ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 20 № 2 2014

УДК 519.854

РЕГУЛЯРИЗАЦИЯ И НОРМАЛЬНЫЕ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ И НЕРАВЕНСТВ[1]

А. И. Голиков, Ю. Г. Евтушенко

Приведены примеры взаимно двойственных задач безусловной оптимизации, которые возникают из регуляризованных задач систем линейных уравнений и/или неравенств. Решение любой из взаимно двойственных задач вычисляется по решению другой по простым формулам. Так как взаимно двойственные задачи отличаются размерностью, то естественно решать задачу безусловной оптимизации меньшей размерности.

Ключевые слова: регуляризация, кусочно-квадратичная функция, безусловная оптимизация, взаимно двойственные задачи, обобщенный метод Ньютона.

A. I. Golikov, Yu. G. Evtushenko. Regularization and normal solutions of systems of linear equations and inequalities.

The paper provides some examples of mutually dual unconstrained optimization problems originating from regularization problems for systems of linear equations and/or inequalities. The solution of each of these mutually dual problems can be found from the solution of the other problem by means of simple formulas. Since mutually dual problems have different dimensions, it is natural to solve the unconstrained optimization problems with smaller dimension.

Keywords: regularization, piecewise quadratic function, unconstrained optimization, mutually dual problems, generalized Newton method.

Введение

И. И. Еремин хорошо известен как создатель теории двойственности несобственных задач линейной оптимизации, он всегда уделял большое внимание выявлению двойственности в задачах, возникающих в различных методах оптимизации. На авторов данной статьи произвело большое впечатление замечание И. И. Еремина, сделанное на одной из конференций "Математическое программирование и приложения", о том, что двойственная задача квадратичного программирования без ограничений может рассматриваться как взаимно двойственная для исходной прямой задачи квадратичного программирования с ограничениями [1; 2].

Формально задачи безусловной минимизации не имеют функции Лагранжа, и, следовательно, для них нельзя непосредственно построить двойственную задачу. Тем не менее с помощью дополнительных переменных можно ввести искусственные ограничения и получить эквивалентную задачу нелинейного программирования, для которой уже стандартным образом определяется двойственная задача. Существует класс таких задач оптимизации, для которых взаимно двойственные задачи являются задачами безусловной оптимизации и решение любой из этих двух задач выражается через решение другой. Это — задачи квадратичного программирования, которые возникают, например, при регуляризации систем линейных уравнений и/или неравенств. Так как взаимно двойственные задачи отличаются размерностью, то естественно решать задачу безусловной оптимизации меньшей размерности. Приводится один типичный результат, возникающий при регуляризации системы линейных уравнений и неравенств [3] и в svm-методе распознавания образов [4].

[1]Работа выполнена при поддержке РФФИ (проект 14-07-00805), программы государственной поддержки ведущих научных школ (НШ-4640.2014.1) и программы Президиума РАН П-18.

В первом разделе рассматривается регуляризованная задача решения системы линейных уравнений. Всюду в статье используется евклидова норма. Для регуляризованной задачи, которая является задачей безусловной минимизации строго выпуклой квадратичной функции, приводится взаимно двойственная задача безусловной максимизации строго вогнутой квадратичной функции и получены простые формулы, по которым решение любой из этих задач находится через решение другой. Рассмотрены некоторые подходы для нахождения нормального решения системы линейных уравнений, отличные от метода регуляризации.

Во втором и третьем разделах рассмотрены аналогичные взаимно двойственные задачи для нахождения нормальных решений систем линейных уравнений с неотрицательными переменными и систем линейных неравенств соответственно. Здесь возникают задачи безусловной оптимизации кусочно-квадратичных функций, для которых особенно эффективен глобально сходящийся за конечное число шагов обобщенный метод Ньютона.

1. Нормальные решения систем линейных уравнений

Рассмотрим совместную систему линейных уравнений

$$Ax = b. \tag{1.1}$$

Здесь A — ненулевая матрица размерности $m \times n$, вектор $b \in \mathbb{R}^m$, $\|b\| \neq 0$. В методе регуляризации рассматривается последовательность задач безусловной минимизации

$$\min_{x \in \mathbb{R}^n} F(x), \qquad F(x) = \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2) \tag{1.2}$$

с положительным параметром ε, стремящимся к нулю. Единственное решение $x(\varepsilon)$ задачи (1.2) при фиксированном ε выражается явно следующей формулой:

$$x(\varepsilon) = (\varepsilon I_n + A^\top A)^{-1} A^\top b. \tag{1.3}$$

Здесь и ниже через I_k обозначена единичная матрица порядка k. При $\varepsilon \to 0$ решение $x(\varepsilon)$ сходится к нормальному решению системы (1.1) [3]. В (1.3) обратная матрица существует при любом ранге матрицы A и любом $\varepsilon > 0$.

Выражение (1.3) для вычисления $x(\varepsilon)$ можно представить в другом виде, используя формулу Шермана — Моррисона — Вудбери [5]

$$x(\varepsilon) = \frac{1}{\varepsilon}\big(I_n - A^\top(\varepsilon I_m + AA^\top)^{-1}A\big)A^\top b. \tag{1.4}$$

Отметим, что в этой формуле следует обращать квадратную матрицу порядка m в отличие от формулы (1.3), в которой обращается матрица порядка n. Ниже будет выведена еще одна формула (1.17) вычисления $x(\varepsilon)$, в которой тоже требуется обращать матрицу порядка m, но с меньшим количеством перемножений матриц, чем в формуле (1.4).

Задачу (1.2) можно рассматривать с разных точек зрения: как метод квадратичных штрафных функций, как регуляризацию задачи линейного программирования с нулевой целевой функцией, как метод наименьших квадратов, как многокритериальную оптимизацию. Так, например, (1.2) есть вспомогательная задача метода штрафов с коэффициентом штрафа при целевой функции для следующей задачи квадратичного программирования:

$$\min_{x \in X} \frac{1}{2}\|x\|^2, \qquad X = \{x \in \mathbb{R}^n \colon Ax = b\}. \tag{1.5}$$

Задача (1.2) эквивалентна следующей регуляризованной задаче линейного программирования:

$$\min_{x \in X}\big\{0_n^\top x + \frac{\varepsilon}{2}\|x\|^2\big\}, \qquad X = \{x \in \mathbb{R}^n \colon Ax = b\}. \tag{1.6}$$

Легко показать, что двойственной к (1.6) является следующая задача безусловной максимизации квадратичной функции [1] (И. И. Ереминым было показано, что эти задачи взаимно двойственны):

$$\max_{u\in\mathbb{R}^n}\{b^\top u - \frac{1}{2\varepsilon}\|A^\top u\|^2\}, \tag{1.7}$$

которая в свою очередь есть оштрафованная задача линейного программирования (ЛП)

$$\max_{u\in U} b^\top u, \qquad U = \{u\in\mathbb{R}^m\colon A^\top u = 0_n\}.$$

Заметим, что из решения $u(\varepsilon)$ задачи безусловной максимизации (1.7) легко вычисляется решение задачи (1.6) по формуле

$$x(\varepsilon) = \frac{1}{\varepsilon}A^\top u(\varepsilon).$$

Согласно [6] и [7] эта формула дает решение задачи (1.5) при любом $\varepsilon > 0$.

При фиксированном параметре ε задачу (1.2) можно рассматривать как метод наименьших квадратов (метод минимизации невязок), примененный к следующей несовместной системе:

$$Ax = b, \qquad -\sqrt{\varepsilon}x = 0_n. \tag{1.8}$$

Вектор $x(\varepsilon)$ есть решение задачи безусловной минимизации (1.2) и псевдорешение системы (1.8). Через $z_1(\varepsilon) = b - Ax(\varepsilon)$ и $z_2(\varepsilon) = \sqrt{\varepsilon}x(\varepsilon)$ обозначим составляющие вектора минимальных невязок $z(\varepsilon)^\top = [z_1(\varepsilon)^\top, z_2(\varepsilon)^\top]$ системы (1.8).

Согласно теореме об альтернативах (см., например, [8; 9]) для несовместной системы (1.8) можно построить совместную альтернативную систему вида

$$A^\top u_1 - \sqrt{\varepsilon}u_2 = 0_n, \qquad b^\top u_1 = \rho > 0. \tag{1.9}$$

Здесь ρ — фиксированная положительная константа, векторы неизвестных $u_1\in\mathbb{R}^m$, $u_2\in\mathbb{R}^n$. Согласно [9] нормальный вектор $\tilde u(\varepsilon)^\top = [\tilde u_1(\varepsilon)^\top, \tilde u_2(\varepsilon)^\top]$ альтернативной системы (1.9) выражается через вектор минимальных невязок $z(\varepsilon)$ по формулам

$$\tilde u_1(\varepsilon) = \frac{\rho z_1(\varepsilon)}{\|z(\varepsilon)\|^2}, \qquad \tilde u_2(\varepsilon) = \frac{\rho z_2(\varepsilon)}{\|z(\varepsilon)\|^2}.$$

Пусть вектор переменных $z\in\mathbb{R}^{m+n}$ имеет разбиение $z^\top = [z_1^\top, z_2^\top]$, где $z_1\in\mathbb{R}^m$, $z_2\in\mathbb{R}^n$. Запишем задачу строго вогнутого квадратичного программирования

$$\max_{z\in Z}\Big\{b^\top z_1 - \frac{1}{2}(\|z_1\|^2 + \|z_1\|^2)\Big\}, \qquad Z = \{z\in\mathbb{R}^{m+n}\colon A^\top z_1 - \sqrt{\varepsilon}z_2 = 0_n\}. \tag{1.10}$$

Эту задачу можно рассматривать как регуляризацию задачи ЛП следующего вида:

$$\max_{z\in Z}\{b^\top z_1 + 0_n^\top z_2\}, \qquad Z = \{z\in\mathbb{R}^{m+n}\colon A^\top z_1 - \sqrt{\varepsilon}z_2 = 0_n\},$$

которая является взаимно двойственной к задаче ЛП

$$\min_{x\in X} 0_n^\top x, \qquad X = \{x\in\mathbb{R}^n\colon Ax = b,\ -\sqrt{\varepsilon}x = 0_n\}. \tag{1.11}$$

При $\varepsilon \neq 0$ и $\|b\| \neq 0$ ограничения в (1.11) несовместны, и задача является несобственной 1-го рода [10]. Задачу (1.2) можно рассматривать как вспомогательную задачу метода штрафных квадратичных функций, примененного к задаче ЛП (1.11). Как известно [1], задача метода штрафных квадратичных функций для задачи ЛП и регуляризованная задача ЛП являются взаимно двойственными, т. е. задачи (1.2) и (1.10) взаимно двойственны.

В задаче (1.10) можно исключить переменные z_2, выразив их через z_1, и подставить $z_2 = (1/\sqrt{\varepsilon})A^\top z_1$ в целевую функцию задачи (1.10). Тогда приходим к следующей эквивалентной задаче безусловной максимизации строго вогнутой квадратичной функции

$$\max_{z_1\in\mathbb{R}^m} H(z_1), \qquad H(z_1) = b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2. \tag{1.12}$$

Таким образом, эта задача и задача (1.2) взаимно двойственны.

Теорема 1.1. *При любом $\varepsilon > 0$ единственное решение $x(\varepsilon) = \operatorname{Arg}\min\limits_{x\in\mathbb{R}^n} F(x)$ задачи* (1.2) *и единственное решение $z_1(\varepsilon) = \operatorname{Arg}\max\limits_{z_1\in\mathbb{R}^m} H(z_1)$ задачи* (1.12) *связаны между собой соотношениями*

$$x(\varepsilon) = \frac{1}{\varepsilon}A^\top z_1(\varepsilon), \tag{1.13}$$

$$z_1(\varepsilon) = b - Ax(\varepsilon), \tag{1.14}$$

и имеет место равенство оптимальных значений целевых функций $F(x(\varepsilon)) = H(z_1(\varepsilon))$.

Д о к а з а т е л ь с т в о. При $\varepsilon > 0$ строго выпуклая квадратичная функция F(x) на $\mathbb{R}^n$ ограничена снизу нулем. Поэтому по теореме Франка — Вульфа [11] задача (1.2) всегда имеет единственное решение.

Максимизируемая квадратичная функция $H(z_1)$ при $\varepsilon > 0$ строго вогнута и ограничена сверху на всем пространстве $\mathbb{R}^m$. Действительно, справедливы следующие выражения:

$$H(z_1) = b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2 = \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b\|^2 + b^\top z_1 - \frac{1}{2}\|z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2$$

$$= \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b - z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 \le \frac{1}{2}\|b\|^2.$$

Следовательно, задача (1.10) при любом $\varepsilon > 0$ всегда имеет единственное решение.

Для взаимно двойственных задач (1.2) и (1.12) по теореме слабой двойственности для любых $z_1 \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $\varepsilon > 0$ справедливо неравенство

$$b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2 \le \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2),$$

и по теореме двойственности оптимальные значения целевых функций этих задач совпадают:

$$b^\top z_1(\varepsilon) - \frac{1}{2\varepsilon}\|A^\top z_1(\varepsilon)\|^2 - \frac{1}{2}\|z_1(\varepsilon)\|^2 = \frac{1}{2}(\|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2). \tag{1.15}$$

Легко убедиться, что выражения (1.13) и (1.14) удовлетворяют необходимым и достаточным условиям оптимальности для задачи (1.2)

$$-A^\top(b - Ax(\varepsilon)) + \varepsilon x(\varepsilon) = 0_n$$

и необходимым и достаточным условиям оптимальности для задачи (1.12)

$$b - z_1(\varepsilon) - \frac{1}{\varepsilon}AA^\top z_1(\varepsilon) = 0_m.$$

Теорема доказана.

Заметим, что равенство (1.15) оптимальных значений целевых функций взаимно двойственных задач (1.2) и (1.12) с учетом связей (1.13), (1.14) между решениями $x(\varepsilon)$ и $z(\varepsilon)$

можно представить в следующих двух видах, в которых присутствуют решения только одной из задач

$$b^\top z_1(\varepsilon) = \|z(\varepsilon)\|^2,$$
$$b^\top (b - Ax(\varepsilon)) = \|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2. \tag{1.16}$$

Из (1.16) следует $\varepsilon\|x(\varepsilon)\|^2 = z_1(\varepsilon)^\top Ax(\varepsilon)$. Отсюда получаем известный результат в методе наименьших квадратов: при $\varepsilon = 0$ имеет место $z_1(0)\perp Ax(0)$. Заметим, что если система (1.1) несовместна, то $z_1(0) \neq 0$.

Задача (1.12) при любом $\varepsilon > 0$ и любом ранге матрицы A может быть явно разрешена:

$$z_1(\varepsilon) = \varepsilon(\varepsilon I_m + AA^\top)^{-1}b.$$

Подставляя это выражение в (1.13), получаем еще одну формулу для вычисления решения $x(\varepsilon)$ задачи (1.2)

$$x(\varepsilon) = A^\top(\varepsilon I_m + AA^\top)^{-1}b. \tag{1.17}$$

В этой формуле требуется обращать матрицу порядка m в отличие от формулы (1.3), в которой требуется обращать матрицу порядка n. Поэтому если в задаче (1.1) $m < n$, то для вычисления $x(\varepsilon)$ целесообразно использовать формулу (1.17) или формулу (1.13). В случае применения последней формулы требуется решать задачу безусловной оптимизации (1.12) с m неизвестными.

Выражение (1.17) для вычисления $x(\varepsilon)$ также можно представить в другом виде с использованием формулы Шермана — Моррисона — Вудбери

$$x(\varepsilon) = \frac{1}{\varepsilon}A^\top(I_m - A(\varepsilon I_n + A^\top A)^{-1}A^\top)b.$$

В этой формуле требуется обращать матрицу порядка n в отличие от формулы (1.17), в которой обращается матрица порядка m.

В [12] показано, что для всякой матрицы A из системы (1.1) следующим образом определена псевдообратная матрица:

$$A^+ = \lim_{\varepsilon\to 0}(\varepsilon I_n + A^\top A)^{-1}A^\top = \lim_{\varepsilon\to 0}A^\top(\varepsilon I_m + AA^\top)^{-1},$$

и при любом векторе b (в том числе и при b, делающим систему (1.1) несовместной) вектор $\tilde{x}_* = A^+b$ является вектором с минимальной нормой среди всех векторов, минимизирующих $\|b - Ax\|^2$.

В [9] предложен другой метод нахождения нормального решения системы (1.1), основанный на применении теорем об альтернативах. При этом для несовместной альтернативной к (1.1) системы

$$A^\top u = 0_n, \qquad b^\top u = \rho \neq 0$$

решается следующая задача минимизации ее невязок

$$\min_{u\in\mathbb{R}^m}\frac{1}{2}\{\|A^\top u\|^2 + (\rho - b^\top u)^2\}. \tag{1.18}$$

Здесь ρ — произвольная ненулевая константа.

Пусть u_* — решение задачи безусловной минимизации (1.18). Тогда нормальное решение исходной системы (1.1) выражается по формуле

$$\tilde{x}_* = \frac{A^\top u_*}{\rho - b^\top u_*}. \tag{1.19}$$

Если в исходной системе ранг матрицы A равен m, то решение задачи (1.18) находится аналитически

$$u_* = \rho(AA^\top + bb^\top)^{-1}b.$$

Подставляя эту формулу в (1.19), получаем еще одно выражение для нормального решения системы (1.1)

$$\tilde{x}_* = \frac{A^\top(AA^\top + bb^\top)^{-1}b}{1 - b^\top(AA^\top + bb^\top)^{-1}b}.$$

2. Регуляризация систем линейных уравнений с неотрицательными переменными

Рассмотрим теперь совместную систему линейных уравнений с неотрицательными переменными

$$Ax = b, \qquad x \geq 0_n,$$

и ее регуляризованную задачу

$$\min_{x \in \mathbb{R}^n_+} F(x), \qquad F(x) = \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2) \tag{2.1}$$

с положительным параметром ε, стремящимся к нулю.

Следующая задача безусловной максимизации

$$\max_{u \in \mathbb{R}^m} H(u), \qquad H(u) = b^\top u - \frac{1}{2\varepsilon}\|(A^\top u)_+\|^2 - \frac{1}{2}\|u\|^2 \tag{2.2}$$

является двойственной к регуляризованной задаче (2.1). Здесь и ниже a_+ обозначает вектор a, у которого все отрицательные компоненты заменены на нули.

Теорема 2.1. *При любом* $\varepsilon > 0$ *единственное решение* $x(\varepsilon) = \operatorname{Arg}\min\limits_{x \in \mathbb{R}^n_+} F(x)$ *задачи* (2.1) *и единственное решение* $u(\varepsilon) = \operatorname{Arg}\max\limits_{u \in \mathbb{R}^m} H(u)$ *задачи* (2.2) *связаны между собой соотношениями*

$$x(\varepsilon) = \frac{1}{\varepsilon}(A^\top u(\varepsilon))_+,$$
$$u(\varepsilon) = b - Ax(\varepsilon),$$

и имеет место равенство оптимальных значений целевых функций $F(x(\varepsilon)) = H(u(\varepsilon))$.

Доказательство. Минимизируемая функция в задаче (2.1) является при $\varepsilon > 0$ строго выпуклой квадратичной функцией, ограниченной снизу нулем на $\mathbb{R}^n_+$. Поэтому по теореме Франка — Вульфа [11] задача (2.1) всегда имеет решение, которое единственно.

В задаче (2.2) максимизируемая кусочно-квадратичная функция является при $\varepsilon > 0$ строго вогнутой и ограниченной сверху на всем пространстве $\mathbb{R}^m$. По теореме Франка — Вульфа задача (2.2) всегда имеет единственное решение.

Вводя дополнительные переменные $y \in \mathbb{R}^m$ и дополнительные ограничения $Ax + y = b$, перепишем задачу (2.1) в следующем эквивалентном виде:

$$\min_{y \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n_+} \frac{1}{2}\{\|y\|^2 + \varepsilon\|x\|^2\}, \tag{2.3}$$
$$Ax + y = b.$$

Для этой задачи введем функцию Лагранжа

$$L(y, x, u) = \frac{1}{2}\|y\|^2 + \frac{\varepsilon}{2}\|x\|^2 + u^T(b - Ax - y).$$

Здесь $u \in \mathbb{R}^m$ — множители Лагранжа для задачи (2.3). Двойственная к (2.3) задача имеет вид

$$\max_{u \in \mathbb{R}^m} \min_{y \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n_+} L(y, x, u). \tag{2.4}$$

Запишем условия минимума по y и x для внутренней задачи в (2.4):

$$L_y(y(\varepsilon), x(\varepsilon), u) = y(\varepsilon) - u = 0_m,$$
$$L_x(y(\varepsilon), x(\varepsilon), u) = \varepsilon x(\varepsilon) - A^\top u \geq 0_n, \qquad x^\top(\varepsilon)(\varepsilon x(\varepsilon) - A^\top u) = 0, \qquad x(\varepsilon) \geq 0_n.$$

Из этих условий легко находим решения внутренней задачи минимизации (2.4)

$$y(\varepsilon) = u \tag{2.5}$$

$$x(\varepsilon) = \frac{1}{\varepsilon}(A^\top u)_+. \tag{2.6}$$

Подставляя решения (2.5) и (2.6) в функцию Лагранжа $L(y, x, u)$, после простых преобразований получаем двойственную функцию для задачи (2.4)

$$H(u) = b^\top u - \frac{1}{2\varepsilon}\|(A^\top u)_+\|^2 - \frac{1}{2}\|u\|^2,$$

т. е. приходим к двойственной задаче (2.2) для задачи (2.3) и, следовательно, для задачи (2.1). Из теории двойственности следует равенство оптимальных значений целевых функций задач (2.1) и (2.2).

Внешняя задача в (2.4) состоит в безусловной максимизации $H(u)$ по u. Необходимое и достаточной условие максимума этой задачи есть

$$H_u(u(\varepsilon)) = b - \frac{1}{\varepsilon}A(A^\top u(\varepsilon))_+ - u(\varepsilon) = b - Ax(\varepsilon) - u(\varepsilon) = 0_m.$$

Отсюда с учетом (2.6) следует

$$u(\varepsilon) = b - \frac{1}{\varepsilon}A(A^\top u(\varepsilon))_+ = b - Ax(\varepsilon).$$

Теорема доказана.

Из теоремы следует, что если в матрице A размерности $m \times n$ число строк $m < n$, то вместо задачи минимизации (2.1) целесообразно решать двойственную задачу (2.2), которая является вогнутой кусочно-квадратичной задачей безусловной максимизации. Для решения (2.2) весьма эффективен обобщенный метод Ньютона. Максимизируемая функция $H(u)$ в задаче (2.2) является вогнутой кусочно-квадратичной и дифференцируемой. Для этой функции обычная матрица Гессе не существует. Действительно, градиент

$$H_u(u) = b - \frac{1}{\varepsilon}A(A^\top u)_+ - u$$

функции $H(u)$ недифференцируем. Но для этой функции можно определить обобщенную матрицу Гессе, которая является невырожденной $(m \times m)$-матрицей следующего вида:

$$H_{uu} = -\Big(\frac{1}{\varepsilon}AD(z)A^\top + I_m\Big),$$

где через $D(z)$ обозначена диагональная $(n \times n)$-матрица с i-м диагональным элементом z^i, равным 1, если $(A^\top u)^i > 0$, и равным 0, если $(A^\top u)^i \leq 0$, $i = 1, \ldots, n$. Доказательство конечной глобальной сходимости обобщенного метода Ньютона с выбором шага по правилу Армихо для безусловной оптимизации кусочно-квадратичной функции можно найти в [4; 13; 14]. Обобщенный метод Ньютона позволяет эффективно решать задачи при $n \approx 10^6$ и $m \approx 10^4$ на однопроцессорных компьютерах и при n порядка десятков миллионов, m порядка сотен тысяч — на многопроцессорных вычислительных комплексах [15].

3. Регуляризация систем линейных неравенств

Регуляризация систем нелинейных неравенств аналогична рассмотренной в предыдущем параграфе. Пусть дана совместная система линейных неравенств

$$Ax \geq b.$$

Регуляризованная задача имеет вид

$$\min_{x \in \mathbb{R}^n} F(x), \qquad F(x) = \frac{1}{2}(\|(b - Ax)_+\|^2 + \varepsilon\|x\|^2) \tag{3.1}$$

с положительным параметром ε, стремящимся к нулю. Тогда следующая задача максимизации на положительном ортанте

$$\max_{u \in \mathbb{R}^m_+} H(u), \qquad H(u) = b^\top u - \frac{1}{2\varepsilon}\|A^\top u\|^2 - \frac{1}{2}\|u\|^2 \tag{3.2}$$

является двойственной к регуляризованной задаче (3.1).

В этом случае справедлива теорема, аналогичная теореме 2.1.

Теорема 3.1. *При любом $\varepsilon > 0$ единственное решение $x(\varepsilon) = \operatorname{Arg}\min\limits_{x \in \mathbb{R}^n} F(x)$ задачи* (3.1) *и единственное решение $u(\varepsilon) = \operatorname{Arg}\max\limits_{u \in \mathbb{R}^m_+} H(u)$ задачи* (3.2) *связаны между собой соотношениями*

$$x(\varepsilon) = \frac{1}{\varepsilon} A^\top u(\varepsilon),$$
$$u(\varepsilon) = (b - Ax(\varepsilon))_+\,,$$

и имеет место равенство оптимальных значений целевых функций $F(x(\varepsilon)) = H(u(\varepsilon))$.

К сожалению, непосредственное применение метода Ньютона к задаче (3.2) затруднительно в отличие от задачи (2.2) [7; 16].

СПИСОК ЛИТЕРАТУРЫ

1. **Еремин И.И.** О квадратичных и полноквадратичных задачах выпуклого программирования // Изв. вузов. Математика. 1998. № 12. С. 22–28.
2. **Еремин И.И.** Теория двойственности в линейной оптимизации. Челябинск: Изд-во Южно-Урал. гос. ун-та, 2005. 195 с.
3. **Тихонов А.Н.** О некорректных задачах линейной алгебры и устойчивом методе их решения // Докл. АН СССР. 1965. Т. 163, № 4. С. 591–594.
4. **Mangasarian O.L.** A finite Newton method for classification // Optim. Methods Softw. 2002. Vol. 17, no. 5. P. 913–929.
5. **Ортега Дж., Рейнболдт В.** Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975. 560 с.

6. **Голиков А.И., Евтушенко Ю.Г.** Нахождение проекции заданной точки на множество решений задач линейного программирования // Тр. Ин-та математики и механики УрО РАН. 2008. Т. 14, № 2. С. 33–47.
7. **Голиков А.И., Евтушенко Ю.Г.** Обобщенный метод Ньютона для задач линейной оптимизации с ограничениями-неравенствами // Тр. Ин-та математики и механики УрО РАН. 2013. Т. 19, №. 2. С. 98–108.
8. **Mangasarian O.L.** Nonlinear programming. New York: McCrow-Hill, 1969. 220 p.
9. **Голиков А.И., Евтушенко Ю.Г.** Теоремы об альтернативах и их применение в численных методах // Журн. вычисл. математики и мат. физики. 2003. Т. 43, № 3. С. 354–375.
10. **Еремин И.И., Мазуров В.Д., Астафьев Н.Н.** Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983. 335 с.
11. **Frank M., Wolfe P.** An algorithm for quadratic programming // Naval Res. Logist. Quart. 1956. Vol. 3. P. 95–110.
12. **Алберт А.** Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977. 224 с.
13. **Mangasarian O.L.** A Newton method for linear programming // J. Optim. Theory Appl. 2004. Vol. 121, no. 1. P. 1–18.
14. **Kanzow C., Qi H., Qi L.** On the minimum norm solution of linear programs // J. Optim. Theory Appl. 2003. Vol. 116, no. 2. P. 333–345.
15. Параллельная реализация метода Ньютона для решения больших задач линейного программирования / В.А. Гаранжа, А.И. Голиков, Ю.Г. Евтушенко, М.Х. Нгуен // Журн. вычисл. математики и мат. физики. 2009. Т. 49, № 8. С. 1369–1384.
16. **Попов Л.Д.** Квадратичная аппроксимация штрафных функций при решении задач линейного программирования большой размерности // Журн. вычисл. математики и мат. физики. 2007. Т. 47, № 2. С. 206–221.

Голиков Александр Ильич
канд. физ.-мат. наук
вед. науч. сотрудник
ВЦ РАН им. А.А.Дородницына

Поступила 14.01.2014

Евтушенко Юрий Гаврилович
д-р физ.-мат. наук, академик РАН
директор
ВЦ РАН им. А.А. Дородницына
e-mail: evt@ccas.ru