ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Июль—декабрь 2005. Серия 2. Том 12, № 2, 3–22

УДК 512.643.8:519.85

ОПТИМАЛЬНАЯ МАТРИЧНАЯ КОРРЕКЦИЯ НЕСОВМЕСТНЫХ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ С БЛОЧНЫМИ МАТРИЦАМИ КОЭФФИЦИЕНТОВ*)

В. А. Горелик, В. И. Ерохин, Р. В. Печенкин

Рассматриваются задачи матричной коррекции матриц (расширенных матриц) несовместных систем линейных алгебраических уравнений с блочной структурой с критериями качества коррекции двух типов: по минимуму суммы квадратов взвешенных евклидовых норм блоков матрицы коррекции и по минимуму квадрата максимальной взвешенной евклидовой нормы блока матрицы коррекции. Задачи обоих типов исследуются как с условием неотрицательности решения скорректированной системы, так и без него. Задачи первого типа сведены к вспомогательным задачам минимизации сумм дробно-квадратичных функций в зависимости от вида исходной задачи, содержащей ограничения в виде системы линейных неравенств или нет. Для целевых функций вспомогательных задач аналитически получены частные производные первого и второго порядка, что позволяет проводить безусловную минимизацию методом Ньютона, а условную — градиентными методами.

Задачи второго типа сведены к вспомогательным задачам поиска минимакса на некотором наборе дробно-квадратичных функций. Они также либо содержат ограничения в виде системы линейных неравенств, либо являются безусловными. Для решения указанных задач предлагается использовать разработанную В.Ф. Демьяновым и В.Н. Малоземовым минимаксную версию метода наискорейшего спуска или её модификацию, учитывающую наличие ограничений в виде системы линейных неравенств.

Для задач безусловной матричной коррекции по минимуму суммы квадратов взвешенных евклидовых норм показано, что необходимым условием их разрешимости является полнота столбцевого ранга корректируемой блочной матрицы.

Все рассмотренные задачи матричной коррекции проиллюстрированы численными примерами.

*) Исследование выполнено при финансовой поддержке гранта по президентской программе поддержки ведущих научных школ РФ (проект НШ 2094.2003.1).

Введение

Проблема коррекции всех коэффициентов (матричная коррекция) несовместных систем линейных алгебраических уравнений и связанные с ней задачи — обобщённый метод наименьших квадратов, несобственные задачи линейного программирования — исследовались в различных постановках (см., например [2, 4, 7, 8]) и в том числе с учетом фиксации (освобождения от коррекции) различных комбинаций строк и столбцов матриц (расширенных матриц) коэффициентов исследуемых линейных систем [5]. Однако все известные результаты не применимы к коррекции линейных систем, матрицы которых имеют блочную структуру. В данной статье получены первые результаты.

Пусть дана несовместная система линейных алгебраических уравнений

$$Ax = b, \tag{1}$$

которая, возможно, дополняется условием неотрицательности решения

$$x \geqslant 0, \tag{2}$$

где $x \in \mathbb{R}^N$, $b \in \mathbb{R}^M$ и $A \in \mathbb{R}^{M\times N}$, причём матрица A имеет следующую блочную структуру:

$$A = \left[\begin{array}{c|c|c|c} \multicolumn{4}{c}{A_0} \\ \hline A_1 & 0 & \cdots & 0 \\ \hline 0 & A_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_k \end{array}\right] \tag{3}$$

где $A_0 \in \mathbb{R}^{m_0\times N}$, $A_i \in \mathbb{R}^{m_i\times n_i}$, $1 \leqslant i \leqslant k$, причём $\sum\limits_{i=0}^{k} m_i = M$ и $\sum\limits_{i=1}^{k} n_i = N$. Для последующих выкладок полезно ввести в соответствии с (3) блочные представления для векторов x и b:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix},$$

где $x_j \in \mathbb{R}^{n_j}$, $1 \leqslant j \leqslant k$, $b_i \in \mathbb{R}^{m_i}$, $1 \leqslant i \leqslant k$. Подсистему

$$A_0 x = b_0 \tag{4}$$

будем полагать совместной.

В настоящей статье исследуются задачи нахождения матриц $H_i \in \mathbb{R}^{m_i \times n_i}$ и векторов $h_i \in \mathbb{R}^{m_i}$ таких, что системы

$$\left[\begin{array}{c|c|c|c} \multicolumn{4}{c}{A_0} \\ \hline A_1+H_1 & 0 & \cdots & 0 \\ \hline 0 & A_2+H_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_k+H_k \end{array}\right] \cdot x = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix}, \tag{5}$$

$$\left[\begin{array}{c|c|c|c} \multicolumn{4}{c}{A_0} \\ \hline A_1+H_1 & 0 & \cdots & 0 \\ \hline 0 & A_2+H_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_k+H_k \end{array}\right] \cdot x = \begin{bmatrix} b_0 \\ b_1+h_1 \\ \vdots \\ b_k+h_k \end{bmatrix}, \tag{6}$$

возможно, дополненные условием (2), являются совместными, а элементы матриц H_i и векторов h_i удовлетворяют естественному для прикладных задач требованию «малости». Это требование можно записать в виде

$$\sum_{i=1}^{k} \left\| L_i H_i R_i \right\|_E^2 \to \min, \tag{7}$$

$$\sum_{i=1}^{k} \left\| L_i \left[H_i \quad -h_i \right] R_i \right\|_E^2 \to \min, \tag{8}$$

$$\max_{i=1,\ldots,k} \left\{ \left\| L_i H_i R_i \right\|_E^2 \right\} \longrightarrow \min \tag{9}$$

или

$$\max_{i=1,\ldots,k} \left\{ \left\| L_i \left[H_i \quad -h_i \right] R_i \right\|_E^2 \right\} \to \min\,. \tag{10}$$

где $L_i \in \mathbb{R}^{n_i \times m_i}$, $R_i \in \mathbb{R}^{(n_i \times n_i)}$ или $R_i \in \mathbb{R}^{(n_i+1)\times(n_i+1)}$ — некоторые невырожденные матрицы при любом $i = 1, ..., k$, а $\| \cdot \|_E$ — евклидова матричная норма.

Структура корректируемых матриц в системах (5), (6) может быть интерпретирована следующим образом. Имеется система, состоящая из k подсистем (например, корпорация из k предприятий). Система в целом должна удовлетворять некоторым условиям устойчивости или координации функционирования подсистем. Эти условия связывают все переменные, записываются в виде (4) и не могут быть подвергнуты коррекции, т. е. являются жёсткими. Коэффициенты же подсистем могут корректироваться. При этом случай (5) может быть интерпретирован как коррекция технологических коэффициентов подсистем, а случай (6) — как одновременная коррекция технологических коэффициентов и ресурсов подсистем. Формулы (7)-(10) означают, что требуется найти минимальные матрицы коррекции, обеспечивающие совместность всех условий исследуемой системы.

Постановки задач (7)–(10) могут быть дополнены требованием коррекции матрицы A_0 (или части ее коэффициентов). Но такие задачи имеют другую содержательную интерпретацию, и они не могут быть решены с помощью принятого в настоящей статье подхода, основанного на использовании явного параметрического представления для множества решений системы (4). Поэтому в данной статье такие задачи не рассматривается.

Задачи (7)–(10) можно рассматривать как обобщения на случай линейных систем с блочной структурой задач многопараметрической (матричной) коррекции несовместных систем линейных алгебраических уравнений и несобственных задач линейного программирования общего вида по минимуму евклидовой нормы, исследованных в работах [2, 4, 5, 7, 8].

1. Сведение задач (7) и (8) к задачам безусловной минимизации

Заметим, что в соответствии со структурой вида (3) матрицы A подсистема (4) является недоопределённой системой линейных алгебраических уравнений и в общем случае имеет бесконечное множество решений. Это множество будем обозначать через $X(A_0, b_0) = \{x \in \mathbb{R}^n | A_0 x_0 \equiv b_0\}$. Для последующих выкладок окажется полезной следующая параметризация множества $X(A_0, b_0)$ (см., например, [1]):

$$X(A_0, b_0) = \{x = \hat{x} + Py\}, \qquad (11)$$

где

$$\hat{x} = A_0^+ b_0, \qquad (12)$$

$$P = I - A_0^+ A_0, \qquad (13)$$

I — единичная матрица порядка N, $A_0^+ \in \mathbb{R}^{N \times m_0}$ — матрица, *псевдообратная** (обобщённая обратная по Муру–Пенроузу [3]) к матрице A и

*Операцию псевдообращения мы будем применять и к векторам, считая их матрицами, состоящими из одного столбца.

$y \in \mathbb{R}^N$ — произвольный вектор. Заметим, что любой вектор x, построенный по формуле (11), является решением подсистемы (4), и наоборот, для любого решения подсистемы (4) справедливо представление (11). Блочные представления матриц A_0, P и векторов $\hat{x}$, x, естественным образом связанные с представлением (3), запишем в виде

$$A_0^+ = \begin{bmatrix} B_1 \\ \vdots \\ B_k \end{bmatrix}, \quad P = \begin{bmatrix} P_1 \\ \vdots \\ P_k \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} \hat{x_1} \\ \vdots \\ \hat{x_k} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \tag{14}$$

где $B_i \in \mathbb{R}^{n_i \times m_0}$, $P_i \in \mathbb{R}^{n_i \times N}$, $\hat{x}_i = B_i b_0$, $x_i = \hat{x}_i + P_i y$, $i = 1, \ldots, k$.

Рассмотрим коррекцию i-й подсистемы системы (1), $1 \leqslant i \leqslant k$. Систему (5) можно переписать как совокупность k систем линейных алгебраических уравнений с неизвестной матрицей H_i:

$$(A_i + H_i)\, x_i = b_i \Leftrightarrow H_i x_i = b_i - A_i x_i. \tag{15}$$

Поскольку в силу сделанных выше предположений матрицы L_i и R_i являются квадратными и невырожденными, эквивалентные преобразования системы (15) можно продолжить следующим образом:

$$H_i x_i = b_i - A_i x_i \Leftrightarrow L_i H_i R_i \cdot R_i^{-1} x_i = L_i (b_i - A_i x_i). \tag{16}$$

Согласно лемме из [9] решение уравнения (16) относительно неизвестной матрицы $L_i H_i R_i$, обладающее минимальной евклидовой нормой, существует при любом $x_i \neq 0$, единственно и задаётся формулой

$$L_i \hat{H}_i R_i = L_i (b_i - A_i x_i)(R_i^{-1} x_i)^+, \tag{17}$$

причём

$$\left\| L_i \hat{H}_i R_i \right\|_E = \frac{\| L_i (b_i - A_i x_i) \|}{\left\| R_i^{-1} x_i \right\|}, \tag{18}$$

где $\| \cdot \|$ — евклидова векторная норма. В свою очередь, из (17)–(18) следует, что единственным решением уравнения (15) относительно неизвестной матрицы H_i, обладающим минимальной взвешенной евклидовой нормой при некотором $x_i \neq 0$ является матрица

$$\hat{H}_i = (b_i - A_i x_i)(R_i^{-1} x_i)^+ R_i^{-1}, \tag{19}$$

для взвешенной евклидовой нормы которой будет справедлива формула (18). Таким образом, задачу (7) можно свести к задаче

$$\sum_{i=1}^{k} \frac{\|L_i(b_i - A_i x_i)\|^2}{\left\|R_i^{-1} x_i\right\|^2} \to \min_{x_i \neq 0,\, i=1,\ldots,k}. \tag{20}$$

Выполнив некоторые преобразования формулы (18) с использованием формул (11), (12) и представлений (14), получим

$$\frac{\|L_i(b_i - A_i x_i)\|}{\left\|R_i^{-1} x_i\right\|} = \frac{\|\tilde{b}_i - \tilde{A}_i y\|}{\|q_i + \tilde{P}_i y\|}, \tag{21}$$

где

$$\tilde{b}_i = L_i(b_i - A_i \hat{x}_i), \tag{22}$$

$$\tilde{A}_i = L_i A_i P_i, \tag{23}$$

$$q_i = R_i^{-1} \hat{x}_i, \tag{24}$$

$$\tilde{P}_i = R_i^{-1} P_i. \tag{25}$$

Тогда задача (20) может быть представлена в виде

$$f(y) = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \tilde{A}_i y\|^2}{\|q_i + \tilde{P}_i y\|^2} = \sum_{i=1}^{k} \varphi_i(y) \to \min_{q_i + \tilde{P}_i y \neq 0,\, i=1,\ldots,k}. \tag{26}$$

Заметим, что задача (26) не является задачей безусловной минимизации в силу условий вида $q_i + \tilde{P}_i y \neq 0$. Тем не менее легко показать, что любой численный метод минимизации, снабжённый дополнительными проверками во вспомогательных процедурах типа одномерного поиска и гарантирующий убывание функции $f(y)$ на каждой итерации, не приведёт к точкам разрыва $f(y)$ при условии, что стартовая точка для $f(y)$ является допустимой. Это означает, что в процессе минимизации функцию $f(y)$ можно считать непрерывной, а задачу (26) — задачей безусловной минимизации.

Итак, задача (7) сведена к задаче (26). Предположим, что $y^* \in \mathbb{R}^N$ — точка минимума задачи (26). Тогда, вычислив вектор $x^* \in \mathbb{R}^N$ по формуле $x^* = \hat{x} + P y^*$, в соответствии с (19) можно построить оптимальные для задачи (7) матрицы коррекции H_i^*, $1 \leqslant i \leqslant k$, по формуле $H_i^* = \left(b_i - A_i x_i^*\right)\left(R_i^{-1} x_i^*\right)^+ R_i^{-1}$. При этом в силу приведённых выше выкладок

$$x^* \in X(A_0, A_1 + H_1^*, \ldots, A_k + H_k^*, b_0, b_1 \ldots, b_k),$$

т. е. вектор x^* принадлежит множеству решений скорректированной системы вида (5).

Обратимся к исследованию задачи (8). Систему (6) можно записать в виде системы линейных алгебраических уравнений с неизвестной матрицей $[H_i \;\; -h_i]$:

$$(A_i + H_i)x_i = b_i + h_i \Leftrightarrow [H_i \;\; -h_i] \cdot \begin{bmatrix} x_i \\ 1 \end{bmatrix} = b_i - A_i x_i. \tag{27}$$

Поскольку по предположению матрицы L_i и R_i являются квадратными и невырожденными, эквивалентные преобразования системы (27) можно продолжить следующим образом:

$$[H_i \;\; -h_i] \cdot \begin{bmatrix} x_i \\ 1 \end{bmatrix} = b_i - A_i x_i \Leftrightarrow L_i[H_i \;\; -h_i]R_i \cdot R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} = L_i(b_i - A_i x_i). \tag{28}$$

Согласно лемме из [9] решение уравнения (28) относительно неизвестной матрицы, обладающее минимальной евклидовой нормой, существует при любом векторе x_i, единственно и задается следующим образом

$$L_i[\hat{H}_i \;\; -\hat{h}_i]R_i = L_i(b_i - A_i x_i) \cdot \left(R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right)^+, \tag{29}$$

причём

$$\|L_i[\hat{H}_i \quad -\hat{h}_i]R_i\|_E = \frac{\|L_i(b_i - A_i x_i)\|}{\left\| R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right\|}. \tag{30}$$

В свою очередь из формул (29)–(30) следует, что единственным решением уравнения (27) относительно неизвестной матрицы $[H_i \;\; -h_i]$, обладающим минимальной взвешенной евклидовой нормой при некотором x_i, является матрица

$$[\hat{H}_i \quad -\hat{h}_i] = (b_i - A_i x_i) \cdot \left(R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right)^+ \cdot R_i^{-1}, \tag{31}$$

для взвешенной евклидовой нормы которой справедлива формула (30). Таким образом, задача (8) сведена к задаче

$$\sum_{i=1}^{k} \frac{\|L_i(b_i - A_i x_i)\|^2}{\left\| R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right\|^2} \to \min_{x_i,\, i=1,\ldots,k}. \tag{32}$$

Преобразовав (30) с использованием (11), (12) и (14), получим

$$\frac{\|L_i(b_i - A_i x_i)\|}{\left\| R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right\|} = \frac{\|\tilde{b}_i - \tilde{A}_i y\|}{\|\breve{q}_i + \breve{P}_i y\|}, \tag{33}$$

где при любом i, $1 \leqslant i \leqslant k$, для $\tilde{b}_i$ и $\tilde{A}_i$ справедливы формулы (22) и (23) соответственно,

$$\breve{q}_i = \mathcal{R}_i \hat{x}_i + \mathfrak{r}_i, \tag{34}$$

$$\breve{P}_i = \mathcal{R}_i P_i, \tag{35}$$

и

$$R_i^{-1} = [\mathcal{R}_i \quad \mathfrak{r}_i], \quad \mathcal{R}_i \in \mathbb{R}^{(n_i+1)\times n_i}, \quad \mathfrak{r}_i \in \mathbb{R}^{n_i+1}. \tag{36}$$

Тогда задача (32) может быть представлена в виде

$$\tilde{f}(y) = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \tilde{A}_i y\|^2}{\|\breve{q}_i + \breve{P}_i y\|^2} = \sum_{i=1}^{k} \tilde{\varphi}_i(y) \to \min. \tag{37}$$

Заметим, что задача (37) очень похожа на задачу (26). Так, целевые функции обеих задач имеют один и тот же вид. Отличие заключается в том, что задача (37) в строгом смысле является задачей безусловной минимизации. Приведём обоснование этого утверждения. Из (11)–(14) и (34)–(36) следует, что

$$\left\| R_i^{-1} \begin{bmatrix} x_i \\ 1 \end{bmatrix} \right\| \equiv \left\| \breve{q}_i + \breve{P}_i y_i \right\|. \tag{38}$$

Но в силу невырожденности матриц R_i, $1 \leqslant i \leqslant k$, левая часть тождества (38) отлична от нуля при любых значениях x_i. Следовательно, правая часть тождества (38) отлична от нуля при любых значениях y_i. Отсюда следует непрерывность функции $\tilde{f}(y)$ в $\mathbb{R}^N$.

Итак, задача (8) сведена к задаче (37). Предположим, что $y^* \in \mathbb{R}^N$ — точка минимума задачи (37). Тогда в соответствии с (11) $x^* = \hat{x} + Py^*$ и в соответствии с (31) можно построить оптимальные для задачи (8) матрицы коррекции $[H_i^* \ \ -h_i^*]$, $1 \leqslant i \leqslant k$, по формуле

$$\left[H_i^* \ \ -h_i^*\right] = \left(b_i - A_i x_i^*\right) \cdot \left(R_i^{-1} \begin{bmatrix} x_i^* \\ 1 \end{bmatrix} \right)^+ \cdot R_i^{-1}.$$

При этом из приведённых выше выкладок следует, что

$$x^* \in X(A_0, A_1 + H_1^*, \ldots, A_k + H_k^*, b_0, b_1 + h_1^*, \ldots, b_k + h_k^*),$$

т. е. x^* принадлежит множеству решений скорректированной системы вида (6).

Лемма 1. *Если* $\operatorname{rank} A < N$, *то минимум в задачах* (7) *и* (8) *не достигается.*

Доказательство. Пусть $\operatorname{rank} A < N$. Следовательно, существует

$$z = \begin{bmatrix} z_1 \in \mathbb{R}^{n_1} \\ \vdots \\ z_k \in \mathbb{R}^{n_k} \end{bmatrix} \quad \text{такое, что} \quad z \neq 0 \text{ и } Az = 0 \Leftrightarrow \begin{cases} A_0 z = 0, \\ A_1 z = 0, \\ \quad \cdots \\ A_k z = 0. \end{cases} \tag{39}$$

Предположим, что минимум в задаче (7) или (8) достигается. Тогда минимум достигается и в задаче (26) или (37). Пусть $y^* \in \mathbb{R}^N$ — некоторая точка минимума задачи (26). Заметим, что $f(y^*) > 0$, так как при $f(y^*) = 0$ все матрицы коррекции H_i имеют нулевую евклидову норму, а значит в силу аксиомы невырожденности сами являются нулевыми, что противоречит предположению о несовместности системы (1). Далее, используя (13), (14), (23)–(26) и (39), нетрудно видеть, что

$$\lim_{\alpha \to \infty} f(\bar{y} + \alpha z) = \inf_{y \in R^N} f(y) = 0, \tag{40}$$

где $\bar{y} \in \mathbb{R}^N$ — произвольный вектор, не являющийся точкой разрыва функции $f(y)$. Действительно,

$$f(\bar{y} + \alpha z) = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \hat{A}_i(\bar{y} + \alpha z)\|^2}{\|q_i + \tilde{P}_i(\bar{y} + \alpha z)\|^2} = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \tilde{A}_i \bar{y} + \alpha L_i A_i P_i z\|^2}{\|q_i + \tilde{P}_i \bar{y} + \alpha R_i^{-1} P_i z\|^2}.$$

Но $P_i z = z_i - B_i A_0 z = z_i$. Поэтому

$$f(\bar{y} + \alpha z) = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \tilde{A}_i \bar{y}\|^2}{\|q_i + \tilde{P}_i \bar{y} + \alpha R_i^{-1} P_i z\|^2}. \tag{41}$$

Теперь заметим, что в силу невырожденности матриц R_i и условия $z \neq 0$ существует i такое, что

$$R_i^{-1} z_i \neq 0. \tag{42}$$

Учитывая (41), (42) и неотрицательность функции $f(y)$, получаем соотношение (40). Далее, предполагая, что $y^* \in \mathbb{R}^N$ — некоторая точка минимума задачи (37), имеем $\tilde{f}(y^*) > 0$, а используя (13), (14), (22), (23), (34)–(37) и (39), убеждаемся в том, что

$$\lim_{\alpha\to\infty} f(\bar{y}+\alpha z) = \inf_{y\in R^N} f(y) = 0, \tag{43}$$

где $\bar{y} \in \mathbb{R}^N$ — произвольный вектор. Действительно,

$$f(\bar{y}+\alpha z) = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \hat{A}_i(\bar{y}+\alpha z)\|^2}{\|q_i + \tilde{P}_i(\bar{y}+\alpha z)\|^2} = \sum_{i=1}^{k} \frac{\|\tilde{b}_i - \tilde{A}_i\bar{y} + \alpha L_i A_i P_i z\|^2}{\|q_i + \tilde{P}_i\bar{y} + \alpha R_i^{-1} P_i z\|^2}$$
$$= \frac{\|\tilde{b}_i - \tilde{A}_i\bar{y}\|^2}{\|q_i + \tilde{P}_i\bar{y} + \alpha R_i^{-1} P_i z\|^2}. \tag{44}$$

Но подматрица $\mathcal{R}_i \in \mathbb{R}^{(n_i+1)\times n_i}$ матрицы $R_i \in \mathbb{R}^{(n_i+1)\times(n_i+1)}$ имеет полный столбцевой ранг (в противном случае матрица R_i^{-1} не существовала бы). Учитывая это условие, а также условие $z \neq 0$, заключаем, что существует такое i, что выполняется (42).

Пользуясь (44), (42) и неотрицательностью функции $\tilde{f}(y)$, получаем соотношение (43). Но условие $f(y^*) > 0$ противоречит (40), а условие $\tilde{f}(y^*) > 0$ противоречит (43). Лемма 1 доказана.

2. Методы решения задач (26) и (37), использующие первые и вторые производные

Пусть в некоторой точке y функция $f(y)$ непрерывна. Тогда функция $f(y)$ является дифференцируемой, причём существуют её частные производные любого порядка. Мы ограничимся вычислением частных производных первого и второго порядка, так как этого будет достаточно для построения как градиентных и квазиньютоновских методов минимизации, так и самого метода Ньютона.

В силу (26) имеем

$$\nabla f(y) = \sum_{i=1}^{k} \nabla\varphi_i(y) \tag{45}$$

и

$$\nabla^2 f(y) = \sum_{i=1}^{k} \nabla^2\varphi_i(y). \tag{46}$$

Сложность построения формулы для $\nabla\varphi_i(y)$ сравнительно невелика. Несколько больше выкладок приходится проделать при выводе формулы для $\nabla^2\varphi_i(y)$. Однако принципиальных математических сложностей нет. Применяя стандартные правила дифференцирования векторно-матричных выражений, получаем

$$\nabla\varphi_i(y) = 2(D_i y - t_i - \varphi_i(y)p_i(y)) \cdot s(y), \tag{47}$$

$$\nabla^2\varphi_i(y) = 2(D_i - p_i^T(y)\nabla\varphi_i(y) - (\nabla\varphi_i(y))^T p_i(y) - \varphi_i(y)G_i) \cdot s_i(y), \tag{48}$$

где

$$D_i = \tilde{A}_i^T\tilde{A}_i, \quad t_i = \tilde{A}_i^T\tilde{b}_i, \tag{49}$$

$$p_i(y) = G_i y + g_i, \quad G_i = \tilde{P}_i^T\tilde{P}_i, \tag{50}$$

$$g_i = \tilde{P}_i^T q_i, \quad s_i(y) = (q_i^T q_i + 2g_i^T y_i + y_i^T G_i y)^{-1}. \tag{51}$$

Теперь займёмся дифференцированием функции $\tilde{f}(y)$. Как отмечалось выше, эта функция определена на всём $\mathbb{R}^N$ и существуют её частные производные требуемого порядка. В силу аналогии между формулами (26) и (37) существуют соответствующие аналогии и в формулах для вычисления градиента и матрицы Гессе функций $f(y)$ и $\tilde{f}(y)$. Так для $\nabla\tilde{f}(y)$, $\nabla^2\tilde{f}(y)$, $\nabla\tilde{\varphi}_i(y)$ и $\nabla^2\tilde{\varphi}_i(y)$ справедливы формулы (45)–(48), в которых по-прежнему можно использовать объекты, определённые формулами из (49), а вместо формул из (50) и (51) следует использовать формулы

$$\tilde{p}_i(y) = \tilde{G}_i y + \tilde{g}_i, \ \tilde{G}_i = \breve{P}_i^T\breve{P}_i,$$
$$\tilde{g}_i = \breve{P}_i^T\breve{q}_i, \ \tilde{s}_i(y) = (\tilde{q}_i^T\tilde{q}_i + 2\tilde{g}_i^T y_i + y_i^T\tilde{G}_i y)^{-1}.$$

3. Решение задач (9) и (10) методом наискорейшего спуска

Опираясь на выкладки, приведённые в разделе 1, и, в частности, используя формулы (18) и (21)–(25), а также (30) и (33)–(36), можно показать, что

$$\min \max_{i=1,\dots,k}\left\{\left\|L_i H_i R_i\right\|_E^2\right\} = \min_y \max_{i=1,\dots,k}\left\{\varphi_i(y)\right\} = \min_y\left\{F(y)\right\} \tag{52}$$

и

$$\min \max_{i=1,\dots,k}\left\{\left\|L_i\left[H_i \quad -h_i\right]R_i\right\|_E^2\right\} = \min_y \max_{i=1,\dots,k}\left\{\tilde{\varphi}_i(y)\right\}$$
$$= \min_y\left\{\tilde{F}(y)\right\}, \tag{53}$$

где $F(y) = \max\limits_{i=1,\ldots,k}\big\{\varphi_i(y)\big\}$, $\tilde{F}(y) = \max\limits_{i=1,\ldots,k}\big\{\tilde{\varphi}_i(y)\big\}$.

Другими словами, задача (9) сводится к задаче (52), а задача (10) — к задаче (53). Последние две задачи принадлежат к классу так называемых «дискретных минимаксных задач», детально рассмотренных в [6]: при этом задачи (52) и (53) тесно связаны с рассмотренными выше задачами (26) и (37), поскольку их целевые функции построены из одинаковых «кирпичиков» — функций $\varphi_i(y)$ и $\tilde{\varphi}_i(y)$.

Приведём три важных теоретических результата из [6]. Для этого введём некоторые дополнительные обозначения, которые для обеспечения компактности изложения сначала будут согласованы только с задачей (52). Заметим, что целевая функция $F(y)$ задачи (52) может иметь точки разрыва, которые совпадают с точками разрыва целевой функции $f(y)$ задачи (26) и являются точками разрыва функций $\varphi_i(y)$. В то же время, как уже отмечалось выше, вне точек разрыва функции $\varphi_i(y)$ имеют частные производные любого необходимого нам порядка. Поэтому условимся, что определяемые ниже объекты *рассматриваются вне точек разрыва функций* $\varphi_i(y)$.

Итак, нам потребуется множество индексов

$$J(y) = \Big\{j \in \big\{j = 1, \ldots, k\big\} \Big| \varphi_j(y) = F(y)\Big\}. \tag{54}$$

Кроме того, потребуются множества

$$G(y) = \Big\{\nabla\varphi_j(y) \Big| j \in J(y)\Big\}$$

и L_j, представляющее собой множество векторов, принадлежащих выпуклой оболочке множества $G(y)$:

$$L(y) = \Big\{z = \sum_{j\in J(y)} \alpha_j \nabla\varphi_j(y) \Big| \alpha_j \geqslant 0, \sum_{j\in J(y)} \alpha_j = 1\Big\}. \tag{55}$$

С использованием (54) и (55) один из фундаментальных результатов работы [6] можно сформулировать в виде следующего утверждения.

Теорема 1. *Для того чтобы функция $F(y)$ достигала минимума при $y = y^*$ необходимо, чтобы выполнялось условие*

$$0 \in L(y^*). \tag{56}$$

Определим *производную функции $F(y)$ в точке $y_0 \in \mathbb{R}^N$ по направлению $u \in \mathbb{R}^N$*, $\|u\|$=1, следующим образом

$$\frac{\partial F(y_0)}{\partial u} = \frac{F(y_0 + \alpha u) - F(y_0)}{\alpha}.$$

Тогда, как было показано в [6], справедлива

Теорема 2. *Пусть функции $\varphi_j(y)$, $1 \leqslant j \leqslant k$, являются непрерывно дифференцируемыми в некоторой окрестности $S(y_0, \delta)$ точки y_0, где*

$$S(y_0, \delta) = \{y \big| \|y - y_0\| < \delta\}, \delta > 0.$$

Тогда функция $F(y)$ дифференцируема в точке y_0 по любому направлению $u, \|u\| = 1$, причём

$$\frac{\partial F(y_0)}{\partial u} = \max_{j \in J(y_0)} \{u^T \varphi_j(y_0)\}.$$

Направлением наискорейшего спуска функции $F(y)$ в точке y_0 будем называть вектор $v(y_0)$ такой, что

$$\|v(y_0)\| = 1, \quad \frac{\partial F(y_0)}{\partial v(y_0)} = \min_{\|u\|=1} \frac{\partial F(y_0)}{\partial u}.$$

Достаточные условия существования и единственности направления наискорейшего спуска функции $F(y)$ в точке y_0 — это третий фундаментальный результат из [6].

Теорема 3. *Если для некоторой точки $y_0 \in \mathbb{R}^N$ выполняется условие*

$$0 \notin L(y_0), \tag{57}$$

то в точке y_0 функция $F(y)$ имеет единственное направление наискорейшего спуска $v(y_0)$, определяемое формулой

$$v(y_0) = -\frac{z^*}{\|z^*\|}, \tag{58}$$

где вектор $z^* \in \mathbb{R}^N$ — ближайшая к началу координат (в евклидовой метрике) точка множества $L(y_0)$, т. е. z^* — решение задачи

$$\|z^*\| \to \min_{z \in L(y_0)}. \tag{59}$$

При этом $\dfrac{\partial F(y_0)}{\partial v(y_0)} = -\|z^*\|$.

Замечание. Можно показать, что в точке y^*, для которой выполняется условие (56), функция $F(y)$ не имеет направления наискорейшего спуска. Таким образом, условие (57) является не только достаточным, но и необходимым условием существования направления наискорейшего спуска функции $F(y)$ в точке y_0.

Теоремы 1–3 и это замечание являются основой для построения вычислительных алгоритмов минимизации функции $F(y)$ с использованием направлений наискорейшего спуска. В частности, таким алгоритмом является *алгоритм наискорейшего спуска*, который может быть записан следующим образом.

Алгоритм наискорейшего спуска для минимизации функции $F(y)$

1. Задать стартовую точку $y^{(0)} \in \mathbb{R}^N$, положить $i = 0$.
2. Найти решение z^* задачи (59). Если $\|z^*\| = 0$, то $y^{(i)}$ — точка локального минимума функции $F(y)$. Работу алгоритма закончить. В противном случае перейти к шагу 3.
3. Вычислить $v(y^{(i)})$ по формуле (58) и решить вспомогательную задачу
$$F\left(y^{(i)} + \alpha v\big(y^{(i)}\big)\right) \longrightarrow \min_{\alpha}. \qquad (60)$$
4. Положить $y^{(i+1)} = y^{(i)} + \alpha^* v\big(y^{(i)}\big)$, где α^* — решение задачи (60), положить $i = i + 1$ и перейти к шагу 2.

Заметим, что в силу неограниченного возрастания функции $\varphi_i(y)$ вблизи точки разрыва будет наблюдаться и неограниченный рост функции $F(y)$. Поэтому справедливо замечание, сделанное при анализе задачи (26): любой численный метод минимизации, очевидным образом доработанный и гарантирующий убывание $F(y)$ на каждой итерации (и конечность $F(y)$ в промежуточных выкладках, например, во время одномерного поиска), не приведёт к точкам разрыва $F(y)$ при условии, что стартовая точка $F(y)$ будет допустимой. Но это означает, что в процессе минимизации $F(y)$ может рассматриваться как непрерывная функция, а задача (52) при сделанных оговорках может считаться задачей безусловной минимизации.

Пусть вспомогательная задача (52) имеет решение, которым является точка y^*. Тогда решение задачи (9) может быть получено тем же способом, каким по решению задачи (26) может быть получено решение задачи (7), а именно: по формуле (11) вычисляем $x^* \in \mathbb{R}^N$ и по формуле (19) строим оптимальные для задачи (9) матрицы коррекции H_i^*, $1 \leqslant i \leqslant k$. При этом

$$x^* \in X(A_0, A_1 + H_1^*, \ldots, A_k + H_k^*, b_0, b_1 \ldots, b_k),$$

т. е. x^* принадлежит множеству решений скорректированной системы вида (5).

Обратимся теперь к исследованию задачи (53). Её запись совпадает с записью задачи (52) с точностью до обозначений ив силу непрерывности и дифференцируемости функций $\tilde{\varphi}_i(y)$ на $\mathbb{R}^N$ в задаче (53) нет ограничений, связанных с задачей (52). Следовательно, решение задачи (53) может осуществляться теми же методами, что и задачи (52). При этом переход от вспомогательной задачи (53) к основной задаче (10) совпадает с переходом от вспомогательной задачи (37) к основной задаче (8), а именно: пусть y^* — решение задачи (53), тогда вычислив $x^* \in \mathbb{R}^N$ по формуле (11), по формуле (33) можно построить оптимальные для задачи (10) матрицы коррекции $\left[H_i^* \ \ -h_i^*\right]$. При этом в силу приведенных в разделе 1 настоящей работы выкладок

$$x^* \in X(A_0, A_1 + H_1^*, \ldots, A_k + H_k^*, b_0, b_1 + h_1^* \ldots, b_k + h_k^*),$$

т. е. x^* принадлежит множеству решений скорректированной системы вида (6).

4. Численные примеры и вычислительные эксперименты

Приведём примеры и некоторые предварительные результаты, характеризующие сходимость по аргументу и по целевой функции алгоритмов решения вспомогательных задач (26), (37), (52) и (53).

Мы рассмотрели небольшое число модельных несовместных систем малой размерности с блочной структурой. Указанные модельные системы были получены из совместных систем линейных алгебраических уравнений после внесения возмущений в вектор правой части. Результаты коррекции и поведение вычислительных алгоритмов во всех случаях были однотипными, в силу чего мы приводим только одну модельную несовместную систему с матрицей коэффициентов A и вектором правой части b вида

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ -1 & 1 & 2 & 1 & 3 \\ 2 & -1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 4 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 3 & 1 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 9 \\ 1 \\ 4 \\ 8 \\ 7 \\ 8 \\ 3 \end{bmatrix},$$

для которой последовательно были решены задачи (7)–(10) без учёта и с учётом условия неотрицательности решения скорректированной системы. При этом весовые матрицы L_i и R_i во всех случаях были единичными. Очевидно, что $A_0 = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ -1 & 1 & 2 & 1 & 3 \end{bmatrix}$, $b_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$, $A_1 = \begin{bmatrix} 2 & -1 \\ 3 & 1 \\ 4 & -2 \end{bmatrix}$, $b_1 = \begin{bmatrix} 4 \\ 8 \\ 7 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 1 & 5 \end{bmatrix}$, $b_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$.

Численное решение задач (7)–(8) производилась с помощью запрограммированного авторами в системе Matlab метода Ньютона. При этом использовалось псевдообращение матрицы Гессе в сингулярном разложении, что позволило преодолеть вырожденность и плохую обусловленность указанной матрицы.

Т а б л и ц а 1

Результаты решения задачи (7)

y^*
0,79032327
−0,28429196
−2,53152817
−1,47799667
1,87188941

H_1^*			x_1^*
0,17651444	0,08576679		2,30907327
−0,01722986	−0,00837183		1,12195804
0,00267078	0,00129771		
H_2^*			x_2^*
−0,01572363	0,00537151	0,01302739	−2,64402817
0,19388520	−0,06623509	−0,16063844	0,90325333
			2,19063941

При решении задач (9)–(10) использовалась функция `fminimax` из оптимизационного пакета системы Matlab, предназначенная для решения минимаксных задач. Результаты расчетов представлены в таблицах 1–4. Первый столбец в этих таблицах содержит вектор y^*, являющийся решением соответствующей вспомогательной задачи (26), (37), (52) или (53).

Т а б л и ц а 2

Результаты решения задачи (8)

y^*
0,78807425
−0,28420326
−2,52458170
−1,47544856
1,86562982

H_1^*			h_1^*	x_1^*
0,15471231	0,07525257		−0,06706723	2,30682425
−0,01293924	−0,00629369		0,00560911	1,12204674
0,00511139	0,00248620		−0,00221577	
H_2^*			h_2^*	x_2^*
−0,01712337	0,00588164	0,01418384	−0,00649330	−2,63708170
0,17840940	−0,06128119	−0,14778226	0,06765410	0,90580144
				2,18437982

Т а б л и ц а 3

Результаты решения задачи (9)

y^*
0,55407557
−0,07123926
−2,23154183
−1,36371113
2,48092351

H_1^*			x_1^*
0,20237566	0,10343203		2,29338157
−0,01807038	−0,00923558		1,17212279
0,05902256	0,03016580		
H_2^*			x_2^*
−0,00894330	0,00286539	0,00732557	−2,72748578
0,17819469	−0,05709274	−0,14596158	0,87387360
			2,23411892

Т а б л и ц а 4

Результаты решения задачи (10)

y^*
0,78497942
−0,18892909
−2,72849263
−1,55161928
2,01845229

H_1^*			h_1^*	x_1^*
0,17693364	0,09294536		−0,07668678	2,30722472
−0,03958364	−0,02079376		0,01715639	1,21201287
0,05777546	0,03035014		−0,02504111	
H_2^*			h_2^*	x_2^*
0,01732845	−0,00500870	−0,01412944	0,00608059	−2,84979598
0,16739569	−0,04838484	−0,13649277	0,05873953	0,82371842
				2,32369514

Результаты решения задач (7)–(10) с дополнительным условием (2) представлены в таблицах 5–8. В первом столбце в них приведён вектор y^*, являющийся решением соответствующей вспомогательной задачи (26), (37), (52) или (53).

Решение задач (7)–(8) производились с помощью процедуры `fmincon` оптимизационного пакета системы Matlab. При решении задач (9)–(10) использовалась функция `fminimax` оптимизационного пакета системы Matlab, предназначенная для минимаксных задач.

Упомянутые процедуры `fmincon` и `fminimax` допускают задание ограничений в виде системы линейных неравенств. В соответствии с (11) указанная система для обеспечения условия (2) имела вид $-Py \leqslant \hat{x}$.

Таблица 5

Результаты решения задачи (7) при ограничении $x \geqslant 0$

y^*
0,68803232
−1,36383771
0,04528361
−0,21469570
−0,05300377

H_1^*		x_1^*
−0,09265286	0	2,09715364
0,81469428	0	0
−0,66214250	0	

H_2^*			x_2^*
0	1,16835686	0,19227709	0
0	−0,36608202	−0,06024631	2,07346169
			0,34123065

Таблица 6

Результаты решения задачи (8) при ограничении $x \geqslant 0$

y^*
0,68482790
−1,36432775
0,04605947
−0,21438640
−0,05324036

H_1^*		h_1^*	x_1^*
−0,07402556	0	0,03533077	2,09521512
0,66641759	0	−0,31806643	0
−0,53677909	0	0,25619283	

H_2^*			h_2^*	x_2^*
0	0,95301576	0,15622483	−0,45931872	0
0	−0,29681978	−0,04865672	0,14305627	2,07484634
				0,34012293

Таблица 7

Результаты решения задачи (9) при ограничении $x \geqslant 0$

y^*
0,41048360
−1,40628227
0,11248423
−0,18790574
−0,07349602

H_1^*		x_1^*
0,07334446	0	1,92925010
1,14668892	0	0
−0,37164720	0	

H_2^*		x_2^*
1,18517838	0,13253777	0
−0,18903877	−0,02114009	2,19339279
		0,24528577

Данные в таблицах 1–4 близки между собой, но не совпадают, что вызвано различиями в постановках задач (7)–(10). Аналогичная ситуация имеет место для данных, приведённых в таблицах 5–8. Мы также видим, что условие (2) приводит к существенному изменению матриц коррекции и вектора решения скорректированных систем.

Т а б л и ц а 8

Результаты решения задачи (10) при ограничении $x \geqslant 0$

y^*
0,41471743
−1,35901169
0,13736175
−0,09946308
−0,04590183

H_1^*			h_1^*	x_1^*
0,07657770	0		−0,04015805	1,90690768
0,93746601	0		−0,49161583	0
−0,25814434	0		0,13537328	
H_2^*			h_2^*	x_2^*
0,00187281	0,98939050	0,10276049	−0,44769829	0,00418320
−0,00026083	−0,13779280	−0,01431149	0,06235111	2,20994926
				0,22953067

Приведём теперь результаты исследования сходимости по аргументу и по целевой функции численных алгоритмов решения вспомогательных задач (26), (37), (52) и (53). При решении задач (26) и (37) без ограничений методом Ньютона (при удачном выборе начального приближения) в обоих случаях наблюдалась квадратичная сходимость, что иллюстрируют приводимые ниже рисунки 1 и 2. Представленные на них данные получены в ходе решения задач (7) и (8) для рассматриваемой в настоящем разделе несовместной системы.

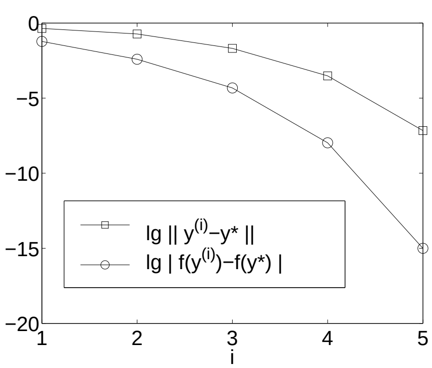

Рис. 1. Сходимость по аргументу и по целевой функции для задачи (26) без ограничений

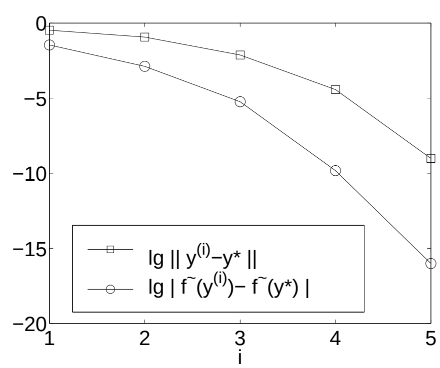

Рис. 2. Сходимость по аргументу и по целевой функции для задачи (37) без ограничений

При решении задач (26) и (37) с ограничениями, а также при решении задач (52) и (53) (с ограничениями и без ограничений) наблюдалась линейная или сверхлинейная сходимость как по аргументу, так и по целевой функции. При этом число итераций колебалось от 10 до 40, а графики сходимости были уже не столь гладкими, как графики, представленные на рис. 1 и 2.

Таким образом, исследование сходимости численных алгоритмов решения вспомогательных задач (26), (37), (52) и (53) дало вполне ожидаемые результаты, хорошо согласующиеся как с теорией решения нелинейных задач математического программирования, так и с имеющимся практическим опытом.

ЛИТЕРАТУРА

1. **Алберт А.** Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977.
2. **Ватолин А. А.** Аппроксимация несобственных задач линейного программирования по критерию евклидовой нормы // Журн. вычисл. матем. и матем. физика. 1984. Т. 24, № 12. С. 1907–1908.
3. **Гантмахер Ф. Р.** Теория матриц. М.: Наука, 1988.
4. **Горелик В. А.** Матричная коррекция задачи линейного программирования с несовместной системой ограничений // Журн. вычисл. матем. и матем. физика. 2001. Т. 41, № 11. С. 1697–1705.
5. **Горелик В. А., Ерохин В. И.** Оптимальная матричная коррекция несовместных систем линейных алгебраических уравнений по минимуму евклидовой нормы. М.: ВЦ РАН, 2004.
6. **Демьянов В. Ф., Малоземов В. Н.** Введение в минимакс. М.: Наука, 1972.
7. **Еремин И. И., Мазуров В. Д., Астафьев Н. Н.** Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983.
8. **Ерохин В. И.** Свойства оптимальной одноранговой коррекции матриц коэффициентов несовместных неоднородных линейных моделей // Дискрет. анализ и исслед. операций. Сер. 2. 2002. Т. 9, № 1. С. 33–60.
9. **Тихонов А. Н.** О нормальных решениях приближенных систем линейных алгебраических уравнений // Докл. АН СССР. 1980. Т. 254, № 3. С. 549–554.

Адрес авторов:

Горелик В.А.,
ВЦ им. А.А. Дородницына РАН,
ул. Вавилова, 40,
119991, Москва,
Россия.
E-mail: gorelik@ccas.ru

Статья поступила
30 мая 2005 г.

Ерохин В.И.,
Борисоглебский гос. пед. институт,
ул. Народная, 43,
397160, Борисоглебск,
Воронежская обл., Россия.
E-mail: erohin_v_i@mail.ru
Печенкин Р.В.,
Моск. пед. гос. ун–т,
ул. Малая Пироговская, д.1,
119992, Москва,
Россия.
E-mail:pecr@rarus.ru