ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 28 № 4 2022

УДК 519.658.4

ОБ УПРАВЛЕНИИ ПАРАМЕТРАМИ В ИТЕРАЦИОННЫХ МЕТОДАХ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ, ОСНОВАННЫХ НА НОВОМ КЛАССЕ ГЛАДКИХ ВНЕШНИХ ШТРАФНЫХ ФУНКЦИЙ[1]

Л. Д. Попов

Приведены новые результаты по конструированию внешних штрафных функций повышенной гладкости в линейном программировании и по построению на их основе итерационных методов с автоматическим согласованием их параметров. Новые конструкции, подобно внутренним штрафным функциям, позволяют применять методы оптимизации второго порядка и в то же время не требуют знания хотя бы одной внутренней допустимой точки исходной задачи для своего старта. Более того, новые штрафные функции могут быть применены и к несобственным задачам линейного программирования (задачам с противоречивыми системами ограничений), для которых они способны вырабатывать обобщенные (компромиссные) решения. Приведены теоремы сходимости и данные численных экспериментов.

Ключевые слова: линейное программирование, несобственные задачи, обобщенные решения, метод штрафных функций, метод Ньютона.

L. D. Popov. On parameter control in iterative linear programming methods based on a new class of smooth exterior penalty functions.

New results are presented on the construction of exterior penalty functions of increased smoothness in linear programming and on the construction of iterative methods on their basis with automatic matching of their parameters. New constructions, similarly to interior penalty functions, make it possible to use second-order optimization methods and at the same time do not require knowledge of at least one interior admissible point of the original problem for the start of the operation. Moreover, the new penalty functions can also be applied to improper linear programming problems (problems with inconsistent constraint systems), for which they can produce generalized (compromise) solutions. Convergence theorems are proved and data of numerical experiments are presented.

Keywords: linear programming, improper (ill-posed) problems, generalized solutions, penalty functions method, Newton method.

MSC: 90C05, 90C51, 90C53

DOI: 10.21538/0134-4889-2022-28-4-191-200

Введение

При решении задач линейного программирования, наряду с классическим симлекс-методом Дж. Данцига [1], часто обращаются к приближенным методам итерационного типа [2–9], позволяющим быстро получить первые характеристики решаемой задачи, выявить возможные погрешности модели и, как некий бонус, найти хорошую стартовую точку для симплекс-метода с тем, чтобы сократить время его работы. Итерационные методы в той или иной степени привязаны к аппарату штрафных функций, которые делятся на точные [5], внешние [2–4; 6] и барьерные (или внутренние) [7–9]. Последние тесно связаны с популярными методами внутренней точки и центрального пути [10]. При этом каждый класс имеет свои сильные и слабые стороны.

Так, точные штрафные функции способны дать решение исходной задачи уже при конечных (хотя и больших) значениях штрафного параметра, но в силу своей недифференцируемости не допускают применения методов оптимизации даже первого порядка.

[1]Работа выполнена в рамках исследований, проводимых в Уральском математическом центре при финансовой поддержке Министерства науки и высшего образования Российской Федерации (номер соглашения 075-02-2022-874).

Напротив, внутренние (или барьерные) штрафные функции позволяют получать решение исходной задачи только асимптотически, при устремлении штрафного параметра к бесконечности, и для их запуска необходимо изначально знать хотя бы одну внутреннюю точку допустимой области решаемой задачи. Однако их высокая гладкость позволяет использовать для их оптимизации высокоэффективные методы второго порядка (методы Ньютона). То же относится к методам внутренней точки и центрального пути.

Промежуточное положение по степени гладкости вспомогательных оптимизационных подзадач занимают внешние штрафные функции. С одной стороны, они позволяют применять для своей оптимизации методы первого порядка (градиентные), но с другой — не вызывают проблем, связанных с поиском стартовой точки вычислительного процесса. Более того, они обладают тем важным свойством, что могут применяться с успехом не только к разрешимым задачам, но и к задачам с такой особенностью, как несовместность системы исходных ограничений [11–13].

В данной работе автор предлагает новые конструкции внешних штрафных функций (повышенной гладкости), которые бы совмещали в себе возможности применения метода Ньютона с простотой начального старта алгоритма и имели бы более широкую сферу применения по сравнению с методами центрального пути. Статья является развитием исследований [14–16].

1. Внешние штрафные функции повышенной гладкости

Рассмотрим задачу линейного программирования (ЛП) вида

$$\max\{(c,x)\colon\ Ax \le b\}; \tag{1.1}$$

где числовая матрица $A=(a_{ij})_{m\times n}$ и векторы c и b заданы; x — вектор прямых переменных (неизвестных), значения которых надо определить; круглые скобки используются для обозначения скалярного произведения; $m > n = \operatorname{rank} A$.

Будем также обозначать строки матрицы A как a_1, a_2 и так далее до a_m.

Следуя [16], поставим в соответствие задаче (1.1) штрафную конструкцию

$$F^{\varepsilon}(x,u) = (c,x) + \varepsilon_2 \sum_{i=1}^{n} \ln\bigl(u_i + b_i - (a_i,x)\bigr) - \frac{1}{2\,\varepsilon_1}\,\|u\|^2, \quad \varepsilon = (\varepsilon_1,\varepsilon_2) > 0. \tag{1.2}$$

Здесь к исходным переменным $x \in \mathbb{R}^n$ прямой задачи добавлены дополнительные переменные $u \in \mathbb{R}^m$, существенно расширяющие область определения обычной барьерной штрафной функции. Дополнительное штрафное слагаемое с нормой добавленного вектора отвечает за минимизацию вносимых в задачу изменений.

Введенная выше функция позволяет связать исходную условно-экстремальную задачу (1.1) с задачей безусловной оптимизации вида

$$F^{\varepsilon}(\hat{x}_\varepsilon, \hat{u}_\varepsilon) = \max_{x,u} F^{\varepsilon}(x,u).$$

Как показано в уже упомянутых работах автора, в предположении ограниченности оптимального множества исходной задачи (и даже при более слабом условии телесности допустимой области двойственной задачи) функция (1.2) имеет однозначно определяемые точки глобального максимума при всех $\varepsilon = (\varepsilon_1,\varepsilon_2) > 0$. При этом вектор $\hat{x}_\varepsilon$ будучи допустимым для скорректированной по правым частям исходной задачи

$$\max\{(c,x)\colon\ Ax \le b + \hat{u}_\varepsilon\} := \operatorname{opt}(\hat{u}_\varepsilon),$$

обладает свойством $\operatorname{opt}(\hat{u}_\varepsilon) - m\,\varepsilon_1 \le (c,\hat{x}_\varepsilon) \le \operatorname{opt}(\hat{u}_\varepsilon)$, т. е. является ее приближенным решением. В свою очередь, в случае разрешимости исходной задачи имеет место сходимость $\hat{u}_\varepsilon \to 0$, а в случае противоречивости ее ограничений — сходимость $\hat{u}_\varepsilon \to u_0$, где u_0 — минимальный по норме вектор коррекции ее правых частей, приводящих систему ограничений задачи к совместному виду. Таким образом, решая задачу (гладкой) безусловной максимизации вспомогательной функции (1.2) при достаточно малых значениях параметров $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$, можно получить аппроксимационное решение исходной задачи со сколь угодно высокой точностью при весьма слабых предположениях.

Наконец, отметим высокие свойства гладкости введенной конструкции (она дважды непрерывно-дифференцируема). Так что здесь применим метод Ньютона. При этом в отличие от обычных барьерных функций найти точки из ее области определения для старта этого метода не составляет труда: можно взять произвольный вектор основных переменных x^0 и уже по нему выставить в качестве дополнительной компоненты u^0 любой вектор просто с достаточно большими координатами, удовлетворяющими неравенствам $u^0 > Ax^0 - b$.

К сожалению, платой за хорошие свойства новых штрафных функций выступает расширение размерности пространства их переменных. И именно борьбе с этим недостатком посвящена данная статья. Но напомним вначале общую канву метода в расширенном пространстве переменных.

2. Метод Ньютона в расширенном пространстве переменных

Выпишем необходимые и достаточные условия того, что пара (x, u) является точкой максимума функции (1.2):

$$\nabla_x F^\varepsilon(x,u) = c - \varepsilon_2\, A^\top \mathrm{diag}(u + b - Ax)^{-1} e = 0, \tag{2.1}$$

$$\nabla_u F^\varepsilon(x,u) = \varepsilon_2\, \mathrm{diag}(u + b - Ax)^{-1} e - \frac{1}{\varepsilon_1} u = 0; \tag{2.2}$$

здесь для краткости $e = (1, 1, \ldots, 1)$.

Представим полученную систему уравнений в нескольких, более удобных для дальнейшего анализа, видах. Для начала введем дополнительную переменную $w = u + b - Ax > 0$. Получим эквивалентную перезапись соотношений (2.1), (2.2):

$$w - u + Ax - b = 0, \quad w > 0, \tag{2.3}$$

$$\varepsilon_2\, A^\top \mathrm{diag}(w)^{-1} e - c = 0, \tag{2.4}$$

$$\varepsilon_2\, \mathrm{diag}(w)^{-1} e - \frac{1}{\varepsilon_1} u = 0.$$

Из последнего соотношения выразим вектор $u = \varepsilon_1 \varepsilon_2\, \mathrm{diag}(w)^{-1} e$ и исключим его из соотношений (2.3), (2.4). Получим

$$w - \varepsilon_1 \varepsilon_2\, \mathrm{diag}(w)^{-1} e + Ax - b = 0, \tag{2.5}$$

$$\varepsilon_2\, A^\top \mathrm{diag}(w)^{-1} e - c = 0. \tag{2.6}$$

Введя еще одну дополнительную переменную $y = \varepsilon_2\, \mathrm{diag}(w)^{-1} e > 0$, преобразуем последние соотношения к виду, наиболее близко напоминающему соотношения метода центрального пути:

$$w - \varepsilon_1\, y + Ax - b = 0, \quad w,\, y > 0, \tag{2.7}$$

$$\mathrm{diag}(y)\, \mathrm{diag}(w) e - \varepsilon_2\, e = 0, \tag{2.8}$$

$$A^\top y - c = 0. \tag{2.9}$$

Наконец, введем вспомогательное отображение $\Phi(w,y,x)$ пространства $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n$ в себя по правилу

$$\Phi(w,y,x) = \begin{pmatrix} w - \varepsilon_1\, y + Ax - b \\ \operatorname{diag}(y)\operatorname{diag}(w)\, e - \varepsilon_2\, e \\ A^\top y - c \end{pmatrix}.$$

Введенное отображение позволяет переписать систему (2.7)–(2.9) совсем кратко:

$$\Phi(w,y,x) = 0. \tag{2.10}$$

Для решения системы (2.10) можно применить метод Ньютона. Последний заключается в построении последовательности приближений $z^k = (w^k, y^k, x^k)$ к решению этой системы по рекуррентным формулам

$$z^{k+1} = z^k - \alpha_k\, p^k, \ \text{ где } \ p^k = \big(\nabla\,\Phi(z^k)\big)^{-1}\Phi(z^k), \quad k = 0,1,2,\dots\,. \tag{2.11}$$

В начальном приближении w^0, y^0 положительные, а параметр шага $\alpha_k = 1$ обычно постоянен.

Основные усилия в рассматриваемом варианте метода Ньютона заключаются в отыскании решения p^k системы уравнений

$$\nabla\,\Phi(z^k)\, p^k = \Phi(z^k),$$

или, в детальном виде, системы

$$\begin{pmatrix} E & -\varepsilon_1 E & A \\ Y & W & 0 \\ 0 & A^\top & 0 \end{pmatrix} \begin{pmatrix} p_w \\ p_y \\ p_x \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}, \tag{2.12}$$

где E — единичная матрица, $Y = \operatorname{diag}(y^k)$, $W = \operatorname{diag}(w^k)$, $p^k = (p_w, p_y, p_x)$. Перед нами блочная система линейных уравнений. Формулы ее решения были найдены в [16].

Утверждение 1. *Решение системы* (2.12) *может быть последовательно получено по формулам*

$$p_x = (A^\top H\, Y\, A)^{-1}\big[h_3 - A^\top H\,(h_2 - Y\, h_1)\big],$$
$$p_y = H\,(h_2 - Y\, h_1) + H\, Y\, A\, p_x,$$
$$p_w = h_1 + \varepsilon_1\, p_y - A\, p_x,$$

где $H = (W + \varepsilon_1\, Y)^{-1}$.

Мы видим, что основной объем вычислений в методе связан с обращением положительно определенной симметрической матрицы $Q = A^\top H\, Y\, A = A^\top (W + \varepsilon_1\, Y)^{-1} Y\, A$ размерности $n \times n$. Более того, вместо вычисления и обращения этой матрицы для решения последней из соответствующих линейных подсистем

$$(A^\top H\, Y\, A)\, p_x = h_3 - A^\top H\,(h_2 - Y\, h_1)$$

можно воспользоваться методом сопряженных направлений. Это позволит эффективно учесть также возможную разреженность исходных матриц.

Приведенный вариант применения новых штрафных функций выглядит несколько громоздко. Мы попробуем уменьшить размерность переменных в нашей системе до размерности вектора исходных переменных решаемой задачи ЛП.

3. Сокращение числа переменных при поиске максимума новой штрафной функции

Вернемся к системе (2.5), (2.6):

$$w - \varepsilon_1\,\varepsilon_2\,\mathrm{diag}(w)^{-1}e + Ax - b = 0,$$

$$\varepsilon_2\,A^\top \mathrm{diag}(w)^{-1}e - c = 0.$$

Заметим, что первое уравнение на самом деле определяет w как неявную функцию $w = w(x)$ переменной x, так что фактически перед нами серия обыкновенных квадратных уравнений относительно компонент w_i. Они получаются, если обе части первого уравнения умножить слева на матрицу $\mathrm{diag}(w)$:

$$w_i^2 + w_i\big((a_i, x) - b_i\big) - \varepsilon_1\,\varepsilon_2 = 0, \quad i = 1, 2, \ldots, m.$$

Выбирая положительные корни этих уравнений, получаем аналитические зависимости

$$w_i(x) = \frac{b_i - (a_i, x) + \sqrt{[b_i - (a_i, x)]^2 + 4\,\varepsilon_1\,\varepsilon_2}}{2}, \quad i = 1, 2, \ldots, m \tag{3.1}$$

и

$$\frac{\partial w_i(x)}{\partial x_j} = -\frac{a_{ij} w_i(x)}{2\,w_i(x) + (a_i, x) - b_i}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n. \tag{3.2}$$

Теперь можно посмотреть на второе соотношение исходной системы

$$\Psi^\varepsilon(w(x)) = \varepsilon_2\,A^\top \mathrm{diag}(w(x))^{-1}\,e - c = 0 \tag{3.3}$$

как на неявное уравнение для отыскания x. В нем зависимости $w_i = w_i(x)$ представлены соотношениями (3.1), (3.2).

Чтобы применить для решения (3.3) метод Ньютона в пространстве переменных x сокращенной размерности, нам надо найти якобиан этой системы по x.

Утверждение 2. *При сделанных выше предположениях*

$$\nabla\Psi^\varepsilon(w(x)) = \varepsilon_2\,A^\top \mathrm{diag}\big(w(x)\big)^{-1}\mathrm{diag}\big(2w(x) + Ax - b\big)^{-1}A. \tag{3.4}$$

Доказательство. Имеем

$$\Psi_j^\varepsilon(w(x)) = \varepsilon_2 \sum_{s=1}^{m} \frac{a_{sj}}{w_s(x)} - c_j, \quad j = 1, 2, \ldots, n,$$

и значит,

$$\frac{\partial \Psi_j^\varepsilon(w(x))}{\partial x_i} = -\varepsilon_2 \sum_{s=1}^{m} \frac{a_{sj}}{w_s(x)^2}\,\frac{\partial w_s(x)}{\partial x_i} = \varepsilon_2 \sum_{s=1}^{m} \frac{a_{sj} a_{si}}{w_s(x)[2\,w_s + (a_s, x) - b_s]},$$

$$i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n,$$

что и требовалось. □

Итак, предлагаемый здесь альтернативный алгоритм Ньютона работает в пространстве только прямых переменных x, хотя и описывается все теми же формулами

$$x^{k+1} = x^k - \alpha_k\big(\nabla\Psi^\varepsilon(w(x^k))\big)^{-1}\Psi^\varepsilon(w(x^k)), \quad k = 0, 1, 2, \ldots\,. \tag{3.5}$$

Существенное снижение размерности пространства переменных нового процесса по сравнению с изложенным в предыдущем разделе сопровождается необходимостью расчета показателей неявной функции $w = w(x)$, однако этот расчет оказывается достаточно простым (см. (3.1), (3.4)) по сравнению с формулами из утверждения 1. Впрочем, тут есть и подводные камни, обсуждение которых пойдет в следующей секции.

4. Управление штрафными параметрами на основе обратной связи с достигнутой точностью решения

Предварительные вычислительные эксперименты выявили различия в поведении методов (2.11) и (3.5), часть которых играет в пользу первого из них, а часть — в пользу второго.

Так, метод (3.5), работающий в пространстве только прямых переменных, требует меньшего объема оперативной памяти компьютера и заметно меньшего объема вычислений на одну итерацию. Вместе с тем выяснилось, что область сходимости исходного метода (2.11) мало зависит от величины штрафных параметров, так что их можно изначально установить на низкий уровень для достаточно широкого диапазона начальных значений переменных задачи. Напротив, штрафная функция из (3.5) имеет более сложную природу, и потому локальная область сходимости метода Ньютона для нее сильно зависит от принятых значений штрафных параметров ε_1, ε_2 и исходных приближений переменных x. Это означает, вообще говоря, что при заданном произвольно начальном приближении штрафные параметры итерационного процесса (3.5) поначалу должны лежать в среднем диапазоне (чтобы обеспечить его локальную сходимость) и лишь постепенно, по мере уменьшения $\|\Psi^\varepsilon(w(x^k))\|$, значение этих параметров может снижаться.

Почему же управление штрафными параметрами на основе принципа обратной связи по отношению к скорости сходимости основного процесса наиболее разумно увязывать с величиной $\|\Psi^\varepsilon(w(x^k))\|$?

Дело в том, что метод (3.5), работая с последовательностью $\{x^k\}$, неявно определяет связанные с ней последовательности $\{y^k\}$ и $\{u^k\}$, где

$$y^k = \varepsilon_2 \operatorname{diag}(w(x^k))^{-1}e, \quad u^k = w(x^k) + Ax^k - b \ge Ax^k - b.$$

При этом вектор x^k является допустимым для прямой задачи

$$\max\{(c,x)\colon Ax \le b + u^k\},$$

а вектор y^k нарушает ограничения двойственной к ней задачи

$$\min\{(b+u^k, y)\colon A^\top y = c,\ y \ge 0\}$$

как раз на величину

$$A^\top y^k - c = \varepsilon_2 A^\top \operatorname{diag}(w(x^k))^{-1}e - c = \Psi^\varepsilon(w(x^k)).$$

В то же время различия в значениях целевых функций выписанных задач в этих точках отличаются на величину

$$\Delta = (b+u^k, y^k) - (c, x^k) = (Ax^k + w(x^k), y^k) - (c, x^k)$$
$$= \varepsilon_2 \operatorname{diag}(w(x^k))^{-1} w(x^k) - (c - A^\top y^k, x^k) = n\varepsilon_2 + (\Psi^\varepsilon(w(x^k)), x^k),$$

т. е. пропорционально значению штрафных параметров и все тому же значению $\|\Psi^\varepsilon(w(x^k))\|$. Тем самым нет смысла брать особо малые значения штрафных параметров до тех пор, пока малым не станет значение невязки системы (3.3).

Итак, стартовые значения штрафных параметров целесообразно выбирать из среднего диапазона чисел, тем больших, чем больше предполагаемое отклонение стартового приближения x^0 от точного решения. Большие значения штрафных параметров обеспечивают более широкую область локальной сходимости метода Ньютона. В ходе вычислений значение $\|\Psi^\varepsilon(w(x^k))\|$ постепенно снижается, и в тот момент, когда оно достигнет приемлемого значения $O(\varepsilon_1, \varepsilon_2)$, можно несколько понизить значения штрафных параметров с тем, чтобы обеспечить одинаковую скорость сходимости к нулю невязок прямой и двойственной задач и разности текущих значений их целевых функций.

Ниже приведен алгоритм для пакета MatLab, с которым проводились эксперименты.

`А л г о р и т м  в сокращенном пространстве для пакета MatLab`

```
mo = 3*no
A  = sprand(mo,no,nzproc/2) - sprand(mo,no,nzproc/2);
A  = A + [speye(no,no);speye(no,no);speye(no,no)];

xo = ones(no,1);
yo = [ones(no,1);zeros(no,1);zeros(no,1)];
b  = A*xo + [zeros(no,1);ones(no,1);ones(no,1)];
c  = A'*yo;
Fo = c'*xo

x = 10*xo;
w = ones(mo,1);
e = ones(mo,1);

for it = 1:ITMAX
    AXB = A*x-b;
    for i=1:mo
        w(i)=(-AXB(i)+sqrt(AXB(i)*AXB(i)+4*eps*eps))/2;
    end
    GR = A'*inv(diag(w))*e-c/eps;
    H  = A'*inv(diag(w)*diag(2*w+AXB))*A;
    x  = x - H\GR;

    [it,eps,norm(GR),norm(x-xo),abs(Fo-c'*x),norm(max(0,AXB))]

    if eps*norm(GR)<xTOL
       eps = max(xEPSmin,eps*0.3);
    end
    if norm(x-xo)< xFIN
       break
    end
end
```

Здесь вначале с помощью датчика случайных чисел генерируется тестовая задача ЛП. Матрица ее ограничений `A` имеет задаваемые пользователем до запуска макроса `no` столбцов и `mo=3no` строк. Также пользователем задается процент заполнения этой матрицы ненулевыми элементами `nzproc`. Правые части ограничений `b` и вектор целевой функции `c` строятся так, чтобы задача имела известные заранее прямое и двойственное решения `xo` и `yo`. Оптимальное значение построенной задачи запоминается в переменной `Fo`.

Так же задаются начальные приближения для `x`.

Затем начинается собственно итерационный процесс. Число итераций ограничено переменной `ITMAX`. Итерация начинается с вычисления вектора текущих невязок ограничений прямой задачи `AXB`. Затем вычисляются корни квадратных уравнений (компоненты вектора `w`). Далее определяются `GR` — масштабированное значение левой части системы уравнений (3.3) и `H` — ее якобиан, после чего происходит переопределение текущего приближения `x` по схеме Ньютона.

Далее идут вывод на экран промежуточных характеристик процесса и блока адаптивной настройки параметров алгоритма. Среди них `eps` — общее значение двух штрафных параметров и `xTOL` — константа, задаваемая пользователем и принимающая среднее числовое значение. Ее предназначение — сделать все невязки, которые мы обсуждали выше, одного порядка малости с `eps`. Также для предупреждения получения вырожденного якобиана задается минимально допустимое значение штрафных параметров `xEPSmin`. Поскольку решение известно заранее, тестируемому алгоритму разрешено завершить работу до истечения заданного числа итераций при условии достижения приемлемой точности решения.

5. Вычислительный эксперимент

Вычислительный эксперимент проводился в среде MATLAB на вычислительном комплексе УРАН ИММ УрО РАН.

Были обработаны результаты 50 тестовых задач ЛП размерности 1000×3000. Их матрицы ограничений генерировались при помощи датчика случайных чисел с равномерным распределением от -1 до 1. Матрицы имели разреженную структуру. Заполнение матриц ненулевыми элементами составляло 10–11%. Правые части ограничений и коэффициенты целевой функции подбирались таким образом, чтобы решение задачи было единственным и совпадало с некоторым заданным заранее (см. приведенный выше алгоритм). Стартовые приближения во всех случаях были одинаковыми (их значения на порядок отличались от оптимальных).

Цель экспериментов состояла в проверке адаптационных свойств алгоритма, т. е. устойчивости его скорости сходимости к вариации стартовых значений параметров `xEPS` и `xTOL`.

Результаты расчетов приведены ниже в таблице. Это число шагов сокращенного метода Ньютона, потребовавшихся для достижения заданной точности решения при различных начальных значениях параметров `xEPS` и `xTOL`. Точность определялась по условию $\|\texttt{x-xo}\|<0.00001$ (при этом точность выполнения ограничений прямой задачи и достижения ее оптимального значения в конечной точке всегда получалась на один-два порядка выше).

Среднее число шагов сокращенного метода Ньютона, потребовавшееся для достижения заданной точности по x

	xEPS = 10	xEPS = 1	xEPS = 0.1	xEPS = 0.01	xEPS = 0.001
xTOL = 0.01	62	47	44	40	–
xTOL = 0.1	55	54	38	38	–
xTOL = 1	54	46	42	30	–
xTOL = 10	42	38	26	37	–
xTOL = 100	33	28	25	29	–
xTOL = 1000	23	25	25	28	–
xTOL = 10000	14	13	15	–	–
xTOL = 100000	16	12	–	–	–

Мы видим, что задавать малые стартовые значения штрафных параметров и, напротив, очень большие значения параметра `xTOL` опасно из-за возможной потери сходимости. При средних значениях этих параметров предлагаемая алгоритмическая схема их взаимоувязки работает достаточно уверенно.

Заключение

В настоящей работе предложены оригинальные алгоритмы решения задач ЛП, основанные на новом классе внешних штрафных функций повышенной гладкости. Алгоритмы включают в себя методы оптимизации второго порядка и процедуры автоматического согласования своих параметров в ходе вычислений. Новые штрафные функции могут быть применены и к несобственным задачам линейного программирования (к задачам с противоречивыми системами ограничений), для которых они способны вырабатывать обобщенные (компромиссные) квазирешения.

СПИСОК ЛИТЕРАТУРЫ

1. **Данциг Дж.** Линейное программирование, его обобщения и применения. М.: Прогресс, 1966. 600 с.
2. **Фиакко А., Мак-Кормик Г.** Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972. 240 с.
3. **Зангвилл У. И.** Нелинейное программирование. Единый подход. М.: Сов. радио, 1973. 312 с.
4. **Полак Э.** Численые методы оптимизации. М.: Мир, 1974. 376 с.
5. **Еремин И. И.** Метод штрафов в выпуклом программировании // Докл. АН СССР. 1967. Т. 173, № 4. С. 748–751.
6. **Моисеев Н. Н., Иванилов Ю. П., Столярова Е. М.** Методы оптимизации. М.: Наука, 1978. 351 с.
7. **Евтушенко Ю. Г.** Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982. 432 с.
8. **Гилл Ф., Мюррей У., Райт М.** Практическая оптимизация. М.: Мир, 1985. 509 с.
9. **Васильев Ф.П.** Численные методы решения экстремальных задач. М.: Наука, 1988. 552 с.
10. **Roos C., Terlaky T., Vial J.-Ph.** Theory and algorithms for linear optimization. Chichester: John Wiley & Sons Ltd, 1997. 484 p.
11. **Еремин И.И.** Двойственность для несобственных задач линейного и выпуклого программирования // Докл. АН СССР. 1981. Т. 256, № 2. С. 272–276.
12. **Еремин И. И., Мазуров Вл. Д., Астафьев Н. Н.** Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983. 336 с.
13. **Кочиков И. В., Матвиенко А. Н., Ягола А. Г.** Обобщенный принцип невязки для решения несовместных уравнений // Журн. вычисл. математики и мат. физики. 1984. Т. 24, № 7. С. 1087–1090.
14. **Попов Л. Д.** Поиск обобщенных решений несобственных задач линейного и выпуклого программирования с помощью барьерных функций // Изв. Иркутского гос. ун-та. Сер. Математика. 2011. Т. 4, № 2. С. 134–146.
15. **Попов Л. Д.** Применение барьерных функций для оптимальной коррекции несобственных задач линейного программирования 1-го рода // Автоматика и телемеханика. 2012. Вып. 3. С. 3–11.
16. **Попов Л. Д.** Об одном приеме повышения гладкости внешних штрафных функций в линейном и выпуклом программировании // Тр. Ин-та математики и механики УрО РАН. 2021. Т. 27, № 4. С. 88–101.

Поступила 19.07.2022
После доработки 21.09.2022
Принята к публикации 26.09.2022

Попов Леонид Денисович
д-р физ.-мат. наук, ведущий науч. сотрудник
Институт математики и механики им. Н. Н. Красовского УрО РАН;
профессор, Институт математики и компьютерных наук
Уральский федеральный университет им. Б. Н. Ельцина
г. Екатеринбург
e-mail: popld@imm.uran.ru

REFERENCES

1. Dantzig George B. Linear programming and extensions. Ser. Princeton Landmarks in Mathematics and Physics, Princeton: Princeton University Press, 1963, 656 p. Translated to Russian under the title *Lineinoe programmirovanie, ego obobshcheniya i primeneniya*, Moscow: Progress Publ., 1966, 600 p.

2. Fiacco A.V., McCormick G.P. *Nonlinear programming: sequential unconstrained minimization techniques*. Classics in applied mathematics, vol. 4. Philadelphia: SIAM, 1987, 226 p. ISBN: 0898712548 . Translated to Russian under the title *Nelineinoe programmirovanie. Metody posledovatel'noi bezuslovnoi minimizatsii*, Moscow: Mir Publ., 1972, 240 p.
3. Zangwill W.I. *Nonlinear programming: a unified approach*. Englewood Cliffs: Prentice-Hall, 1969, 356 p. ISBN: 0136235794 . Translated to Russian under the title *Nelineinoe programmirovanie. Edinyi podkhod*, Moscow: Sov. Radio Publ., 1973, 312 p.
4. Polak E. *Computational methods in optimization*. NY: Acad. Press, 1971, 329 p. ISBN: 9780080960913 . Translated to Russian under the title *Chislenye metody optimizatsii*, Moscow: Mir Publ., 1974, 376 p.
5. Eremin I.I. The "penalty" method in convex programming. *Sov. Math., Dokl.*, 1967, vol. 8, pp. 459–462.
6. Moiseev N.N., Ivanilov Yu.P., Stolyarova E.M. *Metody optimizatsii* [Optimization methods]. Moscow: Nauka Publ., 1978, 351 p.
7. Evtushenko Yu.G. *Metody resheniya ekstremal'nykh zadach i ikh primenenie v sistemakh optimizatsii* [Methods for solving extremal problems and their application in optimization systems]. Moscow: Nauka Publ., 1982, 432 p.
8. Gill P.E., Murray W., Wright M.H. *Practical optimization*. Philadelphia: SIAM, 2019, 401 p. doi: 10.1137/1.9781611975604 . Translated to Russian under the title *Prakticheskaya optimizatsiya*. Moscow: Mir Publ., 1985, 509 p.
9. Vasilyev F.P. *Chislennye metody resheniya ekstremal'nykh zadach* [Numerical methods for solving extremal problems]. Moscow: Nauka Publ., 1988, 552 p. ISBN: 5-02-013796-0 .
10. Roos C., Terlaky T., Vial J.-Ph. *Theory and algorithms for linear optimization*. NY: Wiley, 1997, 454 p. ISBN: 047195676 .
11. Eremin I.I. Duality for improper problems of linear and convex programming. *Sov. Math., Dokl.* 1981, vol. 23, pp. 62–66.
12. Eremin I.I., Mazurov Vl.D., Astaf'ev N.N. *Nesobstvennye zadachi lineinogo i vypuklogo programmirovaniya* [Improper problems of linear and convex programming]. Moscow: Nauka Publ., 1983, 336 p.
13. Kochikov I.V., Matvienko A.N., Yagola A.G. The generalized residual principle for the solution of inconsistent equations. *Comput. Math. Math. Phys.*, 1984, vol. 24, no. 4, pp. 78–80. doi: 10.1016/0041-5553(84)90233-7 .
14. Popov L.D. Search of generalized solutions to improper linear and convex programming problems using barrier functions. *Bull. Irkutsk State Univ. Ser. Math.*, 2011, vol. 4, no. 2, pp. 134–146 (in Russian).
15. Popov L.D. Use of barrier functions for optimal correction of improper problems of linear programming of the 1st kind. *Autom. Remote Control*, 2012, vol. 73, no. 3, pp. 417–424. doi: 10.1134/S0005117912030010 .
16. Popov L.D. On one method of increasing the smoothness of external penalty functions in linear and convex programming. *Trudy Inst. Mat. Mekh. UrO RAN*, 2021, vol. 27, no. 4, pp. 88–101 (in Russian). doi: 10.21538/0134-4889-2021-27-4-88-101 .

Received July 19, 2022
Revised September 21, 2022
Accepted September 26, 2022

Funding Agency: The work was performed as part of research conducted in the Ural Mathematical Center with the financial support of the Ministry of Science and Higher Education of the Russian Federation (Agreement number 075-02-2022-874).

Leonid Denisovich Popov, Dr. Phys.-Math. Sci., Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620108 Russia; Ural Federal University, Yekaterinburg, 620000 Russia, e-mail: popld@imm.uran.ru .

Cite this article as: L. D. Popov. On parameter control in iterative linear programming methods based on a new class of smooth exterior penalty functions. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2022, vol. 28, no. 4, pp. 191–200 .