ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 19 Январь 1979 Февраль № 1

УДК 519.8

О МЕТОДЕ ВТОРОГО ПОРЯДКА ДЛЯ ЗАДАЧИ ДИСКРЕТНОГО МИНИМАКСА

В. М. ПАНИН

(Киев)

Для решения общей задачи математического программирования предлагается итерационный метод второго порядка, использующий квадратичную аппроксимацию ограничений. Доказана сверхлинейная скорость сходимости с показателем, равным $^3/_2$, при отсутствии требования регулярности точки минимума. Расширение области сходимости достигается регулировкой шагового множителя.

Рассматривается задача: найти

$$\min_{x\in\Omega}\max_{i\in I} f_i(x), \qquad \Omega=\{x : g_j(x)\leqslant 0, \quad j\in J\}, \tag{1}$$

где I, J — конечные множества индексов, $x\in E^n$.

Эта задача подробно изучалась, например, в [1–3], где для отыскания решения x_* были построены итерационные методы первого порядка, сходящиеся в лучшем случае со скоростью геометрической прогрессии.

Покажем, что использование квадратичной аппроксимации функций $f_i(x)$, $g_j(x)$ приводит к сверхлинейной скорости сходимости. Введем обозначения

$$f_i^k(x)=f_i(x_k)+(f_i'(x_k), x-x_k)+\frac{1}{2}f_i''(x_k)(x-x_k)^2,$$

$$g_j^k(x)=g_j(x_k)+(g_j'(x_k), x-x_k)+\frac{1}{2}g_j''(x_k)(x-x_k)^2,$$

$$\begin{gathered}\varphi(x)=\max_{i\in I} f_i(x), \qquad F(x)=\max[\max_{j\in J} g_j(x), 0],\\ I_k=\{i\in I : f_i(x_k)\geqslant\varphi(x_k)-\delta\},\\ J_k=\{j\in J : g_j(x_k)\geqslant \max_{j\in J} g_j(x_k)-\gamma\},\end{gathered} \tag{2}$$

$$R(x)=\{i\in I : f_i(x)=\varphi(x)\},$$
$$\bar R(x)=\{j\in J : g_j(x)=\max_{j\in J} g_j(x)\},$$

$$\varphi_k(x)=\max_{i\in I_k} f_i^k(x), \qquad F_k(x)=\max[\max_{j\in J_k} g_j^k(x), 0],$$

где δ, $\gamma>0$ — произвольные постоянные числа.

Сущность предлагаемого метода сводится к решению в каждой точке x_k итерационного процесса следующей вспомогательной задачи: найти

$$(3)\qquad \min_{x\in\Omega_k}\max_{i\in I_k} f_i^k(x), \qquad \Omega_k=\{x: g_j^k(x)\leqslant 0, j\in J_k\},$$

решение которой обозначим через $\bar{x}_k$, а через $(\bar{x}_k, \Lambda^{k+1}, \beta^{k+1})$ — соответствующую точку Куна — Таккера. Здесь $\Lambda^{k+1}=\{\lambda_i^{k+1}\}$, $i\in I_k$, $\beta^{k+1}=\{\beta_j^{k+1}\}$, $j\in J_k$.

Основные предположения. а. Для любых x, $y\in E^n$, $i\in I$, $j\in J$

$$(4)\qquad \begin{gathered} m\|x\|^2\leqslant (f_i''(y)x, x)\leqslant M\|x\|^2, \\ m\|x\|^2\leqslant (g_j''(y)x, x)\leqslant M\|x\|^2, \quad M\geqslant m>0. \end{gathered}$$

б. Для задач (1), (3) выполняется условие Слейтера, причем множители Лагранжа $\beta_j^{k+1}=\beta_j(x_k)$ вспомогательной задачи (3), соответствующие ограничениям $g_j^k(x)\leqslant 0$, равномерно ограничены:

$$(5)\qquad \sum_{j\in J_k}\beta_j(x_k)\leqslant N \quad \forall x_k\in D_N=\{x:\varphi(x)+NF(x)\leqslant\varphi(x_0)+NF(x_0)\}.$$

Прежде чем переходить к формулировке алгоритма, приведем несколько утверждений. Отметим прежде всего, что, в силу (4), функции $\varphi(x)$ и $\varphi_k(x)$ строго выпуклы, причем решения x_*, $\bar{x}_k$ задач (1) и (3) единственны. Запишем необходимые условия экстремума для задачи (3) (см. [1]):

$$(6)\qquad \begin{gathered} \sum_i \lambda_i^{k+1}[f_i'(x_k)+f_i''(x_k)p_k]+\sum_j \beta_j^{k+1}[g_j'(x_k)+g_j''(x_k)p_k]=0, \\ \lambda_i^{k+1}\,[f_i^k(\bar{x}_k)-\varphi_k(\bar{x}_k)]=0, \qquad \beta_j^{k+1}\,g_j^k(\bar{x}_k)=0, \\ \sum_i \lambda_i^{k+1}=1, \qquad \lambda_i^{k+1}\geqslant 0, \qquad \beta_j^{k+1}\geqslant 0, \end{gathered}$$

где $p_k=\bar{x}_k-x_k$, а индексы i, j принимают все значения из соответствующих множеств I_k, J_k.

Предложение 1. *Условие $p_k=0$ является необходимым и достаточным для того, чтобы точка $\bar{x}_k$ была решением задачи* (1).

Доказательство вытекает из сравнения необходимых условий экстремума для задач (1) и (3) и совершенно аналогично доказательству леммы 5.1 из [2].

Предложение 2. *Если $x_k\to x_*$, то найдутся числа $\delta_1\leqslant\delta$ и $\gamma_1\leqslant\gamma$ такие, что для любых постоянных $\delta'>0$ и $\gamma'>0$ $(\delta'\leqslant\delta_1,\ \gamma'\leqslant\gamma_1)$ при достаточно больших k будет*

$$f_i(x_k)\geqslant\varphi(x_k)-\delta' \text{ для } i\in R(x_*)$$

и

$$g_j(x_k)\geqslant\max_{j\in J} g_j(x_k)-\gamma' \text{ для } j\in\bar{R}(x_*).$$

Доказательство см. в [1, 2].

Предложение 3. *Пусть для некоторых подмножеств индексов* $i=1, 2, \ldots, l\subset I$ *и* $j=1, 2, \ldots, l_1\subset J$ *выполняется*

$$\varphi(x)=\max_{1\leqslant i\leqslant l} f_i(x) \quad\text{и}\quad F(x)=\max[\max_{1\leqslant j\leqslant l_1} g_j(x), 0].$$

Тогда

$$\varphi(x)+NF(x)=\max_{\langle\Lambda,\beta\rangle\in Q}\sum_i\sum_j[\lambda_i f_i(x)+\beta_j g_j(x)],$$

где

$$Q=\left\{\langle\Lambda, \beta\rangle\in E^l\times E^{l_1}:\sum_1^l\lambda_i=1,\ \sum_1^{l_1}\beta_j\leqslant N,\ \lambda_i\geqslant 0,\ \beta_j\geqslant 0\right\}.$$

В самом деле, если $\langle\Lambda, \beta\rangle\in Q$, то

$$\max_{\Sigma\beta_j\leqslant N}\sum_1^{l_1}\beta_j g_j(x)=\max_{\Sigma\beta_j=N}\sum_1^{l_1}\beta_j g_j(x)=NF(x)\quad\text{при}\quad F(x)>0$$

и

$$\max_{\Sigma\beta_j\leqslant N}\sum_1^{l_1}\beta_j g_j(x)=0\quad\text{при}\quad F(x)=0.$$

Кроме того,

$$\varphi(x)=\max_{\Sigma\lambda_i=1}\sum_1^l\lambda_i f_i(x)$$

(см. [1]). Отсюда и следует утверждение.

Формулировка алгоритма. Пусть выбрано указанное N и числа γ, δ, ε (γ, $\delta>0$, $0<\varepsilon<1$) и построена точка x_k процесса.

Шаг 1. Находим I_k, J_k и решаем задачу выпуклого программирования (3).

Шаг 2. Полагаем $\alpha=1$ и проверяем неравенство

$$\begin{aligned}\varphi(x_k+\alpha p_k)+NF(x_k+\alpha p_k)&\leqslant\varphi(x_k)+\\+NF(x_k)+\varepsilon\alpha\psi_k(\bar{x}_k),&\qquad 0<\varepsilon<1,\end{aligned}\tag{7}$$

где $\psi_k(\bar{x}_k)=\varphi_k(\bar{x}_k)-[\varphi(x_k)+NF(x_k)]$, $p_k=\bar{x}_k-x_k$. Если (7) выполняется, то, обозначая полученное α через α_k, находим $x_{k+1}=x_k+\alpha_k p_k$. В противном случае находим $\alpha/2$. Вновь обозначая его через α, переходим к проверке неравенства (7).

Теорема. *Пусть для задач* (1), (3) *выполняются предположения* а *и* б. *Тогда для сформулированного алгоритма справедливы утверждения:*

1) $x_k\to x_*$, $\quad k\to\infty$;

2) *если* $f_i''(x)$, $g_j''(x)$ $(i\in R(x_*),\ j\in\bar{R}(x_*))$ *непрерывны в окрестности* x_*, *то начиная с некоторой итерации будет* $\alpha_k=1$ *и*

$$\|p_k\|/\|p_{k-1}\|\to 0,\qquad k\to\infty;\tag{8}$$

3) *если, кроме того,* $f_i''(x)$, $g_j''(x)$ $(i\in R(x_*),\ j\in \bar R(x_*))$ *удовлетворяют условию Липшица, то при* $k\to\infty$

$$\|p_k\|\leqslant C\|p_{k-1}\|^{3/2}, \tag{9}$$

где $C>0$ — *постоянная величина.*

Доказательство проводится по схеме доказательства для метода Ньютона в [2, 4].

Из разложения Тейлора для $f_i(x_k+\alpha p_k)$, $\alpha\leqslant 1$, следует, с учетом (4), что

$$\max_{i\in I_k} f_i(x_k+\alpha p_k)>\varphi(x_k)-\alpha K\|p_k\|, \tag{10}$$

$$f_i(x_k+\alpha p_k)\leqslant\varphi(x_k)-\delta+\alpha(K+M\|p_k\|/2)\|p_k\|,\ \text{если}\ i\notin I_k, \tag{11}$$

где $K>0$ — константа, ограничивающая $\|f_i'(x)\|$, $\|g_j'(x)\|$ на отрезке $[x_k, \bar x_k]$. Если взять

$$\alpha\leqslant\min\{\delta,\gamma\}[(2K+M\|p_k\|)\|p_k\|]^{-1},\qquad \alpha\leqslant 1, \tag{12}$$

то из (10), (11) вытекает, что

$$\max_{i\in I_k} f_i(x_k+\alpha p_k)>\varphi(x_k)-\frac{\delta}{2}$$

и

$$f_i(x_k+\alpha p_k)\leqslant\varphi(x_k)-\frac{\delta}{2}\ \text{для}\ i\notin I_k,$$

откуда

$$\varphi(x_k+\alpha p_k)=\max_{i\in I_k} f_i(x_k+\alpha p_k).$$

Совершенно аналогично

$$F(x_k+\alpha p_k)=\max[\max_{j\in J_k} g_j(x_k+\alpha p_k),0].$$

Таким образом, если α удовлетворяет (12), то, учитывая предложение 3, имеем

$$\varphi(x_k+\alpha p_k)+NF(x_k+\alpha p_k)=$$

$$=\max_{\langle\Lambda,\beta\rangle\in Q_k}\sum_{i\in I_k}\sum_{j\in J_k}[\lambda_i f_i(x_k+\alpha p_k)+\beta_j g_j(x_k+\alpha p_k)],$$

$$Q_k=\left\{\langle\Lambda,\beta\rangle\in E^{I_k}\times E^{J_k}:\ \sum_{i\in I_k}\lambda_i=1,\ \sum_{j\in J_k}\beta_j\leqslant N,\ \lambda_i\geqslant 0,\ \beta_j\geqslant 0\right\}. \tag{13}$$

Обозначив $Q^k=\langle Q_k, 0\rangle\in E^I\times E^J$, можно записать (13) также в виде

$$\varphi(x_k+\alpha p_k)+NF(x_k+\alpha p_k)=\max_{\langle\Lambda,\beta\rangle\in Q^k} L(x_k+\alpha p_k,\Lambda,\beta),$$

где

$$L(x,\Lambda,\beta)=\sum_{i\in I}\lambda_i f_i(x)+\sum_{j\in J}\beta_j g_j(x).$$

Покажем теперь обоснование выбора α_k из условия проверки (7). Если α удовлетворяет (12), то, разлагая в ряд Тейлора $f_i(x_k+\alpha p_k)$, $g_j(x_k+\alpha p_k)$, получаем из (13)

$$\varphi(x_k+\alpha p_k)+NF(x_k+\alpha p_k)\leqslant \max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k+\alpha p_k,\Lambda,\beta)+ \\ +\frac{\alpha^2}{2}(c_k+\bar c_k)\|p_k\|^2,$$

(14)
$$L_k(x,\Lambda,\beta)=\sum_{i\in I_k}\lambda_i f_i^k(x)+\sum_{j\in J_k}\beta_j g_j^k(x),\qquad c_k=\max_{i\in I_k}\|\Phi_k^i\|, \\ \bar c_k=N\max_{j\in J_k}\|\bar\Phi_k^j\|,\qquad \Phi_k^i=f_i''(x_k^{\theta_i})-f_i''(x_k), \\ \bar\Phi_k^j=g_j''(x_k^{\theta_j})-g_j''(x_k), \\ x_k^{\theta_i}=x_k+\theta_i\alpha p_k,\qquad x_k^{\theta_j}=x_k+\theta_j\alpha p_k,\qquad 0\leqslant\theta_i,\ \theta_j\leqslant 1.$$

Функции $L(x,\Lambda,\beta)$ и $L_k(x,\Lambda,\beta)$ являются [1] функциями Лагранжа для задач (1), (3).

Так как функция $\max\limits_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x,\Lambda,\beta)$ выпукла по x, то

(15)
$$\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k+\alpha p_k,\Lambda,\beta)\leqslant \max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k,\Lambda,\beta)+ \\ +\alpha\Big[\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(\bar x_k,\Lambda,\beta)-\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k,\Lambda,\beta)\Big].$$

В точке Куна — Таккера $(\bar x_k,\Lambda^{k+1},\beta^{k+1})$ достигается, как известно, минимакс функции Лагранжа $L_k(x,\Lambda,\beta)$ на $E^n\times\langle\Lambda,\beta\rangle_k^+$, где

$$\langle\Lambda,\beta\rangle_k^+=\Big\{\langle\Lambda,\beta\rangle\in E^{I_k}\times E^{J_k}\colon \sum_{i\in I_k}\lambda_i=1,\quad \lambda_i\geqslant 0,\quad \beta_j\geqslant 0,\quad j\in J_k\Big\}.$$

Но, по предположению б, вектор $\langle\lambda^{k+1},\beta^{k+1}\rangle\in Q_k$. Из сказанного следует, что

(16)
$$L_k(\bar x_k,\Lambda^{k+1},\beta^{k+1})=\min_{x\in E^n}\max_{\langle\Lambda,\beta\rangle_k^+} L_k(x,\Lambda,\beta)= \\ =\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(\bar x_k,\Lambda,\beta)=\varphi_k(\bar x_k),$$

где учитывалось (аналогично предложению 3), что

$$\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x,\Lambda,\beta)=\varphi_k(x)+NF_k(x)$$

и $F_k(\bar x_k)=0$.

Отметим, что $f_i^k(x_k)=f_i(x_k)$, $g_j(x_k)=g_j^k(x_k)$ и поэтому, с учетом (13),

$$\max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k,\Lambda,\beta)=\varphi(x_k)+NF(x_k).$$

Тогда, на основании (16),

$$(17)\qquad L_k(\bar{x}_k, \Lambda^{k+1}, \beta^{k+1}) - \max_{\langle\Lambda,\beta\rangle\in Q_k} L_k(x_k, \Lambda, \beta) = \psi_k(\bar{x}_k).$$

Используя (15) и (17), получаем из (14)

$$(18)\qquad \varphi(x_k+\alpha p_k)+NF(x_k+\alpha p_k)\leqslant\varphi(x_k)+$$

$$+NF(x_k)+\alpha\psi_k(\bar{x}_k)\left[1+\frac{\alpha(c_k+\bar{c}_k)}{2\psi_k(\bar{x}_k)}\|p_k\|^2\right].$$

Докажем, что

$$(19)\qquad \psi_k(\bar{x}_k)\leqslant -m\|p_k\|^2/2.$$

Далее мы не указываем области суммирования I_k и J_k по i и по j, имея в виду, что они определяются соответствующим компактом Q_k. Умножая первое уравнение в (6) на p_k, получаем

$$(20)\qquad (\Sigma\lambda_i^{k+1} f_i'(x_k)+\Sigma\beta_j^{k+1} g_j'(x_k), p_k)\leqslant -[\Sigma\lambda_i^{k+1} f_i''(x_k)+\Sigma\beta_j^{k+1} g_j''(x_k)]p_k^2.$$

С учетом (20) из (17) следует

$$\psi_k(\bar{x}_k)=\Sigma\lambda_i^{k+1} f_i(x_k)+\Sigma\beta_j^{k+1} g_j(x_k)-\max_{\langle\Lambda,\beta\rangle\in Q_k}[\Sigma\lambda_i f_i(x_k)+\Sigma\beta_j g_j(x_k)]+$$

$$+(\Sigma\lambda_i^{k+1} f_i'(x_k)+\Sigma\beta_j^{k+1} g_j'(x_k), p_k)+$$

$$+\frac{1}{2}[\Sigma\lambda_i^{k+1} f_i''(x_k)+\Sigma\beta_j^{k+1} g_j''(x_k)]p_k^2\leqslant-\frac{m}{2}\|p_k\|^2.$$

Оценка (19) получена.

В силу (19), вместо (18) имеем неравенство

$$(21)\qquad \varphi(x_k+\alpha_k p_k)+NF(x_k+\alpha p_k)\leqslant\varphi(x_k)+NF(x_k)+$$

$$+\alpha\psi_k(\bar{x}_k)\left(1-\alpha\frac{c_k+\bar{c}_k}{m}\right),$$

из которого следует (7), если $\alpha_k\leqslant m(1-\varepsilon)/(c_k+\bar{c}_k)$ и если α_k удовлетворяет (12). Обоснование способа выбора α_k показано.

Процедура получения α_k (шаг 2) гарантирует конечное число дроблений единицы на шаге, причем, как нетрудно видеть,

$$(22)\qquad 1\geqslant\alpha_k\geqslant\min\left[\frac{1}{2}\frac{m(1-\varepsilon)}{(c_k+\bar{c}_k)}, \frac{\min(\delta,\gamma)}{2(2K+M\|p_k\|)\|p_k\|}, 1\right].$$

Таким образом, на каждом шаге происходит релаксация функции $\varphi(x)+NF(x)$, откуда $x_k\in D_N$. Но множество D_N ограничено, что легко следует из (4), значит, $\{x_k\}$ — ограниченная последовательность. Поэтому величина K в (22) равномерно ограничена. Величина $\|p_k\|$ также равномерно ограничена, так как в противном случае, умножая верхнее уравнение в (6) на p_k и используя (4), придем к противоречию. Поскольку $c_k+\bar{c}_k\leqslant 2(N+1)M$, то из сказанного вытекает, что правая часть в (22) ограничена снизу числом $\bar{\alpha}>0$, т. е. $\alpha_k \geqslant \bar{\alpha}>0$. Отсюда и из неравенства

$$\varphi(x_{k+1})+NF(x_{k+1}) \leqslant \varphi(x_k)+NF(x_k)-$$
$$-m\varepsilon\alpha_k\|p_k\|^2/2, \qquad 0<\varepsilon<1,$$

вытекающего из (7), (19), следует, что $p_k \to 0$ при $k\to\infty$, так как в противном случае $\varphi(x_k)+NF(x_k)\to-\infty$, что противоречит (4). Используя предложение 1, получаем, что $x_k \to x_*$. Первое утверждение теоремы доказано.

Так как $x_k\to x_*$, то без ограничения общности можно считать δ и γ выбранными так, что для достаточно больших k будет $I_k=R(x_*)$ и $J_k=\bar{R}(x_*)$ (предложение 2). Тогда $Q_k=Q_*$ — фиксированный компакт.

Пусть выполнены условия 2) теоремы. Легко видеть, что c_k, $\bar{c}_k\to 0$ при $k\to\infty$ и, в силу (22), найдется число $T>0$ такое, что $\alpha_k=1$ при $k\geqslant T$.

Для доказательства (8), (9) получим нижнюю оценку для $\psi_k(\bar{x}_k)$. Из (17) имеем при $\alpha_k=1$

$$\psi_k(\bar{x}_k)=\psi_k(x_{k+1})=\max_{\langle\Lambda,\beta\rangle\in Q_*} L_k(x_{k+1},\Lambda,\beta)-$$
$$-\max_{\langle\Lambda,\beta\rangle\in Q_*} L_k(x_k\ \Lambda,\beta)>\max_{\langle\Lambda^k,\beta^k\rangle\in Q_*}\{\Sigma\lambda_i^k[f_i(x_k)+(f_i'(x_k),p_k)]+$$
$$+\Sigma\beta_j^k[g_j(x_k)+(g_j'(x_k),p_k)]\}-\max_{\langle\Lambda,\beta\rangle\in Q_*} L_k(x_k,\Lambda,\beta).$$

Здесь и ниже суммирование по i, j производится по множествам $R(x_*)$, $\bar{R}(x_*)$. Разлагая в ряд Тейлора слагаемые в правой части и группируя, получаем

$$\psi_k(\bar{x}_k)>\max_{\langle\Lambda^k,\beta^k\rangle\in Q_*}\Big\{L_{k-1}(x_k,\Lambda^k,\beta^k)+(\Sigma\lambda_i^k[f_i'(x_{k-1})+$$
$$+f_i''(x_{k-1})p_{k-1}]+\Sigma\beta_j^k[g_j'(x_{k-1})+g_j''(x_{k-1})p_{k-1}],p_k)+$$
$$+([\Sigma\lambda_i^k R_{k-1}^{i}+\Sigma\beta_j^k\bar{R}_{k-1}^{j}]p_{k-1},p_k)+$$
$$+\frac{1}{2}(\Sigma\lambda_i^k\Phi_{k-1}^{i}+\Sigma\beta_j^k\bar{\Phi}_{k-1}^{j})p_{k-1}^2\Big\}-$$
$$-\max_{\langle\Lambda,\beta\rangle\in Q_*} L_{k-1}(x_k,\Lambda\ \beta)-\frac{1}{2}\max_{\langle\Lambda,\beta\rangle\in Q_*}[(\Sigma\lambda_i\Phi_{k-1}^{i}+\Sigma\beta_j\bar{\Phi}_{k-1}^{j})p_{k-1}^2], \tag{23}$$

где

$$R_{k-1}^{i}=f_i''(x_{k-1}^{\omega_i})-f_i''(x_{k-1}), \qquad \bar{R}_{k-1}^{j}=g_j''(x_{k-1}^{\omega_j})-g_j''(x_{k-1}),$$
$$x_{k-1}^{\omega_i}=x_{k-1}+\omega_i p_{k-1}, \qquad x_{k-1}^{\omega_j}=x_{k-1}+\omega_j p_{k-1}, \qquad 0\leqslant\omega_i,\quad \omega_j\leqslant 1,$$

а $L_{k-1}(x,\Lambda,\beta)$, Φ_{k-1}^{i} и т. д. соответствуют выражениям для $L_k(x,\Lambda,\beta)$, Φ_k^{i} и т. д. Так как в силу необходимого условия экстремума

$$\partial L_{k-1}(x_k,\Lambda^k,\beta^k)/\partial p_k\geqslant 0,$$

то из (23) вытекает

$$\begin{aligned}(24)\qquad &\psi_k(\bar{x}_k) > \frac{\partial L_{k-1}(x_k, \Lambda^k, \beta^k)}{\partial p_k} + \\ &+ \min_{\langle \Lambda^k, \beta^k \rangle \in Q_*} \left((\Sigma \lambda_i^k R_{k-1}^i + \Sigma \beta_j^k \bar{R}_{k-1}^j) p_{k-1}, p_k \right) + \\ &+ \frac{1}{2} \min_{\langle \Lambda^k, \beta^k \rangle \in Q_*} \left[(\Sigma \lambda_i^k \Phi_{k-1}^i + \Sigma \beta_j^k \bar{\Phi}_{k-1}^j) p_{k-1}^2 \right] - \\ &- \frac{1}{2} \max_{\langle \Lambda, \beta \rangle \in Q_*} \left[(\Sigma \lambda_i \Phi_{k-1}^i + \Sigma \beta_j \bar{\Phi}_{k-1}^j) p_{k-1}^2 \right] \geqslant \\ &\geqslant -r_{k-1} \|p_k\| \, \|p_{k-1}\| - d_{k-1} \|p_{k-1}\|^2,\end{aligned}$$

где

$$r_{k-1} = (N+1) \max_{i,j} (\|R_{k-1}^i\|, \|\bar{R}_{k-1}^j\|),$$

$$d_{k-1} = (N+1) \max (\|\Phi_{k-1}^i\|, \|\bar{\Phi}_{k-1}^j\|), \qquad i \in R(x_*), \quad j \in \bar{R}(x_*).$$

Из (24) и из (19) следует

$$(25)\qquad \begin{aligned}&m\|p_k\|^2/2 \leqslant r_{k-1} \|p_k\| \, \|p_{k-1}\| + d_{k-1} \|p_{k-1}\|^2, \\ &\|p_k\| \leqslant t_{k-1} \|p_{k-1}\|, \qquad t_{k-1} = [r_{k-1} + (r_{k-1}^2 + 2md_{k-1})^{1/2}]/m.\end{aligned}$$

Ясно, что если выполнены условия 2) теоремы, то r_{k-1}, $d_{k-1} \to 0$. Тогда $t_{k-1} \to 0$ и из (25) следует (8).

Если, наконец, L — константа Липшица для $f_i''(x)$, $g_j''(x)$, $i \in R(x_*)$, $j \in \bar{R}(x_*)$, то $r_{k-1}^2 < r_{k-1} \leqslant L(N+1) \|p_{k-1}\|$, $d_{k-1} \leqslant L(N+1) \|p_{k-1}\|$ и $t_{k-1} \leqslant C \|p_{k-1}\|^{1/2}$. Поэтому из (25) вытекает (9). Теорема доказана.

Остановимся на вопросе выполнения условий, содержащихся в предположении б.

Требование наличия внутренней точки для множества Ω_k может быть нарушено, если точка x_k достаточно далека от Ω. Однако если

$$x_k \in D = \{x_k : \min_{x \in \Omega} \|x - x_k\| \leqslant \bar{\rho}\},$$

где $\bar{\rho} > 0$ — достаточно малое число, множество Ω содержит внутреннюю точку $\bar{x}$, $g_j''(x)$ $(j \in J)$ непрерывны и выполняются (4), то для каждого множества Ω_k найдется внутренняя точка z_k такая, что

$$(26)\qquad g_j^k(z_k) \leqslant -\bar{\varepsilon} < 0 \quad \forall j \in J, \quad \forall x_k \in D,$$

где $\bar{\varepsilon} > 0$ — постоянная величина.

Для доказательства (26) достаточно воспользоваться неравенством $(g_j'(y), \bar{x} - y) \leqslant -d < 0$, $j \in \bar{R}(y)$, для любого

$$y \in \Gamma(\Omega) = \{y : \max_{j \in J} g_j(y) = 0\}$$

и условиями равномерной непрерывности функций. Мы опускаем эти простые, но достаточно громоздкие рассуждения.

Лемма. *Пусть выполнены сформулированные выше условия, при которых* (26) *имеет место. Тогда найдется число* $\rho>0$ *такое, что при выборе начального приближения* x_0 *из окрестности*

$$\rho(\Omega)=\{x_0\colon \min_{x\in\Omega}\|x-x_0\|\leqslant\rho\}\in D$$

указанное в предположении 6 *число* N *существует.*

Доказательство. Покажем сначала, что найдутся числа $\rho>0$ и $\bar N=\bar N(\rho)>0$ такие, что при $N_k\geqslant\bar N$ для любого $x_0\in\rho_1(\Omega)$

$$D_{N_k}=\{x\colon \varphi(x)+N_kF(x)\leqslant\varphi(x_0)+N_kF(x_0)\}\subset D.$$

В самом деле, если $F(x_0)=0$, то для точек $x\in D_{N_k}$

$$0\leqslant F(x)\leqslant[A-\varphi(x)]/N_k,\qquad A=\max_{x_0\in\Omega}\varphi(x_0),$$

т. е. $D_{N_k}\subset\omega=\{x\colon \varphi(x)\leqslant A\}$ — ограниченное множество. Так как последовательность функций $[A-\varphi(x)]/N_k$ сходится к нулю монотонно при $N_k\to\infty$ для любого фиксированного $x\in\omega$, то, по теореме Дини, сходимость будет равномерной, т. е. для любого $\nu>0$ найдется число $N(\nu)>0$ такое, что $F(x)\leqslant[A-\varphi(x)]/N_k\leqslant\nu$ при $x\in\omega$, если $N_k\geqslant N(\nu)$. Таким образом, $D_{N_k}\subset\{x\colon F(x)\leqslant\nu\}\cap\omega$ при $N_k\geqslant N(\nu)$. Отсюда и вытекает существование указанного $\bar N$, когда $x_0\in\Omega$.

Пусть теперь $x_0\notin\Omega$. Если при этом для точек $x\in D_{N_k}$ выполняется $F(x)=0$, то $x\in\Omega\subset D$ при любом $N_k>0$. Рассмотрим поэтому точки $x\in D_{N_k}$, для которых $F(x)>0$. Ясно, что для каждого x найдется точка $y\in\Gamma(\Omega)$ такая, что $x=y+t(y-\bar x)$, $t>0$. Так как $(g_j'(y),\bar x-y)<0$, $j\in\bar R(y)$, то $(g_j'(y),x-y)>0$. Поэтому с учетом (4) из разложения Тейлора следует

$$F(x)\geqslant\max_{j\in\bar R(y)}g_j(x)>m\|x-y\|^2/2.$$

С другой стороны, в силу (4) и ограниченности $\Gamma(\Omega)$ имеем $\varphi(x)\geqslant\geqslant\bar C_1-\bar C_2\|x-y\|+m\|x-y\|^2/2$, где

$$\bar C_2=\max_{i\in I}\max_{y\in\Gamma(\Omega)}\|f_i'(y)\|>0\ \text{и}\ \bar C_1=\min_{x\in E^n}\varphi(x)\leqslant\varphi(x_0).$$

Из условия $x\in D_{N_k}$ и оценок для $\varphi(x)$, $F(x)$ имеем неравенство

$$\bar C_1-\bar C_2\|x-y\|+m(N_k+1)\|x-y\|^2/2\leqslant\varphi(x_0)+N_kF(x_0),$$

решая которое относительно $\|x-y\|$, получаем

$$\|x-y\|\leqslant[2F(x_0)/m]^{1/2}+\mu(N_k,x_0),\tag{27}$$

где функции $\mu(N_k,x_0)$ непрерывно зависят от x_0 и $\mu(N_k,x_0)\to0$ монотонно при $N_k\to\infty$. Поэтому сходимость к нулю $\mu(N_k,x_0)$ будет равномерной на любом фиксированном компакте $\rho(\Omega)$. Выберем, в частности, ρ так, чтобы при $x_0 \in \rho(\Omega)$ выполнялось $[2F(x_0)/m]^{1/2} \leqslant \bar{\rho}/2$. Тогда найдется $\bar{N}=\bar{N}(\rho)$ такое, что из (27) вытекает $\|x-y\| \leqslant \bar{\rho}$, т. е. $D_{N_k} \subset D$ при $N_k \geqslant \bar{N}$.

Покажем теперь равномерную ограниченность $\Sigma\beta_j^{k+1}$ $(j \in J_k)$ в (5). Из теоремы Куна — Таккера о седловой точке для задачи (3) следует

$$\Sigma\lambda_i^{k+1} f_i^k(\bar{x}_k) \leqslant \Sigma\lambda_i^{k+1} f_i^k(z_k) + \Sigma\beta_j^{k+1} g_j^k(z_k), \tag{28}$$

где

$$\Sigma\lambda_i^{k+1}=1, \quad \lambda_i^{k+1} \geqslant 0, \quad \beta_j^{k+1} \geqslant 0, \quad i \in I_k, \quad j \in J_k.$$

Будем считать, что $x_0 \in \rho(\Omega)$ и $N_k > \bar{N}$, так что $x_k \in D_{N_k} \subset D$ и справедливо (26). Так как последовательность $\{x_k\}$ ограничена, то $\{z_k\}$ также ограничена, что вытекает из (26) и (4). Поэтому ограничена и последовательность $\{\bar{x}_k\}$ в силу (28). Воспользовавшись условием $\Sigma\beta_j^{k+1} g_j^k(z_k) \leqslant -\varepsilon\beta_j^{k+1}$, $j \in J_k$, и непрерывностью $f_i^k(x)$, получаем из (28)

$$\beta_j^{k+1} \leqslant \sum_{i \in I_k} \lambda_i^{k+1} [f_i^k(z_k) - f_i^k(\bar{x}_k)]/\varepsilon \leqslant \bar{C} \quad \forall j \in J_k,$$

где $\bar{C}>0$ — постоянная величина. Отсюда для любого $x_k \in D_{N_k} \subset D$ будет

$$\sum_{j \in J_k} \beta_j^{k+1} \leqslant \bar{M},$$

$\bar{M}$ — константа.

Взяв $N=\max\{\bar{N}, \bar{M}\}$, завершаем доказательство леммы. Доказанная лемма гарантирует выполнение предположения б лишь в достаточно малой окрестности множества Ω. Может случиться, конечно, что оно выполнится и в более широкой области, определяемой (26).

Для обеспечения выполнения (5) на практике можно рекомендовать прием, используемый в [2] (стр. 230), а именно: полагать

$$N=2\sum_{j \in J_k} \beta_j(x_k)$$

всякий раз, когда (5) не выполняется. При реализации алгоритма не требуется, таким образом, знания констант m, M, N: нужен сам факт их существования. Использование постоянных δ и γ позволяет учитывать вдали от решения не все, а часть функций $f_i(x)$, $g_j(x)$ при нахождении $\bar{x}_k$.

Отметим, что если $I=i_0$, $\Omega=E^n$, т. е. рассматривается задача безусловной минимизации, то метод совпадает с методом Ньютона с регулировкой шага [2]. В этом случае $d_{k-1}=0$, и тогда из (25) следуют известные оценки скорости сходимости метода Ньютона [2, 4]. Для задачи дискретного минимакса без ограничений ($\Omega=E^n$) описанный метод изучен в [5].

К сожалению, сейчас нет алгоритмов, позволяющих найти решение вспомогательной задачи (3) в общем случае за конечное число шагов. Поэтому в рассматриваемом методе вспомогательная задача решается менее эффективно по сравнению с методом Ньютона для задачи отыскания минимакса, рассмотренным в [6], или с методом Ньютона для задачи выпуклого программирования, использующим только линейную аппроксимацию ограничений [7]. Однако указанные методы в отличие от предложенного сходятся в любом случае локально и только в предположении регулярности точки минимума, обусловливающем невырожденность соответствующего якобиана. Это же замечание относится и к квазиньютоновским методам решения задач выпуклого программирования, сходящимся со сверхлинейной скоростью [7].

Поступила в редакцию 1.07.1977
Переработанный вариант 14.02.1978

Цитированная литература

1. *В. Ф. Демьянов, В. Н. Малоземов.* Введение в минимакс. М., «Наука», 1972.
2. *Б. Н. Пшеничный, Ю. М. Данилин.* Численные методы в экстремальных задачах. М., «Наука», 1975.
3. *А. Б. Певный.* Скорость сходимости некоторых методов нахождения минимакса. Кибернетика, 1972, № 4, 95–98.
4. *Е. С. Левитин, Б. Т. Поляк.* Методы минимизации при наличии ограничений. Ж. вычисл. матем. и матем. физ., 1966, 6, № 5, 787–823.
5. *В. М. Панин.* О методе Ньютона с регулировкой шага для задачи дискретного минимакса. В сб. «Теория оптимальных решений». Киев, ИК АН УССР, 1977, 46–55.
6. *В. Ф. Демьянов.* Ускорение сходимости при решении минимаксных задач. Ж. вычисл. матем. и матем. физ., 1972, 12, № 1, 51–60.
7. Численные методы условной оптимизации. Ред. *Ф. Гилл, У. Мюррей.* М., «Мир», 1977.