

УДК 7.977

МЕТОД НЬЮТОНОВСКОГО ТИПА ДЛЯ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ И НЕРАВЕНСТВ¹⁾

© 2019 г. А. И. Голиков^{1,2,*}, Ю. Г. Евтушенко^{1,2,**}, И. Е. Капорин^{1,2,***}

¹⁾ 119333 Москва, ул. Вавилова, 40, ВЦ ФИЦ ИУ РАН, Россия;

²⁾ 141700 Долгопрудный, М.о., Институтский пер., 9, НИУ МФТИ, Россия)

*e-mail: gol-a@yandex.ru

**e-mail: evt@ccas.ru

***e-mail: igorkaporin@mail.ru

Поступила в редакцию 27.06.2019 г.
Переработанный вариант 27.06.2019 г.
Принята к публикации 05.08.2019 г.

Предложен метод ньютоновского типа для численной минимизации выпуклых кусочно-квадратичных функций и анализируется его сходимость. Ранее аналогичный метод успешно применялся для задач оптимизации, возникающих в генерации численных сеток. Показана применимость метода к вычислению проекции заданной точки на множество неотрицательных решений системы линейных уравнений, а также к нахождению расстояния между двумя выпуклыми многогранниками. Производительность метода протестирована на наборе задач из библиотеки NETLIB. Библ. 24. Табл. 4.

Ключевые слова: системы линейных уравнений и неравенств, регуляризация, метод штрафа, двойственность, проекция точки, кусочно-квадратичная функция, метод Ньютона, предобусловленный метод сопряженных градиентов.

DOI: 10.1134/S0044466919120093

ВВЕДЕНИЕ

Работа посвящена теоретическому и экспериментальному изучению метода ньютоновского типа, использующего приближенные направления, получаемые с применением предобусловленного метода сопряженных градиентов, и является обобщением и уточнением результатов [1]. Ранее подобный метод успешно применялся в задачах оптимизации, возникающих при генерации численных сеток [2]–[4], см. также [5]. Здесь рассматривается применение данного метода для численного решения задач безусловной минимизации выпуклой кусочно-квадратичной функции. К таким задачам относятся нахождение нормального решения задачи линейного программирования (ЛП) [6], [7] и нахождение проекции заданной точки на множество решений задачи ЛП [8], [9]. С задачей ЛП тесно связано нахождение проекции точки на множество неотрицательных решений недоопределенной системы линейных уравнений [10], [11] и вычисление расстояния между двумя выпуклыми многогранниками [12].

Впервые О. Мангасариан показал эффективность применения обобщенного метода Ньютона для минимизации выпуклой кусочно-квадратичной функции. В своих работах [7], [13] он исходил из возможности простого вычисления обобщенной матрицы Гессе, предложенного в [11], так как кусочно-квадратичная функция непрерывно дифференцируема только один раз.

В разд. 1 задача нахождения неотрицательного решения недоопределенной системы линейных равенств рассматривается с точки зрения линейного программирования, что приводит к рассмотрению взаимно двойственных задач, одна из которых (прямая задача) является оштрафованной, а вторая (двойственная задача) – регуляризованной. При этом через решение задачи безусловной оптимизации кусочно-квадратичной функции легко вычисляется нормальное решение исходной линейной системы (проекция точки). В разд. 2 приведены некоторые вспомогательные технические результаты, связанные с кусочно-квадратичной функцией. В разд. 3 описывается метод ньютоновского типа для безусловной оптимизации кусочно-квадратичной

¹⁾ Работа выполнена при частичной финансовой поддержке РФФИ (код проекта 17-07-00510).

функции. Раздел 4 содержит анализ сходимости метода ньютоновского типа с учетом специального правила остановки внутренних итераций предобусловленного метода сопряженных градиентов. В разд. 5 представлены численные результаты для различных модельных задач.

1. ЗАДАЧИ ОПТИМИЗАЦИИ И СИСТЕМА ЛИНЕЙНЫХ УРАВНЕНИЙ И НЕРАВЕНСТВ

Применение теории оптимизации к задачам нахождения решения систем линейных уравнений и неравенств может быть весьма полезным, так как дает возможность упростить исходную задачу. Так, используя теорию двойственности, можно перейти к решению задачи с меньшим числом переменных, т.е. к задаче, более удобной для применения эффективного метода оптимизации. Метод штрафных функций позволяет снять ограничения и использовать методы безусловной оптимизации; метод регуляризации позволяет выделить единственное решение из множества решений исходной задачи (например, найти нормальное решение или проекцию заданной точки). Теория двойственности наглядно демонстрирует связь между методом штрафа и методом регуляризации.

Рассмотрим задачу нахождения неотрицательного решения системы линейных уравнений

$$Ax = b, \quad x \geq 0_n. \tag{1}$$

Здесь матрица $A \in R^{m \times n}$, вектор $b \in R^m$ заданы и $m \leq n$. Через 0_i обозначен i -мерный нулевой вектор. Будем предполагать, что исходная система (1) разрешима. Оказывается, эту систему полезно рассмотреть как задачу линейного программирования (ЛП)

$$P : \min_{x \in X} 0_n^T x, \quad X = \{x \in R^n : Ax = b, x \geq 0_n\}, \tag{2}$$

в которой вектор целевой функции $c = 0_n$. Двойственная к (2) имеет вид

$$D : \max_{u \in U} b^T u, \quad U = \{u \in R^m : A^T u \leq 0_n\}.$$

Прямую задачу линейного программирования P преобразуем следующими двумя способами: 1) с помощью квадратичного штрафа (преобразование pen), 2) используя квадратичную регуляризацию (преобразование reg). Можно продолжить преобразование полученных задач, применяя, соответственно, квадратичную регуляризацию и квадратичный штраф. Последовательность преобразований задачи P можно представить схематически в виде $P_{pr} \xleftarrow{reg} P_p \xleftarrow{pen} P \xrightarrow{reg} P_r \xrightarrow{pen} P_{pr}$. Если применить квадратичный штраф к задаче P , то приходим к минимизации квадратичной функции на неотрицательном ортанте:

$$P_p : \min_{x \in R^n_+} \left(0_n^T x + \frac{1}{2} \|b - Ax\|^2 \right). \tag{3}$$

Вообще говоря, (3) есть задача наименьших квадратов для решения системы (1). В силу постоянства целевой функции в задаче P , которая всегда равна нулю, в задаче P_p не требуется вводить коэффициент штрафа. Отметим, что квадратичная задача (3) всегда разрешима независимо от того, имеет ли решение исходная система (1).

Если к задаче P применить квадратичную регуляризацию (регуляризация по Тихонову [15]), то получим задачу строго выпуклого квадратичного программирования при ограничениях

$$P_r : \min_{x \in X} \left(0_n^T x + \frac{1}{2} \|x\|^2 \right), \quad X = \{x \in R^n : Ax = b, x \geq 0_n\}. \tag{4}$$

В этой задаче также не требуется вводить коэффициент регуляризации.

Задачу минимизации на неотрицательном ортанте P_p можно регуляризовать, и тогда приходим к следующей строго выпуклой квадратичной задаче:

$$P_{pr} : \min_{x \in R^n_+} \left(0_n^T x + \frac{1}{2} \|b - Ax\|^2 + \frac{\epsilon}{2} \|x\|^2 \right), \tag{5}$$

где $\epsilon > 0$ представляет собой коэффициент регуляризации. Решение $x(\epsilon)$ задачи (5) единственно при любом $\epsilon > 0$. При $\epsilon \rightarrow 0$ решение $x(\epsilon)$ сходится к нормальному решению системы (1) ([15]), т.е. к решению задачи (4).

Если в задаче P_r снять ограничения с помощью квадратичного штрафа, то получаем следующую задачу:

$$P_{rp} : \min_{x \in R^n} \left(0_n^T x + \frac{1}{2} \|x\|^2 + \frac{1}{2\varepsilon} \|b - Ax\|^2 \right).$$

Очевидно, что задачи P_{rp} и P_{pr} эквивалентны в смысле совпадения их решений при любом положительном ε .

Выпишем теперь двойственные задачи к приведенным выше оптимизационным задачам. Начнем с задачи P_p , двойственной для которой является задача строго вогнутого квадратичного программирования:

$$D_r : \max_{u \in U} \left(b^T u - \frac{1}{2} \|u\|^2 \right), \quad U = \{u \in R^m : A^T u \leq 0_n\}.$$

Хорошо известно (см., например, [16]), что задачу безусловной оптимизации можно легко свести к задаче с ограничениями с помощью дополнительных переменных и естественно возникающих ограничений. Для такой задачи уже можно ввести функцию Лагранжа и стандартным образом получить двойственную задачу. Задачи P_p и D_r являются взаимно двойственными. По решению x_* задачи P_p легко вычисляется решение задачи D_r по формуле $u_* = b - Ax_*$. Следует отметить, что квадратичные задачи P_p и D_r всегда разрешимы вне зависимости от разрешимости задач линейного программирования P и D . Задача D_r получается применением квадратичной регуляризации к задаче D .

Двойственной к P_r является следующая оштрафованная задача D

$$D_p : \max_{u \in R^m} \left(b^T u - \frac{1}{2} \|(A^T u)_+\|^2 \right).$$

Здесь и ниже a_+ обозначает вектор a , у которого все отрицательные компоненты заменены на нули. По решению u_* задачи D_p находится решение $x_* = (A^T u_*)_+$ задачи P_r , т.е. нормальное решение исходной линейной системы (1).

Приведем двойственную задачу для P_{rp}

$$D_{pr} : \max_{u \in R^m} \left(b^T u - \frac{1}{2} \|(A^T u)_+\|^2 - \frac{\varepsilon}{2} \|u\|^2 \right),$$

которая является задачей безусловной максимизации строго вогнутой кусочно-квадратичной функции. При любом $\varepsilon > 0$ по решению $u(\varepsilon)$ задачи D_{pr} вычисляется решение $x(\varepsilon) = (A^T u(\varepsilon))_+$ задачи P_{rp} . При $\varepsilon \rightarrow 0$ решение $x(\varepsilon)$ сходится к решению задачи P_r , т.е. к нормальному решению исходной системы (1). Итак, оштрафованную задачу P_r , т.е. задачу минимизации P_{rp} на положительном ортанте с n переменными, заменили задачей безусловной максимизации D_{pr} другой размерности с m переменными.

В свою очередь решение $u(\varepsilon)$ задачи D_{pr} находится через решение $x(\varepsilon)$ задачи P_{rp} минимизации строго выпуклой кусочно-квадратичной функции на неотрицательном ортанте по простой формуле $u(\varepsilon) = \frac{1}{\varepsilon} (b - Ax(\varepsilon))$.

Аналогично получается двойственная задача для P_{pr} :

$$D_{rp} : \max_{u \in R^m} \left(b^T u - \frac{1}{2} \|u\|^2 - \frac{1}{2\varepsilon} \|(A^T u)_+\|^2 \right).$$

При любом $\varepsilon > 0$ решения взаимно двойственных задач P_{pr} и D_{rp} выражаются друг через друга. По решению $u(\varepsilon)$ задачи D_{rp} легко находится решение $x(\varepsilon) = \frac{1}{\varepsilon} (A^T u(\varepsilon))_+$ задачи P_{pr} и из решения $x(\varepsilon)$ задачи P_{pr} имеем решение $u(\varepsilon) = b - Ax(\varepsilon)$ задачи D_{rp} .

Приведенную связь оштрафованных и регуляризованных задач с помощью двойственности можно представить в виде следующей схемы:

$$\begin{array}{ccccccccc}
 P_{pr} & \xleftarrow{reg} & P_p & \xleftarrow{pen} & P & \xrightarrow{reg} & P_r & \xrightarrow{pen} & P_{rp} \\
 \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\
 D_{rp} & \xleftarrow{pen} & D_r & \xleftarrow{reg} & D & \xrightarrow{pen} & D_p & \xrightarrow{reg} & D_{pr}
 \end{array}$$

В эту же схему укладываются и задачи нахождения проекции заданной точки \hat{x} на множество решений системы (1). Регуляризованная задача P будет иметь вид

$$P_r : \min_{x \in X} \left(0_n^T x + \frac{1}{2} \|x - \hat{x}\|^2 \right), \quad X = \{x \in R^n : Ax = b, x \geq 0_n\}, \tag{6}$$

взаимно двойственной к ней будет задача

$$D_p : \max_{u \in R^m} \left(b^T u - \frac{1}{2} \|(\hat{x} + A^T u)_+\|^2 \right). \tag{7}$$

При этом через ее решение u_* находится решение $\hat{x}_* = (\hat{x} + A^T u_*)_+$ задачи P_r (6), т.е проекция точки \hat{x} на множество решений системы (1). Задача D_p (7) есть оштрафованная задача D , в которой ограничения сдвинуты на вектор \hat{x} . Отметим, что задача D_p (7) безусловной максимизации от m переменных имеет в общем случае неединственное решение.

Если оштрафовать задачу P_r (6), то приходим к задаче минимизации строго выпуклой квадратичной функции от n переменных на неотрицательном ортанте:

$$P_{rp} : \min_{x \in R_+^n} \left(0_n^T x + \frac{1}{2} \|x - \hat{x}\|^2 \right) + \frac{1}{2\varepsilon} \|b - Ax\|^2, \tag{8}$$

взаимно двойственная к которой имеет вид

$$D_{pr} : \max_{u \in R^m} \left(b^T u - \frac{1}{2} \|(\hat{x} + A^T u)_+\|^2 - \frac{\varepsilon}{2} \|u\|^2 \right). \tag{9}$$

Задача (9) является задачей безусловной максимизации строго вогнутой кусочно-квадратичной функцией от m переменных. Эта задача получается регуляризацией задачи D_p (7). При любом $\varepsilon > 0$ решения $x(\varepsilon)$ и $u(\varepsilon)$ соответственно задач (8) и (9) выражаются по простым формулам $x(\varepsilon) = (\hat{x} + A^T u(\varepsilon))_+$, $u(\varepsilon) = \frac{1}{\varepsilon} (b - Ax(\varepsilon))$.

Далее остановимся на рассмотрении задачи (7), которую перепишем в следующем виде безусловной минимизации кусочно-квадратичной функции

$$u_* = \arg \min_{u \in R^m} \left(\frac{1}{2} \|(\hat{x} + A^T u)_+\|^2 - b^T u \right). \tag{10}$$

Задача (10) является взаимно двойственной к задаче нахождения проекции заданного вектора \hat{x} на множество неотрицательных решений недоопределенной системы линейных уравнений:

$$x_* = \arg \min_{\substack{Ax=b \\ x \geq 0}} \frac{1}{2} \|x - \hat{x}\|^2.$$

Решение x_* этой задачи выражается через решение u_* задачи (10) по простой формуле $x_* = (\hat{x} + A^T u_*)_+$.

Итак, введем в рассмотрение выпуклую дифференцируемую кусочно-квадратичную функцию $\varphi(u) : R^m \rightarrow R^1$:

$$\varphi(u) = \frac{1}{2} \|(\hat{x} + A^T u)_+\|^2 - b^T u. \tag{11}$$

Ее градиент $g(u) = \text{grad}\varphi(u)$ имеет вид

$$g(u) = A(\hat{x} + A^T u)_+ - b, \tag{12}$$

а обобщенную матрицу Гессе [14] определим как

$$H(u) = A \text{Diag}(\text{sign}(\hat{x} + A^T u)_+) A^T. \tag{13}$$

Связь между $H(u)$, $\varphi(u)$ и $g(u)$ будет установлена ниже в Замечании 1.

2. ТЕЙЛОРОВСКОЕ РАЗЛОЖЕНИЕ ФУНКЦИИ $(\cdot)_+^2$

Следующий результат представляет собой специальный случай разложения Тейлора с остаточным членом в интегральной форме.

Лемма 1. Для любых действительных скалярных величин η и ζ справедливо

$$\frac{1}{2}((\eta + \zeta)_+)^2 - \frac{1}{2}(\eta_+)^2 - \zeta\eta_+ = \zeta^2 \int_0^1 \left(\int_0^1 \text{sign}(\eta + st\zeta)_+ ds \right) t dt. \tag{14}$$

Доказательство. Пусть $f(\xi) = \frac{1}{2}(\xi_+)^2$, тогда $f'(\xi) = \xi_+$ и $f''(\xi) = \text{sign}(\xi_+)$ (без потери общности будем полагать $f''(0) = 0$) [14]. Подставляя эти равенства в тейлоровское разложение

$$f(\eta + \zeta) = f(\eta) + \zeta f'(\eta) + \zeta^2 \int_0^1 \left(\int_0^1 f''(\eta + st\zeta) ds \right) t dt,$$

приходим к нужному результату.

Лемма 2. Для любых n -мерных векторов y и z справедливо

$$\frac{1}{2}\|(y + z)_+\|^2 - \frac{1}{2}\|y_+\|^2 - z^T y_+ = \frac{1}{2} z^T \text{Diag}(d) z, \tag{15}$$

где

$$d = \int_0^1 \left(\int_0^1 \text{sign}(y + stz)_+ ds \right) 2t dt. \tag{16}$$

Доказательство. Полагая $\eta = y_j$, $\zeta = z_j$ в (14) и суммируя по всем $j = 1, \dots, n$, приходим к нужной формуле. Заметим, что использование множителя 2 в подынтегральном выражении приводит к оценке $\|\text{Diag}(d)\| \leq 1$.

Лемма 3. Функция (11) и ее градиент (12) удовлетворяют равенству

$$\varphi(u + q) - \varphi(u) - q^T g(u) = \frac{1}{2} q^T A \text{Diag}(d) A^T q, \tag{17}$$

где

$$d = \int_0^1 \left(\int_0^1 \text{sign}(\hat{x} + A^T u + stA^T q)_+ ds \right) 2t dt. \tag{18}$$

Доказательство. Полагая $y = \hat{x} + A^T u$ и $z = A^T q$ в (15) и (16), легко получаем требуемый результат (с учетом приведения линейных членов, включающих b в левой части (17)).

Замечание 1. Как видно из (18), если условие

$$\text{sign}(\hat{x} + A^T u + \vartheta A^T q)_+ = \text{sign}(\hat{x} + A^T u)_+, \tag{19}$$

справедливо для любого $0 \leq \vartheta \leq 1$, то (17) приводится к простейшему виду:

$$\varphi(u + q) - \varphi(u) - q^T g(u) = \frac{1}{2} q^T H(u) q, \tag{20}$$

где обобщенная матрица Гессе $H(u)$ задана формулой (13). Это объясняет ключевую роль $H(u)$ в рассмотренном ниже методе ньютоновского типа. Заметим, что достаточным условием для выполнения (19) является

$$\left| (A^T q)_j \right| \leq \left| (\hat{x} + A^T u)_j \right| \quad \text{всегда, когда} \quad (A^T q)_j (\hat{x} + A^T u)_j < 0; \tag{21}$$

т.е. если некоторые компоненты приращения q относительно малы, то φ в точности квадратична (20) в соответствующей окрестности u .

3. МЕТОД НЬЮТОНОВСКОГО ТИПА ДЛЯ ДВОЙСТВЕННОЙ ЗАДАЧИ (10)

Из условия (19) и его следствия (20) вытекает, что можно численно минимизировать ϕ методом ньютоновского типа $u_{k+1} = u_k - d_k$, где $d_k = H(u_k)^{-1}g(u_k)$. Заметим, что согласно (20) точка минимума $u_* = u_{k+1}$ получается за один шаг, если компоненты d_k достаточно малы и удовлетворяют (21), где $u = u_k$ and $q = -d_k$. Однако начальные u_k могут быть еще далеки от решения и поэтому возможны только постепенные улучшения точности приближений. Во-первых, должен использоваться коэффициент α_k для гарантии монотонной сходимости (в смысле уменьшения $\phi(u_k)$ при увеличении k). Во-вторых, вместо $H(u_k)$ должна использоваться некоторая ее аппроксимация M_k , являющаяся обратимой матрицей с не слишком большой нормой обратной $\|M_k^{-1}\|$. Поэтому предлагается следующий прототип метода:

$$u_{k+1} = u_k - \alpha_k M_k^{-1} g(u_k),$$

где

$$M_k = H(u_k) + \delta \text{Diag}(AA^T). \tag{22}$$

Выбор параметров $0 < \alpha_k \leq 1$ и $0 \leq \delta \ll 1$ существенен для лучшей сходимости. Кроме того, на начальных этапах итераций наиболее эффективной стратегией является использование приближенных ньютоновских направлений $d_k \approx M_k^{-1}g(u_k)$, которые могут быть получены предобусловленным методом сопряженных градиентов, примененным к решению уравнения $M_k d_k = g(u_k)$. Как будет показано ниже, достаточно использовать любой вектор d_k , который удовлетворяет условиям

$$d_k^T g_k = d_k^T M_k d_k = \vartheta_k^2 g_k^T M_k^{-1} g_k \tag{23}$$

с $0 < \vartheta_k < 1$, достаточно хорошо отделенным от нуля. При любом предобусловливании метода сопряженных градиентов такие аппроксимации удовлетворяют (23) (см. п. 4.3 ниже). Для выбора параметра $0 < \alpha \leq 1$ используется правило Армийо

$$\phi(u_k - \alpha d_k) \leq \phi(u_k) - \frac{\alpha}{2} d_k^T g(u_k), \quad \alpha \in \{1, 1/2, 1/4, \dots\}, \tag{24}$$

где выбирается максимально возможное значение $0 < \alpha \leq 1$, при котором выполняется условие (24). Соответствующий алгоритм ньютоновского типа можно представить в виде:

АЛГОРИТМ 1

Input:

$$A \in R^{m \times n}, \quad b \in R^m, \quad \hat{x} \in R^m;$$

Initialization:

$$\delta = 10^{-6}, \quad \varepsilon = 10^{-12}, \quad \tau = 10^{-15};$$

$$k_{\max} = 2000, \quad l_{\max} = 10, \quad u_0 = 0_m;$$

Iterations:

for $k = 0, 1, \dots, k_{\max} - 1$:

$$x_k = (\hat{x} + A^T u_k)_+$$

$$\varphi_k = \frac{1}{2} \|x_k\|^2 - b^T u_k$$

$$g_k = Ax_k - b$$

if $(\|g_k\| \leq \varepsilon \|b\|)$ **return** $\{x_k, u_k, g_k\}$

find $d_k \in R^m$ **such that**

$$d_k^T g_k = d_k^T M_k d_k = \vartheta_k^2 g_k^T M_k^{-1} g_k,$$

where $M_k = A \text{Diag}(\text{sign}(x_k)) A^T + \delta \text{Diag}(AA^T)$

```

 $\alpha^{(0)} = 1$ 
 $u_k^{(0)} = u_k - d_k$ 
for  $l = 0, 1, \dots, l_{\max} - 1$ :
   $x_k^{(l)} = (\hat{x} + A^T u_k^{(l)})_+$ 
   $\varphi_k^{(l)} = \frac{1}{2} \|x_k^{(l)}\|^2 - b^T u_k^{(l)}$ 
   $\zeta_k^{(l)} = \left( \frac{1}{2} \alpha^{(l)} d_k^T g_k + \varphi_k^{(l)} \right) - \varphi_k$ 
  if  $(\zeta_k^{(l)} > \tau |\varphi_k|)$  then
     $\alpha^{(l+1)} = \alpha^{(l)} / 2$ 
     $u_k^{(l+1)} = u_k - \alpha^{(l+1)} d_k$ 
  else
     $u_k^{(l+1)} = u_k^{(l)}$ 
  go to NEXT
end if
end for
NEXT:  $u_{k+1} = u_k^{(l+1)}$ 
end for

```

Далее исследуем свойства сходимости этого алгоритма.

4. АНАЛИЗ СХОДИМОСТИ МЕТОДА НЬУТОНОВСКОГО ТИПА

Оказывается, что анализ сходимости Алгоритма 1 можно выполнить тем же образом, что и в работах [4] и [17] (где, однако, рассматривается нелинейная задача наименьших квадратов для гладких функций). Для полноты изложения и совместимости обозначений мы воспроизводим здесь основные результаты из [4].

4.1. Анализ сходимости метода ньютоновского типа

Основные необходимые предположения для рассматриваемой функции $\varphi(u)$ заключаются в том, что она ограничена снизу, имеет градиент $g(u) \in R^m$ и удовлетворяет оценке

$$\varphi(u + q) - \varphi(u) - q^T g(u) \leq \frac{\gamma}{2} q^T M q \quad (25)$$

для симметричной положительно определенной $m \times m$ матрицы $M = M(u)$, определенной выше в (22) и некоторой константы $\gamma \geq 1$. Предполагается также, что доступна достаточно экономичная процедура умножения матрицы M на произвольный вектор.

Заметим, что знание точного значения γ не требуется для реальных вычислений. Существование γ следует из (22) и леммы 3. Действительно, обозначая $D = (\text{Diag}(AA^T))^{1/2}$ и $\hat{A} = D^{-1}A$, для правой части (17) имеем с учетом $\|\text{Diag}(d)\| \leq 1$ и $H(u) \geq 0$ имеем,

$$A \text{Diag}(d) A^T \leq AA^T \leq \|\hat{A}\|^2 \text{Diag}(AA^T) \leq \frac{\|\hat{A}\|^2}{\delta} (\delta \text{Diag}(AA^T) + H(u)) = \frac{\|\hat{A}\|^2}{\delta} M.$$

Поэтому (25) выполняется с

$$\gamma = \|\hat{A}\|^2 / \delta. \quad (26)$$

Последняя формула объясняет выбор M , который способен учитывать большие вариации норм строк в A (см. примеры и обсуждение в п. 5.1).

Укажем также, что если вектор направления d получен после $i \geq 1$ итераций предобусловленного метода сопряженных градиентов с нулевым начальным приближением, примененным к решению системы линейных уравнений $Md = g$, то выполнено требуемое условие нормировки (23), см. п. 4.3).

Оценим уменьшение значения φ , которое достигается при спуске по направлению $(-d)$, удовлетворяющему (23). Далее используются упрощенные обозначения $u = u_k, \hat{u} = u_{k+1}$ и т.д. Можно дать следующую оценку уменьшения значения целевой функции на каждой итерации $\hat{u} = u - \alpha d$ с выбором $\alpha = 2^{-l}$, где $l = 0, 1, \dots$, по правилу Армийо (24):

$$\varphi(\hat{u}) \leq \varphi(u) - \frac{\vartheta^2}{4\gamma} g^T M^{-1} g. \tag{27}$$

В частности, если значения ϑ^2 отделены от нуля положительной константой ϑ_{\min}^2 (нижняя оценка для ϑ следует из п. 4.3 и верхней границы для $\kappa = \text{cond}(CM)$), то с учетом оценки $M \leq (1 + \delta) \|A\|^2 I$ (вытекающей из (13), (22)) и ограниченности φ снизу, следует

$$\sum_{j=0}^{k-1} g_j^T g_j \leq \frac{4\gamma(1 + \delta) \|A\|^2 (\varphi(u_0) - \varphi(u_*))}{\vartheta_{\min}^2}.$$

Замечая, что правая сторона последней оценки не зависит от k , окончательно получаем

$$\lim_{k \rightarrow \infty} \|g(u_k)\| = 0,$$

где k – номер внешней итерации (см. алгоритм 1).

Оценка (27) может быть установлена следующим образом (обратим внимание, что аналогичный анализ можно найти в [13]). Используя $q = -\beta d$, где $0 < \beta < 2/\gamma$, получаем из (25) и (23) следующую оценку уменьшения φ вдоль направления $(-d)$:

$$\varphi(u - \beta d) = \varphi(u) - \beta d^T g + (\varphi(u - \beta d) - \varphi(u) - \beta d^T g) \leq \varphi(u) - \left(\beta - \frac{\gamma}{2} \beta^2\right) d^T M d. \tag{28}$$

Возможны следующие два случая.

Случай 1. Если условие (24) выполняется сразу при $\alpha = 1$, то с учетом левого равенства (23) это означает, что

$$\varphi(\hat{u}) \leq \varphi(u) - \frac{1}{2} d^T M d.$$

Случай 2. Иначе, если было выполнено по крайней мере одно деление длины шага пополам (и его значение стало равно α), то, используя (28) с $\beta = 2\alpha$, получаем

$$\varphi(u) - \alpha d^T M d < \varphi(u - 2\alpha d) \leq \varphi(u) - (2\alpha - 2\gamma\alpha^2) d^T M d,$$

откуда имеем $\alpha > 1/(2\gamma)$. С другой стороны, справедливо

$$\varphi(u - \alpha d) \leq \varphi(u) - \frac{\alpha}{2} d^T M d,$$

и поэтому верно следующее:

$$\varphi(u - \alpha d) \leq \varphi(u) - \frac{1}{4\gamma} d^T M d. \tag{29}$$

Объединяя рассмотренные два случая, учитывая условие $\gamma \geq 1$ и используя второе равенство (23), приходим к требуемой оценке (27).

Осталось заметить, что когда с увеличением k нормы векторов g_k станут достаточно малыми, получаемые направления d_k также будут иметь малые нормы. Поэтому реализуется ситуация, рассмотренная в замечании 1, и наступит этап весьма быстрой сходимости ньютоновского метода (тем более быстрой, чем меньшим выбрано значение параметра δ в (22)).

4.2. Критерий завершения итераций метода сопряженных градиентов

Ниже будет рассмотрено соотношение между сходимостью итераций метода сопряженных градиентов при вычислении приближений к ньютоновским направлениям и общей эффективностью метода ньютоновского типа, использующего такие приближенные направления. Аналогичные рассуждения проводились в работах [3], [4], [17], [18].

Приближение $d^{(i)}$ к решению задачи $Md = g$, построенное на i -й итерации метода сопряженных градиентов по рекуррентной формуле $d^{(i+1)} = d^{(i)} + s^{(i)}$ (см. ниже алгоритм 2) может быть записано следующим образом (напомним, что здесь всегда выбирается начальное приближение $d^{(0)} = 0$):

$$d^{(i)} = \sum_{j=0}^{i-1} s^{(j)}, \quad (30)$$

где векторы направлений $s^{(j)}$ образуют M -ортогональную систему: $(s^{(j)})^T Ms^{(l)} = 0$, $j \neq l$. Введем также обозначения для M -норм этих направлений $\eta^{(j)} = (s^{(j)})^T Ms^{(j)}$, $j = 0, 1, \dots, i-1$. Тогда, используя (30), можно определить величины

$$\zeta^{(i)} = (d^{(i)})^T Md^{(i)} = \sum_{j=0}^{i-1} \eta^{(j)},$$

так что оценка (29) принимает вид

$$\varphi_{k+1} \leq \varphi_k - \frac{1}{4\gamma} \sum_{j=0}^{i_k-1} \eta_k^{(j)},$$

где через k обозначается номер ньютоновской итерации. Суммируя полученные неравенства для $0 \leq k \leq m-1$, получаем

$$c_0 \equiv 4\gamma(\varphi_0 - \varphi_*) \geq \sum_{k=0}^{m-1} \sum_{j=0}^{i_k-1} \eta_k^{(j)}. \quad (31)$$

С другой стороны, мера затрат, пропорциональная общему времени, требуемому для выполнения m итераций ньютоновского типа (при i_k итерациях метода сопряженных градиентов на каждой из них), может быть оценена как

$$T_m = \sum_{k=0}^{m-1} (\epsilon_{CG}^{-1} + i_k) \leq c_0 \frac{\sum_{k=0}^{m-1} (\epsilon_{CG}^{-1} + i_k)}{\sum_{k=0}^{m-1} \sum_{j=0}^{i_k-1} \eta_k^{(j)}} \leq c_0 \max_{k < m} \frac{\epsilon_{CG}^{-1} + i_k}{\sum_{j=0}^{i_k-1} \eta_k^{(j)}} = c_0 \max_{k < m} \frac{\epsilon_{CG}^{-1} + i_k}{\zeta^{(i_k)}}.$$

Здесь введен малый параметр $\epsilon_{CG} \in (0, 1)$, предназначенный для учета отношения затрат на одну итерацию метода сопряженных градиентов к дополнительным затратам на одну итерацию метода Ньютона (включающим, в частности, затраты на построение предобусловливания и на несколько вычислений φ в процессе выбора длины шага α), а также возможной потери эффективности из-за преждевременного окончания итераций метода сопряженных градиентов. Таким образом, вводя функцию $\psi(i) = (\epsilon_{CG}^{-1} + i)/\zeta^{(i)}$, можно предложить соответствующий критерий остановки итераций по условию $\psi(i) > \psi(i-1)$, которое можно переписать в виде

$$(\epsilon_{CG}^{-1} + i)\eta^{(i-1)} \leq \zeta^{(i)}. \quad (32)$$

Очевидно, использование меньших значений ϵ_{CG} соответствует повышению получаемой границы числа итераций. Заметим, что при использовании указанного критерия остановки метод сопряженных градиентов всегда делает не меньше двух итераций.

Приведем расчетные формулы алгоритма, реализующего предобусловленный метод сопряженных градиентов с использованием предложенного критерия остановки итераций.

Следуя [9], используем предобусловливание Якоби

$$C = (\text{Diag}(M))^{-1}. \quad (33)$$

Кроме того, используем переформулировку [19] метода сопряженных градиентов [20], [21]. Такой выбор может обеспечить возможность более эффективной параллельной реализации, см., например, [9].

Следуя [19], напомним, что на каждой итерации метода сопряженных градиентов M^{-1} -норма $(i + 1)$ -й невязки $r^{(i+1)} = g - Md^{(i+1)}$ минимальна в соответствующем подпространстве Крылова. Используя стандартные рекуррентные соотношения метода (см. ниже п. 4.3), можно найти

$$d^{(i+1)} = d^{(i)} + Cr^{(i)}\alpha^{(i)} + s^{(i-1)}\alpha^{(i)}\beta^{(i-1)}.$$

Поэтому оптимальное приращение $s^{(i)}$ в рекуррентной формуле $d^{(i+1)} = d^{(i)} + s^{(i)}$, где $s^{(i)} = V^{(i)}h^{(i)}$ и $V^{(i)} = [Cr^{(i)} | s^{(i-1)}]$, может быть определено как решение следующей двумерной линейной задачи наименьших квадратов:

$$\begin{bmatrix} \alpha^{(i)} \\ \beta^{(i)} \end{bmatrix} = h^{(i)} = \arg \min_{h \in \mathbb{R}^2} \|g - Md^{(i+1)}\|_{M^{-1}} = \arg \min_{h \in \mathbb{R}^2} \|r^{(i)} - MV^{(i)}h\|_{M^{-1}}.$$

Переопределяя $r^{(i)} := -r^{(i)}$ и вводя векторы $t^{(i)} = Ms^{(i)}$, получаем требуемую переформулировку алгоритма:

АЛГОРИТМ 2

$r^{(0)} = -g, \quad d^{(0)} = s^{(-1)} = t^{(-1)} = 0, \quad \zeta^{(-1)} = 0;$
 $i = 0, 1, \dots, it_{\max} :$
 $w^{(i)} = Cr^{(i)},$
 $z^{(i)} = Mw^{(i)},$
 $\gamma^{(i)} = (r^{(i)})^T w^{(i)}, \quad \xi^{(i)} = (w^{(i)})^T z^{(i)}, \quad \eta^{(i-1)} = (s^{(i-1)})^T t^{(i-1)},$
 $\zeta^{(i)} = \zeta^{(i-1)} + \eta^{(i-1)},$
if $((\varepsilon_{CG}^{-1} + i)\eta^{(i-1)} \leq \zeta^{(i)})$ **or** $(\gamma^{(i)} \leq \varepsilon_{CG}^2 \gamma^{(0)})$ **return** $\{d^{(i)}\};$
if $(k = 0)$ **then**
 $\alpha^{(i)} = -\gamma^{(i)} / \xi^{(i)}, \quad \beta^{(i)} = 0;$
else
 $\delta^{(i)} = \gamma^{(i)} / (\xi^{(i)} \eta^{(i-1)} - (\gamma^{(i)})^2),$
 $\alpha^{(i)} = -\eta^{(i-1)} \delta^{(i)}, \quad \beta^{(i)} = \gamma^{(i)} \delta^{(i)};$
end if
 $t^{(i)} = z^{(i)} \alpha^{(i)} + t^{(i-1)} \beta^{(i)}, \quad r^{(i+1)} = r^{(i)} + t^{(i)},$
 $s^{(i)} = w^{(i)} \alpha^{(i)} + s^{(i-1)} \beta^{(i)}, \quad d^{(i+1)} = d^{(i)} + s^{(i)}.$

Для большей надежности, наряду с новым критерием остановки итераций (32), здесь также применяется и стандартный критерий; однако, почти всегда в расчетах срабатывал именно новый критерий.

Несмотря на некоторое увеличение затрат рабочей памяти и количества векторных операций алгоритма 2 по сравнению со стандартным (см. следующий раздел), он может допускать более эффективную параллельную реализацию операций скалярного произведения при большом числе процессов MPI и значительной латентности операций обмена данными. Так, на каждой итерации алгоритма 2 достаточно применения одной операции сборки/расылки $\text{MPI_AllReduce}(*, *, 3, \dots)$, в то время как стандартный алгоритм требует двух таких операций $\text{MPI_AllReduce}(*, *, 1, \dots)$. Аналогичные алгебраически эквивалентные переформулировки алгоритма предобусловленного метода сопряженных градиентов можно найти в работе [22] и цитированных в ней источниках.

4.3. Оценки сходимости метода сопряженных градиентов

Напомним некоторые использованные выше основные свойства метода сопряженных градиентов (см., например, [21]). Стандартный алгоритм (алгебраически эквивалентный алгоритму 2) для решения задачи $Md = g$ с нулевого начального приближения d_0 выглядит следующим образом:

$$\begin{aligned}
 & d^{(0)} = 0, \quad r^{(0)} = g, \quad s^{(0)} = Cr^{(0)}; \\
 & \text{for } i = 0, 1, \dots, m-1 : \\
 & \quad \alpha^{(i)} = (r^{(i)})^T Cr^{(i)} / (s^{(i)})^T Ms^{(i)}, \\
 & \quad d^{(i+1)} = d^{(i)} + s^{(i)} \alpha^{(i)}, \\
 & \quad r^{(i+1)} = r^{(i)} - Ms^{(i)} \alpha^{(i)}, \\
 & \quad \text{if } ((r^{(i)})^T Cr^{(i+1)} \leq \epsilon_{CG}^2 (r^{(0)})^T Cr^{(0)}) \text{ return } d^{(i+1)} \\
 & \quad \beta^{(i)} = (r^{(i+1)})^T Cr^{(i+1)} / (r^{(i)})^T Cr^{(i)}, \\
 & \quad s^{(i+1)} = Cr^{(i+1)} + s^{(i)} \beta^{(i)}. \\
 & \text{end for}
 \end{aligned} \tag{34}$$

Справедливость условия нормировки (23) $d^T g = d^T Md$ (здесь и далее опущены индексы при d) можно доказать следующим образом. Пусть вектор $d = d^{(i)}$ получен после выполнения i итераций приведенного выше метода сопряженных градиентов. Тогда имеем

$$d \in K_i = \text{span}\{Cg, CM Cg, \dots, (CM)^{i-1} Cg\},$$

и, в силу свойства оптимальности метода, справедливо соотношение

$$d = \arg \min_{d \in K_i} (g - Md)^T M^{-1} (g - Md).$$

Так как $\beta d \in K_i$ для любого скаляра β , получаем

$$(g - \beta Md)^T M^{-1} (g - \beta Md) \geq (g - Md)^T M^{-1} (g - Md).$$

Полагая здесь $\beta = d^T g / d^T Md$, можно легко преобразовать полученное неравенство к виду $0 \geq (-d^T g + d^T Md)^2$, откуда следует (23). Кроме того, согласно известной оценке [21] погрешности метода сопряженных градиентов с использованием многочленов Чебышёва, получаем

$$1 - \theta^2 \equiv (g - Md)^T M^{-1} (g - Md) / g^T M^{-1} g \leq \cosh^{-2}(2i/\sqrt{\kappa}),$$

где

$$\kappa = \text{cond}(CM) \equiv \lambda_{\max}(CM) / \lambda_{\min}(CM).$$

В силу условия нормировки (23), отсюда следует оценка

$$\theta^2 = d^T Md / g^T M^{-1} g \geq \tanh^2(2i/\sqrt{\kappa}). \tag{35}$$

Таким образом, установлено, что $0 < \theta < 1$, причем $\theta^2 \rightarrow 1$ при увеличении числа итераций i .

5. РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ТЕСТОВ

Ниже рассмотрены два семейства тестовых задач, которые могут быть решены сведением к безусловной минимизации кусочно-квадратичных функций. Первое семейство — нахождение проекции заданной точки на множество неотрицательных решений системы линейных уравнений (см. также [10]), второе представлено задачей вычисления расстояния между двумя выпуклыми многогранниками, рассмотренной в работе [12]. Последняя задача важна, например, в таких приложениях, как робототехника и компьютерная анимация.

Таблица 1. Свойства матриц для 11 задач из NETLIB

Задача	m	n	$\frac{nz(A)}{mn}$	$\ x_*\ $	$\min_i(AA^T)_{ii}$	$\max_i(AA^T)_{ii}$
afiro	27	51	7.E-2	634.029569	1.18490000	44.9562810
addlittle	56	138	5.E-2	430.764399	1.00000000	10654.0000
Agg3	516	758	1.E-2	765883.022	1.00000001	179783.783
25fv47	821	1876	7.E-3	3310.45652	0.00000000	88184.0358
Pds_02	2953	7716	7.E-5	160697.180	1.00000000	91.0000000
cre_a	3516	7248	7.E-5	1162.32987	0.00000000	27476.8400
80bau3b	2262	12061	8.E-4	4129.96530	1.00000000	321739.679
Ken_13	28362	42659	8.E-5	25363.3224	1.00000000	170.000000
maros_r7	3136	9408	5.E-3	141313.207	3.05175947	3.37132546
cre_b	9648	77137	3.E-3	624.270129	0.00000000	27476.8400
Osa_14	2337	54797	2.E-3	119582.321	18.0000000	845289.908

5.1. Результаты расчетов для задач из коллекции NETLIB

Входные данные следующих 11 задач линейного программирования (тех же самых задач из NETLIB, которые были использованы для тестирования в [11]) послужили для формирования тестовых задач типа (10). Чтобы обеспечить возможность непосредственного сравнения с результатами расчетов из [11], ниже рассматривается только случай $\hat{x} = 0$. Напомним также обозначение $x_* = (\hat{x} + A^T u_*)_+$. Задачи в табл. 1 даны в порядке возрастания количества ненулевых элементов $nz(A)$ в матрице $A \in R^{m \times n}$. Здесь и далее для вещественных чисел используются обозначения типа $7.E-09 = 7 \times 10^{-9}$.

Как видно из двух последних столбцов табл. 1, три из 11 матриц имеют нулевые строки, и более половины имеют довольно большой разброс строчных норм. Это объясняет использование в настоящей работе регуляризации гессiana в виде (22) вместо ранее предложенной в [10], [11] конструкции $M_k = H(u_k) + \delta I_m$. Последняя более подходит для матриц со строками примерно равной длины, таких как в задаче maros_r7 или для матриц с равномерно распределенными псевдослучайными элементами, использовавшихся для тестирования в [9], [11]. В частности, оценка (26) при $D = I$ имеет вид $\gamma = \|A\|^2 / \delta$, что повышает чувствительность метода к выбору параметра δ .

Некоторые задачи, например 80bau3b, имели матрицы с нулевыми столбцами; однако использованная реализация метода позволила устранить возможные численные затруднения.

В табл. 2 воспроизведены результаты, опубликованные в [11], и представлены аналогичные результаты, полученные с использованием предложенного в настоящей работе метода ньютоновского типа (см. алгоритм 1). Подчеркнем, что для предложенного метода использовался фиксированный набор настроечных параметров

$$\delta = 10^{-6}, \quad \varepsilon = 10^{-12}, \quad \varepsilon_{CG} = 10^{-3}, \quad l_{\max} = 10 \tag{36}$$

для всех 11 задач. В работе [11] выбор параметров, например для процедуры Армийо, не был конкретизирован.

В [11] вычисления проводились на компьютере 5 GHz AMD 64 Athlon X2 Dual Core. В наших расчетах использовалось одно вычислительное ядро компьютера 3.40 GHz X8 Intel (R) Core (TM) i7-3770 CPU, предположительно обладающего меньшей производительностью.

Отметим, что алгоритм, приведенный в [11], основан на непосредственном вычислении матрицы M_k и ее треугольной факторизации, в то время как предлагаемая в настоящей работе реализация алгоритма 1 использует метод сопряженных градиентов с диагональным предобуславливанием (33) для вычисления приближенных ньютоновских направлений (аналогично тому, как было предложено в [9]). Поэтому эффективность алгоритма 1 существенно зависит от сходимости итераций метода сопряженных градиентов, которая иногда оказывается медленной (отчасти это компенсируется использованием нового критерия сходимости (32)). С другой стороны,

Таблица 2. Сравнение методов на 11 задачах из NETLIB

Задача	Метод	Время (с)	$\ Ax - b\ _{\infty}$	#NewtIter	#MVMult
afiro	GNewtEGK	0.001	8.63E-11	17	398
»	ssGNewton	0.06	6.39E-14	—	—
»	cqpMOSEK	0.31	1.13E-13	—	—
addlittle	GNewtEGK	0.003	6.45E-10	22	1050
»	ssGNewton	0.05	2.27E-13	—	—
»	cqpMOSEK	0.35	7.18E-11	—	—
agg3	GNewtEGK	0.14	3.93E-07	116	9234
»	ssGNewton	0.27	3.59E-08	—	—
»	cqpMOSEK	0.40	2.32E-10	—	—
25fv47	GNewtEGK	0.54	7.15E-10	114	32234
»	ssGNewton	1.51	3.43E-09	—	—
»	cqpMOSEK	1.36	1.91E-11	—	—
pds_02	GNewtEGK	0.32	1.55E-08	75	8559
»	ssGNewton	2.30	1.40E-07	—	—
»	cqpMOSEK	0.51	8.20E-06	—	—
cre_a	GNewtEGK	3.36	2.64E-09	219	85737
»	ssGNewton	1.25	4.13E-06	—	—
»	cqpMOSEK	0.61	2.15E-10	—	—
80bau3b	GNewtEGK	0.27	3.33E-09	79	6035
»	ssGNewton	0.95	1.18E-12	—	—
»	cqpMOSEK	0.80	2.90E-07	—	—
ken_13	GNewtEGK	1.41	2.70E-08	55	6285
»	ssGNewton	9.09	4.39E-09	—	—
»	cqpMOSEK	2.09	1.71E-09	—	—
maros_r7	GNewtEGK	0.10	1.18E-09	27	535
»	ssGNewton	2.86	2.54E-11	—	—
»	cqpMOSEK	55.20	3.27E-11	—	—
cre_b	GNewtEGK	9.25	6.66E-10	75	24590
»	ssGNewton	13.20	1.62E-09	—	—
»	cqpMOSEK	2.31	1.61E-06	—	—
osa_14	GNewtEGK	42.59	8.25E-08	767	104874
»	ssGNewton	60.10	4.10E-08	—	—
»	cqpMOSEK	4.40	7.82E-05	—	—

основные вычислительные процедуры алгоритма 1 представлены операциями матрично-векторного умножения типа $x = Ap$ или $q = A^T u$ и элементарными векторными операциями, что способствует его эффективной параллельной реализации.

В табл. 2 сокращение cqpMOSEK относится к использованию пакета MOSEK программ выпуклой квадратичной оптимизации, см. [11]. Сокращение ssGNewton относится к методу, разработанному и протестированному в [11], а через GNewtEGK обозначен метод, предложенный в настоящей работе.

Несмотря на использование менее производительного компьютера, предлагаемый метод GNewtEGK работает быстрее остальных двух для 8 задач из 11. В противном случае, можно заметить, что меньшее время вычислений для cqpMOSEK сопровождается выдачей приближенного решения с гораздо большей невязкой, чем для GNewtEGK (см. результаты решения задач cre_b и osa_14).

Результаты непосредственного сравнения эффективности стандартного критерия остановки итераций метода сопряженных градиентов $\gamma^{(i)} \leq \epsilon_{CG}^2 \gamma^{(0)}$ (см. алгоритм 2) с новым критерием оста-

Таблица 3. Сравнение критериев остановки итераций метода СГ

Критерий	ϵ_{CG}	Средн. геом. время	Средн. арифм. время	Средн. геом. невязка
Стандартный	0.05	1.78	12.37	3.64–09
Стандартный	0.03	1.45	12.21	3.47–09
Стандартный	0.01	1.35	14.91	2.11–09
Стандартный	0.003	1.48	21.14	2.14–09
Стандартный	0.001	1.82	29.12	3.64–09
Новый	0.003	1.66	11.12	3.46–09
Новый	0.002	1.39	11.36	3.64–09
Новый	0.001	1.26	10.81	4.65–09
Новый	0.0003	1.26	11.25	4.23–09
Новый	0.0001	1.33	12.47	3.60–09

новки (32) дают в табл. 3, где приводятся значения времени счета и точности, усредненных по 11 задачам, в зависимости от заданного значения ϵ_{CG} . Видно, что те же уровни нормы невязки $\|Ax - b\|_\infty$ могут быть достигнуты и с новым критерием остановки итераций, причем с меньшими затратами времени и при менее критичной зависимости результата от выбранного значения ϵ_{CG} .

5.2. Вычисление расстояния между двумя выпуклыми многогранниками

Пусть два выпуклых многогранника \mathcal{X}_1 и \mathcal{X}_2 заданы следующими двумя системами линейных неравенств:

$$\mathcal{X}_1 = \{x_1 : A_1^T x_1 \leq b_1\}, \quad \mathcal{X}_2 = \{x_2 : A_2^T x_2 \leq b_2\},$$

где $A_1 \in R^{s \times n_1}$, $A_2 \in R^{s \times n_2}$, и векторы x_1 , x_2 , b_1 , b_2 имеют соответствующие размерности. Исходная задача вычисления расстояния между \mathcal{X}_1 и \mathcal{X}_2 имеет следующий вид:

$$x_* = \arg \min_{A_1^T x_1 \leq b_1, A_2^T x_2 \leq b_2} \|x_1 - x_2\|^2 / 2, \quad \text{где } x = [x_1^T, x_2^T]^T \in R^{2s}.$$

Указанная постановка была предложена в [12], где для решения этой задачи применялся метод проекции градиента.

Будем использовать следующую регуляризованную/штрафную приближенную переформулировку задачи, сводящую ее к безусловной минимизации выпуклой кусочно-квадратичной функции. Вводя матрицы

$$A \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \in R^{2s \times (n_1 + n_2)}, \quad B \begin{bmatrix} I_s & -I_s \\ -I_s & I_s \end{bmatrix} \in R^{2s \times 2s},$$

и вектор $b = [b_1^T, b_2^T]^T \in R^{n_1 + n_2}$, рассмотрим задачу

$$x_*(\epsilon) = \arg \min_{x \in R^{2s}} \left(\frac{\epsilon}{2} \|x\|^2 + \frac{1}{2} x^T Bx + \frac{1}{2\epsilon} \|(A^T x - b)_+\|^2 \right),$$

где регуляризирующий/штрафной параметр ϵ является достаточно малым положительным числом (мы использовали значение $\epsilon = 10^{-4}$). Построенную задачу можно решить, непосредственно используя описанный выше алгоритм 1, полагая в (22) $\delta = 0$ и используя следующие явные выражения для градиента и обобщенного гессиана:

$$g(x) = \epsilon x + Bx + \epsilon^{-1} A(A^T x - b)_+, \quad H(x) = \epsilon I + B + \epsilon^{-1} AD(x)A^T,$$

где $D(x) = \text{Diag}(\text{sign}(A^T x - b)_+)$. При решении практической задачи по определению расстояния между двумя трехмерными выпуклыми многогранниками, заданными своими гранями (так что

Таблица 4. Результаты применения метода ньютоновского типа при $\varepsilon = 10^{-4}$ к задаче определения расстояния между двумя псевдослучайными выпуклыми многогранниками (с $n/2$ гранями у каждого)

n	$\ x_1 - x_2\ _2$	$\ (A^T x_* - b)_+\ _\infty$	time (s)	$\ g(x_*)\ _\infty$	#NewtIter
8	0.001815	9.69–09	<0.001	7.89–13	15
16	0.481528	8.63–05	<0.001	1.27–13	3
32	0.795116	8.80–05	<0.001	1.46–12	28
64	1.102286	1.32–04	<0.001	5.58–13	13
128	1.446262	1.36–04	<0.001	7.12–13	17
256	1.449913	9.54–05	<0.001	4.37–13	11
512	1.460197	1.31–04	0.001	8.16–13	15
1024	1.460063	1.46–04	0.002	1.09–12	14
2048	1.463320	1.04–04	0.005	6.58–13	19
4096	1.463766	1.26–04	0.009	3.59–13	20
8192	1.463879	1.03–04	0.009	8.32–14	12
16384	1.463976	7.58–05	0.009	1.64–12	13
32768	1.464046	3.28–05	0.018	1.54–12	13

$s = 3$), ньютоновские итерации производятся в R^6 , и поэтому затраты на каждую такую итерацию пропорциональны суммарному количеству граней многогранников $n = n_1 + n_2$. В этом случае для вычисления $d = H^{-1}g$ предпочтительно выполнение явного формирования гессиана $H(x)$ и его точной треугольной факторизации.

Тестовые многогранники с $n/2$ гранями каждый, строились вокруг точек $e = [1, 1, 1]^T$ и $-e$ и определялись системами неравенств $A_1^T(x - e) \leq b_1$ и $A_2^T(x + e) \leq b_2$ соответственно, где $b_1 = b_2 = [1, \dots, 1] \in R^{n/2}$. Столбцы матриц A_1 и A_2 были заданы как $n/2$ псевдослучайных единичных 3-векторов, компонентами которых служили члены т. наз. логистической рекуррентной последовательности (см., например, [23] и цитированные там источники) $\xi_0 = 0.4$, $\xi_k = 1 - 2\xi_{k-1}^2$. Использовались формулы $A_1(i, j) = \xi_{20(i-1)+3(j-1)}$ и аналогично для A_2 ; затем их столбцы нормировались к единичной длине. Соответствующие результаты вычислений представляются вполне удовлетворительными как по точности, так и по скорости вычислений, см. табл. 4.

Дополнительно отметим, что для вычисленного расстояния всегда сохраняется верхняя оценка $\|x_1 - x_2\| \leq 2\sqrt{3} - 2 \approx 1.464101$ (в силу того, что многогранники \mathcal{X}_1 и \mathcal{X}_2 являются описанными вокруг сфер $\|x_1 - e\| = 1$ и $\|x_2 + e\| = 1$ соответственно).

Авторы благодарны В.А. Гаранже за многочисленные полезные замечания, позволившие существенно улучшить изложение материала статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Golikov A.I., Kaporin I.E. Inexact Newton Method for Minimization of Convex Piecewise Quadratic Functions // In: Numerical Geometry, Grid Generation and Scientific Computing Proceedings of the 9th International Conference, NUMGRID 2018/Voronoi 150, Celebrating the 150th Anniversary of G.F. Voronoi, Moscow, Russia, December 2018 (Garanzha, V.A., Kamenski, Lennard, Si, Hang, Eds.) Lecture Notes in Computational Science and Engineering, Vol. 131. Springer Nature Switzerland AG (2019) ISBN 978-3-030-23435-5 https://doi.org/10.1007/978-3-030-23436-2_10
2. Гаранжа В.А., Капорин И.Е. Регуляризация барьерного вариационного метода построения расчетных сеток // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 9. С. 1489–1503.
3. Garanzha V., Kaporin I., Konshin I. Truncated Newton type solver with application to grid untangling problem // Numer. Linear Algebra Appls. 2004. V. 11. № 5–6. P. 525–533.

4. *Капорин И.Е.* Использование внутренних итераций метода сопряженных градиентов при решении больших разреженных нелинейных задач оптимизации // Ж. вычисл. матем. и матем. физ. 2003. Т. 43. № 6. С. 802–807.
5. *Garanzha V., Kudryavtseva L.* Hypoelastic Stabilization of Variational Algorithm for Construction of Moving Deforming Meshes // In: Evtushenko Y., Jacimovic M., Khachay M., Kochetov Y., Malkova V., Posypkin M. (eds.) Optimization and Applications. OPTIMA 2018. Communications in Computer and Information Science, Vol. 974. Springer, Cham (2019). P. 497–511.
https://doi.org/10.1007/978-3-030-10934-9_35
6. *Голиков А.И., Евтушенко Ю.Г.* Отыскание нормальных решений в задачах линейного программирования // Ж. вычисл. матем. и матем. физ. 2000. Т. 40. № 12. С. 1766–1786.
7. *Mangasarian O.L.* A Newton method for linear programming // J. Optimizat. Theory and Appl. 2004. V. 121. № 1. P. 1–18.
8. *Голиков А.И., Евтушенко Ю.Г., Моллаверди Н.* Применение метода Ньютона к решению задач линейного программирования большой размерности // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. № 9. С. 1564–1573.
9. *Гаранжа В.А., Голиков А.И., Евтушенко Ю.Г., Нгуен М.Х.* Параллельная реализация метода Ньютона для решения больших задач линейного программирования // Ж. вычисл. матем. и матем. физ. 2009. Т. 49. № 8. С. 1369–1384.
10. *Ганин Б.В., Голиков А.И., Евтушенко Ю.Г.* Проективно-двойственный метод решения систем линейных уравнений с неотрицательными переменными // Ж. вычисл. матем. и матем. физ. 2018. Т. 58. № 2. С. 169–180.
11. *Ketabchi S., Moosaei H., Parandegan M., Navidi H.* Computing minimum norm solution of linear systems of equations by the generalized Newton method // Numerical Algebra, Control and Optimizat. 2017. V. 7. № 2. P. 113–119.
12. *Bobrow J.E.* A direct minimization approach for obtaining the distance between convex polyhedra // The Internat. J. Robotics Research. 1989. V. 8. № 3. P. 65–76.
13. *Mangasarian O.L.* A finite Newton method for classification // Optimizat. Methods and Software. 2002. V. 17. № 5. P. 913–929.
14. *Hiriart-Urruty J.B., Strodjot J.J., Nguyen V.H.* Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data // Applied Mathematics and Optimization. 1984. V. 11. № 1. P. 43–56.
15. *Тихонов А.Н.* О некорректных задачах линейной алгебры и устойчивом методе их решения // Докл. Акад. Наук СССР. 1965. Т. 163. № 4. С. 591–594.
16. *Голиков А.И., Евтушенко Ю.Г.* Регуляризация и нормальные решения систем линейных уравнений и неравенств // Тр. ИММ УрО РАН. 2014. Т. 20. № 2. С. 113–121.
17. *Axelsson O., Kaporin I.E.* Error norm estimation and stopping criteria in preconditioned conjugate gradient iterations // Numer. Linear Algebra Appls. 2001. V. 8. № 4. P. 265–286.
18. *Kaporin I.E., Axelsson O.* On a class of nonlinear equation solvers based on the residual norm reduction over a sequence of affine subspaces // SIAM J. Sci. Comput. 1995. V. 16. № 1. P. 228–249.
19. *Капорин И.Е., Милюкова О.Ю.* Массивно-параллельный алгоритм предобусловленного метода сопряженных градиентов для численного решения систем линейных алгебраических уравнений // Сб. трудов отдела проблем прикладной оптимизации ВЦ РАН / Под ред. В.Г. Жадана. М.: Изд-во ВЦ РАН, 2011. С. 32–49.
20. *Hestenes M.R., Stiefel E.L.* Methods of conjugate gradients for solving linear systems // J. Res. Nat. Bur. Standards. 1952. V. 49. № 1. P. 409–436.
21. *Axelsson O.* A class of iterative methods for finite element equations // Comput. Meth. Appl. Mech. Engrg. 1976. V. 9. P. 123–137.
22. *Dongarra J., Eijkhout V.* Finite-choice algorithm optimization in Conjugate Gradients // Lapack Working Note 159. 2003. University of Tennessee Computer Science Report UT-CS-03-502.
23. *Yu L., Barbot J.P., Zheng G., Sun H.* Compressive sensing with chaotic sequence. // IEEE Signal Proc. Lett. 2010. V. 17. № 8. P. 731–734.