EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. XV KÖIDE
FÜÜSIKA-MATEMAATIKA- JA TEHNIKATEADUSTE SEERIA. 1966, NR. 3

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ XV
СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИХ И ТЕХНИЧЕСКИХ НАУК. 1966, № 3

И. МАУЭР

ПРИБЛИЖЕННЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Задача выпуклого программирования заключается в минимизации выпуклого функционала $f(x)$ при выпуклых ограничениях $g_i(x) \leqslant 0$ $(i=1, \ldots, m)$, где $x \in E^n$. В последние годы разработан целый ряд методов для решения такой задачи. По характеру и времени возникновения их можно группировать на методы лагранжевые, прямые, погружные. В последней группе известны методы [1–4]. Предлагаемый в данной работе метод является видоизменением хорошо известного метода погружения [1].

Рассмотрим задачу выпуклого программирования с многими ограничениями: минимизировать

$$f(x) \tag{1}$$

при ограничениях

$$g_i(x) \leqslant 0 \qquad (i=1, \ldots, m). \tag{2}$$

Предположим, что $f(x)$, $g_i(x)$ $(i=1, \ldots, m)$ — два раза непрерывно дифференцируемые функционалы и что область D, определенная ограничениями задачи, непустая и ограниченная.

Приближенное решение задачи (1), (2) сводим к решению ряда задач $r=1, \ldots, r'$ с одним ограничением: минимизировать

$$f(x) \tag{3}$$

при ограничении

$$\frac{1}{2} \sum_{i=1}^{m} S(g_i(x))\,(g_i(x))^2 - \Delta_r \leqslant 0, \tag{4}$$

где

$$S(t) = \begin{cases} 1, & \text{если } t > 0, \\ 0, & \text{если } t \leqslant 0 \end{cases}$$

и Δ_r — положительное число. Эти задачи отличаются только по значению Δ_r. Как легко доказывается, ограничение (4) представляет собой выпуклый и непрерывно дифференцируемый функционал. А область D_{Δ_r}, определенная ограничением (4), содержит очевидно собственно внутреннюю точку.

Числа $\Delta_r (r=1, \ldots, r')$ выбираем убывающими $\Delta_1 > \Delta_2 > \ldots > \Delta_{r'}$.

Обозначения

Используем в дальнейшем следующие обозначения:

$$G(x) = \frac{1}{2} \sum_{i=1}^{m} S(g_i(x))(g_i(x))^2;$$

$\vec{f}_x$ — градиент функционала $f(x)$ в точке x;

$\bar{x}$ — решение задачи (1), (2);

M — множество всех решений задачи (1), (2);

$\bar{x}_{\Delta_r}$ — решение задачи (3), (4);

$x^k_{\Delta_r}$ — результат, полученный на k-й итерации при решении задачи (3), (4);

x^* — точка, минимизирующая функционал $f(x)$;

x^*_u — точка, минимизирующая функционал $f(x) + uG(x)$ при фиксированном u.

О решении задачи (3), (4)

Решаем задачу (3), (4) с определенным значением Δ.

За исходную точку решения выбираем какую-нибудь точку $x^0_\Delta \in D_\Delta$ и после k-й итерации получим x^k_Δ. Последующую точку x^{k+1}_Δ вычисляем по формуле

$$x^{k+1}_\Delta = x^k_\Delta + \lambda_k \vec{S}^k,$$

где направление движения $\vec{S}^k$ и скаляр λ_k определяем таким образом, что в результате вычислений получим последовательность $\{x^k_\Delta\}$ такую, что

$$\lim_{k \to \infty} x^k_\Delta = \bar{x}_\Delta.$$

Есть разные возможности определения вектора $\vec{S}^k$ и скаляра λ_k. В данном случае считаем целесообразным воспользоваться такими, которые задают движение в пределах области D_Δ и в общем различно определяют направление в краевых точках области, — методами возможных направлений [5–8]. В близких друг от друга краевых точках области D_Δ градиент функционала $G(x)$ может резко отличаться по направлениям и вызывать «зигзагообразное» движение при применении отмеченных методов. Поэтому введем число $\Theta > 0$, известное из работы Г. Зойтендейка [7], и приведем способы определения направления по методу [7] и по методам [5, 6], несколько модифицировав последние.

а) М. Хансон [5]:

$$\vec{S}^k = \begin{cases} -\vec{f}_{x^k_\Delta}, & \text{если } G(x^k_\Delta) < \Delta, \\[2ex] -\vec{f}_{x^k_\Delta} - \Theta \dfrac{|\vec{f}_{x^k_\Delta}|}{|\vec{G}_{x^k_\Delta}|} \vec{G}_{x^k_\Delta}, & \text{если } G(x^k_\Delta) = \Delta. \end{cases}$$

Пользуясь только непрерывной дифференцируемостью функционалов $f(x)$ и $G(x)$, можно показать, что если

$$-\frac{(\vec{f}_{x^k_\Delta},\ \vec{G}_{x^k_\Delta})}{(\vec{G}_{x^k_\Delta},\ \vec{G}_{x^k_\Delta})} < \Theta\,\frac{|\vec{f}_{x^k_\Delta}|}{|\vec{G}_{x^k_\Delta}|} < -\frac{(\vec{f}_{x^k_\Delta},\ \vec{f}_{x^k_\Delta})}{(\vec{f}_{x^k_\Delta},\ \vec{G}_{x^k_\Delta})},$$

то направление $\vec{S}^k$ — подходящее возможное направление.

б) Т. Петржиковский [6]:

$$\vec{S}^k = \begin{cases} -\vec{f}_{x^k_\Delta}, & \text{если } G(x^k_\Delta) < \Delta, \\ u_1\vec{f}_{x^k_\Delta} + u_2\vec{G}_{x^k_\Delta}, & \text{если det. системы (5)} \neq 0 \\ & \text{и } G(x^k_\Delta) = \Delta, \\ 0, & \text{если det. системы (5)} = 0, \end{cases}$$

где u_1 и u_2 определяются как решение системы линейных уравнений

$$\begin{aligned} u_1(\vec{f}_{x^k_\Delta},\ \vec{f}_{x^k_\Delta}) + u_2(\vec{G}_{x^k_\Delta},\ \vec{f}_{x^k_\Delta}) &= -1, \\ u_1(\vec{f}_{x^k_\Delta},\ \vec{G}_{x^k_\Delta}) + u_2(\vec{G}_{x^k_\Delta},\ \vec{G}_{x^k_\Delta}) &= -\Theta. \end{aligned} \tag{5}$$

в) Г. Зойтендейк [7]:

$$\vec{S}^k = \begin{cases} -\vec{f}_{x^k_\Delta}, & \text{если } G(x^k_\Delta) < \Delta, \\ y^k, & \text{если } G(x^k_\Delta) = \Delta, \end{cases}$$

где y^k определяется как решение следующей задачи программирования: максимизировать σ при ограничениях

$$\begin{aligned} \vec{f}_{x^k_\Delta}\, y^k &\leqslant -\sigma \\ \vec{G}_{x^k_\Delta}\, y^k &\leqslant -\Theta\sigma, \\ |y^k| &\leqslant C \ \text{(некоторая константа)}. \end{aligned}$$

Легко сообразить, что в зависимости от значения Θ определяемое направление образует больший или меньший угол с касательной гиперплоскостью функционала $G(x)$ в точке x^k_Δ, и в зависимости от задачи целесообразно в вычислениях выбрать Θ то больше, то меньше единицы.

Выберем достаточно малое положительное число $\varepsilon > 0$, при помощи которого обнаруживаем движение с малым шагом. Если $\lambda_k |\vec{S}^k| > \varepsilon$

($\vec{S}^k$ определено $\Theta = 1$), то продолжаем вычисления с $\Theta = 1$. В противном случае наряду с x_Δ^{k+1} $(\Theta = 1)$ (x_Δ^{k+1} определено $\Theta = 1$) вычисляем еще x_Δ^{k+1} $(\Theta > 1)$ и, если $f(x_\Delta^{k+1}\ (\Theta = 1)) > f(x_\Delta^{k+1}\ (\Theta > 1))$, то $x_\Delta^{k+1} = x_\Delta^{k+1}$ $(\Theta > 1)$. В противном случае вычисляем и x_Δ^{k+1} $(\Theta < 1)$ и за x_Δ^{k+1} принимаем ту из точек x_Δ^{k+1} $(\Theta = 1)$ и x_Δ^{k+1} $(\Theta < 1)$, в которой значение $f(x)$ меньше.

Естественно, чем ближе мы окажемся к точке $\bar{x}_\Delta$, тем меньшим нужно выбирать число $\varepsilon > 0$. Приближение к точке $\bar{x}_\Delta$ обнаруживает число $(-\vec{f}_{x_\Delta^k}, \vec{G}_{x_\Delta^k})/(|\vec{f}_{x_\Delta^k}||\vec{G}_{x_\Delta^k}|)$, которое при $k \to \infty$ стремится к единице.

При всех этих методах λ_k вычисляется одинаковым образом: определяется $\lambda = \lambda_k$, который минимизирует

$$f(x_\Delta^k + \lambda \vec{S}^k)$$

при ограничении

$$G(x_\Delta^k + \lambda \vec{S}^k) - \Delta \leqslant 0.$$

Нетрудно проверить, что для методов **а)**, **б)** $\vec{S}^k = 0$, а для метода **в)** $\vec{x}_\Delta^k = 0$ или $\max \sigma = 0$ являются необходимыми и достаточными условиями для того, чтобы $x_\Delta^k = \bar{x}_\Delta$.

Сходимость метода

Теорема

Соответственно любому числу $\varepsilon > 0$ найдется такое число $\Delta_0 > 0$, что если $0 < \Delta < \Delta_0$, то $\varrho(\bar{x}_\Delta, M) < \varepsilon$.

Доказательство. Если $x^* \in D$, то $x^* = \bar{x} = \bar{x}_\Delta$ и при всех $0 < \Delta < \infty$ $\varrho(\bar{x}_\Delta, M) = 0$

Предположим, что $x^* \notin D$.

Воспользуемся теоремой, которая очевидно равносильна известной теореме Т. Петржиковского [1]: соответственно выбранному числу $\varepsilon > 0$ найдется такое число $u_0 > 0$, что если $u_0 < u < \infty$, то $\varrho(x_u^*, M) < \varepsilon$.

Любому $\bar{u} > 0$ и соответствующему $x_{\bar{u}}^*$ можно сопоставить $\Delta > 0$ и соответствующее $\bar{x}_\Delta$ таким образом, что $x_{\bar{u}}^* = \bar{x}_\Delta$ и $\Delta = G(x_{\bar{u}}^*)$. Действительно, точка $x_{\bar{u}}^*$, минимизирующая функционал $f(x) + \bar{u}G(x)$, минимизирует и функционал $f(x) + \bar{u}(G(x) - G(x_{\bar{u}}^*))$, а, следовательно, $(x_{\bar{u}}^*, \bar{u})$ является седловой точкой функционала Лагранжа $f(x) + u(G(x) - \Delta)$ задачи (3), (4) с $\Delta = G(x_{\bar{u}}^*)$. По теореме Куна—Таккера $x_{\bar{u}}^* = \bar{x}_\Delta$, если $\Delta = G(x_{\bar{u}}^*)$.

Таким образом, для любого $\Delta = G(x_u^*)$ с $u_0 < u < \infty$ $\varrho(\bar{x}_\Delta, M) < \varepsilon$.

Выберем $\Delta_0 = G(x_{u_0}^*)$ и покажем, что для задачи (3), (4) с любым значением $0 < \Delta < \Delta_0$ найдется такое $u > u_0$, что $\Delta = G(x_u^*)$.

Пусть функционал Лагранжа задачи (3), (4) с $\Delta < \Delta_0$ имеет седловую точку (x_u^*, u). Множитель $u > 0$, так как $x^* \notin D_\Delta$ (очевидно $D_\Delta \subset D_{\Delta_0}$ $(\Delta < \Delta_0)$, а $x^* \notin D_{\Delta_0}$ $(u_0 > 0)$). Следовательно, $\Delta = G(x_u^*)$.

Для установления соотношения $u > u_0$ проведем следующее рассуждение.

Так как $\vec{f}_{x_{u_0}^*} + u_0 \vec{G}_{x_{u_0}^*} = 0$ и $\vec{f}_{x_{u_0}^*} + u\, \vec{G}_{x_{u_0}^*} \neq 0$ $(u \neq u_0)$, то

$$f(x_u^*) + u\,(G(x_u^*) - G(x_{u_0}^*)) < f(x_{u_0}^*).$$

С другой стороны, по определению седловой точки

$$f(x_u^*) + u_0\,(G(x_u^*) - G(x_{u_0}^*)) \geqslant f(x_{u_0}^*).$$

Соединяем полученные неравенства:

$$\frac{f(x_{u_0}^*) - f(x_u^*)}{u_0} \leqslant \Delta - \Delta_0 < \frac{f(x_{u_0}^*) - f(x_u^*)}{u}.$$

Учитывая, что $u_0 > 0$, $u > 0$ и что было предположено $\Delta < \Delta_0$, получим соотношение $u > u_0$. Теорема доказана.

Следствие. Если $\lim\limits_{r \to \infty} \Delta_r = 0$, то $\lim\limits_{r \to \infty} \bar{x}_{\Delta_r} = \bar{x}$.

Действительно, по доказанной теореме соответственно выбранному числу $\varepsilon > 0$ найдется ограниченное множество N, содержащее множество M и все точки $\bar{x}_{\Delta_r}$, для которых $\Delta_r < \Delta_0$. Из-за компактности множества N из последовательности $\{\bar{x}_{\Delta_r}\}$ $(\Delta_r < \Delta_0)$ можно выделить сходящуюся подпоследовательность, которую для простоты отметим таким же образом $\{\bar{x}_{\Delta_r}\}$, $\lim\limits_{r \to \infty} \bar{x}_{\Delta_r} = \underline{x}$. Предположим, что $\varrho(\underline{x}, M) = \varepsilon^* > 0$ и выбираем $0 < \varepsilon < \varepsilon^*$. По теореме найдется такое Δ_0, что, если $0 < \Delta_r < \Delta_0$, то $\varrho(\bar{x}_{\Delta_r}, M) < \varepsilon$ и, следовательно, $\varrho(\underline{x}, M) \leqslant \varepsilon < \varepsilon^*$. Полученное противоречие доказывает, что $\varrho(\underline{x}, M) = 0$ и $\underline{x} = \bar{x}$.

Приведем числовой пример:

минимизировать

$$-2x_1 - x_2$$

при ограничениях

$$\begin{aligned} 5x_1^2 + 2x_2^2 + 6x_1x_2 - 4x_1 - 2x_2 - 3 &\leqslant 0 \\ 3x_1 + 2x_2 - 2 &\leqslant 0 \\ -x_1 - x_2 &\leqslant 0. \end{aligned}$$

Приближенное решение задачи найдем как решение следующей задачи: минимизировать

$$-2x_1 - x_2$$

при ограничении

$$S(5x_1^2 + 2x_2^2 + 6x_1x_2 - 4x_1 - 2x_2 - 3)\,(5x_1^2 + 2x_2^2 + 6x_1x_2 - 4x_1 - 2x_2 - 3)^2 +$$
$$+ S(3x_1 + 2x_2 - 2)\,(3x_1 + 2x_2 - 2)^2 + S(-x_1 - x_2)\,(-x_1 - x_2)^2 - \Delta \leqslant 0.$$

Решения задачи $\bar{x}_\Delta = (\bar{x}_{1\Delta}, \bar{x}_{2\Delta})$ при значениях $\Delta = \Delta_r (r = 1, \ldots)$ приведены в нижеследующей таблице.

Δ	$\bar{x}_{1\Delta}$	$\bar{x}_{2\Delta}$
$\Delta_1 = 10^{-2}$	$2 + 0{,}21213 \cdot 10^0$	$-2 - 028284 \cdot 10^0$
$\Delta_2 = 10^{-4}$	$2 + 0{,}21213 \cdot 10^{-1}$	$-2 - 028284 \cdot 10^{-1}$
. . . . .	. . . .	. . . .
$\Delta_r = 10^{-2r}$	$2 + 0{,}21213 \cdot 10^{-r+1}$	$-2 - 028284 \cdot 10^{-r+1}$
. . . . .	. . . .	. . . .
$\lim\limits_{r\to\infty} \Delta_r = 0$	$\lim\limits_{r\to\infty} \bar{x}_{1\Delta_r} = 2$	$\lim\limits_{r\to\infty} \bar{x}_{2\Delta_r} = -2$

Замечание 1. В практических вычислениях за исходную точку решения каждой новой задачи нужно выбирать точку приближенной проекции решения предыдущей задачи на область, определенную ограничением новой задачи.

Замечание 2. Так как не имеется оценки для $\delta(\bar{x}_\Delta, M)$ или $f(\bar{x}_\Delta) - f(\bar{x})$, то требование точности решения можно задать с помощью некоторого вектора $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_m) > 0$: $\bar{x}_\Delta$ считаем решением задачи с требуемой точностью, если

$$g_i(\bar{x}_\Delta) \leqslant \varepsilon_i \qquad (i = 1, \ldots, m),$$

для чего Δ должно быть меньше некоторого Δ_0.

ЛИТЕРАТУРА

1. Pietrzykowski T., Inf. Process 1962, Amsterdam, 1963, pp. 185—189.
2. Carroll Ch. W., Oper. Res., **9**, № 2, 169—185 (1961).
3. Fiacco A. V., Mc Cormick, Manag. Sci., **10**, № 2, 360—366 (1964).
4. Бусленко Н. П., Соколов Г. А., Экономика и матем. методы, **1**, вып. 1, 123—136 (1965).
5. Hanson M. A., Austral. J. Statist., **5**, № 13, 14—19 (1963).
6. Pietrzykowski T., Prace zakl. apar. mat. PAN, A13, 1—22 (1961).
7. Зойтендейк Г., Методы возможных направлений, М., 1963.
8. Зуховицкий С. И., Поляк Р. А., Примак М. Е., Докл. АН СССР, **153**, № 5, 991—994 (1963).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
30/XI 1965

I. MAUER

LIGIKAUDNE MEETOD KUMERPROGRAMMEERIMISÜLESANDE LAHENDAMISEKS

Paljude kitsendavate tingimustega kumerprogrammeerimisülesande ligikaudne lahendamine taandatakse mitme ühe kõrvaltingimusega programmeerimisülesande lahendamisele.

I. MAUER

AN APPROXIMATE METHOD FOR SOLVING CONVEX PROGRAMMING PROBLEMS

The approximate solving of problems of convex programming with many constraints is reduced to solving a number of similar problems with a single one.