

Р. А. ПОЛЯК

О НАИЛУЧШЕМ ВЫПУКЛОМ ЧЕБЫШЕВСКОМ ПРИБЛИЖЕНИИ

(Представлено академиком Л. В. Канторовичем 25 III 1971)

1. Пусть заданы выпуклые по $x \in E^n$ функции $f_j(x)$, $j \in J = \{1, \dots, m\}$. Обозначим $f(x) = \max_{j \in J} f_j(x)$. Под задачей выпуклого чебышевского приближения понимают задачу отыскания вектора x^* из условия

$$f(x^*) = \min_x f(x) = \min_x \max_{j \in J} f_j(x). \quad (1)$$

Вектор x^* называют чебышевским приближением системы $f_j(x)$, $j \in J$.

Минимальным назовем множество $\bar{J} \subset J$, для которого

$$\min_x \max_{j \in \bar{J}} f_j(x) = f(x^*), \text{ а } \min_x \max_{j \in J \setminus \bar{J}} f_j(x) < f(x^*) \quad \forall l \in \bar{J}.$$

Единственность вектора x^* следует, например, из предположения строгой выпуклости любой $f_j(x)$, $j \in \bar{J}$.

Если же указанное условие не выполнено, то множество $\Omega_1 = \{x: f(x) = f(x^*) = \mu_1\}$ может, вообще говоря, состоять более чем из одной точки.

Оказывается, что в этом случае $J(x^*) = \{j: f_j(x^*) = f(x^*)\}$ содержит непустое подмножество $J_1 = \{j: f_j(x) = \mu_1 \quad \forall x \in \Omega_1\}$.

Рассмотрим множество $J \setminus J_1$ и поставим задачу отыскания

$$\mu_2 = \min_{x \in \Omega_1} \max_{j \in J \setminus J_1} f_j(x).$$

Если $\Omega_2 = \{x: x \in \Omega_1, \max_{j \in J \setminus J_1} f_j(x) = \mu_2\}$ состоит более чем из одной точки,

то рассуждая аналогично предыдущему, придем к существованию множества $J_2 = \{j: f_j(x) = \mu_2 \quad \forall x \in \Omega_2\}$. Пусть $J(2) = J_1 \cup J_2$. Процесс выделения множеств J_q продолжаем до тех пор, пока либо для некоторого q множество Ω_q окажется состоящим из одной точки, либо окажется $J(q) = J$.

Каждую точку x^* множества Ω_q будем называть наилучшим чебышевским приближением (н.ч.п.). (Для задачи чебышевского приближения функции на конечном множестве понятие н.ч.п. рассматривалось в (1).) Отметим, что каждое н.ч.п. является решением задачи векторной оптимизации (оптимумом по Парето) вектор-функции $F(x) = (f_1(x), \dots, f_m(x))$ на E^n , т. е.

$$F(x^*) = \min \{d: F(x) \leq d, x \in E^n\}. \quad (2)$$

Неравенство $d^1 \leq d^2$ означает, что $d_j^1 \leq d_j^2$ для $j \in J$.

В настоящей работе строится вычислительный метод для отыскания н.ч.п. и исследуется его сходимость.

Основная трудность, возникающая при решении задачи (1), связана с тем, что $f(x)$ не везде дифференцируема, если даже гладки $f_j(x)$, $j \in J$. Решая задачу (1) методом наискорейшего спуска (2, 3), приходится вводить параметр, который в некотором смысле «сглаживает» функцию $f(x)$ в точ-

ках, где она недифференцируема. Наличие указанного параметра позволяет устранить «заедание» процесса, но при этом усложняется отыскание на каждом шаге направления спуска, так как введение параметра увеличивает размеры экстремальной задачи, решаемой для отыскания указанного направления. Интересный метод для решения задачи типа (1) построен в (4), однако последовательность приближений, получаемых при его использовании, вообще говоря, не сходится к н.ч.п.

В настоящей работе строится вычислительный метод, являющийся вариантом метода наискорейшего спуска, однако вместо локального сглаживания функции $f(x)$ в точках, где она недифференцируема, производится глобальное ее сглаживание во всем пространстве. Это позволяет избежать решения на каждом шаге экстремальной задачи, дает возможность получить последовательность $\{x^h\}$, сходящуюся к н.ч.п., и оценить в некоторых случаях быстроту сходимости метода.

2. Изложению вычислительного метода предположим следующие леммы.

Лемма 1. Пусть функции $f_j(x)$, $j \in J$, выпуклы и положительны в E^n , тогда для любого целого k функция $F_k(x) = (\sum_j f_j^k(x))^{1/k}$ выпукла, причем

$$0 \leq F_k(x) - f(x) \leq mf(x)/k \quad \forall x \in E^n.$$

Лемма 2. Пусть $f_j(x)$, $j \in J$, — положительные дифференцируемые функции, градиенты которых удовлетворяют условию Липшица, тогда для любого целого k функция $F_k(x)$ дифференцируема и существует константа $L > 0$ такая, что градиент $F_k(x)$ удовлетворяет условию Липшица с константой Lk , т. е.

$$\|\nabla F_k(x) - \nabla F_k(y)\| \leq Lk \|x - y\|.$$

Лемма 3. В предположениях леммы 2 относительно функций $f_j(x)$, $j \in J$, для любого целого $k > 0$ имеет место неравенство

$$F_k(x - t\nabla F_k(x)) \leq F_k(x) - \left(t - \frac{kL^2}{2}\right) \|\nabla F_k(x)\|^2,$$

причем $F_k(x - t\nabla F_k(x)) \leq F_k(x) \quad \forall t \in [0, 2/(Lk)]$.

Следствие. Для $\forall t \in [0, 1/(Lk)]$ имеет место

$$F_k(x - t\nabla F_k(x)) \leq F_k(x) - \frac{t}{2} \|\nabla F_k(x)\|^2.$$

Лемма 4. Пусть положительные функции $f_j(x)$, $j \in J$, сильно выпуклы, т. е.

$$f_j\left(\frac{x+y}{2}\right) \leq \frac{1}{2} [f_j(x) + f_j(y)] - \gamma_j \|x - y\|^2, \quad \gamma_j > 0, \quad j \in J.$$

Тогда для любого целого $k > 0$ функция $F_k(x)$ сильно выпукла, т. е.

$$F_k\left(\frac{x+y}{2}\right) \leq \frac{1}{2} [F_k(x) + F_k(y)] - \gamma \|x - y\|^2, \quad \gamma = \min_j \gamma_j.$$

Отметим, что выпуклость (сильная выпуклость) $F_k(x)$ следует из соответствующих свойств функций $f_j(x)$, $j \in J$, и неравенства Минковского*.

Следствие. Если функции $f_j(x)$, $j \in J$, дифференцируемы и удовлетворяют условиям леммы 4, то

$$a) F_k(x) - F_k(x_k^*) \geq \frac{1}{2} \gamma \|x - x_k^*\|^2,$$

$$б) \|\nabla F_k(x)\|^2 \geq 2\gamma (F_k(x) - F_k(x_k^*)),$$

где $F_k(x_k^*) = \min_x F_k(x)$.

Перейдем к изложению вычислительного метода.

3. В качестве исходного приближения выбираем произвольный вектор x^0 . Пусть уже построены приближения $x^1, \dots, x^{h-1}$. Очередное прибли-

* Доказательство леммы 4 в случае дважды дифференцируемых $f_j(x)$, $j \in J$, сообщил автору Г. Д. Майстровский.

жение отыскиваем по формуле

$$x^k = x^{k-1} - t_k \nabla F_k(x^{k-1}), \quad (3)$$

где

$$t_k = \max \left\{ \tau = \frac{t_{k-1} 2^{-v}}{k} : F_k(x^{k-1} - \tau \nabla F_k(x^{k-1})) \leq \right. \\ \left. \leq F_k(x^{k-1}) - \tau \|\nabla F_k(x^{k-1})\|^2, v - \text{целое} \right\},$$

а t_0 — достаточно большое число.

Заметим, что в силу леммы 3, начиная с некоторого номера k , величина kt_k не будет меняться и для определения параметра t_k понадобится вычисление значения функции $F_k(x)$ лишь в одной точке $x^{k-1} - t_k \nabla F_k(x^{k-1})$.

4. Сходимость процесса (3) устанавливает

Теорема 1. Если градиенты положительных функций $f_j(x)$, $j \in J$, удовлетворяют условию Липшица и $\Omega(x^0) = \{x: f(x) \leq f(x^0)\}$ ограничено, то любая сходящаяся подпоследовательность из $\{x^k\}$ имеет своим пределом н.ч.н. $x^* \in \Omega_*$.

Скорость сходимости процесса (3) устанавливает

Теорема 2. Пусть положительные функции $f_j(x)$, $j \in J$, сильно выпуклы, их градиенты удовлетворяют условию Липшица и $\gamma > L$, тогда

$$F_k(x^k) - f(x^*) \leq c/k, \quad \|x^k - x^*\|^2 \leq c_1/k. \quad (4)$$

5. Процесс (3) можно использовать для решения общей задачи выпуклого программирования, состоящий в отыскании

$$\varphi^* = \varphi(x^*) = \min \{ \varphi(x) : \varphi_j(x) \leq 0, j \in J \}, \quad (5)$$

если известно значение φ^* . Легко заметить, что последняя задача эквивалентна задаче (1), если к системе функций $f_j(x) = \varphi_j(x)$, $j \in J$, присоединить $f_0(x) = \varphi_0(x) = \varphi(x) - \varphi^*$.

Для применимости процесса (3) необходимо лишь к каждой функции $\varphi_j(x)$, $j \in J' = J \cup \{0\}$ прибавить такое положительное число, чтобы полученные функции оказались положительными на $\Omega(x^0)$. Будем считать, что это уже сделано и функции $\varphi_j(x)$, $j \in J'$, удовлетворяют условиям теоремы 2. Тогда применив процесс (3) для решения задачи (1) с системой $\varphi_j(x)$, $j \in J'$, получим последовательность $\{x^k\}$, сходящуюся к x^* со скоростью (4), а также окажется, что

$$\lim_{k \rightarrow \infty} [\varphi_j(x^{k-1}) / \Phi_k(x^{k-1})]^{k-1} = v_j^*, \quad j \in J',$$

где $\Phi_k(x) = \left(\sum_{j \in J'} \varphi_j^k(x) \right)^k$.

Если множество $\Omega = \{x: \varphi_j(x) \leq 0, j \in J\}$ удовлетворяет условию Слейтера, то $v_0^* > 0$. Поэтому для множителей Лагранжа, соответствующих ограничениям задачи (5), получим

$$u_j^* = \frac{v_j^*}{v_0^*} = \lim_{k \rightarrow \infty} \left[\frac{\varphi_j(x^{k-1})}{\varphi_0(x^{k-1})} \right]^{k-1}, \quad j \in J.$$

Таким образом, применение процесса (3) к задаче (4) позволяет построить последовательности, сходящиеся к решению прямой и двойственной задачи выпуклого программирования.

В заключение отметим, что процесс (3) можно применить для решения задачи (4), если φ^* неизвестно, а известна лишь оценка снизу для φ^* .

Украинский филиал научно-исследовательского
института планирования и нормативов
при Госплане СССР
Киев

Поступило
26 II 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

¹ J. Descloux, J. Soc. Indust. Appl. Math., 11, № 4 (1963). ² Г. Зойтендейк, Методы возможных направлений, ИЛ, 1963. ³ С. И. Зуховицкий, Р. А. Поляк, М. Е. Прима, ДАН, 153, № 5 (1963). ⁴ Н. З. Шор, Кибернетика, № 3 (1968).