обратиться к ранним ссылкам в [3], основополагающей монографии [4] или недавнему обзору [5].²

Настоящая статья посвящена робастному оцениванию сигнала по “непрямым” линейным наблюдениям. В частности, целью является восстановление линейного образа $w_* = Bx_*$ неизвестного сигнала x_* , который принадлежит заданному выпуклому множеству $\mathcal{X} \subset \mathbf{R}^p$, при известных $B \in \mathbf{R}^{q \times p}$, $A \in \mathbf{R}^{m \times p}$ и имеющемуся зашумленному наблюдению

$$(1) \quad \omega = Ax_* + \eta_* + \xi \in \mathbf{R}^m$$

сигнала x_* , возмущенного смешанным (случайным и детерминированным) шумом $\xi + \eta_*$. Здесь ξ – случайная составляющая шума, а η_* – “адверсариальный” детерминированный шум. В последнее время эта проблема привлекает большое внимание в контексте робастного восстановления разреженного сигнала x_* (с не более чем $s \ll p$ ненулевыми элементами). В частности, разреженная регрессия с акцентом на помеху с загрязнением исследована в [12–16]; методы, основанные на введении штрафов для вектора выбросов, изучены в [17, 18], см. также [19, 20]. Современное состояние этой области описано в монографии [21].

В данной работе акцент сделан на существенно отличных от стандартных предположениях о структуре неизвестного сигнала x_* и о загрязнении η_* , что исключает прямое сравнение с приведенными результатами. А именно, предполагается, что множество $\mathcal{X}$ сигналов является *элитопом* – выпуклым, компактным и симметричным относительно начала координат подмножеством $\mathbf{R}^p$, ограниченным квадратичными поверхностями.³ Интерес авторов к элитопам обусловлен тем, что эти множества сигналов “удобны” для решения задачи оценки неизвестного сигнала x_* по наблюдениям (1) в гауссовском шуме (при отсутствии загрязнения) ($\eta_* = 0$, $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$). В частности, рассмотрим линейную оценку $\hat{w}_{\text{lin}}(\omega) = G_{\text{lin}}^T \omega$ и *полиэдральную оценку* $\hat{w}_{\text{poly}}(\omega) = B\hat{x}_{\text{poly}}(\omega)$, где

$$\hat{x}_{\text{poly}}(\omega) \in \underset{x \in \mathcal{X}}{\text{Argmin}} \|G_{\text{poly}}^T (Ax - \omega)\|_{\infty},$$

для сигнала w_* . Пусть $\mathcal{X}$ – элитоп, а ошибка оценки измеряется в коэлитопической норме $\|\cdot\|$ (т.е. такой, что единичный шар $\mathcal{B}_*$ нормы $\|\cdot\|_*$, сопряженной к $\|\cdot\|$, является элитопом). В этой ситуации можно указать (ср. [1, 2, 22]) эффективно вычислимые *контрастные матрицы* $G_{\text{lin}} \in \mathbf{R}^{m \times q}$ и $G_{\text{poly}} \in \mathbf{R}^{m \times m}$, при которых оценки $\hat{w}_{\text{lin}}(\cdot)$ и $\hat{w}_{\text{poly}}(\cdot)$ становятся почти минимаксными.

² Важным вкладом в развитие робастной статистики стали работы Б. Поляка и Я. Цыпкина по робастным алгоритмам стохастической оптимизации и идентификации систем, см. [6–11].

³ См. [2, раздел 4.2.1] или раздел 1.3 ниже; простейшим примером элитопы является пересечение конечного семейства эллипсоидов/эллиптических цилиндров с общим центром.

Предположим, что возмущение η_* имеет особую структуру $\eta_* = N\nu_*$, где $\nu_* \in \mathcal{N}$, “множество помех” $\mathcal{N} \subset \mathbf{R}^n$ известно, а матрица N размерности $m \times n$ задана. Рассмотрим два типа предположений относительно $\mathcal{N}$: $\mathcal{N}$ является либо 1) (непустым) компактным и выпуклым множеством, либо 2) множеством разреженных возмущений (с не более чем $s \leq n$ ненулевыми элементами). Внимание авторов сосредоточено на построении и анализе эффективности полиэдральной оценки $\hat{w}(\omega)$ для $w_* = Bx_*$, робастной по отношению к загрязняющему сигналу, при этом решение задачи в первой постановке позволяет построить нетривиальное решение для второй.

Хотелось бы подчеркнуть главную особенность предлагаемого авторами подхода: в данной работе A , B и N – произвольные матрицы соответствующей размерности, а $\mathcal{X}$ и $\mathcal{N}$ – достаточно общие множества. Как следствие, здесь используется “операционный” подход⁴, предложенный в [23] и развитый в [1, 2, 24, 25]. В рамках этого подхода и оценки, и их риски являются результатом эффективного вычисления, а не явных аналитических выкладок, которые попросту невозможны в данных обстоятельствах. При этом термин “эффективный” является существенным, а также обуславливает принципиальные ограничения последующих результатов. Прежде всего, он накладывает ограничения на структуру интересующего авторов множества сигналов и на норму, в которой измеряются ошибки оценивания. Как показано в [2], максимум квадратичной формы по эллитопу допускает достаточно точную эффективно вычислимую верхнюю границу, что приводит к качественным границам для риска линейных и полиэдральных оценок, когда множество сигналов является эллитопом. Более того, если в случае выпуклого компактного множества $\mathcal{N}$ загрязнений построение границ риска для полиэдральной оценки $\hat{w}_G(\omega)$, связанной с *заданной* матрицей контрастов G , возможно при довольно слабых предположениях о множестве помех $\mathcal{N}$ (по существу, эффективной вычислимости⁵ этого множества достаточно), то фундаментальная задача *контрастного синтеза* (минимизация этих границ по матрицам контрастов) допускает эффективно вычислимое решение только тогда, когда $\mathcal{N}$ является либо эллитопом, либо “коэллитопом” (полярой к эллитопу).

Здесь хотелось бы упомянуть еще одно направление исследований проблемы оценки сигнала x_* по наблюдениям (1) при детерминированном возмущении (случай $\xi = 0$): стандартную задачу оптимального восстановления [27, 28] и гарантированное оценивание в динамических системах при неслучайных ограниченных возмущениях [29–35]. Настоящая работа может рассматриваться как попытка распространить соответствующие методы на случай, когда присутствуют как детерминированные, так и случайные шумы наблюдения.

⁴ В отличие от классического “описательного” подхода к решению задачи оценивания путем выведения аналитического выражения для риска оценки-кандидата и последующей оптимизации ее параметров.

⁵ Для большинства практических целей эффективная вычислимость означает, что ограничения на множество можно моделировать с помощью CVX [26]. Что это значит на самом деле? Краткий ответ читатель найдет в [2, приложение А].

Структура статьи. В разделе 1.2 даны точная постановка рассматриваемой задачи оценивания и необходимые сведения для последующего анализа. В разделе 2 представлены анализ и построение полиэдральной оценки для случая ограниченного загрязнения. Затем в разделе 3 описано применение предложенного подхода в случае (неограниченного) сингулярного загрязнения с использованием разреженной модели вектора помех. В Приложении 2 приведены некоторые результаты по ℓ_1 -восстановлению, использованные в статье.

Обозначения. В дальнейшем отношения порядка между векторами понимаются поэлементно; например, запись $t \geq t'$ для $t, t' \in \mathbf{R}^n$ означает, что вектор $t - t'$ имеет неотрицательные элементы. Записи $[A; B]$ и $[A, B]$ означают вертикальную и горизонтальную конкатенацию матриц A и B соответствующей размерности. Обозначим через $\mathbf{S}^m$ пространство симметричных матриц размерности $m \times m$, а через $\mathbf{S}_+^m$ – конус неотрицательно определенных матриц в $\mathbf{S}^m$; запись $A \succeq B$ ($A \succ B$) означает, что матрица $A - B$ неотрицательно определена (соответственно положительно определена). В дальнейшем

$$\phi_Z(\zeta) := \max_{z \in Z} \zeta^T z$$

представляет собой опорную функцию непустого компактного множества $Z \subset \mathbf{R}^N$; $\|G\|$ обозначает максимальную евклидову норму столбцов матрицы G .

1.2. Постановка задачи

Интересующая авторов задача оценивания состоит в следующем.

Считается доступным наблюдение (ср. (1))

$$(2) \quad \omega = Ax_* + N\nu_* + \xi \in \mathbf{R}^m,$$

где:

- $N \in \mathbf{R}^{m \times n}$ и $A \in \mathbf{R}^{m \times p}$ – заданные матрицы,
- $\nu_* \in \mathbf{R}^n$ – неизвестный сигнал помехи, $\nu_* \in \mathcal{N}$ (известное подмножество $\mathbf{R}^n$),
- x_* – неизвестный искомый сигнал, принадлежащий заданному выпуклому компактному множеству $\mathcal{X} \subset \mathbf{R}^p$, симметричному относительно начала координат,
- $\xi \sim P_\xi$ – случайный шум наблюдения.

По наблюдению ω требуется восстановить линейный образ $w_* = Bx_*$ сигнала x_* , где B – заданная матрица размерности $q \times p$.

При заданном $\epsilon \in (0, 1)$ будем оценивать качество восстановления $\hat{w}(\cdot)$ его ϵ -риском⁶

$$\text{Risk}_\epsilon[\hat{w}] = \inf \left\{ \rho : \text{Prob}_{\xi \sim P_\xi} \{ \xi : \|Bx_* - \hat{w}(\omega)\| > \rho \} \leq \epsilon \quad \forall (\nu_* \in \mathcal{N} \text{ и } x_* \in \mathcal{X}) \right\},$$

где $\|\cdot\|$ – заданная норма.

⁶ Помимо ϵ и самой оценки, ϵ -риск оценки зависит от “параметров” $\|\cdot\|$, $\mathcal{X}$ и $\mathcal{N}$; эти элементы ясны из контекста и поэтому не упоминаются в обозначении $\text{Risk}_\epsilon[\cdot]$.

Предположение о шуме наблюдения. Шум наблюдения ξ предполагается субгауссовским с параметрами $(0, \sigma^2 I_m)$, т.е.

$$\mathbf{E} \left\{ e^{h^T \xi} \right\} \leq \exp \left(\frac{1}{2} \sigma^2 \|h\|_2^2 \right).$$

1.3. Эллитопы

Геометрия множества сигналов $\mathcal{X}$ и нормы $\|\cdot\|$ играет важную роль в анализе риска полиэдральной оценки. В дальнейшем ограничимся случаем, когда $\mathcal{X}$ и *поляра* $\mathcal{B}_*$ единичного шара нормы $\|\cdot\|$ являются *эллитопами*.

По определению [2, 22] *базовый эллитоп* в $\mathbf{R}^n$ – это множество вида

$$(3) \quad \mathcal{X} = \{x \in \mathbf{R}^n : \exists, t \in \mathcal{T} : x^T T_\ell x \leq t_\ell, \ell \leq L\},$$

где $T_\ell \in \mathbf{S}^k$, $T_\ell \succeq 0$, $\sum_\ell T_\ell \succ 0$, а $\mathcal{T} \subset \mathbf{R}_+^L$ – выпуклое компактное множество с непустой внутренностью, обладающее свойством монотонности: при $0 \leq t' \leq t \in \mathcal{T}$ имеем $t' \in \mathcal{T}$. Любой эллитоп – это образ базового эллитопа при некотором линейном отображении. Будем называть L *эллитопической размерностью* $\mathcal{X}$.

Очевидно, что каждый эллитоп является выпуклым компактным множеством, симметричным относительно начала координат; кроме того, *базовый эллитоп* имеет непустую внутренность.

Примеры.

А. Ограниченное пересечение $\mathcal{X}$ L эллипсоидов/эллиптических цилиндров с центром в начале координат $\{x \in \mathbf{R}^n : x^T T_\ell x \leq 1\}$ [$T_\ell \succeq 0$] является базовым эллитопом:

$$\mathcal{X} = \{x \in \mathbf{R}^n : \exists t \in \mathcal{T} := [0, 1]^L : x^T T_\ell x \leq t_\ell, \ell \leq L\}.$$

В частности, единичный гиперкуб $\{x \in \mathbf{R}^n : \|x\|_\infty \leq 1\}$ является базовым эллитопом.

Б. Единичный шар $\|\cdot\|_p$ -нормы, где $p \in [2, \infty]$, является базовым эллитопом в $\mathbf{R}^n$:

$$\{x \in \mathbf{R}^n : \|x\|_p \leq 1\} = \left\{ x : \exists t \in \mathcal{T} = \{t \in \mathbf{R}_+^n, \|t\|_{p/2} \leq 1\} : \underbrace{x_\ell^2}_{x^T T_\ell x} \leq t_\ell, \ell \leq n \right\}.$$

Эллитопы допускают полностью алгоритмическое “исчисление”: это семейство замкнуто относительно основных операций, сохраняющих выпуклость и симметрию относительно начала координат, например операций конечных пересечений, взятия линейного образа, обратного образа при линейном отображении, прямых произведений и арифметического суммирования (подробнее см. [2, раздел 4.6]).

Основное предположение. В дальнейшем считаем, что множество сигналов $\mathcal{X}$ и поляра $\mathcal{B}_*$ единичного шара нормы $\|\cdot\|$ являются *базовыми эллитопами*:⁷

$$(4a) \quad \mathcal{X} = \{x \in \mathbf{R}^n : \exists t \in \mathcal{T} : x^T T_k x \leq t_k, k \leq K\},$$

$$(4b) \quad \mathcal{B}_* = \{y \in \mathbf{R}^q : \exists s \in \mathcal{S} : y^T S_\ell y \leq s_\ell, \ell \leq L\},$$

где $\mathcal{T} \subset \mathbf{R}_+^K$, $\mathcal{S} \subset \mathbf{R}_+^L$ – монотонные, выпуклые и компактные множества с непустыми внутренностями, $T_k \succeq 0$, $\sum_k T_k \succ 0$, $S_\ell \succeq 0$ и $\sum_\ell S_\ell \succ 0$.

2. Случай ограниченной помехи

В этом разделе рассмотрим случай *ограниченной* неопределенности. В частности, считаем, что $\mathcal{N} \subset \mathbf{R}^n$ – выпуклое и компактное множество, симметричное относительно начала координат. Определим $\pi(\cdot)$ как полунорму в $\mathbf{R}^m$ вида

$$\pi(h) = \sup_u \{(Nu)^T h : u \in \mathcal{N}\}.$$

2.1. Вычисление границ ϵ -риска полиэдральной оценки

В этом разделе полиэдральная оценка задается матрицей контрастов G размерности $m \times I$ следующим образом: при имеющемся наблюдении ω (см. (2)) найдем оптимальное решение $\hat{x}_G(\omega)$ (очевидно разрешимой) задачи оптимизации

$$(5) \quad \min_{x, \nu} \{\|G^T(Ax + N\nu - \omega)\|_\infty : x \in \mathcal{X}, \nu \in \mathcal{N}\}.$$

При заданной матрице $G = [g_1, \dots, g_I]$ размерности $m \times I$ пусть $\Xi_\epsilon[G]$ – множество всех реализаций ξ , удовлетворяющих условию

$$(6) \quad |[g_i^T \xi]_i| \leq \underbrace{\sigma \sqrt{2 \ln [2I/\epsilon]}}_{=:\kappa(\epsilon)} \|g_i\|_2, \quad \forall i \leq I.$$

Заметим, что

$$(7) \quad \text{Prob}_{\xi \sim \mathcal{SG}(0, \sigma^2 I_m)} \{\xi \notin \Xi_\epsilon(G)\} \leq \epsilon.$$

Действительно, $\mathbf{E}_\xi \{e^{\gamma g^T \xi}\} \leq e^{\frac{1}{2} \gamma^2 \|g\|_2^2 \sigma^2}$; значит, для всех $a \geq 0$

$$\text{Prob}\{g^T \xi > a\} \leq \inf_{\gamma > 0} \exp \left\{ \frac{1}{2} \gamma^2 \|g\|_2^2 \sigma^2 - \gamma a \right\} = \exp \left\{ -\frac{1}{2} a^2 \sigma^2 \|g\|_2^2 \right\},$$

⁷ Последующие результаты прямо распространяются на случай, когда $\mathcal{X}$ и $\mathcal{B}_*$ являются “общими” эллитопами; полагаем их базовыми для компактных обозначений.

и в результате имеем

$$\text{Prob} \left\{ \exists i \leq I : |g_i^T \xi| > \varkappa(\epsilon) \|g_i\|_2 \right\} \leq 2I \exp \left\{ -\frac{\varkappa^2(\epsilon)}{2\sigma^2} \right\} \leq \epsilon.$$

При заданной матрице контрастов $G = [g_1, \dots, g_I]$ размерности $m \times I$ рассмотрим задачу оптимизации

$$(8) \quad \text{Opt}[G] = \min_{\lambda, \mu, \gamma} \left\{ f_G(\lambda, \mu, \gamma) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\},$$

где

$$f_G(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + 4\psi^2[G] \sum_i \gamma_i$$

с

$$\psi[G] = \max_i \pi(g_i) + \varkappa(\epsilon) n[G].$$

Предложение 1. Пусть (λ, μ, γ) – допустимое решение задачи (8). Тогда

$$\text{Risk}_{\epsilon}[\hat{w}_G] \leq f_G(\lambda, \mu, \gamma),$$

т.е. $f_G(\lambda, \mu, \gamma)$ – верхняя граница ϵ -риска оценки $\hat{w}_G$.

Доказательство. Зафиксируем $\xi \in \Xi_{\epsilon}[G]$, $x_* \in \mathcal{X}$ и $\eta_* \in \mathcal{N}$. Кроме того, пусть $\hat{x} = \hat{x}_G(\omega)$ является x -компонентой некоторого оптимального решения $[\hat{x}; \hat{\nu}]$, $\hat{\nu} \in \mathcal{N}$, задачи (5) и пусть $\Delta = \hat{x} - x_*$. Очевидно, что пара $[x, \nu] = [x_*, \nu_*]$ допустима для (5) и $\|G^T[Ax_* + N\nu_* - \omega]\|_{\infty} = \|G^T \xi\|_{\infty} \leq \varkappa(\epsilon) n[G]$, из чего следует, что также $\|G^T[A\hat{x} + N\hat{\nu} - \omega]\|_{\infty} \leq \varkappa(\epsilon) n[G]$. Поэтому

$$\|G^T A \Delta\|_{\infty} \leq 2\varkappa(\epsilon) n[G] + \|G^T N[\hat{\nu} - \nu_*]\|_{\infty}.$$

Учитывая, что $\hat{\nu}, \nu_* \in \mathcal{N}$, имеем $\|G^T N[\hat{\nu} - \nu_*]\|_{\infty} \leq 2 \max_i \pi(g_i)$ и приходим к

$$(9) \quad |g_i^T A \Delta| \leq 2\psi[G], \quad i = 1, \dots, I.$$

Теперь пусть $\Delta \in 2\mathcal{X}$, т.е. для некоторого $t \in \mathcal{T}$ и всех k выполняется $\Delta^T T_k \Delta \leq 4t_k$; кроме того, пусть $v \in \mathcal{B}_*$ таким образом, что для некоторого $s \in \mathcal{S}$ и всех ℓ справедливо неравенство $v^T S_{\ell} v \leq s_{\ell}$. Ограничение неотрицательной определенности задачи (8) означает, что

$$v^T B \Delta \leq v^T \left[\sum_{\ell} \lambda_{\ell} S_{\ell} \right] v + \Delta^T \left[\sum_k \mu_k T_k \right] \Delta + [A \Delta]^T \sum_i \gamma_i g_i g_i^T A \Delta \leq \\ \leq \sum_{\ell} \lambda_{\ell} s_{\ell} + 4 \sum_k \mu_k t_k + \sum_i \gamma_i (g_i^T A \Delta)^2 \leq \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \gamma_i (g_i^T A \Delta)^2 \leq \\ \leq \phi_S(\lambda) + 4\phi_T(\mu) + 4\psi^2[G] \sum_i \gamma_i.$$

Максимизируя левую часть полученного неравенства по $v \in \mathcal{B}_*$, в итоге получаем $\|B\Delta\| \leq f_G(\lambda, \mu, \gamma)$.

Заметим, что задача в правой части (8) (очевидно) является выпуклой задачей оптимизации и $\text{Opt}[G]$ эффективно вычислима при условии, что таковы функции ϕ_S, ϕ_T и π . Таким образом, предложение 1 дает эффективно вычисляемую верхнюю границу для ϵ -риска полиэдральной оценки при заданной матрице контрастов G и, по сути, удобный вычислительный инструмент для анализа эффективности полиэдральной оценки. К сожалению, этот инструмент не позволяет *строить* такую оценку с помощью оптимизации матрицы контрастов: попытка сделать G переменной, а не параметром, в правой части (8), приводит к невыпуклой, а значит трудно решаемой, задаче оптимизации. Теперь рассмотрим две ситуации, в которых эту трудность можно преодолеть.

2.2. Синтез матриц контрастов для оценок гарантированного качества

Рассмотрим два типа предположений о множестве $\mathcal{N}$ помех, которые позволяют эффективно вычислять “вполне хорошие” матрицы контрастов. А именно:

- 1) “эллиптический случай”, когда $\mathcal{N}$ является базовым эллипсоидом;
- 2) “коэллиптический случай”, когда множество $N\mathcal{N} = \{N\nu : \nu \in \mathcal{N}\}$ является *полярой* эллипсоида

$$\begin{aligned} \mathcal{N}_* &= \{w \in \mathbf{R}^m : \exists \bar{r} \in \bar{\mathcal{R}} : w^T \bar{R}_j w \leq \bar{r}_j, j \leq \bar{J}\}, \\ \left[\bar{R}_j \succeq 0, \sum_j \bar{R}_j \succ 0; \right. \\ &\quad \left. \bar{\mathcal{R}} \subset \mathbf{R}_+^{\bar{J}}, \text{ int } \bar{\mathcal{R}} \neq \emptyset - \text{монотонный выпуклый компакт} \right]. \end{aligned}$$

Укажем, что $\mathcal{N}_*$ есть единичный шар нормы $\pi(g) = \max_{\nu \in \mathcal{N}} g^T N\nu$.

2.2.1. Эллиптический случай

Отметим, что *эллиптический случай может быть непосредственно сведен к случаю отсутствия помех*. Действительно, когда $\mathcal{N}$ является эллипсоидом, прямое произведение $\bar{\mathcal{X}} = \mathcal{X} \times \mathcal{N}$ — также эллипсоид. Таким образом, при задании $\bar{A}[x; \nu] = Ax + N\nu$ и $\bar{B}[x; \nu] = Bx$ наблюдение (2) принимает вид

$$\omega = \bar{A}\bar{x}_* + \xi, \quad [\bar{x}_* = [x_*; \nu_*] \in \bar{\mathcal{X}}],$$

и задача заключается в восстановлении по такому наблюдению линейного образа $w_* = \bar{B}\bar{x}_*$ нового сигнала $\bar{x}_*$. Построение близких к минимаксным, когда $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$, полиэдральных оценок в этой постановке рассмотрено в [2]. Кратко опишем идею построения, поскольку она будет использоваться ниже.

Рассмотрим случай, когда $\mathcal{N} = \{0\}$, и пусть множество сигналов и норма $\|\cdot\|$ по-прежнему заданы (4). В этой ситуации задача (8) сводится к

$$(10) \quad \text{Opt}[G] = \min_{\lambda, \mu, \gamma} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\kappa^2(\epsilon)n^2[G] \sum_i \gamma_i : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2}B^T} \middle| \frac{\frac{1}{2}B}{\sum_k \mu_k T_k + A^T [\sum_i \gamma_i g_i g_i^T] A} \right] \succeq 0 \right\}.$$

Заметим, что $\text{Opt}[G] = \text{Opt}[\theta G]$ при $\theta > 0$. Действительно, (λ, μ, γ) является допустимым решением задачи, определяющей $\text{Opt}[G]$, тогда и только тогда, когда $(\lambda, \mu, \theta^2 \gamma)$ является допустимым решением задачи, определяющей $\text{Opt}[\theta G]$, причем значения соответствующих целевых функций при этих решениях совпадают. Отсюда следует, что при оптимизации $\text{Opt}[G]$ по G можно без потери качества ограничиться матрицами контрастов G с $\kappa(\epsilon)n[G] = 1$. Другими словами, задавая

$$(11) \quad \theta(g) = \kappa(\epsilon)\|g\|_2$$

и дополняя переменные λ, μ и γ в (10) переменными $g_i, \theta(g_i) \leq 1, i = 1, \dots, I$, приходим к задаче

$$(12) \quad \text{Opt} = \min_{\lambda, \mu, \gamma, \{g_i\}, \rho} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\rho : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \theta(g_i) \leq 1, \sum_i \gamma_i \leq \rho, \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2}B^T} \middle| \frac{\frac{1}{2}B}{\sum_k \mu_k T_k + A^T [\sum_i \gamma_i g_i g_i^T] A} \right] \succeq 0 \right\}.$$

(Напомним, что параметр G рассматривается как переменная для минимизации $\text{Opt}[G]$ по G .) Теперь, объединяя переменные γ и $g_1, \dots, g_I$ в матрицу $\Theta = \sum_i \gamma_i g_i g_i^T$, обозначим через $\mathfrak{T}$ множество пар $(\Theta \in \mathbf{S}_+^m, \rho)$, для которых существует разложение $\Theta = \sum_{i \leq I} \gamma_i g_i g_i^T$ с $\theta(g_i) \leq 1$ и $\gamma_i \geq 0, \sum_i \gamma_i \leq \rho$. Тогда задача (12) может быть записана как задача оптимизации

$$(13) \quad \text{Opt} = \min_{\lambda, \mu, \Theta, \rho} \left\{ \phi_S(\lambda) + 4\phi_T(\mu) + 4\rho : \lambda \geq 0, \mu \geq 0, \right. \\ \left. (\Theta, \rho) \in \mathfrak{T}, \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2}B^T} \middle| \frac{\frac{1}{2}B}{\sum_k \mu_k T_k + A^T \Theta A} \right] \succeq 0 \right\}.$$

Укажем, что если $I \geq m$, то $\mathfrak{T}$ — простой выпуклый конус:

$$\mathfrak{T} = \{(\Theta, \rho) : \Theta \succeq 0, \rho \geq \kappa^2(\epsilon)\text{Tr}(\Theta)\}.$$

Следовательно, (13) является (очевидно разрешимой) задачей выпуклой оптимизации. Чтобы преобразовать оптимальное решение $(\lambda^*, \mu^*, \Theta^*, \rho^*)$ для (13) в оптимальное решение для (12), достаточно разложить Θ^* по собственным значениям, т.е. $\Theta^* = \sum_{i=1}^I v_i e_i e_i^T$ с $\|e_i\|_2 = 1$ и $v_i \geq 0, i \in \{1, \dots, m\}$, и $e_i = 0, v_i = 0, i \in \{m, \dots, I\}$, а также задать $g_i^* = \kappa^{-1}(\epsilon)e_i, \gamma_i^* = \kappa^2(\epsilon)v_i$. В результате приходим к оптимальному решению $(\lambda^*, \mu^*, \{g_i^*, \gamma_i^*\}_{i \leq I}, \rho^*)$ задачи (12).

2.2.2. Коэллитопический случай

Вышеизложенный подход, позволяющий свести невыпуклую задачу (12) (синтез наилучшей матрицы контрастов G в смысле $\text{Opt}[G]$) к (очевидно) выпуклой задаче оптимизации в значительной степени использует тот факт, что единичный шар нормы $\theta(\cdot)$ (ср. (11)) является простым эллитопом, кратным единичному евклидову шару. В силу данного обстоятельства $\mathfrak{T}$ – эффективно вычислимый выпуклый конус. Дальнейшие построения основаны на том, что когда единичный шар нормы $\theta(\cdot)$ является базовым эллитопом, происходит нечто подобное: связанное с ним множество $\mathfrak{T}$, хотя и не обязательно выпуклое и допускающее эффективное вычисление, может быть хорошо аппроксимировано эффективно вычислимым выпуклым конусом. Основной результат ([36, Предложение 3.2], с точностью до принятых обозначений) состоит в следующем:

Предложение 2. Пусть $I \geq m$ и $\mathcal{W} \subset \mathbf{R}^m$ – базовый эллитоп:

$$\mathcal{W} = \{w \in \mathbf{R}^m : \exists r \in \mathcal{R} : w^T R_j w \leq r_j, j \leq J\},$$

$$\left[R_j \succeq 0, \sum_j R_j \succ 0; \mathcal{R} \subset \mathbf{R}_+^J, \text{int} \mathcal{R} \neq \emptyset \text{ — монотонный выпуклый компакт} \right].$$

Свяжем с $\mathcal{W}$ замкнутый выпуклый конус⁸

$$\mathbf{K} = \{(\Theta, \rho) : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq \rho r_j, 1 \leq j \leq J, \Theta \succeq 0, \rho \geq 0\}.$$

Если матрица $\Theta \in \mathbf{S}_+^m$ представима в виде $\sum_i \gamma_i w_i w_i^T$ с $\gamma_i \geq 0$ и $w_i \in \mathcal{W}$, то $(\Theta, \sum_{i=1}^I \gamma_i) \in \mathbf{K}$, и “почти наоборот”: при $(\Theta, \rho) \in \mathbf{K}$ можно эффективно (с помощью рандомизированного алгоритма) найти векторы $w_i \in \mathcal{W}$, а также числа $\gamma_i \geq 0$, $i \leq I$ такие, что $\Theta = \sum_i \gamma_i w_i w_i^T$ и

$$\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 J) \rho.$$

Теперь построим “вполне хороший” контраст в коэллитопическом случае. Положим $R_j = \frac{1}{4} \bar{R}_j$, $j \leq \bar{J}$, $R_{\bar{J}+1} = \frac{\kappa^2(\epsilon)}{4} I_m$ и рассмотрим эллитоп

$$\begin{aligned} \mathcal{W} &= 2 \left[\mathcal{N}_* \cap \{w : \kappa(\epsilon) \|w\|_2 \leq 1\} \right] = \\ (14) \quad &= \{w \in \mathbf{R}^m : \exists r \in \mathcal{R} = \bar{\mathcal{R}} \times [0, 1] : w^T R_j w \leq r_j, j \leq J = \bar{J} + 1\}. \end{aligned}$$

Пусть $\theta(\cdot)$ – норма в $\mathbf{R}^n$ с единичным шаром $\mathcal{W}$. Заметим, что $\theta(\cdot) = 2 \max[\pi(\cdot), \kappa(\epsilon) \|\cdot\|_2]$ и для каждого $G = [g_1, \dots, g_I]$ величина $\psi[G]$ (8) ограничена сверху $\max_i \theta(g_i)$; эта граница точна с точностью до множителя 2.

⁸ $\mathbf{K}$ – это действительно замкнутый выпуклый конус – коническая оболочка выпуклого компактного множества $\{\Theta \succeq 0 : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq r_j, 1 \leq j \leq J\} \times \{1\}$.

Следовательно, согласно предложению 1 ϵ -риск полиэдральной оценки с матрицей контрастов G ограничен сверху величиной

$$(15) \quad \overline{\text{Opt}}[G] = \min_{\lambda, \mu, \gamma} \left\{ \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + 4 \left[\max_i \theta(g_i) \right]^2 \sum_i \gamma_i : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A \end{array} \right] \succeq 0 \right\}$$

и $\text{Opt}[G] \leq \overline{\text{Opt}}[G] \leq 2\text{Opt}[G]$. Как и в предыдущем разделе, задача минимизации $\overline{\text{Opt}}[G]$ по G может быть переформулирована в виде (13). Следующий результат дает вычислительно эффективный способ построить эффективную аппроксимацию оптимального решения этой задачи.

Пусть $I \geq m$, $\alpha = 2\sqrt{2} \ln(4m^2 J)$ и

$$\mathbf{K} = \{(\Theta, \rho) : \exists r \in \mathcal{R} : \text{Tr}(\Theta R_j) \leq \rho r_j, 1 \leq j \leq J, \Theta \succeq 0, \rho \geq 0\}$$

(см. (14)). Рассмотрим выпуклую задачу оптимизации

$$(16) \quad \text{Opt}_* = \min_{\lambda, \mu, \gamma, \Theta, \rho} \left\{ \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + 4\alpha\rho : \lambda \geq 0, \mu \geq 0, \right. \\ \left. (\Theta, \rho) \in \mathbf{K}, \left[\begin{array}{c|c} \sum_{\ell} \lambda_{\ell} S_{\ell} & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T \Theta A \end{array} \right] \succeq 0 \right\}.$$

Теорема 1. Θ -компонента Θ^* оптимального решения (разрешимой) задачи (16) может быть эффективным способом преобразована в матрицу контрастов G^* , удовлетворяющую условию

$$\overline{\text{Opt}}[G^*] \leq \sqrt{\alpha} \min_G \overline{\text{Opt}}[G] \leq 2\sqrt{\alpha} \min_G \text{Opt}[G].$$

В частности, ϵ -риск полиэдральной оценки с матрицей контрастов G^* (этот риск ограничен сверху $\overline{\text{Opt}}[G^*]$) не превышает $2\sqrt{\alpha} \min_G \text{Opt}[G]$.

Доказательство. Повторяя рассуждения предыдущего раздела, приходим к выводу, что $\overline{\text{Opt}} := \inf_G \overline{\text{Opt}}[G]$ равно

$$\inf_{g_1, \dots, g_I} \left\{ \overline{\text{Opt}}([g_1, \dots, g_I]) : \max_i \theta(g_i) = 1 \right\}.$$

Последний $\inf$ очевидно достигается на определенном наборе $g_1^+, \dots, g_I^+$ с $\max_i \theta(g_i^+) = 1$. Пусть $G^+ = [g_1^+, \dots, g_I^+]$ и $\lambda^+, \mu^+, \gamma_i^+, i \leq I$, – оптимальное решение задачи в правой части (15), связанное с $g_i = g_i^+, i \leq I$, и пусть $\Theta^+ = \sum_i \gamma_i^+ [g_i^+][g_i^+]^T$ и $\rho^+ = \sum_i \gamma_i^+$. Ясно, что

$$\overline{\text{Opt}} = \overline{\text{Opt}}[G^+] = \phi_{\mathcal{S}}(\lambda^+) + 4\phi_{\mathcal{T}}(\mu^+) + 4\rho^+.$$

Кроме того, в рассматриваемом случае неравенство $\theta(g) \leq 1$ эквивалентно $g \in \mathcal{W}$, а значит, $(\Theta^+, \rho^+) \in \mathbf{K}$ по первому утверждению предложения 2.

Отсюда следует, что $(\lambda^+, \mu^+, \Theta^+, \rho^+)$ – допустимое решение задачи оптимизации в (16). В силу структуры последней задачи при $\kappa > 0$ набор $(\kappa^{-1}\lambda^+, \kappa\mu^+, \kappa\Theta^+, \kappa\rho^+)$ допустим для (16) с соответствующим значением целевой функции $\kappa^{-1}\phi_S(\lambda^+) + \kappa[\phi_T(\mu^+) + 4\alpha\rho^+]$. Поэтому имеем

$$\begin{aligned} \text{Opt}_* &\leq \inf_{\kappa > 0} [\kappa^{-1}\phi_S(\lambda^+) + \kappa[4\phi_T(\mu^+) + 4\alpha\rho^+]] = \\ &= 2(\phi_S(\lambda^+) \underbrace{[4\phi_T(\mu^+) + 4\alpha\rho^+]_{\leq \alpha[4\phi_T(\mu^+) + 4\alpha\rho^+]}})^{1/2} \leq 2\sqrt{\phi_S(\lambda^+)[4\phi_T(\mu^+) + 4\alpha\rho^+]}\sqrt{\alpha} \leq \\ &\leq \sqrt{\alpha}[\phi_S(\lambda^+) + 4\phi_T(\mu^+) + 4\alpha\rho^+] = \sqrt{\alpha}\overline{\text{Opt}}. \end{aligned}$$

Наконец, пусть $\bar{\lambda}, \bar{\mu}, \bar{\Theta}, \bar{\rho}$ – оптимальное решение (16). Поскольку $(\bar{\Theta}, \bar{\rho}) \in \mathbf{K}$, по второму утверждению предложения 2 существует (и может быть эффективно найдено) разложение $\bar{\Theta} = \sum_i \bar{\gamma}_i [\bar{g}_i] [\bar{g}_i]^T$ с $\bar{g}_i \in \mathcal{W}$ (т.е. $\theta(\bar{g}_i) \leq 1$), $i \leq I$, $\bar{\gamma}_i \geq 0$, и $\sum_i \bar{\gamma}_i \leq \alpha\bar{\rho}$. Тогда ϵ -риск полиэдральной оценки с матрицей контрастов $\bar{G} = [\bar{g}_1, \dots, \bar{g}_I]$ имеет верхнюю границу $\overline{\text{Opt}}[\bar{G}]$. Однако $\bar{\lambda}, \bar{\mu}$ и $\{\bar{\gamma}_i\}$ образуют допустимое решение задачи, определяющей $\overline{\text{Opt}}[\bar{G}]$, и значение целевой функции на этом решении ограничено сверху величиной

$$\phi_S(\bar{\lambda}) + 4\phi_T(\bar{\mu}) + 4 \left[\max_i \theta(\bar{g}_i) \right] \sum_i \bar{\gamma}_i \leq \phi_S(\bar{\lambda}) + 4\phi_T(\bar{\mu}) + 4\alpha\bar{\rho} = \text{Opt}_*.$$

Таким образом, ϵ -риск полиэдральной оценки с матрицей контрастов $\bar{G}$ не превышает

$$\text{Opt}_* \leq \sqrt{\alpha}\overline{\text{Opt}} \leq 2\sqrt{\alpha} \min_G \text{Opt}[G]. \quad \square$$

3. Наблюдения с выбросами

Рассмотрим задачу оценивания, поставленную в разделе 1.2, в ситуации разреженной помехи ν_* в (2), т.е. когда она имеет не более чем s ненулевых элементов.

Построение оценки. Пусть $\epsilon \in (0, 1)$ – заданный допуск надежности. Рассмотрим следующую полиэдральную оценку, определенную двумя матрицами контрастов $H = [h_1, \dots, h_n] \in \mathbf{R}^{m \times n}$ и $G = [g_1, \dots, g_I] \in \mathbf{R}^{n \times I}$. При имеющемся наблюдении ω (см. (2)) решим задачу оптимизации

$$(17) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \begin{aligned} |h_k^T [N\nu + Ax - \omega]| &\leq \overline{\varkappa}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T [N\nu + Ax - \omega]| &\leq \overline{\varkappa}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I \end{aligned} \right\},$$

где

$$\overline{\varkappa}(\epsilon) = \sigma \sqrt{2 \ln[2(n+I)/\epsilon]}.$$

Пусть $(\hat{\nu}, \hat{x}) = (\hat{\nu}(\omega), \hat{x}(\omega))$ – оптимальное решение задачи, если она разрешима; в противном случае положим $(\hat{\nu}, \hat{x}) = (0, 0)$. В качестве оценки $w_* = Bx_*$ будем использовать вектор

$$\hat{w}_{G,H}(\omega) = B\hat{x}(\omega).$$

3.1. Анализ рисков

Обозначим через $\Xi_\epsilon(G, H)$ множество реализаций ξ , удовлетворяющих условию

$$(18) \quad \begin{aligned} |h_k^T \xi| &\leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T \xi| &\leq \overline{\mathfrak{x}}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I, \end{aligned} \quad \forall \xi \in \Xi_\epsilon(G, H).$$

На том же основании, что и в (7), получаем

$$\text{Prob}_{\xi \sim \mathcal{SG}(0, \sigma^2 I_m)}(\Xi_\epsilon(G, H)) \geq 1 - \epsilon.$$

Теперь зафиксируем $x_* \in \mathcal{X}$, s -разреженный сигнал ν_* и $\xi \in \Xi_\epsilon(G, H)$; таким образом, наблюдение принимает вид $\omega = Ax_* + N\nu_* + \xi$.

А. В силу (18) имеем $|h_k^T \xi| \leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2$ и $|g_i^T \xi| \leq \overline{\mathfrak{x}}(\epsilon) \|g_i\|_2$ для всех $k \leq n$ и $i \leq I$. Тогда задача (17) сводится к

$$(19) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \right. \\ \left. \begin{aligned} |h_k^T [N[\nu - \nu_*] + A[x - x_*] - \xi]| &\leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, n, \\ |g_i^T [N[\nu - \nu_*] + A[x - x_*] - \xi]| &\leq \overline{\mathfrak{x}}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I \end{aligned} \right\}.$$

Заключаем, что $(\nu, x) = (\nu_*; x_*)$ — допустимое решение для (17). Таким образом, в рассматриваемом случае $\widehat{\nu}$ и $\widehat{x}$ являются допустимыми для (19) и

$$\|\widehat{\nu}\|_1 \leq \|\nu_*\|_1.$$

Б. В дальнейшем предполагается, что $(H, \|\cdot\|_\infty)$ удовлетворяет условию $Q_\infty(s, \kappa)$ из раздела 2 с $\kappa < \frac{1}{2}$, т.е.⁹

$$(20) \quad \|w\|_\infty \leq \|H^T N w\|_\infty + \frac{\kappa}{s} \|w\|_1 \quad \forall w \in \mathbf{R}^n.$$

Поскольку $\widehat{\nu}$ и $\widehat{x}$ допустимы для (19), имеем

$$|h_k^T [N[\widehat{\nu} - \nu_*] + A[\widehat{x} - x_*] - \xi]| \leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad \forall k \leq n.$$

Задействуя (18) и $A[\widehat{x} - x_*] \in 2A\mathcal{X}$ (в силу симметрии $\mathcal{X}$ относительно начала координат), заключаем, что

$$\|H^T N[\widehat{\nu} - \nu_*]\|_\infty \leq \max_k \left[\overline{\mathfrak{x}}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right];$$

⁹ Условие $Q_\infty(s, \kappa)$ является простейшим (и наиболее ограничительным) в семействе условий $Q_q(s, \kappa)$, $q \in [1, \infty]$, используемых для установления свойств ℓ_1 -восстановления разреженных сигналов. Это условие может быть эффективно проверено, что очень важно здесь. В [2, раздел 1.3] рассмотрены эффективно проверяемые условия в разреженном восстановлении и их связь с другими используемыми условиями (свойство ограниченной изометрии (Restricted Isometry Property, RIP) [37], ограниченное собственное значение (Restricted Eigenvalue, RE) [38], взаимная некогерентность (Mutual Incoherence, MI) [39], и совместимость (Compatibility) [40]).

кроме того, вектор ν_* является s -разреженным и $\|\widehat{\nu}\|_1 \leq \|\nu_*\|_1$. Теперь предположение 5 с ν_* в роли ν дает

$$(21) \quad \|\widehat{\nu} - \nu_*\|_q \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \max_k \left[\overline{\mathcal{P}}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right], \quad 1 \leq q \leq \infty;$$

в частности,

$$(22a) \quad \|\widehat{\nu} - \nu_*\|_\infty \leq \frac{1}{1 - 2\kappa} \max_k \left[\overline{\mathcal{P}}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right] =: \rho_H,$$

$$(22b) \quad \|\widehat{\nu} - \nu_*\|_1 \leq \frac{2s}{1 - 2\kappa} \max_k \left[\overline{\mathcal{P}}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} |h_k^T A x| \right] = 2s\rho_H.$$

Более того, согласно [2, Предложение 1.10] множество $\mathcal{H}$ пар (H, κ) с матрицами H размерности $m \times n$, удовлетворяющими условию $Q_\infty(s, \kappa)$, является эффективно вычислимым выпуклым множеством

$$(23) \quad \mathcal{H} = \left\{ (H, \kappa) \in \mathbf{R}^{m \times n} \times \mathbf{R} : \left| [I_n - N^T H]_{ij} \right| \leq s^{-1} \kappa, 1 \leq i, j \leq n \right\}.$$

В. Поскольку $\widehat{\nu}$ и $\widehat{x}$ допустимы для (19), имеем

$$|g_i^T (N[\widehat{\nu} - \nu_*] + A[\widehat{x} - x_*] - \xi)| \leq \overline{\mathcal{P}}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, I,$$

а $|g_i^T \xi| \leq \overline{\mathcal{P}}(\epsilon) \|g_i\|_2 \forall i$ в силу $\xi \in \Xi_\epsilon(G, H)$. Заключаем, что

$$(24) \quad |g_i^T A[\widehat{x} - x_*]| \leq 2\overline{\mathcal{P}}(\epsilon) \|g_i\|_2 + |g_i^T N[\widehat{\nu} - \nu_*]|, \quad i \leq I.$$

Пусть $\|z\|_{k,1}$, $z \in \mathbf{R}^n$, – сумма $\min[k, n]$ наибольших величин элементов в z ; отметим, что $\|\cdot\|_{k,1}$ является нормой, сопряженной к норме с единичным шаром $\{u : \|u\|_\infty \leq 1, \|u\|_1 \leq k\}$. Следовательно, из (22) получаем, что

$$(25) \quad |g_i^T N[\widehat{\nu} - \nu_*]| \leq \rho_H \|N^T g_i\|_{2s,1}$$

и на основании (24)

$$(26) \quad |g_i^T A[\widehat{x} - x_*]| \leq \psi_H[G], \quad \psi_H[G] = \max_i \left[2\overline{\mathcal{P}}(\epsilon) \|g_i\|_2 + \rho_H \|N^T g_i\|_{2s,1} \right].$$

Пусть

$$f_{G,H}(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + \psi_H^2[G] \sum_i \gamma_i;$$

рассмотрим задачу оптимизации (ср. (8))

$$(27) \quad \text{Opt}[G, H] = \min_{\lambda, \mu, \gamma} \left\{ f_{G,H}(\lambda, \mu, \gamma) : \right. \\ \left. \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2} B^T} \middle| \frac{\frac{1}{2} B}{\sum_k \mu_k T_k + A^T \left[\sum_i \gamma_i g_i g_i^T \right] A} \right] \succeq 0 \right\}.$$

Применяя те же рассуждения, что и при доказательстве предложения (1), с (26) в роли (9), устанавливаем следующий результат.

Предложение 3. В условиях текущего раздела при заданных $\kappa \in (0, 1/2)$ и матрице H размерности $m \times n$, удовлетворяющей $(H, \kappa) \in \mathcal{H}$ (23), пусть (λ, μ, γ) – допустимое решение (27). Тогда

$$\text{Risk}_\epsilon[\hat{w}_{G,H}] \leq f_{G,H}(\lambda, \mu, \gamma).$$

3.2. Синтез матриц контрастов

Цель авторов – получить матрицы контрастов H и G с небольшой верхней границей $\text{Opt}[G, H]$ для ϵ -риска оценки $\hat{w}_{G,H}$.

Г. Построение матрицы контрастов $H \in \mathbf{R}^{m \times n}$ не вызывает затруднений: граница $\text{Opt}[G, H]$ риска зависит от $H = [h_1, \dots, h_n]$ исключительно через величину

$$\rho_H = \frac{1}{1 - 2\kappa} \max_{k \leq n} \left[\overline{\mathcal{P}}(\epsilon) \|h_k\|_2 + 2 \max_{x \in \mathcal{X}} \|h_k^T A x\|_\infty \right]$$

и будет тем меньше, чем ниже ρ_H . Для фиксированного числа $\kappa \in (0, 1/2)$ вполне хороший выбор $H = [h_1, \dots, h_n]$ задается оптимальными решениями n задач оптимизации

$$(28) \quad h_k = \underset{h}{\operatorname{argmin}} \left\{ \overline{\mathcal{P}}(\epsilon) \|h\|_2 + 2 \max_{x \in \mathcal{X}} \|h^T A x\|_\infty : h \in \mathbf{R}^m, \right. \\ \left. \|\text{Col}_i[I_n - N^T h]\|_\infty \leq s^{-1} \kappa \right\}.$$

В силу формы $\mathcal{X}$ и конической двойственности эти задачи эквивалентны задачам

$$h_k = \underset{h, v, \chi}{\operatorname{argmin}} \left\{ \overline{\mathcal{P}}(\epsilon) \|h\|_2 + v + \phi_{\mathcal{T}}(\chi) : h \in \mathbf{R}^m, \chi \geq 0, \right. \\ \left. \left[\begin{array}{c|c} v & h^T A \\ \hline A^T h & \sum_k \chi_k T_k \end{array} \right] \succeq 0, \|\text{Col}_i[I_n - N^T h]\|_\infty \leq s^{-1} \kappa \right\}, \quad 1 \leq k \leq n.$$

Д. Предлагаемое построение G менее простое. Действуем следующим образом. Пусть $G = [G_1, G_2]$, где $G_2, G_1 \in \mathbf{R}^{m \times m}$ (таким образом, $I = 2m$).

Д.1 Заметим, что поскольку $\xi \in \Xi_\epsilon(G, H)$, задача (19) разрешима, а $(\hat{x}, \hat{\nu})$ – ее допустимое решение. Для столбца g из G согласно ограничениям задачи имеем

$$(29) \quad |g^T A[\hat{x} - x_*]| \leq 2\overline{\mathcal{P}}(\epsilon) \|g\|_2 + |g^T N[\hat{\nu} - \nu_*]| \leq 2\overline{\mathcal{P}}(\epsilon) \|g\|_2 + 2s\rho_H \|N^T g\|_\infty$$

(здесь использованы (24) и (25)). Отсюда следует, что

$$(30) \quad (g^T A[\hat{x} - x_*])^2 \leq 2(4\overline{\mathcal{P}}^2(\epsilon) \|g\|_2^2 + 4s^2 \rho_H^2 \|N^T g\|_\infty^2), \quad i = 1, \dots, m.$$

Укажем, что множество

$$\mathcal{M} = \{g \in \mathbf{R}^m : 8\overline{\mathcal{P}}^2(\epsilon) \|g\|_2^2 + 8s^2 \rho_H^2 \|N^T g\|_\infty^2 \leq 1\}$$

представляет собой эллипс: обозначая $N = [n_1, \dots, n_n]$, приходим к

$$\mathcal{M} = \left\{ g \in \mathbf{R}^m : \exists r \in [0, 1]^n : g^T \underbrace{(8\bar{\kappa}^2(\epsilon)I_m + 8s^2\rho_H^2 n_j n_j^T)}_{M_j} g \leq r_j, j = 1, \dots, n \right\}.$$

Д.2 Далее заметим, что если $\xi \in \Xi_\epsilon(G, H)$, на основании (21) получаем

$$\|\hat{\nu} - \nu_*\|_2 \leq \frac{\sqrt{2s}}{1 - 2\kappa} \max_{k \leq n} \left[\bar{\kappa}(\epsilon) \|h_k\|_2 + \max_{x \in \mathcal{X}} |h_k^T A x| \right] = \sqrt{2s} \rho_H.$$

Тогда согласно (29) для столбца g матрицы G имеет место

$$\begin{aligned} (g^T A [\hat{x} - x_*])^2 &\leq (2\bar{\kappa}(\epsilon) \|g\|_2 + |g^T N [\hat{\nu} - \nu_*]|)^2 \leq \\ &\leq \left(2\bar{\kappa}(\epsilon) \|g\|_2 + \sqrt{2s} \rho_H \|N^T g\|_2 \right)^2 \leq \\ (31) \quad &\leq g^T (8\bar{\kappa}^2(\epsilon) I_m + 4s \rho_H^2 N N^T) g. \end{aligned}$$

Теперь примем

$$(32) \quad Q = (8\bar{\kappa}^2(\epsilon) I_m + 4s \rho_H^2 N N^T)^{-1/2}$$

и рассмотрим задачу оптимизации

$$(33a) \quad \text{Opt} = \min_{\lambda, \mu, \Theta_1, \Theta_2, \rho} \left\{ f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho) : \lambda \geq 0, \mu \geq 0, \Theta_1 \succeq 0, \Theta_2 \succeq 0, \right. \\ \left. \text{Tr}(M_j \Theta_1) \leq \rho, j = 1, \dots, n, \right. \\ \left. \left[\begin{array}{c|c} \sum_\ell \lambda_\ell S_\ell & \frac{1}{2} B \\ \hline \frac{1}{2} B^T & \sum_k \mu_k T_k + A^T (\Theta_1 + Q \Theta_2 Q^T) A \end{array} \right] \succeq 0 \right\},$$

где

$$(33b) \quad f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho) = \phi_S(\lambda) + 4\phi_T(\mu) + \text{Tr}(\Theta_2) + 2\sqrt{2} \ln(4m^2 n) \rho.$$

Заметим, что ограничения на Θ_1 и ρ задачи (33a) говорят именно о принадлежности (Θ_1, ρ) конусу $\mathbf{K}$, связанному (см. Предложение 2) с эллипсом $\mathcal{M}$ в роли $\mathcal{W}$.

Теорема 2. При имеющемся допустимом решении $(\lambda, \mu, \tau, \Theta_1, \Theta_2)$ задачи (33) построим матрицы контрастов G_1 и G_2 размерности $m \times m$ следующим образом.

- Для построения G_1 применим вторую часть Предложения 2 к $\Theta_1, \rho, \mathcal{M}$ в роли $\Theta, \rho, \mathcal{W}$, чтобы получить (вычислительно эффективным образом) разложение $\Theta_1 = \sum_{i=1}^m \gamma_i g_{1,i} g_{1,i}^T$ с $g_{1,i} \in \mathcal{M}$ и $\gamma_i \geq 0$, $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$. Положим $G_1 = [g_{1,1}, \dots, g_{1,m}]$.
- Для построения G_2 разложим Θ_2 по собственным значениям, т.е. $\Theta_2 = \Gamma \text{Diag}\{\chi\} \Gamma^T$, и зададим $G_2 = [g_{2,1}, \dots, g_{2,m}] = Q \Gamma$.

Заметим, что $\Theta_1 + Q \Theta_2 Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$.

Для полученной полиэдральной оценки $\widehat{w}_{G,H}$ и для всех $x_* \in \mathcal{X}$, s -разреженных ν_* и $\xi \in \Xi_\epsilon(G, H)$ выполняется

$$(34) \quad \|\widehat{w}_{G,H}(Ax_* + N\nu_* + \xi) - Bx_*\| \leq f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho).$$

Следовательно, ϵ -риск оценки ограничен сверху величиной $f_H(\lambda, \mu, \Theta_1, \Theta_2, \rho)$ (в силу $\xi \in \Xi_\epsilon(G, H)$ с вероятностью не менее $1 - \epsilon$).

Доказательство. Зафиксируем $x_* \in \mathcal{X}$, s -разреженный ν_* и $\xi \in \Xi_\epsilon(G, H)$, и пусть $w = Ax_* + N\nu_* + \xi$. На основании п. **A** задача (17) разрешима, $(\widehat{x}, \widehat{\nu}) = (\widehat{x}(\omega), \widehat{\nu}(\omega))$ – ее оптимальное решение, а $\widehat{w} = B\widehat{x}$ – оценка $\widehat{w}_{G,H}(\omega)$. Пусть $\Delta = \widehat{x} - x_*$, и пусть $e_1, \dots, e_m$ – столбцы ортонормированной матрицы Γ . По построению G_2 имеем для всех $j \leq m$ (см. (31))

$$\begin{aligned} (g_{2,j}^T A \Delta)^2 &\leq g_{2,j}^T (8\overline{\kappa}^2(\epsilon)I_m + 4s\rho_H^2 N N^T) g_{2,j} = \\ &= e_j^T [Q (8\overline{\kappa}^2(\epsilon)I_m + 4s\rho_H^2 N N^T) Q] e_j = e_j^T e_j = 1. \end{aligned}$$

Более того, в силу $g_{1,i} \in \mathcal{M}$ имеем (см. (30))

$$(g_{1,i}^T A \Delta)^2 \leq 8\overline{\kappa}^2(\epsilon) \|g\|_2^2 + 8s^2 \rho_H^2 \|N^T g\|_\infty^2 \leq 1 \quad \forall i \leq m.$$

Теперь согласно ограничению неотрицательной определенности в (33а) и $\Theta_1 + Q\Theta_2Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$ для каждого $v \in \mathcal{B}_*$ получаем

$$\begin{aligned} v^T B \Delta &\leq v^T \left[\sum_\ell \lambda_\ell S_\ell \right] v + \Delta^T \left[\sum_k \mu_k T_k \right] \Delta + \\ &+ [A\Delta]^T \left[\sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T \right] A \Delta \leq \\ &\leq \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + \sum_i \chi_i (g_{1,i}^T A \Delta)^2 + \sum_j \gamma_j (g_{2,j}^T A \Delta)^2 \leq \\ &[v^T S_1 v; \dots; v^T S_L v] \in \mathcal{S} \text{ в силу } v \in \mathcal{B}_* \text{ и} \\ &[\Delta^T T_1 \Delta; \dots; \Delta^T T_L \Delta] \in 4\mathcal{T} \text{ в силу } \Delta \in 2\mathcal{X} \leq \\ &\leq \phi_{\mathcal{S}}(\lambda) + 4\phi_{\mathcal{T}}(\mu) + \sum_i \chi_i + \sum_j \gamma_j \leq f_H(\lambda, \mu, \tau, \Theta_1, \Theta_2) \end{aligned}$$

благодаря $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$ и $\sum_i \chi_i = \text{Tr}(\Theta_2)$. Вычислив в полученном неравенстве супремум по $v \in \mathcal{B}_*$, в итоге приходим к (34). $\square$

3.3. Альтернативное построение

Цель раздела – уточнить границы риска (27) и (33а) и построить улучшенные матрицы контраста. Действуем следующим образом. Во-первых, чтобы

расширить возможный выбор H -контрастов, ответственных за восстановление возмущений, улучшаем границы (22) для точности разреженного восстановления, в частности, позволяя использовать контрасты, не удовлетворяющие условию $Q_\infty(s, \kappa)$. Во-вторых, улучшаем границу для риска оценки $\hat{w}(\omega)$ благодаря учету вклада H -компоненты “полной” матрицы контраста $[H, G]$ при оптимизации G -компоненты контраста.

В дальнейшем рассматриваем оценку, описанную в начале раздела 3, с той лишь разницей, что размерности матриц контрастов G и H отличаются: теперь $H = [h_1, \dots, h_M] \in \mathbf{R}^{m \times M}$, а $G = [g_1, \dots, g_{2m}]$. Таким образом, в текущей постановке при имеющемся наблюдении ω решаем задачу оптимизации

$$(35) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \begin{array}{l} |h_k^T(N\nu + Ax - \omega)| \leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T(N\nu + Ax - \omega)| \leq \overline{\mathfrak{x}}(\epsilon) \|g_{2,i}\|_2, \quad i = 1, \dots, 2m \end{array} \right\}$$

с

$$\overline{\mathfrak{x}}(\epsilon) = \sigma \sqrt{2 \ln[(2M + 4m)/\epsilon]}$$

и указываем $\hat{x}(\omega), \hat{\nu}(\omega)$ в качестве оптимального решения задачи, если задача разрешима; в противном случае задаем $(\hat{x}(\omega), \hat{\nu}(\omega)) = (0, 0)$ и берем $\hat{w}_{G,H}(\omega) = B\hat{x}(\omega)$ в качестве оценки Bx_* .

3.3.1. Анализ риска

Приведенную выше задачу можно эквивалентно записать в виде

$$(36) \quad \min_{\nu, x} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \begin{array}{l} |h_k^T(N[\nu - \nu_*] + A[x - x_*] - \xi)| \leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T(N[\nu - \nu_*] + A[x - x_*] - \xi)| \leq \overline{\mathfrak{x}}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, 2m, \end{array} \right\}.$$

Задавая

$$(37) \quad \Xi_\epsilon(G, H) := \left\{ \xi \in \mathbf{R}^m : \begin{array}{l} |h_k^T \xi| \leq \overline{\mathfrak{x}}(\epsilon) \|h_k\|_2, \quad k = 1, \dots, M, \\ |g_i^T \xi| \leq \overline{\mathfrak{x}}(\epsilon) \|g_i\|_2, \quad i = 1, \dots, 2m \end{array} \right\},$$

получаем

$$\text{Prob}_{\xi \sim \mathcal{SG}(0, \sigma^2 I_m)}(\Xi_\epsilon(G, H)) \geq 1 - \epsilon.$$

Зафиксируем $\xi \in \Xi_\epsilon(G, H)$, и пусть $\omega = Ax_* + N\nu_* + \xi$. Поскольку пара $(\hat{\nu}, \hat{x})$ является допустимой для (36), $\hat{x} := \hat{x}(\omega)$, $\hat{\nu} := \hat{\nu}(\omega)$ также допустима, $\|\hat{\nu}\|_1 \leq \|\nu_*\|_1$. Таким образом, по аналогии с доказательством Предложения 5, для $z = \hat{\nu} - \nu_*$ имеет место

$$\|z\|_1 \leq 2\|z\|_{s,1}$$

и, следовательно,

$$(38) \quad \|z\|_1 \leq 2s\|z\|_\infty, \quad \|z\|_2 \leq \sqrt{2s}\|z\|_\infty.$$

Теперь введем обозначение $\Delta = \hat{x} - x_*$ и рассмотрим n пар выпуклых задач оптимизации

$$(P_2[i]) \quad \text{Opt}_2[i] = \max_{v,t,w} \left\{ \sqrt{w_i t} : v \in 2\mathcal{X}, \right. \\ \left. \|w\|_\infty \leq w_i, \|w\|_1 \leq t, t \leq 2s w_i, \right. \\ \left. |h_k^T(Nw + Av)| \leq 2\overline{\pi}(\epsilon) \|h_k\|_2, k = 1, \dots, M \right\},$$

$$(P_\infty[i]) \quad \text{Opt}_\infty[i] = \max_{v,w} \left\{ w_i : v \in 2\mathcal{X}, \right. \\ \left. \|w\|_\infty \leq w_i, \|w\|_1 \leq 2s w_i, \right. \\ \left. |h_k^T(Nw + Av)| \leq 2\overline{\pi}(\epsilon) \|h_k\|_2, k = 1, \dots, M \right\}.$$

Заметим, что допустимое решение (v, t, w) для $(P_2[i])$ удовлетворяет $\|w\|_\infty \leq w_i$ и $\|w\|_1 \leq t$, откуда

$$(39) \quad \|w\|_2 \leq \sqrt{w_i t} \leq \text{Opt}_2[i].$$

Пусть $\iota = \iota_z$ – индекс наибольшего по абсолютной величине элемента в z . Поскольку $\xi \in \Xi_\epsilon(G, H)$ и $\Delta \in 2\mathcal{X}$, заключаем, что при условии $z_\iota \geq 0$ тройка $(v, t, w) = (\Delta, \|z\|_1, z)$ допустима для $(P_2[\iota])$, а пара $(v, w) = (\Delta, z)$ допустима для $(P_\infty[\iota])$; если $z_\iota < 0$, то же самое справедливо для $(v, t, w) = (-\Delta, \|z\|_1, -z)$ и $(v, w) = (-\Delta, -z)$. Действительно, в первом случае $v = \Delta \in \mathcal{X}$, $|h_k^T[A\hat{x} + N\hat{v} - \omega]| \leq \overline{\pi}(\epsilon) \|h_k\|_2$ и $|h_k^T[Ax_* + N\nu_* - \omega]| \leq \overline{\pi}(\epsilon) \|h_k\|_2$, так как обе пары $(\hat{x}, \hat{v})$ и (x_*, ν_*) допустимы для (35), откуда следуют ограничения из второй строки $(P_2[i])$. Заметим, что в рассматриваемом случае $z_\iota = \|z\|_\infty$, т.е. ограничения из первой строки $(P_2[i])$ выполняются для $w = z$ в силу (38). Таким образом, тройка $(\Delta, \|z\|_1, z)$ действительно допустима для $(P_2[i])$. В качестве сопутствующего результата этих рассуждений получаем, что тройка $(\Delta, \|z\|_1, z)$ допустима для $(P_\infty[i])$. Во втором случае аргументация полностью аналогична.

Теперь положим

$$(40) \quad \text{Opt}_2 = \max_i \text{Opt}_2[i], \quad \text{Opt}_\infty = \max_i \text{Opt}_\infty[i]$$

и напомним, что тройка $(\Delta, \|z\|_1, z)$ или $(-\Delta, \|z\|_1, -z)$ допустима для некоторых задач $(P_2[i])$, а пара (Δ, z) или $(-\Delta, -z)$ допустима для некоторых задач $(P_\infty[i])$. Используя (39), для всех $\xi \in \Xi_\epsilon(G, H)$ получаем

$$\|z\|_\infty \leq \text{Opt}_\infty, \quad \|z\|_2 \leq \text{Opt}_2, \quad \|z\|_1 \leq 2s \text{Opt}_\infty.$$

Как следствие, для всех $d \in \mathbf{R}^m$ имеет место

$$(41) \quad |d^T N z| \leq \max_z \{ d^T N z : \|z\|_\infty \leq \text{Opt}_\infty, \|z\|_2 \leq \text{Opt}_2, \|z\|_1 \leq 2s \text{Opt}_\infty \} = \\ = \underbrace{\min_{u,v,w} \{ \|u\|_1 \text{Opt}_\infty + \|v\|_2 \text{Opt}_2 + 2s \|w\|_\infty \text{Opt}_\infty, u + v + w = N^T d \}}_{=:\overline{\pi}(N^T d)}.$$

Поскольку пара $\widehat{x}, \widehat{\nu}$ допустима для (36) и, кроме того, $\xi \in \Xi_\epsilon(G, H)$, заключаем, что столбцы $d_i, i = 1, \dots, M+2m$, “агрегированной” матрицы контрастов $D = [G, H]$ удовлетворяют

$$|d_i^T A \Delta| \leq |d_i^T N z| + |d_i^T \xi| + \overline{\kappa}(\epsilon) \|g\|_2,$$

откуда

$$(42) \quad |d_i^T A \Delta| \leq \underbrace{\overline{\kappa}(N^T d_i) + 2\overline{\kappa}(\epsilon) \|d_i\|_2}_{=: \psi_H(d_i)}, \quad i \leq M+2m.$$

Далее, пусть

$$\bar{f}_{G,H}(\lambda, \mu, \gamma) = \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \gamma_i \psi_H^2(d_i),$$

и рассмотрим задачу оптимизации (ср. (27))

$$(43) \quad \text{Opt}[G, H] = \min_{\lambda, \mu, \gamma} \left\{ \bar{f}_{G,H}(\lambda, \mu, \gamma) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \right. \\ \left. \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2} B^T} \middle| \frac{\frac{1}{2} B}{\sum_k \mu_k T_k + A^T [\sum_i \gamma_i d_i d_i^T] A} \right] \succeq 0 \right\}.$$

Используя те же рассуждения, что и при доказательстве Предложения 1, с (42) в роли (9), устанавливаем следующий результат.

Предложение 4. В условиях текущего раздела пусть (λ, μ, γ) – допустимое решение для (43). Тогда

$$\text{Risk}_\epsilon[\widehat{w}_{G,H} | \mathcal{X}, \mathcal{N}] \leq \bar{f}_{G,H}(\lambda, \mu, \gamma).$$

3.3.2. Синтез матриц контрастов

Продолжаем анализ оценки $\widehat{w}_{G,H}$ при имеющемся наблюдении $\omega = Ax_* + N\nu_* + \xi$ с $\xi \in \Xi_\epsilon(G, H)$, см. (37). Согласно (41) для $z = \widehat{\nu} - \nu_*$ и всех $g \in \mathbf{R}^m$ имеем

$$|g^T N z| \leq \min \left\{ \|N^T g\|_2 \text{Opt}_2, \sqrt{2s} \|N^T g\|_2 \text{Opt}_\infty, 2s \|N^T g\|_\infty \text{Opt}_\infty \right\}.$$

Отсюда следует (ср. (42)), что для всех $i \leq 2m$ справедливо неравенство

$$(44) \quad |g_i^T A \Delta| \leq 2\overline{\kappa}(\epsilon) \|g_i\|_2 + \\ + \min \left\{ \|N^T g_i\|_2 \text{Opt}_2, \sqrt{2s} \|N^T g_i\|_2 \text{Opt}_\infty, 2s \|N^T g_i\|_\infty \text{Opt}_\infty \right\}.$$

Заметим, что правая часть в (44) невыпукла по g , поэтому описанные выше методы синтеза матриц контрастов не применимы. Чтобы преодолеть эту трудность, используем следующее важное свойство полиэдральных оценок:

“агрегирование” оценок позволяет получить новую оценку с риском, незначительно отличающимся от наименьшего из рисков всех агрегируемых оценок.

В рассматриваемой постановке данный подход работает следующим образом. Разбиваем контраст G размерности $m \times 2m$ на два блока $G_\chi = [g_{\chi,1}, \dots, g_{\chi,m}]$, $\chi = 1, 2$, размерности $m \times m$ и строим блоки, используя соответствующие неравенства, унаследованные от (44), а именно:

$$\begin{aligned} |g_{1,i}^T A \Delta| &\leq 2\overline{\mathcal{K}}(\epsilon) \|g_{1,i}\|_2 + 2s \|N^T g_{1,i}\|_\infty \text{Opt}_\infty, \\ |g_{2,i}^T A \Delta| &\leq 2\overline{\mathcal{K}}(\epsilon) \|g_{2,i}\|_2 + \|N^T g_{2,i}\|_2 \underbrace{\min\{\text{Opt}_2, \sqrt{2s} \text{Opt}_\infty\}}_{=:\varrho_{2,H}}, \end{aligned}$$

где $\Delta = \widehat{x} - x_*$. Ослабим эти неравенства до

$$\begin{aligned} |g_{1,i}^T A \Delta|^2 &\leq \pi_1^2(g_{1,i}), \quad \pi_1(g) = \sqrt{8\overline{\mathcal{K}}^2(\epsilon) \|g\|_2^2 + 8s^2 \text{Opt}_\infty^2 \|N^T g\|_\infty^2}, \\ |g_{2,i}^T A \Delta|^2 &\leq \pi_2^2(g_{2,i}), \quad \pi_2(g) = \sqrt{8\overline{\mathcal{K}}^2(\epsilon) \|g\|_2^2 + 2\varrho_{2,H}^2 \|N^T g\|_2^2}. \end{aligned}$$

Заметим, что нормы π_χ , $\chi = 1, 2$, являются эллитопическими; поэтому в текущей ситуации можно использовать процедуру синтеза из раздела 3.2, придя к аналогу теоремы 2. Для этого обозначим через $n_1, \dots, n_n$ столбцы N и зададим

$$\overline{M}_j = 8\overline{\mathcal{K}}^2(\epsilon) I_m + 8s^2 \text{Opt}_\infty^2 n_j n_j^T, \quad j \leq m,$$

и

$$\overline{Q} = (8\overline{\mathcal{K}}^2(\epsilon) I_m + 2\varrho_{2,H}^2 N N^T)^{-1/2}.$$

Далее заметим, что единичный шар $\pi_1(\cdot)$ является эллитопом

$$\overline{M} = \{w \in \mathbf{R}^m : \exists r \in [0, 1]^M : w^T \overline{M}_j w \leq \rho_j, \quad j = 1, \dots, M\},$$

а единичный шар π_2 – эллипсоидом $w^T \overline{Q}^{-2} w \leq 1$. Теперь рассмотрим задачу оптимизации

$$(45a) \quad \text{Opt} = \min_{\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho} \left\{ \overline{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho) : \lambda \geq 0, \mu \geq 0, \tau \geq 0, \right. \\ \left. \Theta_1 \succeq 0, \Theta_2 \succeq 0, \text{Tr}(\overline{M}_j \Theta_1) \leq \rho, \quad j = 1, \dots, n, \right. \\ \left. \left[\frac{\sum_\ell \lambda_\ell S_\ell}{\frac{1}{2} B^T \sum_k \mu_k T_k + A^T \left[\sum_i \tau_i h_i h_i^T \right] A + A^T (\Theta_1 + Q \Theta_2 Q^T) A} \right] \Big| \frac{1}{2} B \right] \succeq 0 \right\},$$

где

$$(45b) \quad \begin{aligned} &\overline{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho) = \\ &= \phi_S(\lambda) + 4\phi_T(\mu) + \sum_i \tau_i \psi_H^2(h_i) + \text{Tr}(\Theta_2) + 2\sqrt{2} \ln(4m^2 n) \rho. \end{aligned}$$

Обращаем внимание, что в силу ограничений на Θ_1 и ρ в этой задаче, (Θ_1, ρ) принадлежит конусу $\mathbf{K}$, связанному с эллитопом $\overline{M}$ в роли $\mathcal{W}$ (Предложение 2).

Теорема 3. При заданном допустимом решении $(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho)$ задачи (45) построим матрицы контрастов G_1 и G_2 размерности $m \times m$ следующим образом.

- Для построения G_1 применим вторую часть Предложения 2 к $(\Theta_1, \rho, \overline{M})$ в роли (Θ, ρ, W) , чтобы эффективно вычислить разложение $\Theta_1 = \sum_{i=1}^m \gamma_i g_{1,i} g_{1,i}^T$ с $g_{1,i} \in \overline{M}$ и $\gamma_i \geq 0$, $\sum_i \gamma_i \leq 2\sqrt{2} \ln(4m^2 n) \rho$. Зададим $G_1 = [g_{1,1}, \dots, g_{1,m}]$.
- Для построения G_2 разложим Θ_2 по собственным значениям, т.е. для $\Theta_2 = \Gamma \text{Diag}\{\chi\} \Gamma^T$ зададим $G_2 = [g_{2,1}, \dots, g_{2,m}] = Q \Gamma$.

Заметим, что $\Theta_1 + Q \Theta_2 Q = \sum_i \gamma_i g_{1,i} g_{1,i}^T + \sum_i \chi_i g_{2,i} g_{2,i}^T$.

Для полученной полиэдральной оценки $\hat{w}_{G,H}$ и всех $x_* \in \mathcal{X}$, s -разреженных ν_* и $\xi \in \Xi_\epsilon(G, H)$ выполняется

$$(46) \quad \|\hat{w}_{G,H}(Ax_* + N\nu_* + \xi) - Bx_*\| \leq \bar{f}_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho).$$

Следовательно, ϵ -риск оценки ограничен сверху величиной $f_H(\lambda, \mu, \tau, \Theta_1, \Theta_2, \rho)$ (в силу $\xi \in \Xi_\epsilon(G, H)$ с вероятностью не менее $1 - \epsilon$).

Доказательство аналогично доказательству теоремы 2 и опускается.

3.4. Итоговый результат

Наконец, можно “собрать” полиэдральные оценки из разделов 3.2 и 3.3 в следующей конструкции (ср. [2, раздел 5.1.6]):

Положим

$$\overline{\alpha}(\epsilon) = \sigma(2 \ln[(2n + 2M + 8m)/\epsilon])^{1/2},$$

и пусть $\tilde{H} = [\tilde{h}_1, \dots, \tilde{h}_n] \in \mathbf{R}^{m \times n}$, $\tilde{G} = [\tilde{g}_1, \dots, \tilde{g}_{2m}] \in \mathbf{R}^{m \times 2m}$ и $\overline{H} = [\overline{h}_1, \dots, \overline{h}_M] \in \mathbf{R}^{m \times M}$, $\overline{G} = [\overline{g}_1, \dots, \overline{g}_{2m}] \in \mathbf{R}^{m \times 2m}$ – матрицы контрастов, полученные с помощью процедуры синтеза из разделов 3.2 и 3.3 соответственно. Определим агрегированную оценку $\hat{w}$ для w_* как $\hat{w}(\omega) = B\hat{x}(\omega)$, где $\hat{x}(\omega)$ – x -компонента вектора

$$(\hat{x}(\omega), \hat{\nu}(\omega)) \in \underset{x, \nu}{\text{Argmin}} \left\{ \|\nu\|_1 : x \in \mathcal{X}, \right. \\ \left. \begin{aligned} |\tilde{h}_k^T[N\nu + Ax - \omega]| &\leq \overline{\alpha}(\epsilon) \|\tilde{h}_k\|_2, & k = 1, \dots, n, \\ |\overline{h}_k^T[N\nu + Ax - \omega]| &\leq \overline{\alpha}(\epsilon) \|\overline{h}_k\|_2, & k = 1, \dots, M, \\ |\tilde{g}_i^T[N\nu + Ax - \omega]|_\infty &\leq \overline{\alpha}(\epsilon) \|\tilde{g}_i\|_2, & i = 1, \dots, 2m, \\ |\overline{g}_i^T[N\nu + Ax - \omega]|_\infty &\leq \overline{\alpha}(\epsilon) \|\overline{g}_i\|_2, & i = 1, \dots, 2m \end{aligned} \right\},$$

когда задача разрешима, и $\hat{x}(\omega) = 0$ в противном случае.

Введем обозначение $G = [\tilde{G}, \overline{G}] \in \mathbf{R}^{m \times 4m}$; кроме того, пусть $(\tilde{\lambda}, \tilde{\mu}, \tilde{\gamma})$ – допустимое решение задачи (27) при $H = \tilde{H}$, а $(\overline{\lambda}, \overline{\mu}, \overline{\gamma})$ – допустимое решение задачи (43) при $H = \overline{H}$. Пусть $f_{G,H}$ и $\bar{f}_{G,H}$ заданы формулами (27) и (43) соответственно. Из Предложений 3 и 4 непосредственно следует, что для каждого

s -разреженного ν_* и каждого $x_* \in \mathcal{X}$ справедлива верхняя граница риска

$$(47) \quad \text{Risk}_\epsilon[\hat{w}(\cdot)|\mathcal{X}, \mathcal{N}] \leq \min[f_{G, \tilde{H}}(\tilde{\lambda}, \tilde{\mu}, \tilde{\gamma}), \bar{f}_{G, \bar{H}}(\bar{\lambda}, \bar{\mu}, \bar{\gamma})].$$

Заметим, что полученную оценку можно эффективно оптимизировать по всем параметрам, кроме $\bar{H}$, задав:

- $\tilde{H}$ как (почти) минимум для $\rho[H]$ по $H \in \mathcal{H}$ (23);
- $\tilde{G}$ как результат разложения (Θ_1, Θ_2) -компоненты (почти) оптимального решения задачи (33a), (33b) (см. теорему 2), связанной с $\tilde{H}$;
- $\bar{G}$ как результат разложения (Θ_1, Θ_2) -компоненты (почти) оптимального решения задачи (45) (см. теорему 3), связанной с $\bar{H}$.

3.5. Иллюстративный пример

В данном эксперименте для подтверждения предложенной концепции сравним три оценки x_* : 1) оценку $\hat{x}_{HG}$ с матрицей контрастов $[H, G]$, вычисленной с помощью процедуры из раздела 3.2, 2) оценку $\hat{x}_{IG}$ с контрастом $[\bar{H}, \bar{G}] = [I_m, \bar{G}]$ с $\bar{G}$, полученной с помощью процедуры синтеза из раздела 3.3.2, и 3) “агрегированную” оценку $\hat{x}_{HIG}$ с комбинированным контрастом $[H, I_m, G, \bar{G}]$. Решаем задачи оптимизации (28) и (33a), (33b) для вычисления контрастов H и G оценки $\hat{x}_{HG}$, а также решаем (45) для построения контраста $\bar{G}$ оценки $\hat{x}_{IG}$. Например, при вычислении контраста $\bar{G}$ зададим $\bar{\kappa}(\epsilon) = \sqrt{2} \text{erfcinv}(\frac{\epsilon}{2n})$, где $\text{erfcinv}(\cdot)$ – обратная дополнительная гауссова функция ошибок. При численном решении задачи (45) Θ_1 была задана 0; полученная задача может быть записана в виде

$$(48) \quad \text{Opt} = \min_{\lambda, \mu, \gamma, \Theta} \left\{ \lambda + 4 \sum_{k=1}^p \mu_k + \sum_{j=1}^n \gamma_j + \text{Tr}(\Theta) : \lambda \geq 0, \mu \geq 0, \gamma \geq 0, \Theta \succeq 0, \right. \\ \left. \left[\frac{\lambda I_p}{\frac{1}{2} I_p} \middle| \frac{\frac{1}{2} I_p}{\bar{A}^T \Theta \bar{A} + P^T \text{Diag}\{\mu\} P + \bar{A}^T \text{Diag}\{\gamma\} \bar{A}} \right] \succeq 0 \right\},$$

где $\bar{A} = A/\rho_2$ с $\rho_2 = 2\bar{\kappa}(\epsilon) + \varrho_{2, \bar{H}}$, последующие элементы в Pz равны $z_1, [z_2 - z_1]/h, [z_{i-2} - 2z_{i-1} + z_i]/h^2, 3 \leq i \leq p$, с $h = 2\pi/p$. Соответствующие границы риска получены путем вычисления решений (43). Задачи оптимизации решены с помощью коммерческого решателя Mosek [41] с использованием CVX [26].¹⁰

В иллюстративном примере:

- $m = n = 256$, $q = p = 32$, $N = I_n$, $B = I_p$, а A – случайная матрица размерности $n \times p$ с гауссовыми элементами, удовлетворяющая соотношению $A^T A = I_p$;

¹⁰ Код MATLAB для этого эксперимента доступен в репозитории GitHub по адресу: <https://github.com/ai1-fr/poly-robust>.

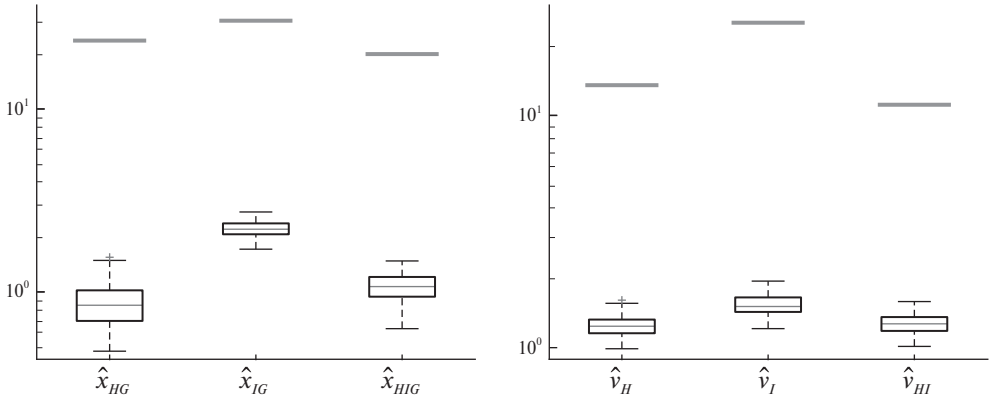


Рис. 1. Левый график: распределения $\|\cdot\|_2$ -ошибок восстановления x_* и теоретические верхние границы для $\text{Risk}_{0,05}$ (жирные горизонтальные линии); правый график: распределения $\|\cdot\|_2$ -ошибок восстановления ν_* и теоретические верхние границы для $\text{Risk}_{0,05}$.

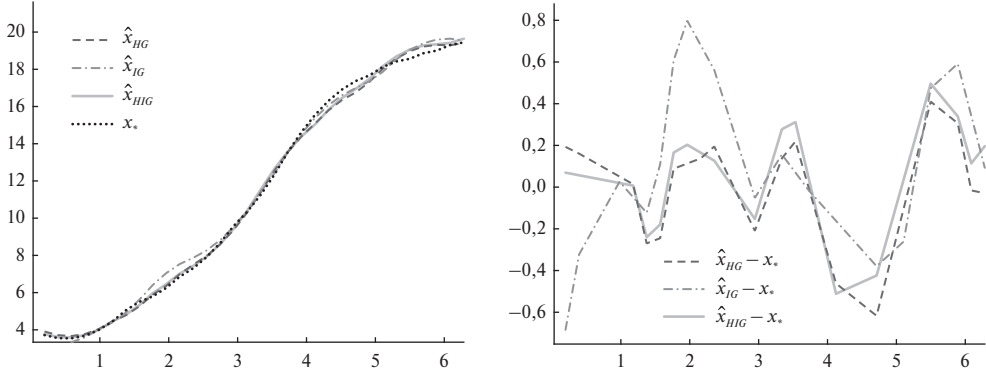


Рис. 2. Типичная реализация сигнала, оценок и ошибок восстановления.

- $\mathcal{X}$ – ограничение на равномерную p -точечную сетку на отрезке $\Delta = [0, 2\pi]$ функций f , удовлетворяющих $|f(0)| \leq 4$, $|f'(0)| \leq 1$ и $|f''(t)| \leq 4$, $t \in \Delta$;
- ошибка восстановления определяется стандартной евклидовой нормой $\|\cdot\|$ на $\mathbf{R}^p$;
- $\xi \sim \mathcal{N}(0, \sigma^2 I_m)$ с $\sigma = 0,1$, $\epsilon = 0,05$ и $s = 8$.

На рис. 1 показаны результаты вычислений. В каждом эксперименте построены $n_S = 100$ восстановлений $\hat{x}_{HG}$, $\hat{x}_{IG}$ и $\hat{x}_{HIG}$ случайно выбранных сигналов $x_* \in \mathcal{X}$ со сгенерированными случайным образом разреженными помехами ν_* . Результаты представлены на левом графике рис. 1. На правом графике даны ошибки восстановления помех ν_* вместе с верхней границей риска Opt_2 (40). На рис. 2 показана типовая реализация сигнала и ошибок восстановления; значения $\|\cdot\|_2$ -ошибок восстановления составляют $\|\hat{x}_{HG} - x_*\|_2 = 1,48 \dots$, $\|\hat{x}_{IG} - x_*\|_2 = 2,02 \dots$ и $\|\hat{x}_{HIG} - x_*\|_2 = 1,43 \dots$, а соответствующая норма сигнала равна $\|x_*\|_2 = 72,2 \dots$.

4. Заключение

В работе проанализированы различные постановки задачи восстановления сигналов по непрямым наблюдениям со случайными и адверсариальными шумами. Рассмотрены процедуры оценки для работы с двумя типами загрязнения – детерминированным возмущением, которое принадлежит заданному (известному) выпуклому множеству, и разреженным (сингулярным) возмущением. Важной особенностью описанных выше *полиэдральных оценок* является их “операционная природа”: они дают “вполне хорошие” оценки сигнала при очень общих предположениях относительно параметров задачи – линейных отображениях и заданных классах сигналов. При этом полиэдральные оценки не задаются аналитическими выражениями и не имеют аналитических формул эффективности. Напротив, соответствующие границы риска и сами оценки доступны численно, как эффективно вычисляемые решения задач выпуклой оптимизации.

ПРИЛОЖЕНИЕ. Граница ошибки для ℓ_1 -восстановления

Условие $\mathbf{Q}_\infty(s, \kappa)$

При заданной наблюдательной матрице N размерности $m \times n$, положительном целом числе $s \leq n$ и $\kappa \in (0, 1/2)$ будем говорить, что матрица H размерности $m \times p$ удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$, если

$$(П.1) \quad \|w\|_\infty \leq \|H^T N w\|_\infty + \frac{\kappa}{s} \|w\|_1 \quad \forall w \in \mathbf{R}^n.$$

Обозначим через y^s вектор, полученный из вектора $y \in \mathbf{R}^n$ обнулением всех элементов, кроме s наибольших по абсолютной величине.

Предложение 5. При заданных N и целом числе $s > 0$ предположим, что H удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$ с $\kappa < \frac{1}{2}$. Тогда для всех $\nu, \hat{\nu} \in \mathbf{R}^n$ таких, что $\|\hat{\nu}\|_1 \leq \|\nu\|_1$, справедливо неравенство

$$(П.2) \quad \|\hat{\nu} - \nu\|_q \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \left[\|H^T N [\hat{\nu} - \nu]\|_\infty + \frac{\|\nu - \nu^s\|_1}{s} \right], \quad 1 \leq q \leq \infty.$$

Доказательство. Введем обозначение $\rho = \|H^T N [\hat{\nu} - \nu]\|_\infty$ и положим $z = \hat{\nu} - \nu$.

1°. Пусть $I \subset \{1, \dots, n\}$ с не более чем s элементами, и пусть $\bar{I}$ – его дополнение в $\{1, \dots, n\}$. Обозначим через x_I вектор, полученный из вектора x обнулением всех элементов с индексами, не принадлежащими I . Тогда

$$\|\hat{\nu}_{\bar{I}}\|_1 = \|\hat{\nu}\|_1 - \|\hat{\nu}_I\|_1 \leq \|\nu\|_1 - \|\hat{\nu}_I\|_1 = \|\nu_I\|_1 + \|\nu_{\bar{I}}\|_1 - \|\hat{\nu}_I\|_1 \leq \|z_I\|_1 + \|\nu_{\bar{I}}\|_1,$$

и поэтому

$$\|z_{\bar{I}}\|_1 \leq \|\hat{\nu}_{\bar{I}}\|_1 + \|\nu_{\bar{I}}\|_1 \leq \|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Отсюда следует, что

$$(П.3) \quad \|z\|_1 = \|z_I\|_1 + \|z_{\bar{I}}\|_1 \leq 2\|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Кроме того, по определению ρ имеем

$$(П.4) \quad \|H^T N z\|_\infty \leq \rho.$$

2°. Поскольку H удовлетворяет $\mathbf{Q}_\infty(s, \kappa)$, получаем

$$\|z\|_{s,1} \leq s\|H^T N z\|_\infty + \kappa\|z\|_1,$$

где $\|z\|_{s,1}$ — ℓ_1 -норма s -мерного вектора, состоящего из s наибольших по абсолютной величине элементов z . В силу (П.4) отсюда следует, что $\|z\|_{s,1} \leq s\rho + \kappa\|z\|_1$. Вместе с очевидным неравенством $\|z_I\| \leq \|z\|_{s,1}$ (напомним, что $\text{Card}(I) = s$) и (П.3) это дает

$$\|z\|_1 \leq 2\|z_I\|_1 + 2\|\nu_{\bar{I}}\|_1 \leq 2s\rho + 2\kappa\|z\|_1 + 2\|\nu_{\bar{I}}\|_1.$$

Поэтому (напомним, что $\kappa \leq \frac{1}{2}$) можно записать

$$(П.5) \quad \|z\|_1 \leq \frac{2s\rho + 2\|\nu_{\bar{I}}\|_1}{1 - 2\kappa}.$$

С другой стороны, поскольку H удовлетворяет условию $\mathbf{Q}_\infty(s, \kappa)$, также имеем

$$\|z\|_\infty \leq \|H^T N z\|_\infty + \frac{\kappa}{s}\|z\|_1.$$

На основании (П.5) и (П.4) отсюда следует, что

$$(П.6) \quad \|z\|_\infty \leq \rho + \frac{\kappa}{s} \frac{2s\rho + 2\|\nu_{\bar{I}}\|_1}{1 - 2\kappa} = (1 - 2\kappa)^{-1} \left[\rho + \frac{\|\nu_{\bar{I}}\|_1}{s} \right].$$

Закключаем, что для всех $1 \leq q \leq \infty$ справедливо неравенство

$$\|z\|_p \leq \|z\|_\infty^{\frac{q-1}{q}} \|z\|_1^{\frac{1}{q}} \leq \frac{(2s)^{\frac{1}{q}}}{1 - 2\kappa} \left[\rho + \frac{\|\nu_{\bar{I}}\|_1}{s} \right]. \quad \square$$

СПИСОК ЛИТЕРАТУРЫ

1. *Juditsky A., Nemirovski A.* On polyhedral estimation of signals via indirect observations // Electronic Journal of Statistics. 2020. V. 14. No. 1. P. 458–502.
2. *Juditsky A., Nemirovski A.* Statistical Inference via Convex Optimization. Princeton: Princeton University Press, 2020.
3. *Tukey J.W.* A survey of sampling from contaminated distributions / Contributions to Probability and Statistics. 1960. P. 448–485.
4. *Huber P.J.* Robust Statistics. New York: Wiley, 1981.

5. Yu C., Yao W. Robust linear regression: A review and comparison // Communications in Statistics – Simulation and Computation. 2017. V. 46. No. 8. P. 6261–6282.
6. Поляк Б.Т., Цыпкин Я.З. Адаптивные алгоритмы оценивания (сходимость, оптимальность, стабильность) // АиТ. 1979. № 3. С. 71–84.
7. Polyak B.T., Tsyppkin Ya.Z. Robust identification // Automatica. 1980. V. 16. No. 1. P. 53–63.
8. Поляк Б.Т., Цыпкин Я.З. Робастные псевдоградиентные алгоритмы адаптации // АиТ. 1980. № 10. С. 91–97.
9. Polyak B.T., Tsyppkin Ya.Z. Optimal and robust methods for unconditional optimization // IFAC Proceedings Volumes. 1981. V. 14. No. 2. P. 519–523.
10. Поляк Б.Т., Цыпкин Я.З. Критериальные алгоритмы стохастической оптимизации // АиТ. 1984. № 6. С. 95–104.
11. Polyak B.T., Tsyppkin Ya.Z. Optimal recurrent algorithms for identification of non-stationary plants // Computers & Electrical Engineering. 1992. V. 18. No. 5. P. 365–371.
12. Chen Y., Caramanis C., Mannor Sh. Robust sparse regression under adversarial corruption // International Conference on Machine Learning. PMLR, 2013. P. 774–782.
13. Balakrishnan S., Du Simon S., Li J., Singh A. Computationally efficient robust sparse estimation in high dimensions // Conference on Learning Theory. PMLR, 2017. P. 169–212.
14. Diakonikolas I., Kong W., Stewart A. Efficient algorithms and lower bounds for robust linear regression // Proceedings of the Thirtieth Annual ACM–SIAM Symposium on Discrete Algorithms. SIAM, 2019. P. 2745–2754.
15. Liu L., Shen Y., Li T., Caramanis C. High dimensional robust sparse regression // International Conference on Artificial Intelligence and Statistics. PMLR, 2020. P. 411–421.
16. Minsker S., Ndaoud M., Wang L. Robust and tuning-free sparse linear regression via square-root slope // SIAM J. Math. Data Sci. 2024. V. 6. No. 2. P. 428–453
17. Foygel R., Mackey L. Corrupted sensing: Novel guarantees for separating structured signals // IEEE Transactions on Information Theory. 2014. V. 60. No. 2. P. 1223–1247.
18. Dalalyan A., Thompson Ph. Outlier-robust estimation of a sparse linear model using l_1 -penalized Huber’s M -estimator // Advances in Neural Information Processing Systems. 2019. V. 32.
19. Bruce A.G., Donoho D.L., Gao H.-Y., Martin R.D. Denoising and robust nonlinear wavelet analysis // Wavelet Applications. SPIE, 1994. V. 2242. P. 325–336.
20. Sardy S., Tseng P., Bruce A. Robust wavelet denoising // IEEE Transactions on Signal Processing. 2001. V. 49. No. 6. P. 1146–1152.
21. Diakonikolas I., Kane D.M. Algorithmic High-Dimensional Robust Statistics. Cambridge: Cambridge University Press, 2023.
22. Juditsky A., Nemirovski A. Near-optimality of linear recovery from indirect observations // Mathematical Statistics and Learning. 2018. V. 1. No. 2. P. 171–225.
23. Donoho D.L. Statistical estimation and optimal recovery // The Annals of Statistics. 1994. V. 22. No. 1. P. 238–270.
24. Juditsky A.B., Nemirovski A.S. Nonparametric estimation by convex programming // The Annals of Statistics. 2009. V. 37. No. 5A. P. 2278–2300.

25. *Juditsky A., Nemirovski A.* Near-optimality of linear recovery in Gaussian observation scheme under $\|\cdot\|_2^2$ -loss // *The Annals of Statistics*. 2018. V. 46. No. 4. P. 1603–1629.
26. *Grant M., Boyd S.* The CVX Users' Guide. Release 2.1, 2014. <https://web.cvxr.com/cvx/doc/CVX.pdf>
27. *Micchelli C.A., Rivlin T.J.* A Survey of optimal recovery / *Optimal Estimation in Approximation Theory*. Micchelli C.A., Rivlin T.J. (Eds.). Boston, MA: Springer, 1977. P. 1–54.
28. *Micchelli C.A., Rivlin T.J.* Lectures on optimal recovery / *Numerical Analysis* Lancaster 1984. *Lecture Notes in Mathematics*. Turner P.R. (Ed.). Berlin–Heidelberg: Springer, 1985. V. 1129. P. 21–93.
29. *Черноусько Ф.Л.* Оценивание фазового состояния динамических систем. М.: Наука, 1988.
30. *Fogel E., Huang Y.-F.* On the value of information in system identification-bounded noise case // *Automatica*. 1982. V. 18. No. 2. P. 229–238.
31. *Граничин О.Н., Поляк Б.Т.* Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003.
32. *Kurzhanski A.B.* Identification – a theory of guaranteed estimates // *From Data to Model*. Willems J.C. (Ed.). Berlin–Heidelberg: Springer, 1989. P. 135–214.
33. *Kurzhanski A., Vályi I.* Ellipsoidal Calculus for Estimation and Control. Boston, MA: Birkhäuser, 1997.
34. *Milanese M., Vicino A.* Optimal estimation theory for dynamic systems with set membership uncertainty: An overview // *Automatica*. 1991. V. 27. No. 6. P. 997–1009.
35. *Schweppe F.C.* Uncertain Dynamic Systems. Englewood Cliffs, NJ: Prentice-Hall, 1973.
36. *Juditsky A., Nemirovski A.* On design of polyhedral estimates in linear inverse problems // *SIAM Journal on Mathematics of Data Science*. 2024. V. 6. No. 1. P. 76–96.
37. *Candes E.J., Tao T.* Decoding by linear programming // *IEEE Transactions on Information Theory*. 2005. V. 51. No. 12. P. 4203–4215.
38. *Bickel P.J., Ritov Ya., Tsybakov A.B.* Simultaneous analysis of Lasso and Dantzig selector // *The Annals of Statistics*. 2009. V. 37. No. 4. P. 1705–1732.
39. *Donoho D.L., Huo X.* Uncertainty principles and ideal atomic decomposition // *IEEE Transactions on Information Theory*. 2001. V. 47. No. 7. P. 2845–2862.
40. *van de Geer S.* Estimation and Testing under Sparsity. Cham: Springer, 2016.
41. Mosek ApS. The MOSEK optimization toolbox for MATLAB manual. Version 8.0, 2015. <http://docs.mosek.com/8.0/toolbox/>

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025