

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

А. В. Чернов

Задачи энтропийно-линейного программирования часто возникают в различных приложениях (транспортные задачи, исследования химических реакций и др.). Такие задачи формулируются обычно как задачи максимизации энтропии (или минимизации минус энтропии) с аффинными ограничениями и линейными ограничениями-неравенствами. В работе исследуется метод решения такой задачи, в основе которого лежит решение двойственной задачи с восстановлением решения прямой задачи: каждой точке в двойственном пространстве, вычисляемой методом, ставится в соответствие определенная точка в прямом. Для указанного метода получена верхняя оценка числа итераций, необходимого для достижения решения с заданной точностью. Изложенный метод применим также к более широкому классу сильно выпуклых функционалов с аналогичным допустимым множеством.

Ключевые слова: энтропийно-линейное программирование, быстрый градиентный метод, двойственная задача, прямо-двойственный метод.

Введение

Данная статья посвящена исследованию метода решения задачи энтропийно-линейного программирования (ЭЛП), в кото-

ром двойственная задача решается с помощью быстрого градиентного метода (БГМ) [1]. Получены оценки скорости его сходимости и приведены результаты численных экспериментов.

Предлагаемый метод является прямо-двойственным, т.е. при решении строятся одновременно последовательности точек в двойственном и прямом пространствах. Условием окончания работы метода является выполнение определенного количества итераций (которое оценивается заранее) с последующей проверкой зазора двойственности в найденных точках (разница между значением прямой функции и двойственной функции не должна превышать требуемую величину) и условия, когда суммарная невязка ограничений не превышает заданного значения.

Задача ЭЛП часто возникает при исследовании равновесных распределений в транспортных задачах. Классическим примером формулирования задачи ЭЛП является парадокс Эренфестов [2, 3], где выделяется задача ЭЛП, как задача поиска равновесного распределения на допустимом множестве с линейными ограничениями типа равенства. В указанных работах рассматриваются только ограничения типа равенства, т.е. их допустимое множество является аффинным. Однако, существует также и потребность в решении задач, когда допустимое множество имеет более сложный вид и содержит помимо ограничений-равенств также и ограничениями неравенства вида $Cx \leq b$. Отличительной особенностью данных задач является их размерность, которая может достигать 1000000 и более (например, задача поиска равновесного распределения транспортных потоков в г. Москва). Формулировки таких задач можно найти, например, в [4, 5, 6].

В работе [7] предложено решение задачи ЭЛП с аффинными ограничениями с помощью регуляризации по Тихонову построенной двойственной функции и последующего решения двойственной задачи с восстановлением решения прямой задачи.

В данной работе рассматривается задача ЭЛП с допустимым множеством, которое определяется как ограничениями типа равенства так и ограничениями типа неравенства.

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

Последующая часть статьи имеет следующая имеет следующую структуру:

- постановка задачи;
- формулировка и доказательство вспомогательных утверждений;
- описание алгоритма поиска решения задачи, формулировка и доказательство основных утверждений;
- результаты численных экспериментов и сравнение предлагаемого метода с тем, что предложен в [7].

Постановка задачи

Рассмотрим на n -мерном вероятностном симплексе $S_n(1) = \left\{ x \in R_{++}^n : \sum_{i=1}^n x_i = 1 \right\}$ задачу ЭЛП (1).

$$f(x) = \sum_{i=1}^n x_i \ln x_i / \xi_i \rightarrow \min_{x \in G}; \quad (1)$$

$$G = \{x \in S_n(1) : C_1 x - b_1 \leq 0; C_2 x - b_2 = 0\}.$$

Здесь $\xi_i (i = 1, \dots, n)$ – параметры задачи, $C_1 \in R^{n \times m_1}$, $b_1 \in R^{m_1}$, $C_2 \in R^{n \times m_2}$, $b_2 \in R^{m_2}$.

Отметим, что задача (1) выпуклая, причем функция $f(x)$ сильно выпукла на R_{++}^n , что означает существование единственного решения.

Введем следующие обозначения для упрощения дальнейших выкладок:

$$C = [C_1; C_2] \in R^{n \times (m_1 + m_2)}; \quad b = [b_1; b_2] \in R^{m_1 + m_2}. \quad (2)$$

Используя введенные обозначения, несложно выписать двойственную задачу (3), где простым множеством является симплекс $S_n(1)$.

$$\begin{aligned} \psi(y) &\rightarrow \max_{y \in Q}, \text{ где } Q = R_+^{m_1} \times R^{m_2}; \\ \psi(y) &= \min_{x \in S_n(1)} \{f(x) + \langle y, Cx - b \rangle\} = f(x(y)) + \langle y, Cx(y) - b \rangle = \\ &= -\langle y, b \rangle - \ln \left[\sum_{i=1}^n \xi_i \exp(-[C^T y]_i) \right]. \end{aligned} \quad (3)$$

При этом функция $x(y)$, полученная при построении двойственной задачи, вычисляется по формуле

$$x_i(y) = \frac{\xi_i \exp(-[C^T y]_i)}{\sum_{j=1}^n \xi_j \exp(-[C^T y]_j)}. \quad (4)$$

Построенная задача эквивалентна задаче минимизации (5) для функции $\phi(y) \equiv -\psi(y)$, которая и будет рассматриваться далее.

$$\phi(y) = \langle y, b \rangle + \ln \left[\sum_{i=1}^n \xi_i \exp(-[C^T y]_i) \right] \rightarrow \min_{y \in Q}, \quad (5)$$

$$\text{где } Q = R_+^{m_1} \times R^{m_2}$$

Отметим, что функция $\phi(y)$ является сильно выпуклой на любом выпуклом подмножестве из

$$Q \setminus \{y \in Q : (C^T y)_i = (C^T y)_j \forall (i, j)\},$$

поэтому решение задачи (5) существует и единственно.

Пусть $D = D(x) = \|d_{i,j}(x)\|$ – диагональная матрица размерности $n \times n$, на диагонали которой стоят компоненты вектора x , т.е. $d_{i,i}(x) = x_i$ и $d_{i,j}(x) = 0$ при $i \neq j$. Несложно показать, что градиент и гессиан функции $\phi(y)$ вычисляются по формулам (6).

$$\nabla \phi(y) = b - Cx(y); \quad \phi''(y) = CD(x(y))C^T - Cx(y)(Cx(y))^T. \quad (6)$$

Введем также следующую функцию-срезку

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

$$z_+ = \begin{cases} z, & \text{если } z \geq 0; \\ 0, & \text{если } z < 0. \end{cases}$$

Определение 1. Под $(\varepsilon_f, \varepsilon_g)$ -решением задачи ЭЛП (1) будем понимать такую точку $x_t \in R^n$, что $|f(x_t) - f^*| < \varepsilon_f$ и $\Delta(x_t, G) < \varepsilon_g$, где f^* - точное решение задачи, а $\Delta(x, G) = \|(C_1x - b_1)_+\| + \|C_2x - b_2\|$ - невязка точки x для множества G .

Здесь и далее под нормой понимается евклидова норма, $\|z\| = \sqrt{z^T z}$, $z \in R^n$.

Вспомогательные результаты

Лемма 1 ([8]). *Имеет место следующее неравенство*

$$\|\nabla\phi(y_1) - \nabla\phi(y_2)\| \leq L\|y_1 - y_2\|,$$

где $L = \max_{1 \leq i \leq n} \|[C]^{(i)}\|_2^2$, $[C]^{(i)}$ - i -ый столбец матрицы C .

Далее будет рассматриваться решение задачи ЭЛП с помощью модификации быстрого градиентного метода (БГМ) [1]. Известно, что данный метод характеризуется следующими последовательностями ($k = 0, 1, 2, \dots$):

1) последовательности чисел α_i , A_k , τ_k :

$$\alpha_0 \in (0, 1]; \quad A_k = \sum_{i=0}^k \alpha_i; \quad \alpha_k^2 \leq A_k; \quad \tau_k = \frac{\alpha_{k+1}}{A_{k+1}}. \quad (7)$$

2) последовательности точек в двойственном пространстве \tilde{y}_k , \check{y}_k , y_k и соответствующая им последовательность точек в прямом пространстве x_k , определяемая на основе уравне-

ния (4)

$$\begin{aligned}\tilde{y}_k &= \arg \min_{y \in Q} \left\{ \frac{L}{2} \|y - y_k\|_2^2 + \phi(y_k) + \langle \nabla \phi(y_k), y - y_k \rangle \right\}; \\ \check{y}_k &= \arg \min_{y \in Q} \left\{ \frac{L}{2} \|y - y_0\|_2^2 + \sum_{i=0}^k \alpha_i (\phi(y_i) + \langle \nabla \phi(y_i), y - y_i \rangle) \right\}; \\ y_{k+1} &= \tau_k \check{y}_k + (1 - \tau_k) \tilde{y}_k; \quad x_k = \sum_{i=0}^k \frac{\alpha_i}{A_k} x(y_i).\end{aligned}\tag{8}$$

3) последовательности функций l_k , Ψ_k , а также их минимумы: l_k^* , Ψ_k^* :

$$\begin{aligned}l_k(y) &= \sum_{i=0}^k \alpha_i (\phi(y_i) + \langle \nabla \phi(y_i), y - y_i \rangle); \quad l_k^* = \min_{y \in Q} l_k(y); \\ \Psi_k(y) &= \frac{L}{2} \|y - y_0\|_2^2 + l_k; \quad \Psi_k^* = \min_{y \in Q} \Psi_k(y).\end{aligned}\tag{9}$$

В указанных выше обозначениях верны леммы, сформулированные далее.

Лемма 2. *Если множество Q представимо в виде*

$$Q = R_+^{m_1} \times R^{m_2},$$

то последовательности точек \tilde{y}_k , \check{y}_k , y_k можно записать в виде (10).

$$\begin{aligned}\tilde{y}_k^j &= \begin{cases} (y_k - \frac{1}{L} \nabla \phi(y_k))_+^j & \text{при } j \leq m_1; \\ (y_k - \frac{1}{L} \nabla \phi(y_k))^j & \text{при } j > m_1; \end{cases} \\ \check{y}_k^j &= \begin{cases} (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))_+^j & \text{при } j \leq m_1; \\ (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))^j & \text{при } j > m_1.\end{cases}\end{aligned}\tag{10}$$

Доказательство. Рассмотрим доказательство только для \tilde{y}_k . Доказательство для \check{y}_k получается совершенно аналогично тривиальной модификацией доказательства для \check{y}_k , приводимого далее.

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

Задача поиска $\arg \min$ на множестве Q для \check{y}_k эквивалентна задаче

$$\check{y}_k = \arg \min_{y \in Q} \left\{ \frac{L}{2} \langle y, y \rangle + \langle -Ly_0 + \sum_{i=0}^k \alpha_i \nabla \phi(y_i), y \rangle \right\}.$$

Очевидно, что полученная функция является сильно выпуклой с константой сильной выпуклости L , а условия Куна-Такера принимают вид

$$\begin{aligned} y - (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i) + \nu) &= 0; \\ \nu_j y^j &= 0; \quad \nu_j \geq 0 \text{ при } j \in [1, m_1]; \\ \nu_j &\equiv 0 \text{ при } j > m_1. \end{aligned}$$

Отсюда находим искомое выражение для \check{y}_k :

$$\check{y}_k^j = \begin{cases} (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))_+^j & \text{при } j \leq m_1; \\ (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))^j & \text{при } j > m_1. \end{cases}$$

□

Лемма 3 ([8]). Пусть $\phi(y)$ - выпуклая функция на выпуклом множестве Q , градиент которой удовлетворяет условию Липшица с константой L , функция $\Psi(y)$ определяется (9), а последовательность $\{\alpha_i\}_{i=0}^{\infty}$ удовлетворяет условию (7), тогда выполнется

$$A_k \phi(\check{y}_k) \leq \Psi_k^* = \min_{y \in Q} \Psi_k(y).$$

Лемма 4 ([9]). Если последовательность точек y_k , определяемая (8), сходится к некоторой точке y^* , то эта последовательность будет лежать в шаре

$$U_r(y^*) = \{y \in R_+^{m_1} \times R^{m_2} : \|y - y^*\| \leq r\},$$

где $r = O(\|y_0 - y^*\|)$.

Замечание. Отметим, что в экспериментах $\|y_k - y^*\| \leq c\|y_0 - y^*\|$, причем $c \leq 10$, а во многих случаях не превосходит $c \leq 3$.

Основные результаты

Используя обозначения и леммы определенные ранее, решение задачи (1) можно получить с помощью следующего представления БГМ

Алгоритм 1: Быстрый градиентный метод

Input: Последовательность $\{\alpha_i\}_{i=0}^\infty$, точка y_0

Output: Последовательности точек \tilde{y}_k, x_k

repeat

 Вычисляем

$$\tilde{y}_k^j = \begin{cases} (y_k - \frac{1}{L} \nabla \phi(y_k))_+^j & \text{при } j \leq m_1; \\ (y_k - \frac{1}{L} \nabla \phi(y_k))^j & \text{при } j > m_1; \end{cases}$$

$$\check{y}_k^j = \begin{cases} (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))_+^j & \text{при } j \leq m_1; \\ (y_0 - \frac{1}{L} \sum_{i=0}^k \alpha_i \nabla \phi(y_i))^j & \text{при } j > m_1; \end{cases}$$

$$y_{k+1} = \tau_k \tilde{y}_k + (1 - \tau_k) \check{y}_k; \quad x_k = \sum_{i=0}^k \frac{\alpha_i}{A_k} x(y_i).$$

until $f(x_k) + \phi(y_k) > \varepsilon_f$ или $\Delta(x_k, G) > \varepsilon_g$;

Отметим, что выход из алгоритма осуществляется при выполнении условий

$$f(x_k) + \phi(y_k) = f(x_k) - \psi(y_k) \leq \varepsilon_f.$$

Величина $f(x_k) - \psi(y_k)$ упоминалась ранее как зазор двойственности в точке (x_k, y_k) .

Приводимая ниже теорема уже была сформулирована ранее и доказана для ограниченных множеств [1] и для множеств, совпадающих со всем пространством. В работе [9] оценки, полученные в работе [1], распространяются на случай неогра-

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

нических допустимых множеств определяемых ограничениями-равенствами. Эти результаты можно распространить и на ограничения-неравенства, что и утверждается в теореме 1.

Теорема 1. *Предположим, что решение задачи существует и лежит в некотором шаре (в двойственном пространстве) радиуса R ($\|y^*\| \leq R$), который также содержит и все остальные точки последовательности y_k . Тогда*

$$\Delta(x_k, G) \leq \frac{2LR}{A_k}; \quad |f(x^*) - f(x_k)| \leq \frac{2LR^2}{A_k}; \quad f(x_k) + \phi(\tilde{y}_k) \leq \frac{2LR^2}{A_k}. \quad (11)$$

Доказательство. Из условия теоремы следует, что $\|y_k - y_0\| \leq 2R$ и $\|y_k - y_*\| \leq 2R$. Таким образом, исходная задача поиска минимума двойственной функции $\phi(y)$ на множестве Q эквивалентна задаче поиска минимума на ограниченном замкнутом множестве $\tilde{Q} = Q \cap U_{2R}(y_0)$. Значит, с учетом леммы 3, можно записать:

$$\begin{aligned} A_k \phi(\tilde{y}_k) &\leq \Psi_k^* = \min_{y \in \tilde{Q}} \Psi_k = \min_{y \in \tilde{Q}} \left\{ l_k(y) + \frac{L}{2} \|y - y_0\|_2^2 \right\} \leq \\ &\leq \min_{y \in \tilde{Q}} \left\{ l_k(y) + \max_{y \in \tilde{Q}} \frac{L}{2} \|y - y_0\|_2^2 \right\} \leq l_k^* + 2LR^2. \end{aligned}$$

Следовательно, выполняется

$$\phi(\tilde{y}_k) - \frac{l_k^*}{A_k} \leq \frac{2LR^2}{A_k}.$$

Из определения функции ϕ ($\phi = \max_{x \in S_n(1)} \{\langle y, b - Cx \rangle - f(x)\} = \langle y, b - Cx(y) \rangle - f(x(y))$) в силу (6) несложно получить, что для некоторой точки $y' \in Q$ выполняется

$$\phi(y') + \langle \nabla \phi(y'), y - y' \rangle = -f(x(y')) + \langle b - Cx(y'), y \rangle.$$

Поэтому, можно записать

А. В. Чернов

$$\begin{aligned}
 \frac{1}{A_k} l_k^* &= \min_{y \in \tilde{Q}} \sum_{i=0}^k \frac{\alpha_i}{A_k} (\phi(y_i) + \langle \nabla \phi(y_i), y - y_i \rangle) = \\
 &= \min_{y \in \tilde{Q}} \sum_{i=0}^k \frac{\alpha_i}{A_k} (-f(x(y_i)) + \langle b - Cx(y_i), y \rangle) \leq \\
 &\leq -f(x_k) + \min_{y \in \tilde{Q}} \langle b - Cx_k, y \rangle \stackrel{def}{=} \tilde{f}(x_k).
 \end{aligned}$$

Значит, на каждом шаге k выполняется неравенство

$$\phi(\tilde{y}_k) - \tilde{f}(x_k) \leq \phi(\tilde{y}_k) - \frac{1}{A_k} l_k^* \leq \frac{2LR^2}{A_k}.$$

Отметим, что

$$\begin{aligned}
 \max_{y \in \tilde{Q}} \langle Cx_k - b, y \rangle &= \max_{u \in U_{2R}^{m_1}(0) \cap R_+^{m_1}} \langle C_1 x_k - b_1, u \rangle + \\
 + \max_{v \in U_{2R}^{m_2}(0)} \langle C_2 x_k - b_2, v \rangle &= 2R \|C_2 x_k - b_2\| + 2R \|(C_1 x_k - b_1)_+\|.
 \end{aligned}$$

Поэтому

$$\phi(\tilde{y}_k) + f(x_k) + 2R \|C_2 x_k - b_2\| + 2R \|(C_1 x_k - b_1)_+\| \leq \frac{2LR^2}{A_k}.$$

В силу $\langle b - Cx^*, y^* \rangle = 0$ и теоремы двойственности ($-f(x^*) = \tilde{f}(x^*) \leq \phi(y)$) находим, что

$$f(x_k) - f(x^*) \leq f(x_k) + \phi(y^*) \leq f(x_k) + \phi(\tilde{y}_k) \leq \frac{2LR^2}{A_k}.$$

С другой стороны

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

$$\begin{aligned}
 -f(x^*) &= \langle y^*, Cx^* - b \rangle - f(x^*) = \phi(y^*) \geq \\
 &\geq \langle y^*, Cx_k - b \rangle - f(x_k) \Rightarrow f(x^*) - f(x_k) \leq -\langle y^*, Cx_k - b \rangle \leq \\
 &\leq R\|C_2x_k - b_2\| + R\|(C_1x_k - b_1)_+\| \Rightarrow -R\|C_2x_k - b_2\| - \\
 &-R\|(C_1x_k - b_1)_+\| \leq f(x_k) - f(x^*) \leq f(x_k) + \phi(y_k) \Rightarrow \\
 &\Rightarrow 0 \leq f(x_k) + \phi(y_k) + R\|C_2x_k - b_2\| + R\|(C_1x_k - b_1)_+\|.
 \end{aligned}$$

Таким образом, получаем, что на k -м шаге будут выполняться условия

$$\begin{aligned}
 \|C_2x_k - b_2\| + \|(C_1x_k - b_1)_+\| &\leq \frac{2LR}{A_k}; \quad |f(x^*) - f(x_k)| \leq \frac{2LR^2}{A_k}; \\
 |f(x_k) + \phi(y_k)| &\leq \frac{2LR^2}{A_k}.
 \end{aligned}$$

□

Следствие 1. Пусть $\alpha_i = (i + 1)/2$ (для такой последовательности выполняются соотношения (7)) и $y_0 = 0$, тогда количество итераций N , которое достаточно выполнить для достижения заданной точности $(\varepsilon_f, \varepsilon_g)$, будет определяться

$$N = \max \left\{ \sqrt{\frac{8LR}{\varepsilon_g}}, \sqrt{\frac{8LR^2}{\varepsilon_f}} \right\}. \quad (12)$$

Доказательство. Т.к. $y_0 = 0$, то в силу леммы 4 следует, что $\|y_k - y^*\| = O(\|y^*\|)$. Тогда все точки y_k , найденные предлагаемым методом, будут лежать в шаре радиуса R ($R = O(\|y^*\|)$). Из теоремы 1 следует, что для того чтобы найденная точка (x_N, y_N) была $(\varepsilon_f, \varepsilon_g)$ -решением, достаточно выполнения в этой точке условия

$$\frac{2LR}{A_N} \leq \varepsilon_g \text{ и } \frac{2LR^2}{A_N} \leq \varepsilon_f.$$

С другой стороны, т.к. $\alpha_i = (i + 1)/2$, то

$$A_N = \sum_{i=1}^N \alpha_i = N/2 + N(N + 1)/4 \geq N^2/4.$$

Поэтому

$$\frac{2LR}{A_N} \leq 8LR/N^2.$$

Из найденных неравенств тривиально выводится утверждение следствия. \square

Несложно видеть, что алгоритм 1 требует для вычисления k -ого шага наличия информации обо всех предыдущих шагах и, исходя из этих данных, вычисления соответствующих точек в прямом и двойственном пространствах. Однако, данное требование устраняется, если в алгоритм 1 ввести дополнительную переменную, которая будет хранить сумму предыдущих градиентов с соответствующими коэффициентами, а также выразить значение переменной x исключительно через её значение на предыдущем шаге и значение переменной y на текущем шаге. Тогда шаги алгоритма 1 могут быть записаны в следующем виде:

$$\begin{aligned} \tilde{y}_k^j &= \begin{cases} (y_k - \frac{1}{L}\nabla\phi(y_k))_+^j & \text{при } j \leq m_1; \\ (y_k - \frac{1}{L}\nabla\phi(y_k))^j & \text{при } j > m_1; \end{cases} \\ \check{y}_k^j &= \begin{cases} (y_0 - \frac{1}{L}u_k)_+^j & \text{при } j \leq m_1; \\ (y_0 - \frac{1}{L}u_k)^j & \text{при } j > m_1; \end{cases} \\ y_{k+1} &= \tau_k \tilde{y}_k + (1 - \tau_k) \check{y}_k; \quad u_{k+1} = u_k + \alpha_k \nabla\phi(y_{k+1}); \\ x_{k+1} &= (x_k \cdot A_k + \alpha_{k+1} \cdot x(y_{k+1}))/A_k. \end{aligned}$$

Причем здесь $u_0 = \nabla\phi(y_0)$, $x_0 = x(y_0)$.

Численные эксперименты

Заметим, что следствие 1, являющееся тривиальным следствием теоремы 1, дает возможность выполнить N итераций,

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

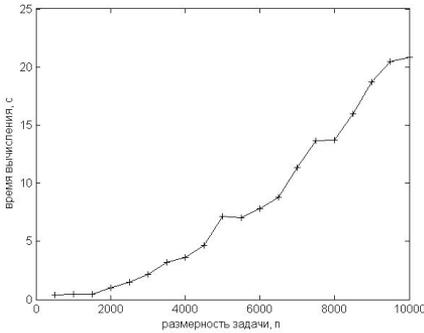
определяемых формулой (12), без проверки условия завершения работы алгоритма и только после этого единожды проверить выполнено ли требуемое условие на окончание алгоритма (зазор двойственности и невязка). Однако, для этого требуется начальное знание об области (радиус шара в двойственном пространстве), в котором находится решение задачи, что не всегда известно на практике. Предлагаемый в статье алгоритм позволяет осуществить постепенный поиск такого шара, иными словами алгоритм состоит из следующих шагов:

- 1) Выбираем исходный размер шара R , в котором должно находиться решение.
- 2) Вычисляем согласно (12) предположительное число итераций N , необходимое для достижения заданной точности при текущей оценке шара.
- 3) Выполняется предписанное число итераций N и далее шаг 4.
- 4) Проверяем условие выхода из цикла (зазор двойственности и ошибка ограничений меньше чем $(\varepsilon_f, \varepsilon_g)$). Если заданная точность достигнута, то возвращаем результат. В противном случае увеличиваем размер шара для поиска решения и переходим к шагу 2.

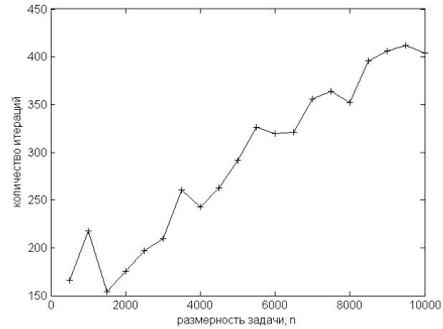
Такой алгоритм был реализован на ЭВМ ASUS N55S с процессором Intel Core i5, оперативной памятью 2Гб под управлением операционной системы MS Windows 7. Полученные результаты вычислений позволяют утверждать о довольно неплохой скорости сходимости метода, а также подтверждают его работоспособность как для ограничений типа равенства так и для ограничений-неравенств. В частности, на рисунке 1 изображены графики времени и количества итераций, которое требуется для достижения заданной точности $(\varepsilon_f, \varepsilon_g)$.

Для указанных вычислений исходные параметры выбирались следующим образом:

- 1) требуемая относительная ошибка значения функции $\varepsilon_f^{rel} = 0.01$; требуемая относительная невязка решения $\varepsilon_g^{rel} = 0.01$;



(а) Требуемое время решения



(б) Затраченное количество итераций

Рис. 1. Продолжительность работы БГМ в зависимости от размерности задачи

- 2) исходная точка в двойственном пространстве $y_0 = 0$;
- 3) исходная оценка радиуса шара $R = 0.5$;
- 4) последовательность чисел $\alpha_i = (i + 1)/2$.

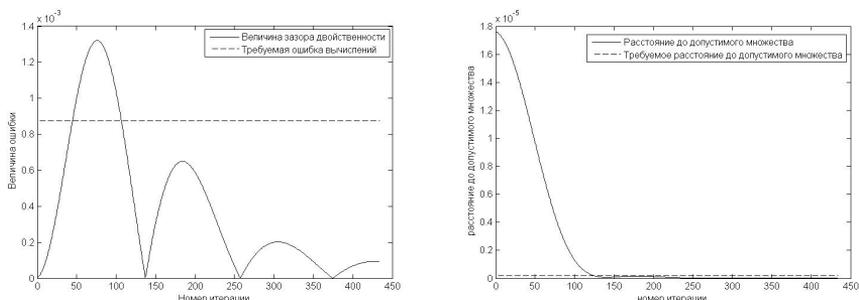
При этом требуемые абсолютная ошибка значения функции и допустимое значение невязки решения вычисляются по формулам

$$\begin{aligned} \varepsilon_f &= \varepsilon_f^{abs} = f(x(y_0)) \cdot \varepsilon_f^{rel}; & \varepsilon_g &= \varepsilon_g^{abs} = \\ &= (\|C_1 x(y_0) - b_1\|_2 + \|C_2 x(y_0) - b_2\|_2) \cdot \varepsilon_g^{rel}. \end{aligned}$$

На рисунке 2 проиллюстрирована сходимость метода к решению задачи, т.е. изменение зазора двойственности от итерации и невязки текущей точки. Размерность исследуемой задачи ЭЛП в данном случае была 10000, исходная точка в двойственном пространстве осталась той же ($y = 0$), относительная ошибка вычисления функции была $\varepsilon_f^{rel} = 0.0001$, а допустимая невязка решения $\varepsilon_g^{rel} = 0.01$.

На графике (б) наблюдается монотонное приближение точки к допустимому множеству (причем достаточно быстрое на первых шагах метода), а на графике (а) отражается периодическое

Прямо-двойственный метод решения задачи энтропийно-линейного программирования



(a) Изменение зазора двойственности

(b) Изменение невязки решения

Рис. 2. Сходимость метода

изменение зазора двойственности, причем амплитуда каждого такого цикла также монотонно убывает, при этом структуру цикла сохраняется: первоначальный резкий рост от минимального значения и последующий более пологий спуск.

Отметим, что требуемая точность задачи достигается существенно раньше, чем расчетное значение, что, прежде всего, связано с отсутствием точной оценки радиуса шара с решением задачи при оценке требуемого количества шагов.

На рисунке 3 показана зависимость количества итераций, необходимых для достижения требуемой точности (здесь исходная точка та же, размерность равна 10000, $\varepsilon_f^{rel} = \varepsilon_g^{rel}$ и меняются от 0.0005 до 0.01 с шагом 0.0005).

На рисунке 4 отражено изменение значений прямой и двойственной функции по итерациям. Здесь размерность задачи равна 10000, исходная точка в двойственном пространстве ($y = 0$), относительная ошибка вычисления функции была $\varepsilon_f^{rel} = 0.01$, а допустимая невязка решения $\varepsilon_g^{rel} = 0.01$. Отметим монотонность изменения двойственной функций и "затухающие колебания" значений прямой функции (что также можно увидеть и на рисунке 2). Немонотность $f(x_k)$ объясняется тем, что $x_k \notin G$ в процессе решения.

В статье [7] был предложен метод решения задачи ЭЛП с аффинными ограничениями $Ax = b, b \in R^m$ с помощью регу-

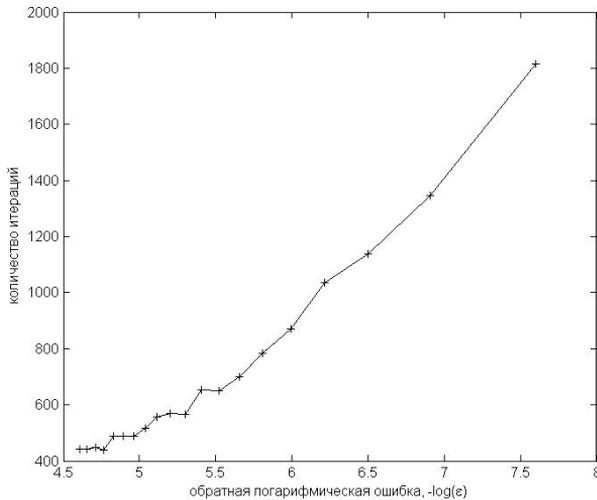


Рис. 3. Зависимость количества итераций для решения задачи от требуемой точности

ляризации по Тихонову двойственной функции. Численные эксперименты показали, что предложенный в данной статье метод работает в среднем в 100 раз быстрее. В частности, на типовых примерах при одних и тех же условиях и ограничениях текущий метод работал 2 секунды, а метод с регуляризацией 3-4 минуты.

Сравнению предлагаемого метода с другими планируется посвятить отдельную публикацию, однако уже сейчас можно отметить, что метод одинаково хорошо работает, как для задач в которых присутствуют только ограничения-равенства (т.е. двойственная задача является задачей безусловной оптимизации) так и с ограничениями-неравенствами, что подтверждается результатами численных экспериментов, в отличие от многих других методов, которые для задачи условной минимизации требуют специальной модификации.

Автор выражает благодарность Бирюкову А.Г., Гасникову А.В. и Двуреченскому П.Е. за ряд ценных замечаний.

Прямо-двойственный метод решения задачи энтропийно-линейного программирования

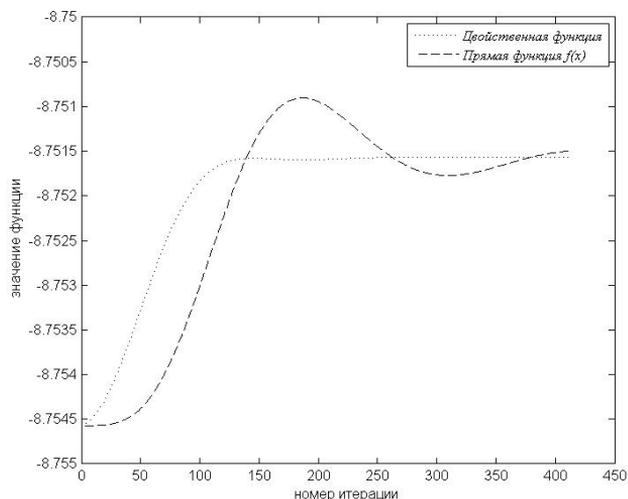


Рис. 4. График изменения значений прямой и двойственной функции

Исследование выполнено при поддержке гранта РФФИ 14-01-00722-а

Список литературы

- [1] Нестеров Ю.Е. Введение в выпуклую оптимизацию. // М.: МЦНМО, 2010.
- [2] Кац М. Вероятность и смежные вопросы в физике. // М.: Мир, 1965.
- [3] Гасников А.В., Гасникова Е.В., М.А. Мендель М.А., Чепурченко К.В. Эволюционные выводы энтропийной модели расчета матрицы корреспонденций. // Математическое моделирование. 2016. Т. 28. arXiv:1508.01077
- [4] Fang S.-C., Rajasekera J.R., Tsao H.-S.J. Entropy optimization and mathematical programming. // Kluwer's International Series, 1997.

- [5] Попков Ю.С. Теория макросистем: Равновесные модели // М.: УРСС, 2013.
- [6] Гасников А.В., Гасникова Е.В., Двуреченский П.Е., Ершов Е.И., Лагуновская А.А. Поиск стохастических равновесий в транспортных моделях равновесного распределения потоков // Труды МФТИ. 2015. Т. 7. № 4. С. 114–128. arXiv:1505.07492.
- [7] Гасников А.В., Гасникова Е.В., Нестеров Е.С., Чернов А.В. Об эффективных численных методах решения задач энтропийно-линейного программирования // ЖВМ и МФ. 2016. Т. 56. № 4. С. 17–28. arXiv:1410.7719.
- [8] Nesterov Yu.E. Smooth minimization of non-smooth function. // Math. Program Ser. A. 2005. V. 103. No , P. 127-152.
- [9] Аникин А.С., Гасников А.В., Тюрин А.И., Чернов А.В. Двойственные подходы к задачам минимизации сильно выпуклых функционалов простой структуры при аффинных ограничениях. // ЖВМ и МФ. 2016. Т. 56. arXiv:1602.01686.