О МЕТОДЕ ДИКИНА
ДЛЯ ЗАДАЧИ КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ*

Г. Ш. Тамасян

grigoriytamasjan@mail.ru

31 октября 2024 г.

В докладе рассматривается задача выпуклого квадратичного программирования. Приводятся необходимые теоретические сведения. Демонстрируются практические приемы сведения общей квадратичной задачи к первой канонической форме, в частности, преобразование Хилдрета. Основное внимание уделяется применению итеративного метода Дикина для задачи с двусторонними ограничениями. Дается подробное его описание согласно работам [1, 2]. С вопросами обоснования и сходимости более подробно можно ознакомится в работах [2, 3]. Приведены результаты вычислительных экспериментов для двумерного случая. В приложении дается общая схема для генерации задач квадратичного программирования (второй канонической формы) с двусторонними ограничениями на переменные.

1°. Предварительные сведения. Пусть N, N_1, M_1 и M_2 — индексные множества, такие, что $N \neq \emptyset$, $N_1 \subset N$, $M_1 \cap M_2 = \emptyset$. Рассмотрим общую задачу квадратичного программирования

$$\begin{gathered} f(x) := \tfrac{1}{2} x^{\mathsf{T}} D x + x^{\top} c \longrightarrow \min, \\ A[M_1, N] \cdot x[N] \geqslant b[M_1], \\ A[M_2, N] \cdot x[N] = b[M_2], \\ x[N_1] \geqslant \mathbf{0}, \end{gathered} \tag{1}$$

где $D = D[N, N]$ — симметричная неотрицательно определенная матрица.

Будем считать, что задача (1) имеет решение (см. [4, с. 111–112]), т. е. множество её планов Ω непусто и целевая функция f ограничена снизу на Ω.

Напомним критерий оптимальности (см. [2, с. 17]). Положим $N_2 = N \setminus N_1$, $M = M_1 \cup M_2$.

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» `http://oml.cmlaboratory.com/`

ТЕОРЕМА 1. *Для того чтобы план $x_* = x_*[N]$ задачи (1) был оптимальным, необходимо и достаточно, чтобы нашлись векторы $y_* = y_*[N]$, $u_* = u_*[M]$, $v_* = v_*[M]$ удовлетворяющие системе равенств и неравенств*:

$$\begin{gathered} Dx_* - A^{\mathsf{T}}u_* - y_* = -c, \\ Ax_* - v_* = b, \\ y_*[N_2] = \mathbf{0}, \quad v_*[M_2] = \mathbf{0}, \\ x_*[N_1] \geqslant \mathbf{0}, \quad y_*[N_1] \geqslant \mathbf{0}, \quad u_*[M_1] \geqslant \mathbf{0}, \quad v_*[M_1] \geqslant \mathbf{0}, \\ x_*[i] \cdot y_*[i] = 0 \ \ \forall\, i \in N_1, \quad u_*[j] \cdot v_*[j] = 0 \ \ \forall\, j \in M_1. \end{gathered} \tag{2}$$

Замечание 1. Из условий Куна – Таккера (2) следует, что

$$f(x_*) = \tfrac{1}{2}(x_*^{\mathsf{T}}c + u_*^{\mathsf{T}}b).$$

2°. Введем двойственную задачу для задачи квадратичного программирования (1):

$$\begin{gathered} g(z) := -\tfrac{1}{2}v^{\mathsf{T}}Dv + u^{\mathsf{T}}b \to \max, \\ D[N_1, N] \cdot v[N] - A^{\mathsf{T}}[N_1, M] \cdot u[M] \geqslant -c[N_1], \\ D[N_2, N] \cdot v[N] - A^{\mathsf{T}}[N_2, M] \cdot u[M] = -c[N_2], \\ u[M_1] \geqslant \mathbf{0}\,. \end{gathered} \tag{3}$$

Здесь $z = (v, u)$. Множество планов задачи (3) обозначим P.

Справедливы следующие две теоремы двойственности (см. [4, с. 113–116]).

ТЕОРЕМА 2 (первая теорема двойственности)**.** *Если одна из двойственных задач квадратичного программирования разрешима, то разрешима и другая задача. При этом справедливо соотношение двойственности*

$$\min_{x \in \Omega} f(x) = \max_{z \in P} g(z).$$

ТЕОРЕМА 3 (вторая теорема двойственности)**.** *Пусть x_0 и $z_0 = (v_0, u_0)$ — планы задач (1) и (3) соответственно. Для того чтобы эти планы были оптимальными, необходимо и достаточно, чтобы выполнялись условия*

$$\big(A[i, N] \cdot x_0[N] - b[i]\big) \cdot u_0[i] = 0 \qquad \text{для всех } \ i \in M_1, \tag{4}$$

$$\big(D[j, N] \cdot v_0[N] - A^{\mathsf{T}}[j, M] \cdot u_0[M] + c[j]\big) \cdot x_0[j] = 0 \qquad \text{для всех } \ j \in N_1, \tag{5}$$

$$D(x_0 - v_0) = \mathbf{0}\,. \tag{6}$$

Замечание 2. Отметим, что *условия дополнительности* (4) записываются на индексах, соответствующих ограничениям-неравенствам в задаче (1), а *условия дополнительности* (5) — на индексах в задаче (3).

3°. Преобразование Хилдрета. Далее всюду будем считать, что матрица D положительно определена. Рассмотрим задачу квадратичного программирования вида

$$\begin{aligned} f(x) := \tfrac{1}{2}x^{\mathsf{T}}Dx + x^{\mathsf{T}}c \longrightarrow \min, \\ A[M_1, N] \cdot x[N] \geqslant b[M_1], \\ A[M_2, N] \cdot x[N] = b[M_2]. \end{aligned} \tag{7}$$

Задача (7) отличается от общей задачи (1) тем, что отсутствуют явные знаковые ограничения на переменные. Это всегда можно сделать, если произвести включение знаковых ограничений на переменные в общие ограничения-неравенства.

Запишем к ней двойственную задачу (см. (3)):

$$\begin{aligned} -\tfrac{1}{2}v^{\mathsf{T}}Dv + u^{\mathsf{T}}b \to \max, \\ Dv - A^{\mathsf{T}}u = -c, \\ u[M_1] \geqslant \mathbf{0}\,. \end{aligned} \tag{8}$$

Заметим, что задачу (8) можно упростить, исключив переменную v. Действительно, из ограничений-равенств задачи (8) следует, что

$$v = D^{-1}(A^{\mathsf{T}}u - c).$$

При этом целевая функция принимает вид

$$-\tfrac{1}{2}(A^{\mathsf{T}}u - c)^{\mathsf{T}}D^{-1}(A^{\mathsf{T}}u - c) + u^{\mathsf{T}}b.$$

Поменяв знак у целевой функции и отбросив константу, придем к вспомогательной задаче квадратичного программирования:

$$\begin{aligned} \tfrac{1}{2}u^{\mathsf{T}}Qu + u^{\mathsf{T}}q \longrightarrow \min, \\ u[M_1] \geqslant \mathbf{0}, \end{aligned} \tag{9}$$

где $Q = AD^{-1}A^{\mathsf{T}}$, $q = -(AD^{-1}c + b)$.

Из первой теоремы двойственности следует, что задача (9) имеет решение. Пусть u_* — решение задачи (9), тогда вектор $z_* = (v_*, u_*)$, где

$$v_* = D^{-1}(A^{\mathsf{T}}u_* - c)$$

будет решением задачи (8).

В силу второй теоремы двойственности, строгой выпуклости целевой функции $f(x)$ на всем пространстве и выпуклости множества Ω имеем, что единственным решением задачи (7) является вектор

$$x_* = v_* = D^{-1}(A^{\mathsf{T}}u_* - c).$$

З а м е ч а н и е 3. Переход от задачи (7) к задаче (9) называется преобразованием Хилдрета.

4°. Метод Дикина. Рассматриваемый далее метод Дикина (см. [2, с. 114]) разработан для задач выпуклого квадратичного программирования вида:

$$\begin{gathered} f(x) = \tfrac{1}{2}x^{\mathsf T}Dx + x^{\mathsf T}c \longrightarrow \min_{x\in\Omega}, \\ \Omega = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geqslant \mathbf{0}\}. \end{gathered} \tag{10}$$

Метод применим при следующих предположениях:

- множество планов Ω содержит относительно внутреннюю точку, т. е. не пусто множество

$$\Omega_0 := \{x \in \mathbb{R}^n \mid Ax = b,\ x > \mathbf{0}\}; \tag{11}$$

- матрица A — полного строчного ранга, т. е. ранг матрицы $A[M,N]$ равен $|M|$;

- целевая функция ограничена снизу на множестве планов Ω.

Для задачи (10) запишем условия Куна – Таккера (см. (2)):

$$\begin{gathered} Dx - A^{\mathsf T}u - y = -c, \\ Ax = b, \\ x \geqslant \mathbf{0}, \quad y \geqslant \mathbf{0}, \\ x[i]\cdot y[i] = 0 \ \ \forall\, i \in N. \end{gathered} \tag{12}$$

Здесь мы воспользовались тем, что индексные множества принимают вид $N_1 = N$, $N_2 = \emptyset$, $M_1 = \emptyset$, $M_2 = M$.

Обозначим $g(x) = Dx + c$. Из первого равенства системы (12) выразим y:

$$y = g(x) - A^{\mathsf T}u.$$

Введем в рассмотрение функцию

$$\Phi(x,u) := \sum_{i\in N}(x[i]\cdot y[i])^2.$$

Справедливо следующее утверждение.

ТЕОРЕМА 4. *Вектор $x_* \in \Omega$ является решением задачи* (10) *тогда и только тогда, когда существует вектор u_*, такой, что*

$$\begin{gathered} g(x_*) - A^{\mathsf T}u_* \geqslant \mathbf{0}, \\ \Phi(x_*,u_*) = 0. \end{gathered}$$

Опишем принципиальную схему метода Дикина. Метод заключается в построении последовательностей точек $x_k \in \Omega_0$ и u_k до тех пор, пока $\Phi(x_k, u_k)$ не станет равной нулю.

Пусть уже имеется k-е приближение $x_k \in \Omega_0$.

(I) Вычислим $g_k = Dx_k + c$ и найдем единственное решение u_k системы

$$A\Lambda_k^2 A^{\mathsf{T}} u = A\Lambda_k^2 g_k,$$

где $\Lambda_k = \operatorname{diag}(x_k)$.

(II) Находим $y_k = g_k - A^{\mathsf{T}} u_k$, $\Phi_k = \sum_{i \in N} (x_k[i] y_k[i])^2$. Если $\Phi_k = 0$, то x_k — решение исходной задачи. В противном случае переходим к следующему пункту.

(III) Определим направление спуска $s_k = -\Lambda_k^2 y_k$.

(IV) Вычисляем шаг $\lambda_k = \min\left\{\dfrac{\alpha}{\sqrt{\Phi_k}}, \dfrac{\Phi_k}{s_k^{\mathsf{T}} D s_k}\right\}$, где параметр $0 < \alpha < 1$.

(V) Строим новое приближение $x_{k+1} = x_k + \lambda_k s_k$.

Описание алгоритма завершено.

З а м е ч а н и е 4. Отметим ряд особенностей алгоритма.

- ◇ Наиболее трудоемкой частью алгоритма является пункт (I), а именно решение системы линейных уравнений на каждом шаге.
- ◇ При всех k очередное приближение x_k принадлежит множеству Ω_0 и последовательность $\{f(x_k)\}$ строго убывает.
- ◇ Параметр α в процессе вычислений можно менять, а именно, рекомендуется устремлять к единице.
- ◇ Отсутствует накопление ошибок.
- ◇ Можно поменять местами пункты (II) и (III), тогда упростится вычисление $\Phi_k = -y_k^{\mathsf{T}} s_k$.

5°. Задача с двусторонними ограничениями. Исследуем подробно следующую задачу квадратичного программирования:

$$\begin{gathered} \tfrac{1}{2}x^{\mathsf T}Dx + x^{\mathsf T}c \longrightarrow \min, \\ \ell \leqslant x \leqslant w. \end{gathered} \tag{13}$$

Здесь D — симметричная положительно определенная матрица, $\ell < w$.

Очевидно, что при этих предположениях задача (13) имеет единственное решение. Положим $A = \begin{pmatrix} E \\ -E \end{pmatrix}$, $b = \begin{pmatrix} \ell \\ -w \end{pmatrix}$. Перепишем задачу (13) в матричной форме:

$$\begin{gathered} \tfrac{1}{2}x^{\mathsf T}Dx + x^{\mathsf T}c \longrightarrow \min, \\ Ax \geqslant b. \end{gathered} \tag{14}$$

Условия Куна – Таккера (см. (2)) для этой задачи принимают вид

$$\begin{cases} Dx - A^{\mathsf T}u = -c, \\ Ax - v = b, \\ v \geqslant \mathbf{0}, \quad u \geqslant \mathbf{0}, \\ v[j] \cdot u[j] = 0 \ \ \forall\, j \in M_1, \\ \ell \leqslant x \leqslant w. \end{cases}$$

Как известно, решение квадратичной задачи равносильно решению соответствующей системы Куна – Таккера. Для поиска решения воспользуемся методом Дикина.

Рассмотрим два подхода.

Подход I (преобразование Хилдрета). Применим результаты из раздела 3°, а именно, сведем задачу (14) к проблеме вида (9), что является частным случаем задачи (10). Получим

$$\begin{gathered} \frac{1}{2}u^{\mathsf T}Qu + u^{\mathsf T}q \longrightarrow \min, \\ u \geqslant \mathbf{0}, \end{gathered} \tag{15}$$

где

$$Q = AD^{-1}A^{\mathsf T} = \begin{pmatrix} D^{-1} & -D^{-1} \\ -D^{-1} & D^{-1} \end{pmatrix},$$

$$q = -(b + AD^{-1}c) = \begin{pmatrix} -\ell - D^{-1}c \\ w + D^{-1}c \end{pmatrix},$$

$$A = \begin{pmatrix} E \\ -E \end{pmatrix}, \quad b = \begin{pmatrix} \ell \\ -w \end{pmatrix}.$$

Пусть u_* — решение задачи (15). Тогда решение исходной задачи (14) определяется по формуле

$$x_* = D^{-1}(A^{\mathsf{T}} u_* - c). \tag{16}$$

В этом случае реализация метода Дикина существенно упрощается, т. к. в задаче (15) отсутствуют ограничения типа равенств. Таким образом, не требуется на каждом шаге решать систему из п. (I), а в п. (II) надо положить $y_k = Qu_k + q$.

Также не вызывает проблем выбор начального приближения $u_0 \in \Omega_0$. Это может быть произвольный элемент из положительного ортанта. Либо можно воспользоваться равенством (16), а именно $(E \ \ {-E})u = Dx + c$: вычислим правую часть $g := Dx_0 + c$ для произвольного $x_0 \in (\ell, w)$, тогда в качестве u_0 возьмем вектор вида $\begin{pmatrix} \max\{\mathbf{0}, g\} \\ \max\{\mathbf{0}, -g\} \end{pmatrix}$.

Подход II. С помощью элементарных преобразований сведем задачу (13) к виду (10). Произведем замену переменной вида $x = z_1 + \ell$, получим

$$\begin{cases} \frac{1}{2} z_1^{\mathsf{T}} D z_1 + z_1^{\mathsf{T}} (D\ell + c) \longrightarrow \min, \\ \mathbf{0} \leqslant z_1 \leqslant w - \ell. \end{cases}$$

Теперь разобьем двойное неравенство и введем новую переменную z_2 так, чтобы второе неравенство стало равенством, имеем

$$\begin{cases} \frac{1}{2} z_1^{\mathsf{T}} D z_1 + z_1^{\mathsf{T}} (D\ell + c) \longrightarrow \min \\ z_1 + z_2 = w - \ell, \\ z_1 \geqslant \mathbf{0},\ z_2 \geqslant \mathbf{0}\,. \end{cases}$$

Осталось переписать в матричной форме. Получим

$$\begin{cases} \frac{1}{2} z^{\mathsf{T}} Q z + z^{\mathsf{T}} q \longrightarrow \min, \\ Az = b, \\ z \geqslant \mathbf{0}, \end{cases} \tag{17}$$

где

$$Q = \begin{pmatrix} D & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}, \qquad q = \begin{pmatrix} D\ell + c \\ \mathbf{0} \end{pmatrix}, \qquad z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

$$A = \begin{pmatrix} E & E \end{pmatrix}, \qquad b = w - \ell.$$

Пусть $z_* = (z_1^*, z_2^*)$ — решение задачи (17). Тогда решение исходной задачи (13) определяется по формуле $x_* = z_1^* + \ell$.

В качестве начального приближения для метода Дикина можно выбрать, например, вектор $z = \begin{pmatrix} \gamma b \\ (1-\gamma)b \end{pmatrix} > \mathbf{0}$ для любого $\gamma \in (0,1)$. Очевидно, что такое z принадлежит Ω_0 (см. (11)).

Подробней разберем пункт (I) метода Дикина. Зафиксируем $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \Omega_0$ и рассмотрим систему линейных уравнений вида

$$A\Lambda^2 A^{\mathsf{T}} u = A\Lambda^2 g.$$

Здесь $\Lambda = \begin{pmatrix} \Lambda_1 & \mathbf{O} \\ \mathbf{O} & \Lambda_2 \end{pmatrix} = \mathrm{diag}(z)$, $g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = Qz + q$ при этом $g_1 = Dz_1 + (D\ell + c)$, $g_2 = \mathbf{0}$. Таким образом, система уравнений примет вид

$$(\Lambda_1^2 + \Lambda_2^2)u = \Lambda_1^2 g_1.$$

Нетрудно понять, что единственное решение определяется по формуле

$$u = (\Lambda_1^2 + \Lambda_2^2)^{-1}\Lambda_1^2 g_1.$$

6°. Вычислительные эксперименты. Приведем результаты вычислительных экспериментов по решению задачи (13) для случая, когда $\mathrm{n} = 2$. Были реализованы оба подхода описанные в разд. 5°, а именно подход **I** как решение задачи (15), подход **II** — (17).

На нижеприведенных графиках 1–3 представлены траектории приближений (пунктирная линия черного цвета для **I** подхода, ломаная красного цвета — **II**), а также линии уровня целевой функции. На легенде каждого графика приводятся количество итераций каждого из подходов, а через Δ близость последнего приближения до оптимального решения. Особенно отметим, что на графиках изображены приближения в исходных переменных x из $\mathbb{R}^2$.

Генерация задач осуществлялась согласно алгоритму описанному в приложении на 11 с.

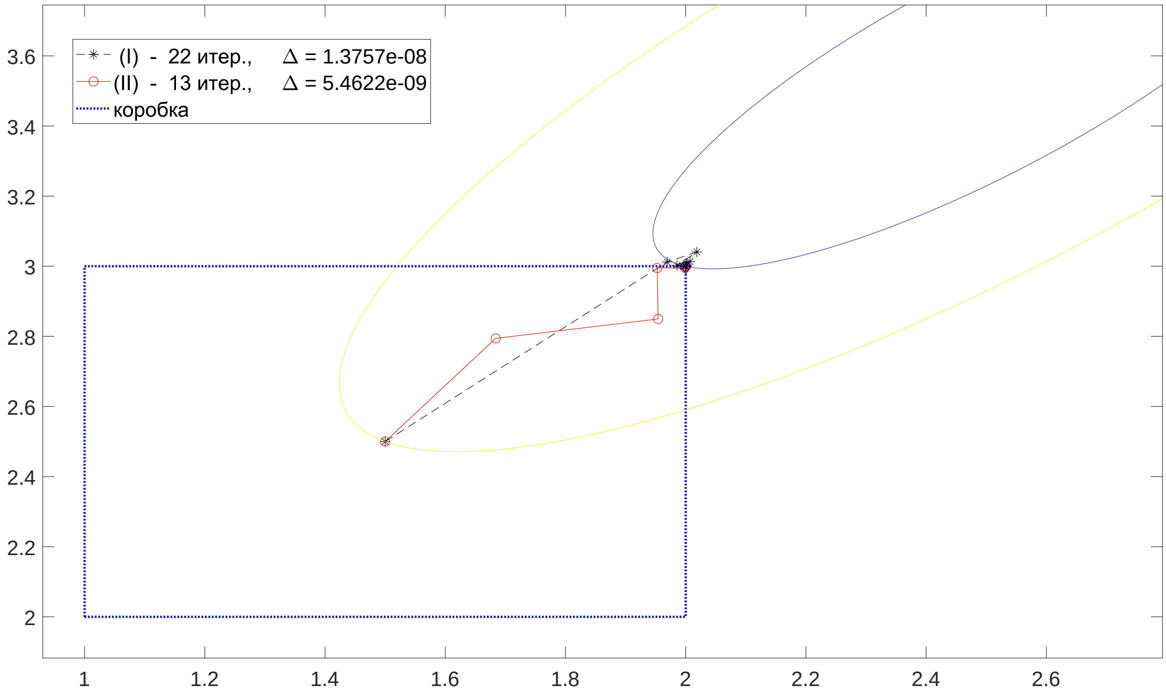

Рис. 1. Решение в угловой точке ограничений

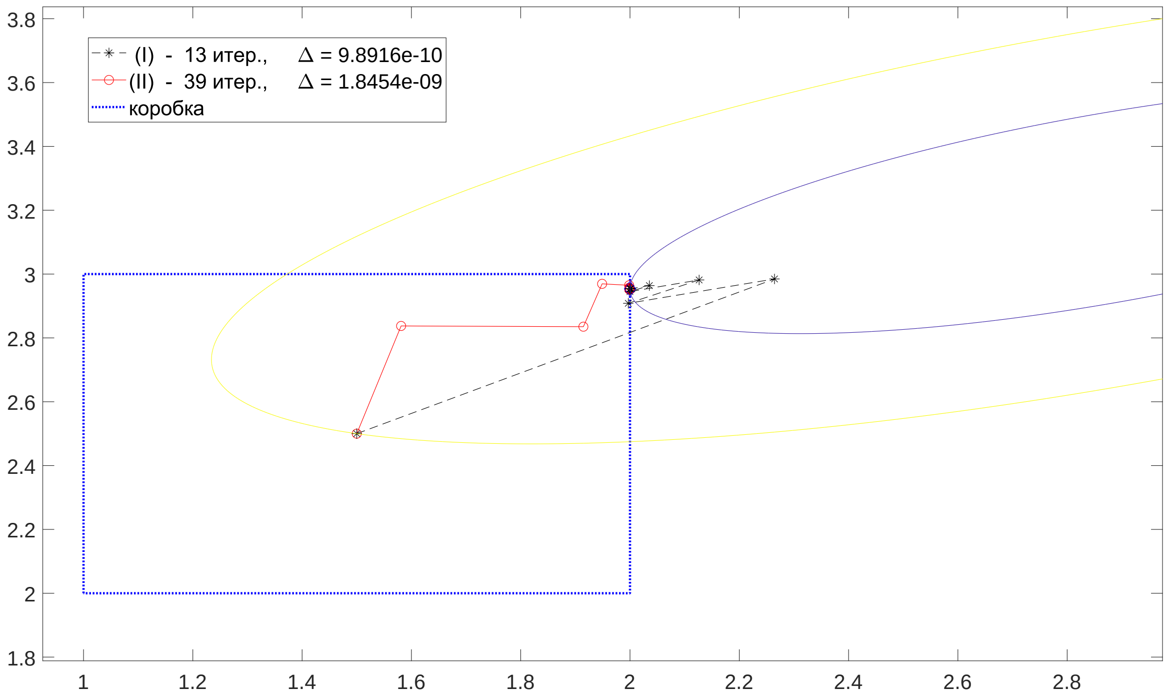

Рис. 2. Решение на одной из границ ограничений

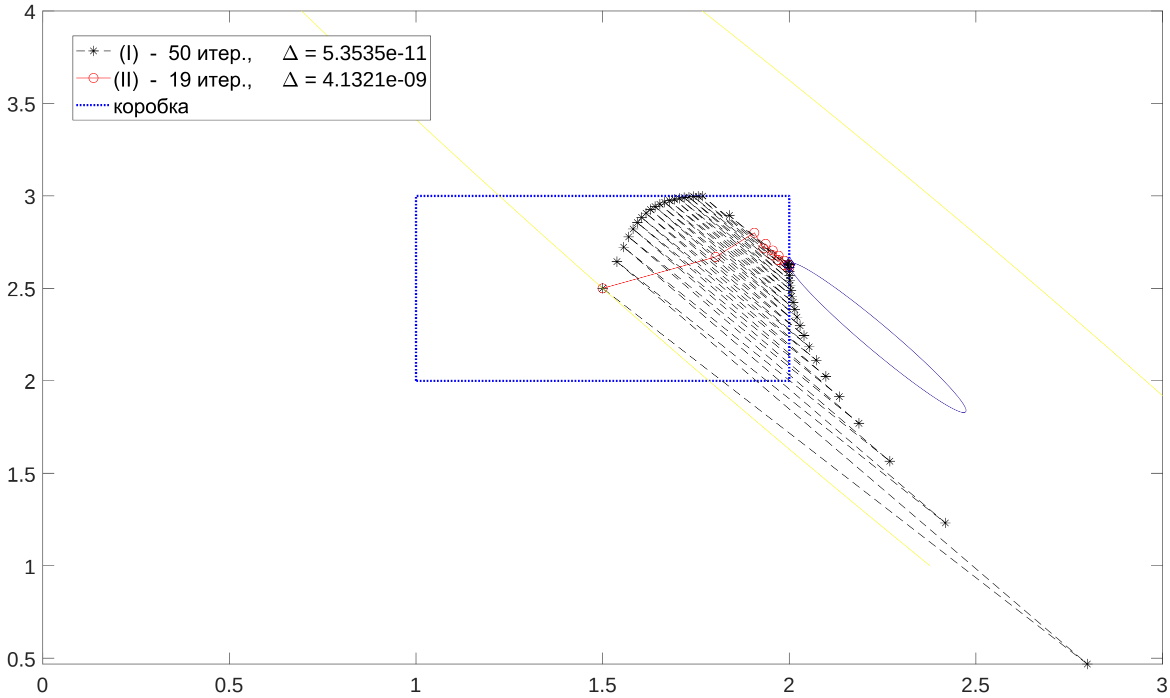

Рис. 3. Случай плохообусловленной матрицы D

ЛИТЕРАТУРА

1. Дикин И. И. *Итеративное решение задач линейного и квадратичного программирования* // ДАН СССР, 1967, Т. 174, № 4, С. 747–748.

2. Даугавет В. А. Численные методы квадратичного программирования. СПб.: Изд-во С.-Петерб. ун-та, 2004. 128 с.

3. Дикин И. И., Зоркальцев В. И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). Новосибирск: Наука, 1980. 143 с.

4. Гавурин М. К., Малоземов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984. 176 с.

5. Малоземов В. Н., Плоткин А. В. Двойственность в квадратичном программировании. Задача Сильвестра // Семинар «CNSA & NDO». Избранные доклады. 8 декабря 2021 г. (http://cnsa.cmlaboratory.com/reps21.shtml#1208)

ПРИЛОЖЕНИЕ

Ниже приведена схема для генерации задач вида (14). Для этого воспользуемся системой Куна – Таккера:

$$\begin{cases} Ax - v = b, \\ Dx - A^{\mathsf{T}}u = -c, \\ v \geqslant \mathbf{0},\ u \geqslant \mathbf{0},\ v^{\mathsf{T}}u = 0, \\ \ell \leqslant x \leqslant w, \end{cases} \tag{18}$$

где $A = \begin{pmatrix} E \\ -E \end{pmatrix}$, $b = \begin{pmatrix} \ell \\ -w \end{pmatrix}$.

Генерация задачи

1) Фиксируем размерность задачи $n \in \mathbb{N}$. Генерируем симметричную положительно определенную $(n \times n)$-матрицу D.

2) Генерируем два вектора ℓ и w из R^n такие, что $\ell < w$.

3) Фиксируем оптимальное решение x_* такое, что $\ell \leqslant x_* \leqslant w$.

4) Вычисляем $v = Ax_* - b = \begin{pmatrix} x_* - \ell \\ w - x_* \end{pmatrix} \geqslant \mathbf{0}$.

5) Генерируем неотрицательный вектор $u \in R^{2n}$. Корректируем компоненты вектора u по правилу:

$$\begin{cases} u[j] = 0, & \text{если } v[j] > 0, \\ u[j], & \text{если } v[j] = 0. \end{cases}$$

6) Вычисляем вектор $c = A^{\mathsf{T}}u - Dx_* = u_1 - u_2 - Dx_*$, где $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$.

Непосредственной проверкой можно убедиться, что построенные по этой схеме векторы x_*, u и v удовлетворяют условиям Куна – Таккера (18).