EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. XVIII KÖIDE
FÜÜSIKA * MATEMAATIKA. 1969, Nr. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ XVIII
ФИЗИКА * МАТЕМАТИКА. 1969, № 4

https://doi.org/10.3176/phys.math.1969.4.03

И. МАУЭР

ДВА МЕТОДА РЕШЕНИЯ ЗАДАЧИ СТРОГО ВЫПУКЛОГО КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ

Существует много методов решения задачи строго выпуклого квадратичного программирования. Некоторые из них приводятся в [1]. В настоящей статье рассматриваются два новых метода. Первый является модификацией метода [2] на случай задачи с ограничениями в виде неравенства. Оригинальной частью метода является способ определения длины шага на каждой итерации. Второй метод разработан на основе ряда новых идей. Практические результаты вычисления показывают, что этот метод по сравнению с первым имеет более быструю сходимость.

1. Постановка задачи. Рассмотрим задачу строго выпуклого квадратичного программирования: найти среди решений системы неравенств

$$Ax-b\leqslant 0$$

решение $\tilde{x}$, которое минимизирует функцию

$$p^T x+\frac{1}{2}x^T Cx.$$

Здесь x и p — n-мерные векторы; b — m-мерный вектор; A — матрица размерности $m\times n$; C — положительно определенная симметричная матрица размерности $n\times n$.

Введем m-мерный вектор u и составим для указанной задачи функцию Лагранжа:

$$\varphi(x,u)=p^T x+\frac{1}{2}x^T Cx+u^T(Ax-b).$$

Как известно [3], решение исходной задачи можно заменить решением задачи, двойственной ей, заключающейся в отыскании такой пары векторов $\tilde{x},\tilde{u}$, для которой выполняется условие

$$\varphi(\tilde{x},\tilde{u})=\max_{u\geqslant 0}\min_{x}\varphi(x,u). \tag{1}$$

Нетрудно показать, что в (1) функция

$$\mu(u)=\min_{x}\varphi(x,u)$$

является строго вогнутой.

Если предположить, что существует конечный $\tilde{x}$, то, согласно теореме Куна—Таккера [3], существует и конечный вектор $\tilde{u}\geqslant 0$, который и определен однозначно ввиду строгой вогнутости $\mu(u)$.

Пусть $\tilde{x}$ — конечный.

2. **Метод I.** Вместо исходной задачи решаем задачу (1), т. е. максимизируем функцию $\mu(u)$ при условии $u \geqslant 0$.

Пусть

$$\varphi(x(u), u) = \mu(u).$$

Если на $k-1$-й итерации мы получим $u^k \geqslant 0$, то k-я итерация предлагаемого метода определяется равенством

$$u^{k+1} = u^k + \varrho_k s^k,$$

где

$$s^k = \max\{0, u^k + R(Ax(u^k) - b)\} - u^k,$$

а скаляр ϱ_k удовлетворяет соотношению

$$\varphi(x(u^{k+1}), u^{k+1}) = \max_{\substack{\rho \\ u^k + \rho s^k \geqslant 0}} \mu(u^k + \varrho s^k).$$

Здесь исходный вектор $u^1 \geqslant 0$ выбираем произвольно, а подходящую величину числа $R > 0$ устанавливаем практически.

При применении метода образуется последовательность $\{x(u^k)\}$, которая при $k \to \infty$ сходится к решению исходной задачи [4].

Чтобы применить изложенный метод, необходимо знать способ определения ϱ_k. С этой целью приводим формулы. Пусть $u^k \neq \tilde{u}$.

Сначала определяем скаляр $\bar{\varrho}_k$:

$$\varphi(x(u^k + \bar{\varrho}_k s^k),\ u^k + \bar{\varrho}_k s^k) = \max_{\rho} \mu(u^k + \varrho s^k).$$

Согласно известным теоремам существует конечный $\bar{\varrho}_k$, который является единственным, так как $\mu(u^k + \varrho s^k)$ — строго вогнутая функция и множество $Q = \{\varrho \mid \mu(u^k + \varrho s^k) \geqslant \mu(u^k)\}$ — ограниченное, замкнутое и выпуклое. Поскольку $\varphi(x, u^k)$ — выпуклая функция, то по теореме Куна—Таккера имеем соотношение

$$\min_{\substack{x \\ s^{kT}(Ax-b)=0}} \varphi(x, u^k) = \max_{\rho} \mu(u^k + \varrho s^k) = \varphi(x(u^k + \bar{\varrho}_k s^k), u^k + \bar{\varrho}_k s^k)$$

и, следовательно,

$$s^{kT}(Ax(u^k + \bar{\varrho}_k s^k) - b) = 0.$$

Используя формулу

$$x(u^k) = -C^{-1}(A^T u^k + p), \tag{2}$$

получим теперь, что

$$s^{kT}(-AC^{-1}(A^T u^k + \bar{\varrho}_k A^T s^k + p) - b) = 0, \tag{3}$$

откуда

$$\bar{\varrho}_k = \frac{s^{kT}(-AC^{-1}(A^T u^k + p) - b)}{s^{kT} AC^{-1} A^T s^k}. \tag{4}$$

Заметим, что (4) содержит правую часть (2). Поэтому

$$\bar{\varrho}_k = \frac{s^{kT}(Ax(u^k) - b)}{s^{kT} AC^{-1} A^T s^k}.$$

Отметим, что $s^{kT} AC^{-1} A^T s^k \neq 0$. Это вытекает из (3), так как $\bar{\varrho}_k$ — конечное и единственное.

Скаляр ϱ_k определим следующим образом:

$$\varrho_k = \begin{cases} \bar{\varrho}_k, & \text{если} \quad u^k + \bar{\varrho}_k s^k \geqslant 0, \\ \max(\varrho \mid u^k + \varrho s^k \geqslant 0) & \text{в остальных случаях.} \end{cases}$$

3. Метод II. Вместо исходной задачи решаем также задачу (1), т. е. максимизируем функцию $\mu(u)$ при условии $u \geqslant 0$, но для ускорения сходимости метода 1 реализуем в нем некоторые новые идеи.

Обозначим

$$\tilde{\mu}(u) = \min_{u^T(Ax-b)=0} \left(p^T x + \frac{1}{2} x^T C x \right).$$

Если на k — 1-й итерации мы получим $u^k \geqslant 0$, то k-я итерация предлагаемого метода определяется равенством

$$u^{k+1} = \tilde{\varrho}_k (u^k + \varrho_k s^k),$$

где

$$s^k = \max\{0, u^k + R(Ax(u^k) - b)\} - u^k,$$

а числа ϱ_k и $\tilde{\varrho}_k$ удовлетворяют соотношениям

$$\tilde{\mu}(u^k + \varrho_k s^k) = \sup_{\substack{\varrho \\ u^k + \varrho s^k \geqslant 0}} \tilde{\mu}(u^k + \varrho s^k)$$

и

$$\varphi(x(u^{k+1}), u^{k+1}) = \tilde{\mu}(u^k + \varrho_k s^k),$$

если ϱ_k — конечное. При $\varrho_k = \infty$ в определении

$$u^{k+1} = \tilde{\varrho}_k s^k,$$

где $\tilde{\varrho}_k$ вытекает из соотношения

$$\varphi(x(\tilde{\varrho}_k s^k), \tilde{\varrho}_k s^k) = \tilde{\mu}(s^k).$$

Для этого метода исходный вектор $u^1 = 0$ и R — выбранное подходящее положительное число.

При применении метода образуется последовательность $\{x(u^k)\}$, которая при $k \to \infty$ сходится к решению исходной задачи. Последнее доказывается так же, как и сходимость метода I [4]. Следует лишь учесть, что $\tilde{\mu}(u) \geqslant \mu(u)$.

Приступим к определению ϱ_k и $\tilde{\varrho}_k$. Для этого выводим аналитические формулы функций $\tilde{\mu}(u^k + \varrho s^k)$ и $\partial\tilde{\mu}(u^k + \varrho s^k)/\partial\varrho$. Пусть $u^k \neq \tilde{u}$.

По теореме Куна—Таккера для любого числа ϱ существует число $\varrho_k(\varrho)$ такое, что

$$\tilde{\mu}(u^k + \varrho s^k) = \mu(\varrho_k(\varrho)(u^k + \varrho s^k)),$$

и, следовательно,

$$(u^k + \varrho s^k)^T (Ax(\varrho_k(\varrho)(u^k + \varrho s^k)) - b) = 0. \tag{5}$$

Если использовать (2), то из (5) в случае $u^k + \varrho s^k \neq 0$ вытекает, что

$$\varrho_k(\varrho) = \frac{(u^k + \varrho s^k)^T (AC^{-1}p + b)}{-(u^k + \varrho s^k)^T AC^{-1}A^T (u^k + \varrho s^k)}. \tag{6}$$

Знаменатель в (6) $-(u^k+\varrho s^k)^T AC^{-1}A^T(u^k+\varrho s^k)\neq 0$. При фиксированном $\varrho=\varrho'$ это получается из (5), так как $\varrho_k(\varrho')$ является конечным и единственным ($\mu(r(u^k+\varrho' s^k))$ — строго вогнутая функция, $Q'=\{r\mid \mu(r(u^k+\varrho' s^k))\geqslant \mu(u^k+\varrho' s^k)\}$ — ограниченное, замкнутое и выпуклое множество (r — переменная)).

Отметим, что если в методе выбрать $u^1=0$, то u^k и $s^k (k>1)$ будут только такими, что при изменении ϱ $u^k+\varrho s^k\neq 0$. Действительно, так как в методе $u^k (k>1)$ является вектором, для которого $\mu(u^k)\geqslant\mu(\varrho u^k)$, то при изменении ϱ $u^k+\varrho s^k$ уже не может равняться нулю, так как найдется $\widehat{\varrho}_k$ такое, что $\mu(u^k+\widehat{\varrho}_k s^k)>\mu(u^k)$.

Пусть $k>1$.

Используя (2), (6) и обозначения

$$\begin{aligned}
A_k &= u^{kT}AC^{-1}A^Tu^k,\\
B_k &= u^{kT}AC^{-1}A^Ts^k,\\
C_k &= s^{kT}AC^{-1}A^Ts^k,\\
K_k &= u^{kT}(AC^{-1}p+b),\\
L_k &= s^{kT}(AC^{-1}p+b),
\end{aligned}$$

выведем формулу для $\tilde{\mu}(u^k+\varrho s^k)$:

$$\tilde{\mu}(u^k+\varrho s^k)=\frac{1}{2}\left(\frac{(K_k+\varrho L_k)^2}{A_k+2\varrho B_k+\varrho^2 C_k}-p^TC^{-1}p\right).$$

По (6) и (2) $\varrho_k(\varrho)$ и $x(\varrho_k(\varrho)(u^k+\varrho s^k))$, следовательно и $\tilde{\mu}(u^k+\varrho s^k)$, являются непрерывными. Поэтому $A_k+2\varrho B_k+\varrho^2C_k\neq 0$. Оказывается, что $\tilde{\mu}(u^k+\varrho s^k)$ является и непрерывно дифференцируемой функцией. Действительно,

$$\frac{\partial\tilde{\mu}(u^k+\varrho s^k)}{\partial\varrho}=\frac{L_k(B_kL_k-C_kK_k)\varrho^2+(A_kL_k^2-C_kK_k^2)\varrho+K_k(A_kL_k-B_kK_k)}{(A_k+2\varrho B_k+\varrho^2C_k)^2}.$$

Для определения ϱ_k проведем анализ стационарных точек $\tilde{\mu}(u^k+\varrho s^k)$, которыми являются нули функции $\partial\tilde{\mu}(u^k+\varrho s^k)/\partial\varrho$, т. е. решения уравнения

$$a_k\varrho^2+b_k\varrho+c_k=0, \tag{7}$$

где

$$\begin{aligned}
a_k &= L_k(B_kL_k-C_kK_k),\\
b_k &= A_kL_k^2-C_kK_k^2,\\
c_k &= K_k(A_kL_k-B_kK_k),
\end{aligned}$$

и найдем $\bar{\varrho}_k$:

$$\tilde{\mu}(u^k+\bar{\varrho}_k s^k)=\sup_{\varrho}\tilde{\mu}(u^k+\varrho s^k).$$

При этом используем следующие свойства $\tilde{\mu}(u^k+\varrho s^k)$:

а) $\tilde{\mu}(u^k+\varrho s^k)$ является непрерывно дифференцируемой функцией;

б) $\lim\limits_{\varrho\to+\infty}\tilde{\mu}(u^k+\varrho s^k)=\lim\limits_{\varrho\to-\infty}\tilde{\mu}(u^k+\varrho s^k)$;

в) существует $\hat{\varrho}_k$ такое, что при возрастании ϱ от 0 до $\hat{\varrho}_k$ возрастает $\tilde{\mu}(u^k+\varrho s^k)$;

г) в стационарной точке $\tilde{\mu}(u^k + \varrho s^k)$ знак $\partial^2\tilde{\mu}(u^k + \varrho s^k)/\partial^2\varrho$ совпадает со знаком $2a_k\varrho + b_k$.

Рассмотрим отдельно случаи $k = 1$ и $k > 1$.

1. $k = 1$. Так как $u^1 = 0$, то $\tilde{\mu}(\varrho s^1) = \mathrm{const}$ для всех $\varrho \neq 0$. Следовательно, за $\bar{\varrho}_1$ можно принять любое отличное от нуля число.

2. $k > 1$. Если $a_k = 0$, то $b_k \neq 0$, поскольку в противном случае и $c_k = 0$, что возможно лишь в случае $k = 1$. Следовательно, если $a_k = 0$, то

$$\bar{\varrho}_k = \begin{cases} -\dfrac{c_k}{b_k}, & \text{если } b_k < 0, \\ \infty, & \text{если } b_k > 0. \end{cases}$$

Если $a_k \neq 0$, то решениями уравнения (7) являются ϱ_{k1} и ϱ_{k2}, не равные друг другу. Действительно, если $\varrho_{k1} = \varrho_{k2}$, то дискриминант уравнения равнялся бы нулю и $\varrho_{k1} = \varrho_{k2} = -\dfrac{b_k}{2a_k}$, откуда $2a_k\varrho_{k1} + b_k = 0$. Но это означает, что ϱ_{k1} является точкой поворота для $\tilde{\mu}(u^k + \varrho s^k)$, что невозможно (свойства а) — г)). Одновременно заметим, что $2a_k\varrho_{k1} + b_k \neq \neq 0$, так как в противном случае $\varrho_{k1} = \varrho_{k2}$. Также и $2a_k\varrho_{k2} + b_k \neq 0$. Следовательно, если $a_k \neq 0$, то

$$\bar{\varrho}_k = \begin{cases} \varrho_{k1}, & \text{если } 2a_k\varrho_{k1} + b_k < 0, \\ \varrho_{k2}, & \text{если } 2a_k\varrho_{k1} + b_k > 0. \end{cases}$$

Если $k = 1$, то за ϱ_1 можно принять любое отличное от нуля число.

Если $k > 1$, то при $u^k + \bar{\varrho}_k s^k$ хотя бы с одним отрицательным компонентом и $\bar{\varrho}_k < 0$ упрощаем метод:

$$\varrho_k = \begin{cases} \infty, & \text{если } s_k \geqslant 0, \\ \max(\varrho \mid u^k + \varrho s^k \geqslant 0), & \text{при остальных значениях } s_k. \end{cases}$$

хотя может оказаться, что $\tilde{\mu}(u^k - \varrho' s^k) > \tilde{\mu}(u^{k+1})$, где $\varrho' = \max(\varrho/u^k - \varrho s^k \geqslant 0)$ и u^{k+1} определен, как это предусмотрено методом в случае $\varrho_k = \infty$ или для ϱ — конечное при $s_k \geqslant 0$ или при s_k хотя бы с одним отрицательным компонентом соответственно. Введем это упрощение потому, что рассматриваемый случай представляется практически маловероятным.

Во всех остальных случаях

$$\varrho_k = \begin{cases} \bar{\varrho}_k, & \text{если } u^k + \bar{\varrho}_k s^k \geqslant 0, \\ \max(\varrho \mid u^k + \varrho s^k \geqslant 0), & \text{при остальных значениях } u^k + \bar{\varrho}_k s^k. \end{cases}$$

Если ϱ_k известно, то по (6) устанавливаем $\tilde{\varrho}_k$:

$$\tilde{\varrho}_k = \begin{cases} -\dfrac{K_k + \varrho_k L_k}{A_k + 2\varrho_k B_k + \varrho_k^2 C_k}, & \text{если } \varrho_k \text{ — конечное,} \\ -\dfrac{L_k}{C_k}, & \text{если } \varrho_k = \infty. \end{cases}$$

Сделаем некоторые замечания:

1. A_{k+1}, K_{k+1} выражаются через A_k, B_k, C_k, K_k, L_k $(A_1 = 0,\ K_1 = 0)$.

2. В вычислениях $\varrho_k = \infty$ можно заменить на $\varrho_k = T$ (T — выбранное большое положительное число) и всегда следовать предписанию метода для случая ϱ_k — конечное.

3. Метод в обработке для задачи строго выпуклого квадратичного программирования с ограничениями $Ax - b = 0$ во многом проще:

$$s^k = u^k + R(Ax - b),\ \varrho_k = \bar{\varrho}_k.$$

4. Проиллюстрируем изложенные здесь методы геометрически $(m = 2)$. Концентрические линии с центром $\tilde{u}$, вычерченные на рис. 1 и 2, являются линиями постоянного значения $\mu(u)$.

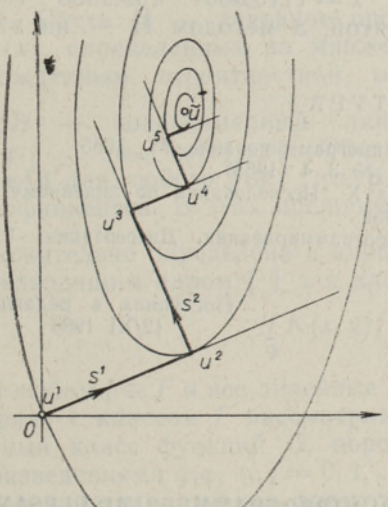

Рис. 1

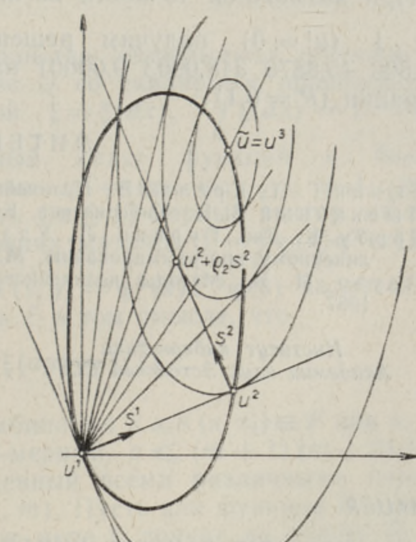

Рис. 2.

На рис. 1 изображены точки $u^1, \ldots, u^5$, полученные методом I.

На рис. 2 изображены точки u^1, u^2, $u^3 = \tilde{u}$, полученные методом II. Жирная линия на этом рисунке является геометрическим местом точек $\varrho_k(\varrho)(u^k + \varrho s^k)$. Так как $\tilde{\mu}(u^k + \varrho s^k) = \mu(\varrho_k(\varrho)(u^k + \varrho s^k))$, то при изменении ϱ на рисунке можно хорошо проследить за изменением $\tilde{\mu}(u^2 + \varrho s^2)$. Так, $\tilde{\mu}(u^2 + \varrho s^2)$ принимает при $\varrho = \varrho_2$ свое максимальное значение, а при $\varrho \to \infty$ и $\varrho \to -\infty$ приближается к своему минимальному значению. В рассматриваемом случае при $k = 2$ уравнение (7) имеет одно решение.

Изложенными выше методами решена задача, для которой

$$p^T = (4\ -1\ -20\ 1\ -35\ 3\ 0),$$

$$C=\begin{pmatrix} 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 70 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 140 \end{pmatrix},$$

$$-A=\begin{pmatrix} 5 & 3 & 0 & 10 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 1 & 3 \\ 1 & 0 & 5 & 1 & 0 & 0 & 1 \\ 7 & 1 & 0 & 0 & 4 & 0 & 0 \end{pmatrix},$$

$$-b^T=(360 \quad 800 \quad 70 \quad 30 \quad 100).$$

При возможной точности вычисления на ЭЦВМ «Минск-22» методом I ($u^1=0$) получим решение $\tilde{x}=(115{,}8654 \quad 139{,}9385 \quad 10{,}2741 \quad -0{,}500 \quad 45{,}8870 \quad 348{,}0963 \quad 0{,}0000)$ на пятой, а методом II — на второй итерации ($R=0{,}1$).

ЛИТЕРАТУРА

1. Кюнци Г. П., Крелле В., Нелинейное программирование, М., 1965.
2. Пшеничный Б. Н., Кибернетика, Киев, № 3, 4 (1965).
3. Эрроу К. Дж., Гурвиц Л., Удзава Х., Исследования по линейному и нелинейному программированию, М., 1962.
4. Мауэр И. В., Методы нелинейного программирования, Диссертация, Таллин, 1967.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
12/XI 1968

I. MAUER

KAKS MEETODIT RANGELT KUMERA RUUTPROGRAMMEERIMISÜLESANDE LAHENDAMISEKS

Esitatakse kaks uut meetodit rangelt kumera ruutprogrammeerimisülesande lahendamiseks. Esimene neist on tuntud meetodi [2] töötlusjuhuks, kui lähteülesande kõrvaltingimusteks on võrratused. Teine meetod on üles ehitatud reale uutele ideedele. Praktiliste arvutustulemuste põhjal võib öelda, et esimesega võrreldes tagab teine meetod suurema koonduvuskiiruse.

I. MAUER

TWO METHODS FOR SOLVING A PROBLEM OF STRICTLY CONVEX PROGRAMMING

In this paper, two new methods for solving a problem of strictly convex quadratic programming are given. The first of them is a modification of the method [2] for the case of restrictions expressed in the terms of inequalities. The second method includes some new ideas.

The computing results show the greater convergence rate of the second method compared with the first one.