Доклады Академии наук СССР
1976. Том 229, № 2

УДК 519.8

МАТЕМАТИКА

В. Ф. ДЕМЬЯНОВ

К МЕТОДУ ЭКСТРЕМАЛЬНОГО БАЗИСА

(Представлено академиком Л. В. Канторовичем 4 III 1976)

1°. Вначале рассмотрим задачу математического программирования: найти

$$\min_{X\in\Omega} f(X), \tag{1}$$

где

$$\Omega=\{X\in E_n \mid h(X,\ Y)\leqslant 0 \quad \forall Y\in\omega\}, \tag{2}$$

$\omega\subset E_p$ — компактное множество, функция $h(X, Y)$ выпукла по X и непрерывна на $E_n\times\omega$ вместе с $h_X'(X, Y)$. Функция $f(X)$ строго выпукла на E_n, т. е. для всех X_1 и X_2 $(X_1\neq X_2)$ и для всех $\alpha\in(0, 1)$ имеет место неравенство

$$f(\alpha X_1+(1-\alpha)X_2)<\alpha f(X_1)+(1-\alpha)f(X_2). \tag{3}$$

Предполагается выполненным условие Слейтера: существует такая точка $\overline{X}\in E_n$, что

$$\max_{Y\in\omega} h(\overline{X},\ Y)=-\mu<0 \tag{4}$$

и что существует точка $\tilde{X}\in\Omega$ такая, что множество

$$M(\tilde{X})=\{X\in E_n \mid f(X)\leqslant f(\tilde{X})\} \tag{5}$$

ограничено.

Множество Ω в (2) может быть задано очень большим или даже бесконечным числом ограничений (если ω — множество, содержащее бесконечное число точек). Ниже описан метод сведения задачи (1), (2) к последовательности задач минимизации $f(X)$ на множествах, заданных небольшим числом ограничений.

Пусть $X\in\Omega$. Введем множества

$$Q(X)=\{Y\in\omega \mid h(X,\ Y)=0\},$$

$$H(X)=\{f_X'(X)\}\cup\{h_X'(X,\ Y)\mid Y\in Q(X)\},\quad L(X)=\operatorname{co} H(X),$$

$$\Gamma^+(X)=\left\{Z=-\sum_{Y_i\in Q(X)}\alpha_i h_X'(X, Y_i)\mid \alpha_i\geqslant 0\right\}.$$

Хорошо известно условие (см., например, [1, 2]): для того чтобы точка $X^*\in\Omega$ была точкой минимума функции $f(X)$ на Ω, необходимо и достаточно, чтобы

$$f_X'(X^*)\in\Gamma^+(X^*). \tag{6}$$

Условие (6) эквивалентно условию

$$0\in L(X^*). \tag{7}$$

Набор точек $\sigma=\{Y_0, Y_1, \ldots, Y_n\}$, $Y_i\in\omega$, назовем базисом.

Опишем вначале «принципиальный» алгоритм (по терминологии [3]). Выберем произвольный базис $\sigma_0=\{Y_{00}, Y_{01}, \ldots, Y_{0n}\}$, $Y_{0i}\in\omega$. Пусть уже найден базис $\sigma_k=\{Y_{k0}, Y_{k1}, \ldots, Y_{kn}\}$. Рассмотрим множество

$$\Omega_k=\{x\in E_n \mid h(X,\ Y)\leqslant 0 \ \ \forall Y\in\sigma_k\} \tag{8}$$

и найдем точку $X_k\in\Omega_k$ такую, что

$$f(X_k)=\min_{X\in\Omega_k} f(X). \tag{9}$$

Точка X_k удовлетворяет условию, аналогичному (6) (или, что то же, (7)), т. е.

$$0 \in L_k, \tag{10}$$

где

$$L_k = \operatorname{co} H_k, \quad H_k = \{f_X'(X_k)\} \cup \{h_X'(X_k, Y) \mid Y \in Q_k\},$$
$$Q_k = \{Y \in \sigma_k \mid h(X_k, Y) = 0\}.$$

Если $X_k \in \Omega$, то X_k — точка минимума $f(X)$ на Ω; в противном случае найдется точка $\overline{Y}_k \in \Omega$ такая, что

$$h(X_k, \overline{Y}_k) = \max_{Y \in \omega} h(X_k, Y) \equiv \lambda_k. \tag{11}$$

Так как любая точка выпуклой оболочки в E_n может быть представлена в виде выпуклой комбинации не более чем $n+1$ точек, то начало координат в (9) может быть записано в виде выпуклой комбинации вектора $f_X'(X_k)$ (он обязательно входит в представление начала координат в силу условия Слейтера (4)) и не более чем n векторов из $H_k' = \{h_X'(X_k, Y) \mid Y \in Q_k\}$. Значит, существует точка $Y_{ki_k} \in \sigma_k$ такая, что начало координат в (10) может быть представлено в виде выпуклой комбинации векторов из H_k, в которой вектор $h_X'(X_k, Y_{ki_k})$ не участвует. Строим теперь новый базис

$$\sigma_{k+1} = \{Y_{k+1,\,0},\ Y_{k+1,\,1}, \ldots, Y_{k+1,\,n}\},$$

где

$$Y_{k+1,\,i} = \begin{cases} Y_{ki}, & i \neq i_k, \\ \overline{Y}_k, & i = i_k. \end{cases}$$

Далее продолжаем аналогично. В результате получим последовательность базисов $\{\sigma_k\}$ и последовательность точек $\{X_k\}$. Если эти последовательности состоят из конечного числа точек, то последняя полученная точка последовательности $\{X_k\}$ является точкой минимума $f(X)$ на Ω; в противном случае справедлива

Теорема 1. *Последовательность $\{X_k\}$ сходится к точке минимума функции $f(X)$ на множестве Ω.*

Замечание 1. Предположение о строгой выпуклости функции $f(X)$ существенно. Алгоритм можно применять и для выпуклых функций, но тогда возможно зацикливание, с которым необходимо бороться так же, как в линейном программировании борются с вырожденностью.

2^0. Практическое применение изложенного выше метода затруднено тем, что необходимо точно решать задачи (9) и (11). Предположим теперь, что вместо условия (3) выполнено более сильное условие: функция f дважды непрерывно дифференцируема, причем

$$(f_{XX}''(X)Z, Z) \geqslant m\|Z\|^2 \quad \forall X, Z \in E_n,$$

где $m > 0$. Кроме того, существует $K < \infty$ такое, что

$$\|h_X'(X, Y)\| \leqslant K \quad \forall X \in E_n, \quad Y \in \omega.$$

Ограниченность множества $M(\bar{X})$ (см. (5)) можно теперь не требовать. Пусть $\varepsilon > 0$ фиксировано и последовательность $\{\rho_k\}$ такова, что $\rho_k > 0$, $\rho_k \to 0$, $\sum_{k=0}^{\infty} \rho_k < \infty$, $\rho_0 < \mu$. В качестве начального выберем произвольный базис σ_0. Пусть уже найден базис $\sigma_k = \{Y_{k0}, \ldots, Y_{kn}\}$. Снова рассмотрим множество Ω_k (см. (8)). В конечное число шагов (см., например, [2]) можно найти точку $X_k \in \Omega_k$ такую, что

$$d_k = \min_{Z \in \Gamma_{k\varepsilon}^+} \|f_X'(X_k) - Z\| = \|f_X'(X_k) - Z_k\| \leqslant \rho_k, \tag{12}$$

где

$$\Gamma_{k\varepsilon}^+ = \left\{ Z = -\sum_{Y_i \in Q_{k\varepsilon}} \alpha_i h_X'(X_k, Y_i) \mid \alpha_i \geqslant 0 \right\}, \quad Q_{k\varepsilon} = \{Y \in \sigma_k \mid -\varepsilon \leqslant h(X_k, Y) \leqslant 0\}.$$

Точка X_k является (ρ_k, ε)-стационарной (см. [2], гл. 5) точкой функции $f(X)$ на множестве Ω_k. Так как любая точка конуса $\Gamma_{k\varepsilon}^+$ может быть представлена в виде неотрицательной комбинации не более чем n векторов, то существует такая точка $Y_{ki_k} \in \sigma_k$ и такое представление точки Z_k (см. (12)), в котором вектор $h_X'(X_k, Y_{ki_k})$ не участвует.

Если $\lambda_k = 0$ (см. (11)), то $X_k - (\rho_k, \varepsilon)$-стационарная точка функции $f(X)$ на множестве Ω; тогда полагаем $\sigma_{k+1} = \sigma_k$ и продолжаем процесс. Если же $\lambda_k > 0$, то найдем такую точку $Y_k' \in \omega$, что $h(X_k, Y_k') > \beta\lambda_k$ (где $\beta \in (0, 1)$ фиксировано) и положим $\sigma_{k+1} = \{Y_{k+1,0}, Y_{k+1,1}, \ldots, Y_{k+1,n}\}$, где

$$Y_{k+1,i} = \begin{cases} Y_{ki}, & i \neq i_k, \\ Y_k', & i = i_k. \end{cases}$$

Далее продолжаем аналогично. В результате получим последовательность точек $\{X_k\}$.

Теорема 2. *Всякая предельная точка последовательности $\{X_k\}$ является ε-стационарной точкой $f(X)$ на Ω, т. е. если $X_{k_s} \to X^*$, то $0 \in L_\varepsilon(X^*)$, где*

$$L_\varepsilon(X^*) = \operatorname{co} H_\varepsilon(X^*), \quad H_\varepsilon(X^*) = \{f_X'(X^*)\} \cup \{h_X'(X^*, Y) \mid Y \in Q_\varepsilon(X^*)\},$$

$$Q_\varepsilon(X^*) = \{Y \in \omega \mid -\varepsilon \leqslant h(X^*, Y) \leqslant 0\}.$$

Замечание 2. Если одновременно менять и ε, устремляя его к нулю, то последовательность $\{X_k\}$ будет стремиться к точке минимума функции $f(X)$ на Ω.

3°. Изложенное выше может быть применено для минимизации функции

$$\varphi(X) = \max_{Y \in G} f(X, Y)$$

на множестве

$$\Omega = \{X \in E_n \mid h(X, Z) \leqslant 0 \ \ \forall Z \in \omega\},$$

где функция $f(X, Y)$ строго выпукла по X, дифференцируема по X и непрерывна по Y. Множество Ω удовлетворяет тем же условиям, что и в п. 1°.

Опишем принципиальную схему алгоритма. Выберем в качестве начального произвольный базис $\sigma_0 = \sigma_0' \cup \sigma_0''$, где

$$\sigma_0' = \{Y_{00}, \ldots, Y_{0s_0}\}, \ Y_{0i} \in G, \quad \sigma_0'' = \{Z_{00}, \ldots, Z_{0t_0}\}, \ Z_{0j} \in \omega,$$

причем $s_0 \geqslant 0$, $s_0 + t_0 = n+1$ (таким образом, базис σ_0 состоит из $n+3$ точек). Пусть уже построен базис $\sigma_k = \sigma_k' \cup \sigma_k''$, где

$$\sigma_k' = \{Y_{k0}, \ldots, Y_{ks_k}\}, \ Y_{ki} \in G, \quad \sigma_k'' = \{Z_{k0}, \ldots, Z_{kt_k}\}, \ Z_{kj} \in \omega,$$

причем $s_k \geqslant 0$, $s_k + t_k = n+1$ (множество σ_k'' может оказаться и пустым). Положим

$$\varphi_k = \max_{Y \in \sigma_k'} f(X, Y), \quad \Omega_k = \{X \in E_n \mid h(X, Z) \leqslant 0 \ \ \forall Z \in \sigma_k''\}$$

и найдем точку $X_k \in \Omega_k$ такую, что $\varphi_k(X_k) = \min_{X \in \Omega_k} \varphi_k(X)$. В точке X_k выполнено условие

$$0 \in L_k, \tag{13}$$

где

$$L_k = \operatorname{co} H_k, \quad H_k = H_k' \cup H_k'', \quad H_k' = \{f_X'(X_k, Y) \mid Y \in R_k\},$$

$$H_k'' = \{h_X'(X_k, Z) \mid Z \in Q_k\}, \quad R_k = \{Y \in \sigma_k' \mid f(X_k, Y) = \varphi_k(X_k)\},$$

$$Q_k = \{Z \in \sigma_k'' \mid h(X_k, Z) = 0\}.$$

Существует представление начала координат в (13), содержащее не более $n+1$ вектора из H_k; при этом в такое представление входит хотя бы один вектор из H_k' (это следует из условия Слейтера (4)). Таким образом, в базисе σ_k есть по крайней мере две точки V_{k1}, V_{k2}, которым соответствуют точки из H_k, не входящие в представление начала коорди-

нат в (13). Теперь найдем точки $Y_k \in G$ и $Z_k' \in \Omega$ такие, что

$$f(X_k, Y_k') = \varphi(X_k); \tag{14}$$

$$h(X_k, Z_k') = \max_{Z \in \omega} h(X_k, Z) = \lambda_k, \tag{15}$$

и построим базис σ_{k+1}, в котором точки V_{k1} и V_{k2} заменены точками Y_k' и Z_k'. Если $\varphi(X_k) = \varphi_k(X_k)$, то точку V_k не меняем, а если $\lambda_k = 0$, то не меняем точку V_{k2}; если же одновременно $\varphi(X_k) = \varphi_k(X_k)$ и $\lambda_k = 0$, то точка X_k — точка минимума функции $\varphi(X)$ на множестве Ω. Если последовательность $\{X_k\}$ содержит бесконечное число точек, то справедлива

Теорема 3. *Последовательность точек* $\{X_k\}$ *сходится к точке минимума функции* $\varphi(X)$ *на множестве* Ω.

Замечание 3. Как и в п. 2°, можно разработать метод, в котором задачи (14) и (15) решаются приближенно.

Замечание 4. Ряд методов линейного и выпуклого программирования основан на последовательном уточнении базиса, состоящего из n точек (для задачи п. 1°) или из $n+1$ точек (для задачи п. 3°) (см., например, ([4], [5])). Аналогичная идея в задаче чебышевской аппроксимации проведена в методе Ремеза ([6]). Изложенные выше методы тоже по существу заключаются в поиске экстремального базиса, но делается это с помощью избыточных точек (одной в методах п.п. 1° и 2° и двух в методе п. 3°).

Замечание 5. Изложенный метод может быть применен для минимизации произвольной выпуклой функции $f(X)$ на произвольном выпуклом множестве $\Omega \subset S \subset E_n$, где S — ограниченное множество. Поскольку

$$f(X) = \max_{Y \in S} [(A(Y), X - Y) + f(Y)], \tag{16}$$

где $A(Y)$ — какой-нибудь субградиент функции f в точке Y, а

$$\Omega = \{X \mid (B(Z), X - Z) \geqslant 0 \quad \forall Z \in \Gamma_\Omega\}, \tag{17}$$

где $B(Z) \in \Gamma^+(Z)$, $\|B(Z)\| = 1$, $\Gamma^+(Z)$ — конус, сопряженный конусу возможных направлений множества Ω в точке Z, Γ_Ω — множество граничных точек множества Ω, то мы получили задачу минимизации функции максимума на множестве, заданном линейными неравенствами. Метод п. 3° в данном случае приведет к последовательности задач линейного программирования небольшой размерности. Так как функции в правой части (16) не являются строго выпуклыми, то, как отмечено выше, возможно заедание, с которым можно бороться, например так: если сходимость замедлилась, надо сделать несколько шагов (любым методом) для решения задачи

$$\min_{X \in \Omega_k} \max_{Y \in \sigma_k} [(A(Y), X - Y) + f(Y) + \varepsilon (X - Y)^2],$$

где $\varepsilon > 0$ — достаточно малое число. Здесь минимизируемая функция строго выпукла.

Представление (16) и (17) было использовано в методе Келли ([7]), где с каждым шагом количество линейных функций и ограничений увеличивалось.

Автор признателен Л. В. Васильеву, В. Н. Малоземову, А. Б. Певному и А. М. Рубинову за обсуждение результатов.

Ленинградский государственный университет
им. А. А. Жданова

Поступило
4 III 1976

ЛИТЕРАТУРА

[1] *Б. Н. Пшеничный*, Необходимые условия экстремума, «Наука», 1969. [2] *В. Ф. Демьянов, В. Н. Малоземов*, Введение в минимакс, «Наука», 1972. [3] *Э. Полак*, Численные методы оптимизации (единый подход), М., «Мир», 1974. [4] *Л. В. Канторович*, Математические методы организации и планирования производства, Л., Изд-во ЛГУ, 1939. [5] *Дж. Данциг*, Линейное программирование, его обобщения и применения, «Прогресс», 1966. [6] *Е. Я. Ремез*, Основы численных методов чебышевского приближения, Киев, «Наукова думка», 1969. [7] *J. E. Kelley*, SIAM J. on Appl. Math., v. 8, № 703 (1960).