ОБОБЩЕНИЯ РЕГУЛЯРИЗОВАННОГО МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ А. Н. ТИХОНОВА*

В. И. Ерохин — erohin_v_i@mail.ru

В. В. Волков — volkov@fizmat.net

16 октября 2014 г.

Аннотация. Рассматривается задача регуляризованного метода наименьших квадратов А. Н. Тихонова и её обобщения на нормы отличные от евклидовой, в том числе, на полиэдральные нормы. В качестве «инструментальных» приводятся обобщения леммы А. Н. Тихонова о минимальном по некоторой норме решении системы линейных алгебраических уравнений относительно неизвестной матрицы. Рассматривается редукции задач регуляризованного метода наименьших квадратов к задачам математического программирования. В качестве проблем, требующих дальнейшего исследования, указываются задачи построения методов и алгоритмов решения редуцированных задач математического программирования, целевые функции и допустимые области которых построены с использованием полиэдральных векторных норм.

1. Регуляризованный метод наименьших квадратов А. Н. Тихонова

Исходной в настоящем исследовании является задача, сформулированная и исследованная А. Н. Тихоновым [1, 2] и названная им впоследствии [3, 4] регуляризованным методом наименьших квадратов (РМНК). Заметим, что указанная задача не стала широко известной и не получила, как задумывалось её автором, распространения в качестве инструмента обработки экспериментальных данных.

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

Задача РМНК: Пусть существует точная совместная система линейных алгебраических уравнения (СЛАУ)

$$A_0 x = b_0, \tag{1}$$

где $A_0 \in \mathbb{R}^{m\times n}$, $b_0 \in \mathbb{R}^m$, $b_0 \neq 0$, соотношения между размерами A_0 и ее рангом не оговариваются, $x_0 \in \mathbb{R}^n$ — решение системы (1) с минимальной евклидовой нормой (нормальное решение). Численные значения A_0 и b_0 неизвестны, а вместо них заданы приближенные матрица $A \in \mathbb{R}^{m\times n}$ и вектор $b \in \mathbb{R}^m$, $b \neq 0$, такие, что выполняются условия

$$\|A_0 - A\| \leqslant \mu,$$
$$\|b_0 - b\| \leqslant \delta < \|b\|,$$

где $\mu \geqslant 0$ и $\delta \geqslant 0$ — известные параметры, символ $\|\cdot\|$ означает в зависимости от контекста матричную или векторную евклидову норму. Полнота ранга матрицы A и совместность системы $Ax = b$ в общем случае не предполагаются.

Требуется найти $A_1 \in \mathbb{R}^{m\times n}$, $b_1 \in \mathbb{R}^m$, $x_1 \in \mathbb{R}^n$ такие, что $\|A - A_1\| \leqslant \mu$, $\|b - b_1\| \leqslant \delta$, $A_1 x_1 = b_1$, $\|x_1\| \to \min$.

Как было показано в работах [1, 2], вектор x_1 является устойчивым приближением к вектору x_0, т. е. $\lim\limits_{\mu,\delta\to 0} x_1 = x_0$.

Лемма А. Н. Тихонова («основная лемма»), приведённая ниже, в работах [1, 2] использовалась для решения задачи РМНК – построения нормального решения приближенной системы линейных алгебраических уравнений (СЛАУ).

ЛЕММА. *Система уравнений $Ax = b$, рассматриваемая относительно неизвестной матрицы $A \in \mathbb{R}^{m\times n}$ при заданных $x \in \mathbb{R}^n$, $x \neq 0$, $b \in \mathbb{R}^m$, разрешима. Решение $\hat{A}$ этой системы, минимальное по евклидовой норме, единственно и дается формулой $\hat{A} = \frac{bx^\top}{x^\top x}$, причем $\|\hat{A}\| = \frac{\|b\|}{\|x\|}$.*

Результаты, приведенные в работах [1, 2], можно трактовать как редукцию задачи РМНК к следующей задаче математического программирования:

Задача $R_E(\mu, \delta)$:

$$\|x\| \to \min_{\|b - Ax\| = \mu\|x\| + \delta}.$$

УТВЕРЖДЕНИЕ ([1, 2]).

1. Задача $R_E(\mu,\delta)$ имеет решение тогда и только тогда, когда существует хотя бы одна совместная СЛАУ $\breve{A}x = \breve{b}$, для которой $\|\breve{A} - A\| \leqslant \mu$, $\|\breve{b} - b\| \leqslant \delta$.

2. Если решение задачи $R_E(\mu,\delta)$ существует, то оно является единственным при $\mu,\delta \to 0$.

3. Если $x_{\mu\delta}$ — решение задачи $R_E(\mu,\delta)$, то существует единственная СЛАУ $A_1x = b_1$, для которой $x_{\mu\delta}$ является решением с минимальной евклидовой нормой и выполняются условия $\|A_1 - A\| = \mu$, $\|b_1 - b\| = \delta$. При этом A_1 и b_1 однозначно определяются через $x_{\mu\delta}$ по формулам

$$b_1 = b - \frac{\delta}{\|b - Ax_{\mu\delta}\|}(b - Ax_{\mu\delta}), \quad A_1 = A + (b_1 - Ax_{\mu\delta})\frac{x_{\mu\delta}^\top}{x_{\mu\delta}^\top x_{\mu\delta}}.$$

5. $\lim\limits_{\mu,\delta\to 0} x_{\mu\delta} = x_0$.

2. Обобщенная задача РМНК

Как удалось показать, лемма А. Н. Тихонова и задача РМНК допускают обобщение на другие векторные и матричные нормы.

Пусть $\varphi(\cdot)$, $\psi(\cdot)$ и $\xi(\cdot)$ — некоторые векторные нормы, $\|\cdot\|_p$ — норма Гёльдера с показателем $p \geqslant 1$, для n-мерного вектора задаваемая формулой

$$\|x\|_p := \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}},$$

$\|\cdot\|_p^*$ — норма, двойственная к норме Гёльдера с показателем $p \geqslant 1$ (для которой, как известно, справедливо представление $\|x\|_p^* = \|x\|_q$, где $q = p/(1-p)$).

Рассмотрим также два типа в (общем случае только аддитивных [5]) матричных норм. Пусть $\|\cdot\|_{\varphi,\psi}$ — матричная норма, определенная как

$$\|A\|_{\varphi,\psi} := \max_{x\neq 0} \frac{\psi(Ax)}{\varphi(x)},$$

$\|\cdot\|_{\ell_p}$ — гёльдеровая норма с показателем $p \geqslant 1$ для $(m \times n)$-матрицы [5], определенная как

$$\|A\|_{\ell_p} := \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^p\right)^{\frac{1}{p}}.$$

Приведем некоторые примеры.

ПРИМЕР 1. [10]

$$\|A\|_{\ell_\infty} = \lim_{p\to+\infty}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{1p} = \max_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n}}|a_{ij}| = \|A\|_{1,\infty}.$$

В этом случае $\varphi(\cdot) = \|\cdot\|_1$, $\psi(\cdot) = \|\cdot\|_\infty$.

ПРИМЕР 2. Пусть $\varphi(\cdot) = \psi(\cdot) = \|\cdot\|_2$. Тогда $\|A\|_{\varphi,\psi} = \|A\|_{2,2} = \max\limits_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2}$ — спектральная норма матрицы [5]. Заметим, что если $\operatorname{rank} A = 1$, то $\|A\|_{2,2} = \|A\|_{\ell_2} = \left(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}$ — евклидова норма матрицы [10].

ПРИМЕР 3. Пусть $\varphi(\cdot) = \|\cdot\|_\infty$, $\psi(\cdot) = \|\cdot\|_1$, $\operatorname{rank} A = 1$. Тогда

$$\|A\|_{\varphi,\psi} = \|A\|_{\infty,1} = \max_{x\neq 0}\frac{\|Ax\|_1}{\|x\|_\infty} = \|A\|_{\ell_1} = \sum_{i,j}|a_{ij}|.$$

ПРИМЕР 4. Пусть $\varphi(\cdot) = \psi(\cdot) = \|\cdot\|_1$. Тогда $\|A\|_{\varphi,\psi} = \|A\|_{1,1} = \max\limits_{j=1,\ldots,n}\sum\limits_{i=1}^{m}|a_{ij}|$.

ПРИМЕР 5. Пусть $\varphi(\cdot) = \psi(\cdot) = \|\cdot\|_\infty$. Тогда $\|A\|_{\varphi,\psi} = \|A\|_{\infty,\infty} = \max\limits_{i=1,\ldots,m}\sum\limits_{j=1}^{n}|a_{ij}|$.

В терминах перечисленных выше векторных и матричных норм сформулируем ряд задач:

2.1 Обобщённая задача РМНК в $\|\cdot\|_{\varphi,\psi}$-норме

Пусть существует точная совместная СЛАУ вида (1), $x_0 \in \mathbb{R}^n$ — решение указанной системы с минимальной $\varphi(\cdot)$-нормой. Численные значения A_0 и b_0 неизвестны, а вместо них заданы приближенные матрица $A \in \mathbb{R}^{m\times n}$ и вектор $b \in \mathbb{R}^m$, $b \neq 0$, такие, что выполняются условия

$$\begin{gathered}\|A_0 - A\|_{\varphi,\psi} \leqslant \mu,\\ \psi(b_0 - b) \leqslant \delta < \psi(b),\end{gathered}$$

где $\mu \geqslant 0$ и $\delta \geqslant 0$ — известные параметры. Полнота ранга матрицы A и совместность системы $Ax = b$ в общем случае не предполагаются.

Требуется найти $A_1 \in \mathbb{R}^{m\times n}$, $b_1 \in \mathbb{R}^m$, $x_1 \in \mathbb{R}^n$ *такие, что* $\|A - A_1\|_{\varphi,\psi} \leqslant \mu$, $\psi(b - b_1) \leqslant \delta$, $A_1 x_1 = b_1$, $\varphi(x_1) \to \min$.

Будем обозначать указанную задачу символом $Z_{\varphi,\psi}(\mu,\delta)$. Ниже будет показано, что задача $Z_{\varphi,\psi}(\mu,\delta)$ сводится к задаче математического программирования:

$$R_{\varphi,\psi}(\mu,\delta): \qquad \varphi(x) \to \min_{\psi(b-Ax)=\mu\cdot\varphi(x)+\delta} (=: \chi_{\varphi,\psi}).$$

Впервые данный результат (без доказательства), был представлен в материале [6].

Символом $\mathbf{X}_{\varphi,\psi}$ будем обозначать допустимую область задачи $R_{\varphi,\psi}(\mu,\delta)$.

2.2 Обобщённая задача РМНК в $\|\cdot\|_{\ell_p}$-норме

Пусть существует точная совместная СЛАУ вида (1), $x_0 \in \mathbb{R}^n$ — решение указанной системы с минимальной $\|x\|_p^*$-нормой. Численные значения A_0 и b_0 неизвестны, а вместо них заданы приближенные матрица $A \in \mathbb{R}^{m\times n}$ и вектор $b \in \mathbb{R}^m$, $b \neq 0$, такие, что выполняются условия

$$\|A_0 - A\|_{\ell_p} \leqslant \mu,$$
$$\|b_0 - b\|_p \leqslant \delta \leqslant \|b\|_p\,,$$

где $\mu \geqslant 0$ и $\delta \geqslant 0$ — известные параметры. Полнота ранга матрицы A и совместность системы $Ax = b$ в общем случае не предполагаются.

Требуется найти $A_1 \in \mathbb{R}^{m\times n}$, $b_1 \in \mathbb{R}^m$, $x_1 \in \mathbb{R}^n$ *такие, что* $\|A - A_1\|_{\ell_p} \leqslant \mu$, $\|b - b_1\|_p \leqslant \delta$, $A_1 x_1 = b_1$, $\|x_1\|_p^* \to \min$.

Будем обозначать указанную задачу символом $Z_{\ell_p}(\mu,\delta)$. Ниже будет показано, что задача $Z_{\ell_p}(\mu,\delta)$ сводится к задаче математического программирования:

$$R_{\ell_p}(\mu,\delta): \qquad \|x\|_p^* \to \min_{\|b-Ax\|_p=\mu\|x\|_p^*+\delta} \left(=: \chi_{\ell_p}\right).$$

3. «Инструментарий»

Используя определение $\|A\|_{\varphi,\psi}$-нормы, несложно убедиться, что для любой матрицы $A \in \mathbb{R}^{m\times n}$ и любого вектора $x \in \mathbb{R}^n$ выполняется неравенство

$$\psi(Ax) \leqslant \|A\|_{\varphi,\psi} \cdot \varphi(x). \tag{2}$$

ОПРЕДЕЛЕНИЕ 1. Векторной нормой, *двойственной* к норме $\varphi(\cdot)$ относительно скалярного произведения [5], будем называть величину

$$\varphi^*(x) := \max_{y \neq 0} \frac{\left|x^\top y\right|}{\varphi(y)}. \tag{3}$$

Аналог свойства (2) для $\|\cdot\|_{\ell_p}$-нормы дает следующая

ЛЕММА 1. *Для любой матрицы $A \in \mathbb{R}^{m\times n}$ и любого вектора $x \in \mathbb{R}^n$ выполняется неравенство*

$$\|Ax\|_p \leqslant \|A\|_{\ell_p} \cdot \|x\|_p^*. \tag{4}$$

Доказательство. Обозначим символом $a_{i\bullet}$ строку матрицы A с номером i. С учетом введенного обозначения можно записать

$$\|Ax\|_p = \left(\sum_{i=1}^m |a_{i\bullet}x|^p\right)^{1/p}.$$

Но в силу (3), $|a_{i\bullet}x| \leqslant \|a_{i\bullet}\|_p \cdot \|x\|_p^*$, откуда и получаем

$$\|Ax\|_p \leqslant \left(\sum_{i=1}^m \left(\|a_{i\bullet}\|_p\right)^p\right)^{1/p} \cdot \|x\|_p^* = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p\right)^{1/p} \cdot \|x\|_p^* = \|A\|_{\ell_p} \cdot \|x\|_p^*.$$

□

ОПРЕДЕЛЕНИЕ 2. Множеством векторов, *двойственных* к заданному вектору $x \neq 0$ относительно нормы $\varphi(\cdot)$ [5], будем называть множество

$$\mathcal{D}(x, \varphi(\cdot)) := \left\{y \;\middle|\; y^\top x = \varphi^*(y) \cdot \varphi(x) = 1\right\}.$$

ЛЕММА 2. *Для того, чтобы множество $\mathcal{D}(x, \varphi(\cdot))$ состояло из единственного вектора достаточно, чтобы единичный шар нормы $\varphi^*(\cdot)$ являлся строго выпуклым множеством.*

Доказательство. Будем рассуждать «от противного»: пусть единичный шар нормы $\varphi^*(\cdot)$ — строго выпуклое множество, но $\mathcal{D}(x, \varphi(\cdot))$ содержит два различных вектора y и z. Без потери общности можно предположить, что $\varphi(x) = 1$, откуда следует $\varphi^*(y) = \varphi^*(z) = 1$. Пусть $p = (y+z)/2$. Очевидно, что $p^\top x = 1$. В силу строгой выпуклости единичного шара нормы $\varphi^*(\cdot)$ имеем $\varphi^*(p) < 1$. Но тогда $p^\top x/\varphi^*(p) > (\varphi(x) = 1)$, что противоречит вычислению $\varphi(x)$ как нормы, двойственной к $\varphi^*(x)$:

$$\varphi(x) = \max_{y \neq 0} \frac{\left|x^\top y\right|}{\varphi^*(y)}.$$

□

Приведенная ниже теорема может рассматриваться как обобщение леммы А. Н. Тихонова на неевклидовы нормы. Предпосылки, близкие по идеям и технике доказательства результаты были приведены в работе [7]. Непосредственно формулировку и доказательство теоремы можно найти в работах [8, 9], но мы его все-таки приводим для удобства читателей.

ТЕОРЕМА 1. *Система уравнений $Ax = b$, рассматриваемая относительно неизвестной матрицы $A \in \mathbb{R}^{m\times n}$, при заданных $x \in \mathbb{R}^n$, $x \neq 0$, $b \in \mathbb{R}^m$, разрешима. Решение $\hat{A}$ этой системы, минимальное по $\|\cdot\|_{\varphi,\psi}$-норме, допускает представление*

$$\hat{A} = by^\top, \tag{5}$$

причём

$$\|\hat{A}\|_{\varphi,\psi} = \frac{\psi(b)}{\varphi(x)}, \tag{6}$$

где $y \in \mathcal{D}\left(x, \varphi\left(\cdot\right)\right)$.

Доказательство.

1. Покажем, что матрица $\hat{A}$ является решением системы $Ax = b$ при фиксированных векторах $x \neq 0$ и b.

Действительно, $y^\top x = 1$ в силу условия $y \in \mathcal{D}\left(x, \varphi\left(\cdot\right)\right)$, поэтому

$$\hat{A}x = by^\top x \equiv b.$$

2. Покажем справедливость формулы (6).

Действительно, в силу определений $\|\cdot\|_{\varphi,\psi}$-нормы, $\varphi^*\left(\cdot\right)$-нормы и свойства $y \in \mathcal{D}\left(x, \varphi\left(\cdot\right)\right)$ имеем

$$\left\|by^\top\right\|_{\varphi,\psi} = \max_{x\neq 0} \frac{\psi\left(b\cdot y^\top x\right)}{\varphi\left(x\right)} = \psi\left(b\right)\cdot \max_{x\neq 0}\frac{\left|y^\top x\right|}{\varphi\left(x\right)} = \psi\left(b\right)\cdot\varphi^*\left(y\right) = \frac{\psi\left(b\right)}{\varphi\left(x\right)}.$$

3. Покажем, что для любой матрицы $A \in \mathbb{R}^{m\times n}$, являющейся решением системы $Ax = b$, выполняется условие

$$\|A\|_{\varphi,\psi} \geqslant \frac{\psi(b)}{\varphi(x)}.$$

Действительно, поскольку $\|A\|_{\varphi,\psi} = \max\limits_{x\neq 0}\frac{\psi(Ax)}{\varphi(x)}$, то

$$\psi(Ax) \leqslant \|A\|_{\varphi,\psi}\cdot\varphi(x) \Leftrightarrow \psi(b) \leqslant \|A\|_{\varphi,\psi}\cdot\varphi(x) \Leftrightarrow \|A\|_{\varphi,\psi} \geqslant \frac{\psi(b)}{\varphi(x)}.$$

Теорема доказана. $\square$

Замечание. В отличие от «основной леммы» минимальная по $\|\cdot\|_{\varphi,\psi}$-норме матрица $\hat{A}$ в общем случае не является единственной. Можно только утверждать, что при выполнении условий леммы 2 представление $\hat{A}$ вида (5) является единственным.

ТЕОРЕМА 2 ([9]). *Система уравнений $Ax = b$, рассматриваемая относительно неизвестной матрицы $A \in \mathbb{R}^{m\times n}$, при заданных $x \in \mathbb{R}^n$, $x \neq 0$, $b \in \mathbb{R}^m$, разрешима. Решение $\hat{A}$ этой системы, минимальное по $\|\cdot\|_{\ell_p}$-норме, допускает представление*

$$\hat{A} = by^\top, \tag{7}$$

причём

$$\left\|\hat{A}\right\|_{\ell_p} = \frac{\|b\|_p}{\|x\|_p^*}, \tag{8}$$

где $y \in \mathcal{D}\left(x, \|\cdot\|_p^*\right)$.

Доказательство.

1. Покажем, что матрица $\hat{A}$ является решением системы $Ax = b$ при фиксированных векторах $x \neq 0$ и b.

Действительно, $y^\top x = 1$ в силу условия $y \in \mathcal{D}\left(x, \|\cdot\|_p^*\right)$, поэтому

$$\hat{A}x = by^\top x \equiv b.$$

2. Покажем справедливость формулы (8).

Действительно, в силу определений $\|\cdot\|_{\ell_p}$-нормы и свойства $y \in \mathcal{D}\left(x, \|\cdot\|_p^*\right)$ имеем

$$\left\|by^\top\right\|_{\ell_p} = \left(\sum_{i=1}^m \sum_{j=1}^n |b_i y_j|^p\right)^{1/p} = \left(\sum_{i=1}^m \sum_{j=1}^n |b_i|^p \cdot |y_j|^p\right)^{1/p} = \left(\sum_{i=1}^m |b_i|^p \cdot \sum_{j=1}^n |y_j|^p\right)^{1/p} =$$

$$= \|b\|_p \cdot \|y\|_p = \frac{\|b\|_p}{\|x\|_p^*}.$$

3. Покажем, что для любой матрицы $A \in \mathbb{R}^{m\times n}$, являющейся решением системы $Ax = b$, выполняется условие

$$\|A\|_{\ell_p} \geqslant \frac{\|b\|_p}{\|x\|_p^*}.$$

Действительно, в силу леммы 1 и условия $Ax = b$

$$\|Ax\|_p = \|b\|_p \leqslant \|A\|_{\ell_p} \cdot \|x\|_p^* \Rightarrow \|A\|_{\ell_p} \geqslant \frac{\|b\|_p}{\|x\|_p^*}.$$

Теорема доказана. $\square$

З а м е ч а н и е 1. При $1 < p < \infty$ в силу строгой выпуклости $\|\cdot\|_{\ell_p}$-нормы матрица $\hat{A}$ является единственной, что несложно показать рассуждениями «от противного».

З а м е ч а н и е 2. Как несложно заметить, результат, полученный в теореме 2, является частным случаем результата теоремы 1, соответствующим выбору норм $\psi(\cdot) = \|\cdot\|_p$, $\varphi(\cdot) = \|\cdot\|_p^*$. Но это, в свою очередь, означает, что из двух представленных выше задач, названных «Обобщенная задача РМНК в $\|\cdot\|_{\varphi,\psi}$-норме» и «Обобщенная задача РМНК в $\|\cdot\|_{\ell_p}$- норме» самостоятельный интерес представляет только первая задача, которую в дальнейшем будем просто называть «Обобщенная задача РМНК».

4. Свойства обобщенной задачи РМНК

ЛЕММА 3. *Если существует совместная система вида $A_0x = b_0$ такая, что*

$$\|A_0 - A\|_{\varphi,\psi} \leqslant \mu,$$
$$\psi(b_0 - b) \leqslant \delta < \psi(b),$$

то обобщённая задача РМНК имеет решение.

Д о к а з а т е л ь с т в о. Предположим, что существует совместная система вида $A_0x = b_0$ такая, что условия леммы выполняются. В этом случае допустимое множество обобщённой задачи РМНК не пусто. Заметим, что в силу условий леммы допустимое множество обобщённой задачи РМНК компактно по параметрам A_0 и b_0. В то же время, для произвольной точки $x^0 \in \mathbb{R}^n$ найдётся такое число $M > 0$, что

$$\varphi(x) \geqslant M \Rightarrow \varphi(x) \geqslant \varphi(x^0).$$

Следовательно, при решении обобщённой задачи РМНК параметр x можно также рассматривать заданным на компактном множестве $\{x \in \mathbb{R}^n \mid \varphi(x) \leqslant \leqslant M\}$. Таким образом, обобщённая задача РМНК имеет оптимальное решение в силу теоремы Вейерштрасса.

Лемма доказана. $\square$

ЛЕММА 4. *Для любого решения $\{A_1, b_1, x_1\}$ обобщённой задачи РМНК выполняется условие $x^* \neq 0$.*

Д о к а з а т е л ь с т в о. Предположим противное, а именно: пусть решение обобщённой задачи РМНК $\{A_1, b_1, x_1\}$ существует, но $x_1 = 0$. В этом случае, в силу аксиомы невырожденности, $\varphi(x) = 0$. Одновременно это означает, что другого значения x в множестве $\{A_1, b_1, x_1\}$ быть не может. Так как система $A_1 x_1 = b_1$ совместна, то $b_1 = 0$, откуда следует $\psi(b_1 - b) = \psi(b)$, что противоречит условию $\psi(b_1 - b) < \psi(b)$.

Получили противоречие. Лемма доказана. □

ЛЕММА 5. *Пусть $\{\breve{A}, \breve{b}, \breve{x}\}$ — решение обобщённой задачи РМНК. Тогда выполняются условия*

$$\|\breve{A} - A\|_{\varphi,\psi} = \mu,$$
$$\psi(\breve{b} - b) = \delta.$$

Доказательство. Заметим, что в силу леммы 4 $\breve{x} \neq 0$, откуда следует, что множество векторов, двойственных вектору $\breve{x}$ относительно нормы $\varphi(\cdot)$, непусто.

Предположим, что условия леммы не выполняются. Рассмотрим три возможных случая.

Случай A: $\|\breve{A} - A\|_{\varphi,\psi} < \mu$, $\psi(\breve{b} - b) = \delta$.

Пусть $\Delta A = \alpha \breve{b} y^\top$, где y — вектор, двойственный вектору $\breve{x}$ относительно нормы $\varphi(\cdot)$), $\alpha > 0$ — скалярный параметр. Положим $z = \frac{1}{1+\alpha}\breve{x}$. Несложно убедиться, что вектор z принадлежит множеству решений системы $(\breve{A}+\Delta A)x = \breve{b}$. В силу того, что $\|\breve{A}+\Delta A - A\|_{\varphi,\psi} \leqslant \|\breve{A} - A\|_{\varphi,\psi} + \|\Delta A\|_{\varphi,\psi} = \|\breve{A} - A\|_{\varphi,\psi} + \alpha\frac{\psi(\breve{b})}{\varphi(\breve{x})}$, подходящим выбором значения $\alpha > 0$ можно добиться выполнения условия $\|\breve{A} + \Delta A - A\|_{\varphi,\psi} \leqslant \mu$. Таким образом, существует набор $\{\breve{A} + \Delta A, \breve{b}, z\}$, который принадлежит допустимому множеству обобщённой задачи РМНК и, в то же время, $\varphi(z) = \frac{1}{1+\alpha}\varphi(\breve{x}) < \varphi(\breve{x})$, что противоречит предположению об оптимальности $\{\breve{A}, \breve{b}, \breve{x}\}$.

Случай B: $\|\breve{A} - A\|_{\varphi,\psi} = \mu$, $\psi(\breve{b} - b) < \delta$.

Положим $z = \frac{1}{1+\beta}\breve{x}$, $\Delta b = -\frac{\beta}{1+\beta}\breve{b}$, где $\beta > 0$ — скалярный параметр. Заметим, что вектор z принадлежит множеству решений системы $\breve{A}x = \breve{b} + \Delta b$. В силу того, что $\psi(\breve{b} + \Delta b - b) \leqslant \psi(\breve{b} - b) + \psi(\Delta b) = \psi(\breve{b} - b) + \frac{\beta}{1+\beta}\psi(\breve{b})$, подходящим выбором значения $\beta > 0$ можно добиться выполнения условия $\psi(\breve{b} + \Delta b - b) \leqslant \delta$. Таким образом, существует набор $\{\breve{A}, \breve{b} + \Delta b, z\}$, который принадлежит допустимому множеству обобщённой задачи РМНК и, в то же время, $\varphi(z) = \frac{1}{1+\beta}\varphi(\breve{x}) < \varphi(\breve{x})$, что противоречит предположению об оптимальности $\{\breve{A}, \breve{b}, \breve{x}\}$.

Случай C: $\|\breve{A} - A\|_{\varphi,\psi} < \mu$, $\psi(\breve{b} - b) < \delta$.

Положим $\Delta A = \alpha \breve{b} y^\top$, $\Delta b = -\frac{\beta}{1+\beta}\breve{b}$, $z = \frac{1}{(1+\alpha)\cdot(a+\beta)}\breve{x}$, где y — вектор, двойственный вектору $\breve{x}$ относительно нормы $\varphi(\cdot)$). Несложно убедиться, что вектор z принадлежит множеству решений системы $(\breve{A}+\Delta A)x = \breve{b}+\Delta b$. Поскольку

$$\|\breve{A} + \Delta A - A\|_{\varphi,\psi} \leqslant \|\breve{A} - A\|_{\varphi,\psi} + \|\Delta A\|_{\varphi,\psi} = \|\breve{A} - A\|_{\varphi,\psi} + \alpha\frac{\psi(\breve{b})}{\varphi(\breve{x})},$$

$$\psi(\breve{b} + \Delta b - b) \leqslant \psi(\breve{b} - b) + \psi(\Delta b) = \psi(\breve{b} - b) + \frac{\beta}{1+\beta}\psi(\breve{b}),$$

подходящим выбором значений $\alpha > 0$, $\beta > 0$ можно добиться выполнения условий $\|\breve{A} + \Delta A - A\|_{\varphi,\psi} \leqslant \mu$, $\psi(\breve{b} + \Delta b - b) \leqslant \delta$. Таким образом, существует набор $\{\breve{A} + \Delta A, \breve{b} + \Delta b, z\}$, который принадлежит допустимому множеству обобщённой задачи РМНК и, в то же время, $\varphi(z) = \frac{1}{(1+\alpha)\cdot(1+\beta)}\varphi(\breve{x}) < \varphi(\breve{x})$, что противоречит предположению об оптимальности $\{\breve{A}, \breve{b}, \breve{x}\}$.

Во всех трёх случаях получили противоречие.

Лемма доказана. □

ЛЕММА 6. *Если при некоторых $\mu > 0$, $\delta > 0$ уравнение*

$$\psi(b - Ax) - \mu \cdot \varphi(x) = \delta$$

имеет решение $x \neq 0$, то существует матрица $\bar{A}$ и вектор $\bar{b}$ такие, что $\bar{A}x = \bar{b}$, $\|A_0 - A\|_{\varphi,\psi} = \mu$, $\psi(b_0 - b) = \delta$.

Доказательство. Заметим, что в силу условий леммы, неотрицательности и невырожденности векторных норм, получаем $\psi(b - Ax) > 0$. Пусть

$$\breve{b} = b - \frac{\delta}{\psi(b - Ax)}(b - Ax), \quad \breve{A} = A + (\breve{b} - Ax)y^\top,$$

где y — вектор, двойственный вектору x относительно нормы $\varphi(\cdot)$. Несложно убедиться, что $\psi(\breve{b} - b) = \delta$, а вектор x принадлежит множеству решений системы $\breve{A}x = \breve{b}$.

Кроме того, $\|\breve{A} - A\|_{\varphi,\psi} = \mu$. Действительно,

$$\|\breve{A} - A\|_{\varphi,\psi} = \|(\breve{b} - Ax)y^\top\|_{\varphi,\psi} = \frac{\psi(\breve{b} - Ax)}{\varphi(x)}.$$

Но

$$\psi(\breve{b} - Ax) = \psi\left(b - Ax - \frac{\delta}{\psi(b - Ax)}(\breve{b} - Ax)\right) =$$
$$= \left|1 - \frac{\delta}{\psi(b - Ax)}\right| \cdot \psi(b - Ax) = \Big|\psi(b - Ax) - \delta\Big|.$$

В свою очередь, из условий леммы следует, что

$$\psi(b - Ax) - \delta = \mu \cdot \varphi(x) > 0,$$

откуда и получаем

$$\psi(\breve{b} - Ax) = \mu \cdot \varphi(x),$$
$$\|\breve{A} - A\|_{\varphi,\psi} = \frac{\psi(\breve{b} - Ax)}{\varphi(x)} = \mu.$$

Лемма доказана. □

ТЕОРЕМА 3. *Обобщённая задача РМНК имеет непустое множество решений тогда и только тогда, когда имеет решение задача $R_{\varphi,\psi}(\mu,\delta)$. Оптимальные значения целевых функций обеих задач совпадают. Если обобщённая задача РМНК имеет непустое множество решений, то к нему принадлежит решение, полученное следующим образом: x^* — решение задачи $R_{\varphi,\psi}(\mu,\delta)$,*

$$y \in \mathcal{D}\left(x^*, \varphi\left(\cdot\right)\right), \tag{9}$$
$$b^* = b - \frac{\delta}{\psi(b - Ax^*)} \cdot (b - Ax^*), \tag{10}$$
$$A^* = A + (b^* - Ax^*)y^\top. \tag{11}$$

Доказательство.

1. Покажем, что если обобщённая задача РМНК имеет некоторое решение $\{\breve{A}, \breve{b}, \breve{x}\}$, то имеет решение и задача $R_{\varphi,\psi}(\mu,\delta)$, причём

$$\min_{\psi(b-Ax)-\mu\cdot\varphi(x)=\delta} \varphi(x) = \varphi(\breve{x}) > 0. \tag{12}$$

Действительно, пусть набор $\{\breve{A} = A + \Delta A, \breve{b} = b + \Delta b, \breve{x}\}$ — решение обобщённой задачи РМНК. Тогда, в силу формулировки задачи и лемм 4 и 5 выполняются условия

$$\|\Delta A\|_{\varphi,\psi} = \mu,$$
$$\psi(\Delta b) = \delta,$$

множество решений системы $\breve{A}x = \breve{b}$ непусто, а векторы решений этой системы ненулевые.

Покажем, что выполняется условие

$$\psi(b - A\breve{x}) \leqslant \mu \cdot \varphi(\breve{x}) + \delta.$$

Действительно, т. к. $\breve{x}$ — решение системы $\breve{A}x = \breve{b}$, то $b - A\breve{x} = \Delta A\breve{x} - \Delta b$, откуда следует, что

$$\psi(b - A\breve{x}) \leqslant \psi(\Delta A\breve{x}) + \psi(\Delta b) \leqslant \|\Delta A\|_{\varphi,\psi} \cdot \varphi(\breve{x}) + \psi(\Delta b) \leqslant \mu \cdot \varphi(\breve{x}) + \delta.$$

Покажем, что рассматриваемое нестрогое неравенство выполняется как равенство. Для этого предположим противное, а именно, пусть

$$\psi(b - A\breve{x}) < \mu \cdot \varphi(\breve{x}) + \delta.$$

Сформируем вектор h как

$$h = -\frac{\delta}{\psi(b - A\breve{x})}(b - A\breve{x})$$

и матрицу H как $H = (b + h - A\breve{x})y^\top$, где y — вектор, двойственный вектору x^* относительно нормы $\varphi(\cdot)$.

Как несложно заметить, вектор $\breve{x}$ является решением системы $(A+H)x = b + h$. Кроме того,

$$\psi(h) = \delta,$$
$$H = \left(\frac{\psi(b - A\breve{x} - \delta)}{\psi(b - A\breve{x})}(b - A\breve{x})\right)y^\top,$$
$$\|H\|_{\varphi,\psi} = \frac{\psi(b - A\breve{x} - \delta)}{\varphi(\breve{x})}.$$

Но из предположения и последнего условия следует, что $\|H\|_{\varphi,\psi} < \mu$. Таким образом, набор $\{A + H, b + h, \breve{x}\}$ также является решением обобщённой задачи РМНК. Осталось заметить, что условие $\|H\|_{\varphi,\psi} < \mu$ противоречит лемме 5.

Таким образом, условие $\psi(b - A\breve{x}) \leqslant \mu \cdot \varphi(\breve{x}) + \delta$ выполняется как равенство, откуда следует, что задача $R_{\varphi,\psi}(\mu, \delta)$ действительно имеет решение. Кроме того,

$$\min_{\psi(b - Ax) - \mu\cdot\varphi(x) = \delta} \varphi(x) \leqslant \varphi(\breve{x}) = \min_{\breve{A}x = \breve{b}, \|\breve{A} - A\|_{\varphi,\psi} \leqslant \mu, \psi(\breve{b} - b) \leqslant \delta} \varphi(x). \tag{13}$$

2. Покажем, что если задача $R_{\varphi,\psi}(\mu, \delta)$ имеет некоторое решение x^*, то и обобщённая задача РМНК имеет решение $\{A^*, b^*, x^*\}$, причём матрица A^* и вектор b^* могут быть определены по формулам (9)–(11).

Сперва покажем, что объекты, задаваемыми формулами (9)–(11), существуют для любого x^*, являющегося решением задачи $R_{\varphi,\psi}(\mu, \delta)$. Действительно, в силу (12) выполняется условие

$$\varphi(x^*) > 0 \Leftrightarrow x^* \neq 0,$$

из которого следует, что существует вектор y, двойственный вектору x^* относительно нормы $\varphi(\cdot)$. В то же время, из условий $\delta > 0$, $\mu > 0$, $\varphi(x^*) > 0$ и $\psi(b - Ax^*) - \mu \cdot \varphi(x^*) = \delta$ следует, что $\psi(b - Ax^*) > 0$. Таким образом, вектор b^* и матрица A^*, задаваемые формулами теоремы 3, действительно существуют.

Покажем, что выполняются условия

$$\|A^* - A\|_{\varphi,\psi} \leqslant \mu,$$
$$\psi(b^* - b) \leqslant \delta.$$

Действительно, поскольку x^* — решение задачи $R_{\varphi,\psi}(\mu, \delta)$, указанный вектор является решением уравнения из леммы 6. Тогда, в силу леммы 6 и формул (9)–(11), выполняются условия $\|A^* - A\|_{\varphi,\psi} = \mu$, $\psi(b^* - b) = \delta$, являющиеся частным случаем рассматриваемых условий-неравенств.

Таким образом, действительно, если задача $R_{\varphi,\psi}(\mu, \delta)$ имеет непустое множество решений, то набор $\{A^*, b^*, x^*\}$, где x^* принадлежит множеству решений указанной задачи, матрица A^* и вектор b^* построены по формулам (9)–(11), принадлежит допустимому множеству обобщённой задачи РМНК. Следовательно, в силу леммы 3, обобщённая задача РМНК имеет решение. При этом

$$\min_{\breve{A}x = \breve{b}, \|\breve{A} - A\|_{\varphi,\psi} \leqslant \mu, \psi(\breve{b} - b) \leqslant \delta} \varphi(x) \leqslant \varphi(x^*) = \min_{\psi(b - Ax) - \mu \cdot \varphi(x) = \delta} \varphi(x). \tag{14}$$

3. Сопоставляя условия (13) и (14), заключаем, что

$$\min_{\breve{A}x = \breve{b}, \|\breve{A} - A\|_{\varphi,\psi} \leqslant \mu, \psi(\breve{b} - b) \leqslant \delta} \varphi(x) = \min_{\psi(b - Ax) - \mu \cdot \varphi(x) = \delta} \varphi(x).$$

Теорема доказана. □

З а м е ч а н и е 3. Совместность системы $Ax = b$ не является необходимым условием существования решения обобщенной задачи РМНК.

Рассмотрим задачу

$$R'_{\varphi,\psi}(\mu, \delta): \qquad \varphi(x) \to \min_{\psi(b - Ax) \leqslant \mu \cdot \varphi(x) + \delta} \left(=: {\chi'}_{\varphi,\psi}\right).$$

Символом ${\mathbf{X}'}_{\varphi,\psi}$ будем обозначать допустимую область задачи ${R'}_{\varphi,\psi}(\mu, \delta)$.

Справедлива

ТЕОРЕМА 4. *Задачи $R_{\varphi,\psi}(\mu,\delta)$ и $R'_{\varphi,\psi}(\mu,\delta)$ эквивалентны: они разрешимы или не разрешимы одновременно, причем в случае их разрешимости $\chi_{\varphi,\psi}={\chi'}_{\varphi,\psi}$, $\mathbf{X}_{\varphi,\psi}={\mathbf{X}'}_{\varphi,\psi}$.*

Доказательство. 1. Пусть задача $R_{\varphi,\psi}(\mu,\delta)$ имеет решение, т.е. $\mathbf{X}_{\varphi,\psi}\neq\varnothing$. Очевидно, что в этом случае $\mathbf{X}_{\varphi,\psi}\subset{\mathbf{X}'}_{\varphi,\psi}\Rightarrow{\mathbf{X}'}_{\varphi,\psi}\neq\varnothing$, ${\chi'}_{\varphi,\psi}\leqslant\chi_{\varphi,\psi}$.

Предположим теперь, что мощность ${\mathbf{X}'}_{\varphi,\psi}$ выше, чем ${\mathbf{X}'}_{\varphi,\psi}$, т.е.

$$\exists z\in{\mathbf{X}'}_{\varphi,\psi}\left|z\notin\mathbf{X}_{\varphi,\psi}\Rightarrow\psi\left(b-Az\right)<\mu\cdot{\chi'}_{\varphi,\psi}+\delta\right. .$$

Однако в силу условия

$$\psi\left(b-Az\right)<\mu\cdot{\chi'}_{\varphi,\psi}+\delta \tag{15}$$

каждая из величин $\chi_{\varphi,\psi}$, ${\chi'}_{\varphi,\psi}$ может быть уменьшена, что противоречит допущению об её оптимальности. Таким образом, $\mathbf{X}_{\varphi,\psi}=\mathbf{X}_{\varphi,\psi}$, и, следовательно, $\chi_{\varphi,\psi}={\chi'}_{\varphi,\psi}$.

2. Пусть задача ${R'}_{\varphi,\psi}(\mu,\delta)$ имеет решение, а задача $R_{\varphi,\psi}(\mu,\delta)$ решения не имеет. Очевидно, что в этом случае существует $z\in{\mathbf{X}'}_{\varphi,\psi}$ такой, что выполняется условие (15), которое, как уже было показано выше, противоречит предположению об оптимальности значения ${\chi'}_{\varphi,\psi}$.

Теорема доказана. □

5. Обобщенные задачи РМНК в полиэдральных нормах

В п. 2 было дано определение задачи $Z_{\varphi,\psi}(\mu,\delta)$ — обобщённой задачи РМНК в $\|\cdot\|_{\varphi,\psi}$-норме. При выборе в качестве норм $\varphi(\cdot)$, $\psi(\cdot)$ норм Гёльдера с показателем $p=1,\infty$, в силу теорем 3 и 4 получаем следующий набор обобщенных задач РМНК и их редукций к задачам математического программирования, допустимые области и целевые функции которых построены с использованием *полиэдральных* [5] норм (используемый ниже знак «$\mapsto$» — синоним фразы «сводится к»):

$$Z_{1,1}(\mu,\delta)\mapsto\|b-Ax\|_1\leqslant\mu\cdot\|x\|_1+\delta,\ \|x\|_1\to\min, \tag{16}$$

$$Z_{1,\infty}(\mu,\delta)\mapsto\|b-Ax\|_\infty\leqslant\mu\cdot\|x\|_1+\delta,\ \|x\|_1\to\min, \tag{17}$$

$$Z_{\infty,1}(\mu,\delta)\mapsto\|b-Ax\|_1\leqslant\mu\cdot\|x\|_\infty+\delta,\ \|x\|_\infty\to\min, \tag{18}$$

$$Z_{\infty,\infty}(\mu,\delta)\mapsto\|b-Ax\|_\infty\leqslant\mu\cdot\|x\|_\infty+\delta,\ \|x\|_\infty\to\min. \tag{19}$$

Задачи (16)-(19), несмотря на кажущуюся простоту формулировок, не имеют в общем случае очевидных методов и алгоритмов решения. Их поиск и построение с использованием методов негладкого анализа и недифференцируемой оптимизации вполне может быть предметом дальнейших исследований.

ГИПОТЕЗА 1. *Допустимые области задач* (16)-(19) *в общем случае не являются выпуклыми множествами.*

ГИПОТЕЗА 2. *Если* $\forall\ \widehat{x} \in \operatorname{Argmin} ||b - Ax||_r \Rightarrow ||b - A\widehat{x}||_r \leqslant \mu \cdot ||\widehat{x}||_s + \delta$, *где* $r, s = 1, \infty$, *то допустимое множество задачи* $Z_{s,r}(\mu, \delta)$ *ограничено и является многогранником, в общем случае невыпуклым. Тем не менее, в этом случае задача* $Z_{s,r}(\mu, \delta)$ *может быть сведена к задаче линейного программирования.*

ГИПОТЕЗА 3. *Если* $\forall\ \widehat{x} \in \operatorname{Argmin} ||b - Ax||_r \Rightarrow ||b - A\widehat{x}||_r > \mu \cdot ||\widehat{x}||_s + \delta$, *где* $r, s = 1, \infty$, *то допустимое множество задачи* $Z_{s,r}(\mu, \delta)$ *неограничено и является дополнением пространства* $\mathbb{R}^n$ *до в общем случае невыпуклого многогранника, заданного условием* $||b - Ax||_r \geqslant \mu \cdot ||x||_s + \delta$.

ЛИТЕРАТУРА

1. *Тихонов А. Н.* О приближенных системах линейных алгебраических уравнений // Ж. вычисл. матем. и матем. физики, 1980. Т. 20. №6. С. 1373–1383.

2. *Тихонов А. Н.* О нормальных решениях приближенных систем линейных алгебраических уравнений // Доклады АН СССР, 1980. Т. 254. №3. С. 549–554.

3. *Тихонов А. Н.* О методах автоматизации обработки наблюдений // Вестн. АН СССР, 1983. №1. С. 14–25.

4. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. М.: Наука, 1986. – 288 с.

5. *Хорн Р., Джонсон Ч.* Матричный анализ. М.: Мир, 1989. – 655 с.

6. *Ерохин В. И.* Лемма А.Н. Тихонова и ее обобщения // Тихонов и современная математика: Обратные и некорректно поставленные задачи: Тез. докл. Междунар. конф. - М.: ВМиК МГУ, 2006. - С. 52-53.

7. *Ерохин В. И.* Оптимальная матричная коррекция и регуляризация несовместных линейных моделей // Дискретн. анализ и исслед. опер., 2002. Сер. 2. Т. 9. № 2. С. 41–77.

8. *Горелик В. А., Ерохин В. И., Печенкин Р. В.* Минимаксная матричная коррекция несовместимых систем линейных алгебраических уравнений с блочными матрицами коэффициентов // Изв. РАН. ТИСУ, 2006. № 5. С. 52-62.

9. *Горелик В. А., Ерохин В. И., Печенкин Р. В.* Численные методы коррекции несобственных задач линейного программирования и структурных систем уравнений. М.: ВЦ РАН, 2006. – 153 с.

10. *Икрамов Х. Д.* Задачник по линейной алгебре. М.: Наука, 1975. – 320 с.