Доклады Академии наук СССР
1975. Том 222, № 2

УДК 518.9 *МАТЕМАТИКА*

С. И. ЗУХОВИЦКИЙ, М. Е. ПРИМАК

О СХОДИМОСТИ МЕТОДА ЧЕБЫШЕВСКИХ ЦЕНТРОВ И МЕТОДА ЦЕНТРИРОВАННЫХ СЕЧЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

(Представлено академиком Л. В. Канторовичем 25 XI 1974)

1. Для решения задачи выпуклого программирования, заключающейся в отыскании минимума (максимума) выпуклой (вогнутой) функции на многограннике, образованном системой линейных неравенств, в ([1]) предложен специальный метод секущей плоскости, названный там методом центрированных сечений *. Хотя с вычислительной точки зрения этот метод трудно реализуем, он привлекает геометрической наглядностью и отсутствием требования дифференцируемости минимизируемой функции. Однако в ([1]) не доказана его сходимость. С другой стороны, в ([2]) для решения задачи выпуклого программирования в связи с задачей отыскания точки равновесия выпуклой игры многих лиц попутно указан (при некоторых ограничениях на минимизируемую функцию) близкий к методу ([1]) метод чебышевских центров **, вычислительно просто реализуемый.

Ниже приводится в виде теоремы 1 специальная форма достаточного условия сходимости последовательности приближений к решению задачи выпуклого программирования, на основании которого исследуется сходимость каждого из этих методов и указывается прием выбрасывания лишних ограничений в задачах линейного программирования, решаемых на каждом шаге в методе чебышевских центров.

2. Достаточное условие сходимости. Пусть в E^n задана вогнутая функция $\varphi(x)=\varphi(\xi_1, \ldots, \xi_n)$ и выпуклый многогранник $\Omega_0=\{x\in E^n\colon (a_i, x)\geqslant\alpha_i,\ i=1, \ldots, m\}$ с непустой внутренностью $(\operatorname{int}\Omega_0\neq\varnothing)$. Рассматривается задача отыскания точки $x^*\in\Omega_0$, для которой

$$\max\{\varphi(x)\colon x\in\Omega_0\}=\varphi(x^*). \tag{1}$$

В силу вогнутости функции $\varphi(x)$, в E^n определена вектор-функция $\partial\varphi(x)$, являющаяся субградиентом функции $\varphi(x)$, т. е. такая, что

$$\varphi(y)-\varphi(x)\leqslant(\partial\varphi(x), y-x)\quad \forall y\in E^n. \tag{2}$$

Теорема 1. *Пусть задана произвольная последовательность $\{x_k\}\subset\Omega_0$ и пусть при ее помощи продолжены конечные последовательности $a_1, \ldots, a_m$ и $\alpha_1, \ldots, \alpha_m$ следующим образом:*

$$a_{m+k}=\partial\varphi(x_k),\quad \alpha_{m+k}=(\partial\varphi(x_k), x_k),\quad k=1, 2, \ldots$$

Тогда, если последовательность чебышевских радиусов $\{\rho_k\}$ систем линейных неравенств

$$(a_1, x)-\alpha_1\geqslant0, \ldots, (a_{m+k}, x)-\alpha_{m+k}\geqslant0,$$

стремится к нулю, т. е.

$$\max_{x\in E^n}\ \min_{1\leqslant i\leqslant m+k}[(a_i, x)-\alpha_i]=\rho_k\to0.$$

* См. ниже п. 5.
** См. ниже п. 3.

то существует подпоследовательность $\{x_{i_k}\}$ *последовательности* $\{x_k\}$ *такая, что* $\varphi(x_{i_{k-1}})\leqslant\varphi(x_{i_k})$ *и* $\varphi(x_i)<\varphi(x_{i_{k-1}})$, $i=i_{k-1}+1, \ldots, i_k-1$, *и каждая предельная точка этой подпоследовательности является решением задачи* (1).

В доказательстве используется тот простой факт, из $\rho_k\to 0$ следует также

$$\max_{x\in\Omega_0}\ \min_{m+1\leqslant i\leqslant m+k}\ [(a_i, x)-\alpha_i]\to 0. \tag{3}$$

Отсюда с учетом (2) устанавливается, что некоторые предельные точки последовательности $\{x_k\}$ являются решениями задачи (1), из чего потом непосредственно следует существование подпоследовательности $\{x_{i_k}\}$ такой, что $\varphi(x_{i_{k-1}})\leqslant\varphi(x_{i_k})$, и сходимость каждой ее сходящейся подпоследовательности к решению задачи (1).

Ниже приводятся два способа построения последовательности $\{x_k\}$, удовлетворяющей условиям этой теоремы.

3. Метод чебышевских центров. Пусть y_1 — произвольный вектор из Ω_0. Полагаем $a_{m+1}=\partial\varphi(y_1)$, $\alpha_{m+1}=(\partial\varphi(y_1), y_1)$. В качестве второго элемента последовательности $\{y_k\}$ берем вектор y_2, являющийся решением задачи

$$\max_{x\in E^n}\ \min_{1\leqslant i\leqslant m+1}\ [(a_i, x)-\alpha_i]=\mu_1\geqslant 0,$$

т. е. являющийся чебышевским центром системы линейных неравенств

$$(a_i, x)-\alpha_i\geqslant 0, \quad i=1, \ldots, m+1.$$

Пусть уже построены векторы $y_1, \ldots, y_k$. Находим

$$a_{m+k}=\partial\varphi(y_k), \quad \alpha_{m+k}=(\partial\varphi(y_k), y_k)$$

и в качестве $(k+1)$-го элемента искомой последовательности берем вектор y_{k+1}, являющийся чебышевским центром системы $(a_i, x)-\alpha_i\geqslant 0$, $i=1, \ldots, m+k$, т. е. являющийся решением задачи

$$\max_{x\in E^n}\ \min_{1\leqslant i\leqslant m+k}\ [(a_i, x)-\alpha_i]=\mu_k, \tag{4}$$

которая сводится к следующей задаче линейного программирования:

$$\max\{\xi_{n+1}\colon \xi_{n+1}\leqslant(a_i, x)-\alpha_i,\ i=1, \ldots, m+k\}. \tag{4'}$$

Нетрудно убедиться в том, что в предположении ограниченности $\|\partial\varphi(x)\|$ на Ω_0 имеем $\lim\mu_k=0$, так что из теоремы 1 следует, что каждая предельная точка подпоследовательности $\{y_{i_k}\}$, для которой $\varphi(y_{i_{k-1}})\leqslant\varphi(y_{i_k})$, является решением задачи 1.

Отметим, что в отличие от ([2]) в этой заметке метод чебышевских центров обосновывается без требования гладкости функции $\varphi(x)$ и без требования ее равномерной вогнутости.

4. Освобождение от лишних ограничений и оценка скорости сходимости. Основным недостатком приведенного метода чебышевских центров является присоединение на каждом шаге дополнительного ограничения $(a_{k+1}, x)-\alpha_{k+1}\geqslant 0$. В связи с этим приведем один прием, который позволяет не накапливать неограниченно эти ограничения при определении μ_{k+1}.

Пусть $\{\delta_k\}$ — монотонно убывающая последовательность положительных чисел, сходящаяся к нулю, и в (4) для некоторого j_1 получили $\mu_{j_1}\leqslant\delta_1$. Тогда множество индексов $\{1, \ldots, m+j_1\}$ «очищается» следующим образом. Находим подмножество индексов $I_{j_1}\subset\{1, \ldots, m+j_1\}$ такое, что

$$\max_{x\in\Omega_0}\ \min_{i\in I_{j_1}}\ [(a_i, x)-\alpha_i]=\mu_{j_1},$$

$$\max_{x\in\Omega_0}\ \min_{i\in I_{j_1}\setminus\{j\}}\ [(a_i, x)-\alpha_i]>\mu_{j_1} \quad \forall j\in I_{j_1}.$$

Другими словами, в очищенное множество I_{j_1} включаются номера тех линейных функций, которые при решении задачи (4) образуют оптимальную вершину в соответствующей задаче (4′) линейного программирования.

На следующем шаге множество I_{j_1} расширяем, т. е. рассматриваем множество $I_{j_1} \cup \{m+j_1+1\}$ и находим

$$\max_{x \in \Omega_0} \min_{i \in I_{j_1} \cup \{m+j_1+1\}} [(a_i, x) - \alpha_i] = \mu_{j_1+1}.$$

Процесс шага расширения множеств продолжаем до получения на некотором j_2 неравенства $\mu_{j_2} \leqslant \delta_2$. В этом случае новое множество индексов снова очищается, т. е. строится множество I_{j_2} и вычислительный процесс продолжается. Ясно, что при $\{\delta_k\}$, быстро сходящейся к нулю, очистка множеств будет осуществляться реже, чем при $\{\delta_k\}$, медленно сходящейся к нулю.

Выбором управляющей последовательности $\{\delta_k\}$ можно не только регулировать процесс очистки, но и скорость сходимости к нулю выделяемой подпоследовательности $\{\mu_{j_k}\}$.

Сравнительно просто получаемая оценка $\mu_k = o(k^{-1/n})$ является довольно грубой.

Точнее оценить скорость сходимости μ_k к нулю не удается. Однако можно оценить скорость сходимости некоторой модификации метода чебышевских центров, более трудоемкой с вычислительной точки зрения, чем исходный метод. В этой модификации вместо задачи (4) решается задача отыскания

$$\nu_k = \max\{\nu\colon \nu \leqslant \max_{x \in \Omega_0} \min_{m+1 \leqslant i \leqslant m+k} [(a_i, x) - \alpha_i - \nu(x, x)]\}.$$

Из самого определения ν_k ясно, что

$$\nu_k = \max_{x \in \Omega_0} \min_{m+1 \leqslant i \leqslant m+k} [(a_i, x) - \alpha_i - \nu_k(x, x)] =$$

$$= \min_{m+1 \leqslant i \leqslant m+k} [(a_i, y_{k+1}) - \alpha_i - \nu_k(y_{k+1}, y_{k+1})]. \qquad (4'')$$

Так как реализация условия (4″) является трудоемкой вычислительной задачей, то модифицированный метод следует рассматривать лишь как способ для оценивания скорости сходимости.

Имеет место следующая

Лемма. *Если функция $\|\varphi(x)\|$ ограничена на Ω_0, то $\nu_k^2 = O(1/k)$.*

Не останавливаясь на доказательстве, заметим, что согласно этой лемме целесообразно в качестве управляющей последовательности $\{\delta_k\}$ в процессе выбрасывания лишних ограничений брать $\delta_k = O(k^{-1/2})$. При таком выборе $\{\delta_k\}$ вероятно, что на каждом шаге будет происходить очистка. Описанным способом выбрасывания лишних ограничений можно пользоваться также и в методе секущей плоскости.

Основываясь на процессе (4″) и лемме можно построить эффективно численно реализуемый алгоритм, на каждом шаге которого осуществляется очистка, а оценка $O(1/k)$ скорости сходимости при этом сохраняется.

5. Метод центрированных сечений. Пусть теперь последовательность $\{z_k\}$ строится методом центрированных сечений ([1]), т. е. z_1 является центром тяжести многогранника Ω_0, z_2 являются центром тяжести многогранника $\Omega_1 = \{x \in \Omega_0\colon (a_{m+1}, x) \geqslant \alpha_{m+1}\}$, где $a_{m+1} = \partial\varphi(z_1)$, $\alpha_{m+1} = (\partial\varphi(z_1), z_1)$, z_{k+1} является центром тяжести многогранника $\Omega_k = \{x \in \Omega_{k-1}\colon (a_{m+k}, x) \geqslant \alpha_{m+k}\}$, где $a_{m+k} = \partial\varphi(z_k)$, $\alpha_{m+k} = (\partial\varphi(z_k), z_k)$ и т. д.

Для объемов V_k многогранников Ω_k, как известно, справедливо неравенство $V_k \leqslant V_0 q^k$, $0 < q < 1$.

Пусть

$$c_k = \max_{x \in E^n} \min_{1 \leqslant i \leqslant m+k} [(a_i, x) - \alpha_i] / \|a_i\|.$$

Ясно, что чебышевский радиус c_k является радиусом некоторого шара, целиком принадлежащего Ω_k, так что его объем не превосходит V_k и, значит, $c_k \leqslant L_0 q_0^k$, где q_0, $0 < q_0 < 1$, зависит от n.

Таким образом, $c_k \to 0$, т. е. выполнены условия теоремы 1 и верна

Теорема 2. *Если функция* $\|\partial\varphi(x)\|$ *ограничена на* Ω_0 *и подпоследовательность* $\{z_{i_k}\}$ *последовательности* $\{z_k\}$, *построенной методом центрированных сечений, такова, что* $\varphi(z_{i_{k-1}}) \leqslant \varphi(z_{i_k})$ *и* $\varphi(z_i) < \varphi(z_{i_{k-1}})$, $i = i_{k-1}+1, \ldots, i_k - 1$, *то каждая ее предельная точка является решением задачи* (1) *и* $\varphi(z_{i_k})$ *сходится к* $\varphi(x^*)$ *со скоростью геометрической прогрессии или точнее*

$$0 \leqslant \varphi(x^*) - \varphi(z_{i_k}) \leqslant L_1 q_1^{i_k}. \tag{5}$$

В доказательстве нуждается лишь оценка (5). Для этого существенно используется существование положительных чисел q_1, λ_1 таких, что

$$\max_{x \in \Omega_0} \min_{m+1 \leqslant i \leqslant m+k} [(a_i, x) - \alpha_i]/\|a_i\| = \gamma_k \leqslant \lambda_1 q_1^k, \tag{6}$$

и вогнутость функции $\varphi(x)$.

Отметим, наконец, что если функция $\varphi(x)$ строго вогнута, то $z_{i_k} \to x^*$, если же $\varphi(x)$ сильно вогнута, то z_{ik} сходится к x^* со скоростью геометрической прогрессии.

Московский инженерно-строительный институт им. В. В. Куйбышева

Поступило 24 X 1974

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] *А. Ю. Левин*, ДАН, т. 160, № 6 (1965). [2] *С. И. Зуховицкий, Р. А. Поляк, М. Е. Примак*, ДАН, т. 185, № 1 (1969).