УДК 519.3

О РЕШЕНИИ НЕЛИНЕЙНЫХ НЕРАВЕНСТВ И ЭКСТРЕМАЛЬНЫХ ЗАДАЧ МЕТОДАМИ ПОВЫШЕННОЙ ТОЧНОСТИ

В.А.Булавский

На возможность применения метода Ньютона для решения экстремальных задач и нелинейных неравенств уже было обращено внимание [1,2,3,6]. В этой статье подобные методы изучаются в более общей ситуации, чем в указанных работах, и при несколько ослабленных предположениях. Строятся также методы типа модифицированного метода Ньютона, позволяющие избежать вычисления производных на каждом шаге, а также дающие возможность в ряде случаев на очередном шаге решение соответствующей линеаризованной задачи начинать с подправки решения аналогичной задачи предыдущего шага. Работа является развитием ранее опубликованных результатов автора [4].

§ 1. Общая схема получения минимизирующей последовательности

Во избежание повторений формулировок и доказательств для различных методов, рассматриваемых в следующих параграфах, мы здесь сформулируем в общем виде итеративные процессы, которые будут использованы в дальнейшем, и докажем две теоремы о сходимости.

Пусть на замкнутом выпуклом множестве Q банахова пространства X задана вещественная непрерывная функция φ, для которой требуется построить минимизирующую последовательность, то есть такую последовательность $\{x_n\} \subset Q$, что $\varphi(x_{n+1}) \geq \varphi(x_n)$ для всех n и $\lim \varphi(x_n) = \inf \{\varphi(x) \mid x \in Q\}$. В общем случае не предполагается, что этот инфимум конечный.

Предположим, что для каждого $x \in Q$ заданы два множества $B_0(x)$ и $B(x)$, содержащиеся в $Q - x$, а также непрерывная вещественная функция ψ, определенная на $Q \times X$. Обозначим через $\mu(x, \xi)$ наибольшее λ из $[0,1]$, при котором $\varphi(x) - \varphi(x + \mu\xi) \geq \mu\psi(x, \xi)$ для всех $\mu \in [0, \lambda]$. Пусть $-\infty \leq a_0 = \inf \{\varphi(x) \mid x \in Q\} < +\infty$ и при всяком $\bar{x} \in Q$ таком, что $\varphi(\bar{x}) > a_0$ выполнены следующие три условия.

1^0. $\delta(\varphi(\bar{x})) = \inf\{\psi(x, \xi) \mid x \in Q, \varphi(x) \geq \varphi(\bar{x}), \xi \in B_0(x) \cup B(x)\} > 0$;

2^0. $M(\varphi(\bar{x})) = \sup\{\|\xi\| \mid x \in Q, \varphi(x) \leq \varphi(\bar{x}), \xi \in B_0(x)\} < +\infty$;

3^0. $\nu(\bar{x}) = \lim\limits_{\varepsilon \to +0} \inf\{\mu(x, \xi) \mid \xi \in B_0(x), x \in Q, \|x - \bar{x}\| \leq \varepsilon\} > 0$.

Возьмем произвольно $x_0 \in Q$ и построим последовательность $\{x_n\}$, где $x_{n+1} = x_n + \lambda_n \xi_n$, следующим образом. Выберем $\xi \in B(x_n)$ и, если окажется $\varphi(x_n) - \varphi(x_n + \xi) \geq \psi(x_n, \xi)$, положим $\xi_n = \xi$ и $\lambda_n = 1$. В противном случае выберем $\xi_n \in B_0(x_n)$, а λ_n (длину шага) определим одним из двух способов, приводимых ниже.

I. В качестве λ_n возьмем наибольшее из чисел q^k, $k = 0, 1, \ldots$, при котором окажется $\varphi(x_n) - \varphi(x_n + \lambda_n \xi_n) \geq \lambda_n \psi(x_n, \xi_n)$. Здесь число $q \in (0, 1)$ и не зависит от n.

II. Если $\varphi(x_n) - \varphi(x_n + \xi_n) \geq \psi(x_n, \xi_n)$, то положим $\lambda_n = 1$. В противном случае в качестве λ_n возьмем наибольший корень уравнения $\varphi(x_n) - \varphi(x_n + \lambda\xi_n) = \lambda\psi(x_n, \xi_n)$.

Заметим, что если $\varphi(x_n) > a_0$ и $B_0(x_n) \neq \emptyset$, то в любом случае окажется $\lambda_n > 0$. Действительно, $\mu(x_n, \xi_n) \geq \nu(x_n) > 0$, и потому в способе I получим $\lambda_n \geq q\,\nu(x_n) > 0$, а в способе II $\lambda_n \geq \nu(x_n) > 0$. Заметим также, что последовательность $\{\varphi(x_n)\}$ убывающая.

ТЕОРЕМА I. Если последовательность $\{x_n\}$ бесконечная, то $\lim \varphi(x_n) \leq a_0$. Если к тому же функция φ ограничена снизу на Q, а $M(a) \leq C\delta(a)$ при $\varphi(x_0) \geq a > a_0$, где константа C не зависит от a, то последовательность $\{x_n\}$ сходится.

ДОКАЗАТЕЛЬСТВО. Предположим, что $\lim \varphi(x_n) = a > a_0$. Тогда $\psi(x_n, \xi_n) \geq \delta(a) > 0$, и поэтому для некоторого N

окажется $\xi_n \in B_0(x_n)$ при $n \geq N$, а λ_n выбрано по способу I или по способу II. Следовательно, $\varphi(x_n)-\varphi(x_{n+1}) > \geq \lambda_n \psi(x_n,\xi_n) \geq \lambda_n \delta(a)$ и ряд $\sum \lambda_n$ сходится. Таким образом, $\lim \lambda_n = 0$, и при достаточно больших n

$$\varphi(x_n)-\varphi(x_n+\mu_n\xi_n) < \mu_n \psi(x_n,\xi_n),$$

где $\mu_n = \lambda_n / q$. Так как $\|\xi_n\| \leq M(\varphi(x_0))$ при $n > N$, то ряд $\sum \lambda_n \xi_n$ сходится, то есть существует $\lim x_n = \bar{x}$. Согласно условию 3^0 для достаточно больших n окажется $\varphi(x_n)-\varphi(x_n+\mu\xi_n) \geq \mu\psi(x_n,\xi_n)$ при всех $\mu \in [0, \nu(\bar{x})/2]$ и, следовательно, $\mu_n > \nu(\bar{x})/2$, то есть $\lambda_n > q\,\nu(\bar{x})/2$. Полученное противоречие показывает, что $\lim \varphi(x_n) \leq a_0$.

Пусть теперь $a_0 > -\infty$ и $M(a) \leq C\delta(a)$ при $a \geq a_0$. Тогда $C(\varphi(x_n)-\varphi(x_{n+1})) \geq C\lambda_n \psi(x_n,\xi_n) \geq \lambda_n \|\xi_n\|$, откуда и следует сходимость последовательности $\{x_n\}$. Теорема доказана.

Пусть теперь помимо способов I и II определения длины шага допускается еще способ III, состоящий в том, что при $\varphi(x_n)- -\varphi(x_n+\xi_n) \geq \psi(x_n,\xi_n)$ полагаем $\lambda_n = 1$, а в противном случае в качестве λ_n берем точку минимума функции $\varphi(x_n+\lambda\xi_n)$ на отрезке $[0,1]$. Если $\varphi(x_n) > a_0$ и $B_0(x_n) \neq \emptyset$, то

$$\min\{\varphi(x_n+\lambda\xi_n) \mid \lambda \in [0,1]\} \leq \varphi(x_n+\nu(x_n)\xi_n) \leq \varphi(x_n)-\nu(x_n)\delta(\varphi(x_n)).$$

Относительно последовательности $\{x_n\}$, полученной таким образом, справедлива следующая теорема.

ТЕОРЕМА 2. Если последовательность $\{x_n\}$ имеет предельную точку $\bar{x}$ (скажем, множество Q компактно), то $\varphi(\bar{x}) = a_0$. Если к тому же $M(a) \leq C\cdot\delta(a)$ при $\varphi(x_0) \geq a \geq a_0$, то $\bar{x} = \lim x_n$.

ДОКАЗАТЕЛЬСТВО. Так как последовательность $\{\varphi(x_n)\}$ убывающая, то $\varphi(\bar{x}) = \lim \varphi(x_n)$ и $\varphi(x_n) \geq \varphi(\bar{x})$. Если $\varphi(\bar{x}) > a_0$, то существует такая последовательность $\{x_{n_k}\}$, что $\mu(x_{n_k},\xi_{n_k}) \geq \nu(\bar{x})/2 > 0$. Оценим убывание функции φ на шаге с номером n_k. Если сразу оказалось, что $\varphi(x_{n_k})- -\varphi(x_{n_k}+\xi_{n_k}) \geq \psi(x_{n_k},\xi_{n_k})$ и $\lambda_{n_k} = 1$, то убывание функции φ на этом шаге не меньше $\delta(\varphi(\bar{x}))$. Если λ_{n_k} выбрано по способу I или II, то $\lambda_{n_k}/q \geq \nu(\bar{x})/2$. Поэтому убывание φ на данном шаге не меньше $q\,\nu(\bar{x})\,\delta(\varphi(\bar{x}))/2$. Наконец, если λ_{n_k} определено по способу III, то $\varphi(x_{n_k}+\lambda_{n_k}\xi_{n_k}) \leq \leq \varphi(x_{n_k})-\nu(\bar{x})\,\delta(\varphi(\bar{x}))/2$. В любом случае на шаге с номером

n_k функция φ убывает не меньше чем на $q^{\nu(\bar{x})}\delta^{\circ}(\varphi(\bar{x}))/2$, что противоречит конечности $\varphi(\bar{x})$. Это показывает, что $\varphi(\bar{x}) = a_0$. Сходимость последовательности $\{x_n\}$ в случае, когда $M(a) \le C\delta(a)$, получается так же, как в теореме I. Теорема доказана.

§ 2. Решение нелинейных неравенств

Пусть X и Y — вещественные банаховы пространства, причем в Y выделен замкнутый конус положительных элементов K_y, и пусть $T: X \to Y$ — нелинейный дифференцируемый по Фреше оператор, производная которого удовлетворяет условию Липшица с показателем α:

$$\| T'(x_1) - T'(x_2) \| \le \mathcal{L} \| x_1 - x_2 \|^{\alpha}.$$

Нас будет интересовать решение следующей системы:

$$\begin{aligned} T(x) &\le 0, \\ x &\in Q, \end{aligned} \tag{I}$$

где Q - замкнутое выпуклое множество в X. В пространстве Y определим сублинейный функционал h равенством:

$$h(y) = \sup\{g(y) \mid g \in K_y^*, \|g\| \le 1\}, y \in Y.$$

Тогда неравенство (I) можно заменить системой

$$\begin{aligned} h(T(x)) &= 0, \\ x &\in Q. \end{aligned}$$

При построении методов типа метода Ньютона по системе (I) строится линеаризованная система

$$\begin{aligned} T(x) + T'(x)\xi &\le 0, \\ x + \xi &\in Q, \end{aligned}$$

причем направление ξ среди решений этой системы следует выбирать по возможности с меньшей нормой. Мы предположим, что направление смещения выбирается среди решений приближенной линеаризованной системы

$$\begin{aligned} h(T(x) + T'(x)\xi) &\le q_1 \min\{h(T(x)), [h(x))]^{\tau(1+\alpha)}\} \\ x + \xi &\in Q, \\ \|\xi\| &\le \mathcal{D}[h(T(x))]^{\tau}, \end{aligned} \tag{2}$$

где константы $\mathcal{D}$, τ и q_1 выбраны так, что $0 \le \tau \le 1$, $0 \le q_1 < 1$, $\mathcal{D} > 0$.

Обозначим через $B_0(x)$ множество решений системы (2) и положим $B(x) = B_0(x)$ при всех $x \in Q$. Положим также $\varphi(x) = h(T(x))$, $\psi(x,\xi) = (1-q_2)\varphi(x)$, где $q_1 < q_2 < 1$. Так как $\varphi(x) \ge 0$, то $\inf\{\varphi(x) \mid x \in Q\} \ge 0$ и равен нулю, если система (I) совместна. При $a > 0$ имеем: $\delta(a) = (1-q_2)\,a > 0$, $M(a) \le \mathcal{D}a^{\tau} < +\infty$. Кроме того, при $x \in Q$ и $\xi \in B_0(x)$ получаем:

$$\varphi(x+\lambda\xi) = h\left(T(x) + \lambda\int_0^1 T'(x+\theta\lambda\xi)\xi\, d\theta\right) =$$

$$= h\left((1-\lambda)T(x) + \lambda(T(x)+T'(x)\xi) + \lambda\int_0^1 (T'(x+\theta\lambda\xi) - T'(x))\xi\, d\theta\right) \le$$

$$\le (1-\lambda)\varphi(x) + \lambda q_1 \min\{\varphi(x), [\varphi(x)]^{\tau(1+\alpha)}\} + \frac{L}{1+\alpha}\|\xi\|^{1+\alpha}\lambda^{1+\alpha},$$

или, учитывая оценку $\|\xi\|$ из (2),

$$\varphi(x+\lambda\xi) \le (1-\lambda)\varphi(x) + \lambda q_1 \min\{\varphi(x), [\varphi(x)]^{\tau(1+\alpha)}\} + \frac{L}{1+\alpha}\lambda^{1+\alpha}\mathcal{D}^{1+\alpha}[\varphi(x)]^{\tau(1+\alpha)}. \quad (3)$$

Поэтому неравенство $\varphi(x) - \varphi(x+\lambda\xi) \ge \lambda\psi(x,\xi)$ окажется выполненным, если

$$q_2\varphi(x) - q_1 \min\{\varphi(x), [\varphi(x)]^{\tau(1+\alpha)}\} - \lambda^{\alpha}\frac{L}{1+\alpha}\mathcal{D}^{1+\alpha}[\varphi(x)]^{\tau(1+\alpha)} \ge 0.$$

Таким образом,

$$\nu(x) \ge \min\left\{1;\ \left[\frac{(q_2\varphi(x) - q_1\min\{\varphi(x), [\varphi(x)]^{\tau(1+\alpha)}\})(1+\alpha)}{L\,\mathcal{D}^{1+\alpha}[\varphi(x)]^{\tau(1+\alpha)}}\right]^{1/\alpha}\right\}.$$

Построим последовательность $\{x_n\}$ так, как указано в первом параграфе, используя способы I и II выбора длины шага, и назовем полученный метод методом Ньютона для задачи (I). Согласно теореме I, либо последовательность $\{x_n\}$ оборвется на некотором шаге с номером m ввиду несовместности системы (2) при $x = x_m$, либо $\lim \varphi(x_n) = 0$. Как здесь, так и повсюду в дальнейшем мы предполагаем, что система (I) совместна. Даже и в этом случае, если неравенство (I) не является выпуклым, линеаризованное неравенство в (2) может оказаться неразрешимым на каком-нибудь шаге. Ниже при некоторых дополнительных условиях мы установим разрешимость системы (2) в достаточной близости от множества решений системы (I). Если $\tau = 1$

то $M(a) \le \mathcal{D}a \le (\mathcal{D}/(1-q_2))\delta(a)$, и последовательность $\{x_n\}$ сходится к некоторой точке $\bar{x}$, причем по непрерывности φ окажется $\varphi(\bar{x})=0$, то есть $\bar{x}$ является решением системы (I). Рассмотрим, однако, вопрос о сходимости подробнее. Пусть N настолько велико, что $\varphi(x_N)<1$. Тогда при $n \ge N$ и $\tau(1+\alpha) \ge 1$

$$\nu(x_n) \ge \min\left\{1;\left[\frac{(q_2-q_1[\varphi(x_n)]^{\tau(1+\alpha)-1})(1+\alpha)}{\mathcal{L}\mathcal{D}^{1+\alpha}[\varphi(x_n)]^{\tau(1+\alpha)-1}}\right]^{1/\alpha}\right\}. \quad (4)$$

Если $\tau(1+\alpha)=1$, то при $n \ge N$

$$\lambda_n \ge q\bar{\lambda} = q\min\left\{1;\ \left[\frac{(q_2-q_1)(1+\alpha)}{\mathcal{L}\mathcal{D}^{1+\alpha}}\right]^{1/\alpha}\right\}.$$

Но $\varphi(x_n)-\varphi(x_{n+1}) \ge \lambda_n(1-q_2)\varphi(x_n) \ge q\bar{\lambda}(1-q_2)\varphi(x_n)$, так что $\varphi(x_{n+1}) \le (1-q\bar{\lambda}(1-q_2))\varphi(x_n)$ и $\varphi(x_{N+m}) \le \bar{q}^m\varphi(x_N)$, где $\bar{q}=1-q\bar{\lambda}(1-q_2) \in (0,1)$. При этом $\|\xi_{N+m}\| \le \le \mathcal{D}\bar{q}^{m/1+\alpha}[\varphi(x_N)]^{1/1+\alpha}$. Поэтому последовательность $\{x_n\}$ сходится к некоторому $\bar{x}$, причем $\|x_n-\bar{x}\| \le \mathcal{D}\sum_{k=n}^{\infty}[\varphi(x_k)]^{1/1+\alpha} \le \le \frac{\mathcal{D}}{1-\bar{q}^{1/1+\alpha}}[\varphi(x_n)]^{1/1+\alpha}$. Таким образом, получаем следующую характеристику скорости сходимости:

$$\overline{\lim}\ \frac{\varphi(x_{n+1})}{\varphi(x_n)} \le 1-q\bar{\lambda}(1-q_2), \quad (5)$$

$$\overline{\lim}\ \frac{\|x_n-\bar{x}\|}{[\varphi(x_n)]^{\tau}} \le \frac{\mathcal{D}}{1-(1-q\bar{\lambda}(1-q_2))^{1/1+\alpha}} \quad (6)$$

$$\bar{\lambda}=\min\left\{1;\left[\frac{(q_2-q_1)(1+\alpha)}{\mathcal{L}\mathcal{D}^{1+\alpha}}\right]^{1/\alpha}\right\}. \quad (7)$$

Если $\tau(1+\alpha)>1$, то можно взять N настолько большим, что правая часть в (4) будет равна единице при $n > N$ и, следовательно, $\lambda_n=1$. Положив в (3) $\lambda=1$, $x=x_n$, $\xi=\xi_n$, и учитывая, что $\varphi(x_n)<1$ при $n>N$, получим

$$\varphi(x_{n+1}) \le q_1[\varphi(x_n)]^{\tau(1+\alpha)} + \frac{\mathcal{L}}{1+\alpha}\mathcal{D}^{1+\alpha}[\varphi(x_n)]^{\tau(1+\alpha)},$$

или

$$\varphi(x_{n+1}) / [\varphi(x_n)]^{\tau(1+\alpha)} \leq q_1 + \frac{L}{1+\alpha} \mathcal{D}^{1+\alpha}. \quad (8)$$

Так как $\varphi(x_n)$ стремится к нулю, то ввиду оценки для $\|\xi_n\|$ из (2) последовательность $\{x_n\}$ сходится к решению $\bar{x}$ системы (I), причем

$$\overline{\lim} \frac{\|x_n - \bar{x}\|}{[\varphi(x_n)]^{\tau}} \leq \mathcal{D}. \quad (9)$$

Мы доказали следующую теорему.

ТЕОРЕМА 3. Если последовательность $\{x_n\}$ бесконечная (то есть $B(x_n) \neq \emptyset$ при всех n), то $\lim h(T(x_n)) = 0$. Если $\tau(1+\alpha) \geq 1$, то существует $\lim x_n = \bar{x}$, являющийся решением системы (I). При этом

а) если $\tau(1+\alpha) = 1$, то сходимость имеет скорость геометрической прогрессии, и эта скорость определяется формулами (5), (6), (7);

б) если $\tau(1+\alpha) > 1$, то скорость сходимости имеет порядок $\tau(1+\alpha)$ и определяется формулами (8), (9).

ЗАМЕЧАНИЕ I. Если X – вещественное гильбертово пространство, а Q совпадает со всем X, то для решения линеаризованной системы (2) может быть использован метод, аналогичный описанному [5] для конечномерного случая.

Предположим, что множество Q_0 решений системы (I) не пусто, и что на множестве $\bar{Q} = \{x \in Q \setminus Q_0 \mid \varphi(x) \leq \varphi(x_0)\}$ оператор T имеет ограниченное сжатие. Именно, существует такая постоянная $\sigma > 0$, что при $x \in \bar{Q}$ и $\delta = \inf\{\|z - x\| \mid z \in Q_0\}$ шар $S_{\sigma\delta}(T(x)) = \{y \in Y \mid \|y - T(x)\| < \sigma\delta\}$ не пересекается с конусом $(-K_y)$. В этом случае (так как $\delta > 0$ в силу замкнутости Q_0) существует такой положительный функционал g, что $\|g\| = 1$ и $\inf\{g(y) \mid \|y - T(x)\| \leq \sigma\delta\} \geq 0$. При этом $g(T(x)) \geq \sigma\delta$, то есть $\varphi(x) = h(T(x)) \geq \sigma\delta$. С другой стороны, для $\varepsilon > 0$ существует $\bar{x} \in Q_0$, при котором $\|x - \bar{x}\| \leq (1+\varepsilon)\delta$. Тогда

$$h(T(x) + T'(x)(\bar{x} - x)) = h(T(\bar{x}) + \int_0^1 [T'(\bar{x} + \lambda(x - \bar{x})) -$$

$$- T'(x)] d\lambda) \leq \frac{L}{1+\alpha} \|x - \bar{x}\|^{1+\alpha} \leq \frac{L}{1+\alpha} \left[\frac{1+\varepsilon}{\sigma}\right]^{1+\alpha} [\varphi(x)]^{1+\alpha}.$$

В то же время $\|x-\bar{x}\| \le \frac{1+\varepsilon}{\sigma}\varphi(x)$. Таким образом, можно гарантировать разрешимость системы (2) при любом $\tau < 1$ и достаточно малом $\varphi(x_0)$. Именно, достаточно x_0 выбрать так, что $\varphi(x_0) < \min\{\beta^{1/1-\tau}, \beta^{\frac{1+\alpha}{\tau}}\}$, где $\beta = \sigma\left[\frac{q_1(1+\alpha)}{\mathcal{L}}\right]^{1/1+\alpha}$. При этом следует взять $\mathcal{D} > [\varphi(x_0)]^{1-\tau}/\sigma$. Предположение о малости $\varphi(x_0)$ не является, конечно, необходимым для разрешимости системы (2) на каждом шаге. Если дополнительно предположить, что оператор T выпуклый, то $T(x)+T'(x)(\bar{x}-x) \le$ $\le T(\bar{x}) \le 0$. Поэтому разрешимость системы (2), если система (I) совместна, можно гарантировать при любой начальной точке x_0, любом $\tau \le 1$ и достаточно большом $\mathcal{D}$. Последнее можно уточнять в процессе счета, если на некотором шаге окажется, что оно было взято слишком малым.

Видоизменим несколько метод. Обозначим через $B(x,z)$ множество решений системы

$$
\begin{aligned}
h(T(x)+T'(z)\xi) &\le q_1 h(T(x)),\\
x+\xi &\in Q, \qquad\qquad (10)\\
\|\xi\| &\le \mathcal{D}h(T(x)).
\end{aligned}
$$

В качестве $B_0(x)$, как и раньше, возьмем множество решений системы (2), а в качестве $B(x)$ - объединение $B(x,z)$ по $z \in X$. Выберем некоторое $x_0 \in Q$, положим $z_0 = x_0$ и построим последовательности $\{x_n\}$ и $\{z_n\}$ следующим образом. Найдем $\xi \in B(x_n, z_n)$. Если $B(x_n, z_n)$ не пусто и для найденного ξ оказалось $\varphi(x_n) - \varphi(x_n+\xi) \ge \psi(x_n,\xi) = (1-q_2)\varphi(x_n)$, то положим $\xi_n = \xi$, $\lambda_n = 1$, $x_{n+1} = x_n + \xi$, $z_{n+1} = z_n$. В противном случае выберем $\xi_n \in B_0(x_n)$ и совершим шаг метода Ньютона для системы (I). При этом положим $z_{n+1} = x_n$. Полученный метод будем называть модифицированным методом Ньютона для задачи (I). По теореме I снова заключаем, что либо последовательность $\{x_n\}$ оборвется, либо $\lim \varphi(x_n) = 0$. В последнем случае, если $\tau(1+\alpha) \ge 1$, на каждом шаге с достаточно большим номером либо $\varphi(x_{n+1}) \le q_2\varphi(x_n)$, либо $\varphi(x_{n+1}) \le$ $\le (1-q\bar{\lambda}(1-q_2))\varphi(x_n)$, где $\bar{\lambda}$ определено формулой (7). Поэтому последовательность $\{x_n\}$ сходится к некоторому решению $\bar{x}$ системы (I).

Предположим, что оператор T имеет ограниченное сжатие, и покажем, что последовательность $\{z_n\}$ стабилизируется, начиная с некоторого места. Действительно, в противном случае $\lim z_n = \bar{x}$. Возьмем $\varepsilon > 0$ и выберем $\tilde{x}_n \in Q_0$ так, что $\|x_n - \tilde{x}_n\| \le (1+\varepsilon)\cdot \inf\{\|x_n - z\| \mid z \in Q_0\}$. Тогда

$$h(T(x_n) + T'(z_n)(\tilde{x}_n - x)) \le$$

$$\le \mathcal{L}\|\tilde{x}_n + \theta(x_n - \tilde{x}_n) - z_n\|^{\alpha}\|\tilde{x}_n - x_n\| \le \frac{\mathcal{L}(1+\varepsilon)^{1+\alpha}}{\sigma}(\|x_n - \bar{x}\| + \|x_n - z_n\|)^{\alpha}\varphi(x_n).$$

Так как $\lim\|x_n - \bar{x}\| = \lim\|x_n - z_n\| = 0$, то при достаточно больших n множества $B(x_n, z_n)$ не пустые. Если $\xi \in B(x_n, z_n)$, то $\varphi(x_n + \xi) = h(T(x_n + \xi)) = h(T(x_n) + T'(z_n)\xi + \int_0^1 (T'(x_n + \theta\xi) - T'(z_n))\xi\, d\theta) \le (q_1 + \mathcal{L}\mathcal{D}\|x_n + \bar{\theta}\xi - z_n\|^{\alpha})\,\varphi(x_n)$.
Так как $q_1 < q_2$, а $\|x_n + \bar{\theta}\xi - z_n\|^{\alpha} \le (\bar{\theta}\mathcal{D}\varphi(x_n) + \|x_n - z_n\|)^{\alpha}$, то найдется такое n_0, что $\varphi(x_n + \xi) \le q_2\,\varphi(x_n)$ при $n \ge n_0$. При этих n получим, что $z_n = z_{n_0}$. Таким образом, последовательность $\{z_n\}$ стабилизируется, и $\varphi(x_{n+1}) \le q_2\,\varphi(x_n)$ при $n \ge n_0$, а $\|x_n - \bar{x}\| \le \frac{\mathcal{D}}{1 - q_2}\varphi(x_n)$. Мы доказали следующую теорему.

ТЕОРЕМА 4. Если последовательность $\{x_n\}$ бесконечная и $\tau(1+\alpha) \ge 1$, то последовательность $\{x_n\}$ сходится к решению системы (I). Если к тому же оператор T имеет ограниченное сжатие на $\bar{Q}$, то в модифицированном методе Ньютона, начиная с некоторого шага $\lambda_n = 1$ и стабилизируется последовательность $\{z_n\}$ (то есть не нужно вычислять $T'(x_n)$), и сходимость имеет скорость геометрической прогрессии со знаменателем q_2.

ЗАМЕЧАНИЕ 2. Если $z_{n+1} = z_n$, то решение системы (I0) для очередного шага в некоторых случаях можно облегчить, используя информацию, полученную при решении этой системы на предыдущем шаге.

§ 3. Задача выпуклого программирования

Пусть Q — непустое компактное выпуклое множество в вещественном гильбертовом пространстве X, а φ — выпуклая дважды дифференцируемая на всем X функция. Требуется решить задачу

$$\min \{\varphi(x) \mid x \in Q\} = \varphi_{min}. \tag{11}$$

Предположим, что φ'' удовлетворяет на Q условию Липшица с показателем α :

$$\|\varphi''(x_1) - \varphi''(x_2)\| \le \mathcal{L} \|x_1 - x_2\|^{\alpha}.$$

Обозначим через x^* для каждой точки $x \in Q$ ближайшую к ней точку минимума функции φ на Q. Обозначим также через H линейную оболочку множества $\{x - y \mid \varphi(x) = \varphi(y) = \varphi_{min},\ x \in Q,\ y \in Q\}$ и сделаем следующие два предположения о невырожденности.

1^{00}. Существует $\gamma > 0$ такое, что для всякого $x \in Q$ оказывается $\varphi'(x^*)(x - x^*) + \varphi''(x^*)(x - x^*)(x - x^*) \ge \gamma \|x - x^*\|^2$.

2^{00}. Если $x_1 - x_2 \in H$, то $\varphi(x_1) = \varphi(x_2)$.

Для x и z из Q через $B(x, z)$ обозначим множество решений ξ задачи

$$\min \{\varphi'(x)\xi + \tfrac{1}{2}\varphi''(z)\xi\xi \mid x + \xi \in Q\}. \tag{12}$$

В случае, когда Q является многогранником, для задачи (12) разработаны точные методы решения, но могут быть и другие случаи, когда задача (12) решается проще, чем задача (11).

ЛЕММА I. Если $\|z - z^*\|^{\alpha} \le \gamma/\mathcal{L}$ и $\xi \in B(x, z)$, то при $\tilde{x} = x + \xi$

$$\|\tilde{x} - \tilde{x}^*\| \le \frac{\frac{\mathcal{L}}{1+\alpha}\|x - x^*\|^{\alpha} + \mathcal{L}\|z - z^*\|^{\alpha}}{\gamma - \mathcal{L}\|z - z^*\|^{\alpha}} \|x - x^*\|, \tag{13}$$

а если $z = x$, то

$$\|\tilde{x} - \tilde{x}^*\| \le \frac{\frac{\mathcal{L}}{1+\alpha}\|x - x^*\|^{1+\alpha}}{\gamma - \mathcal{L}\|x - x^*\|^{\alpha}}. \tag{14}$$

ДОКАЗАТЕЛЬСТВО. Поскольку $\xi \in B(x, z)$, то $(\varphi'(x) + \varphi''(z)\xi)\ell \ge 0$ при $x + \xi + \ell \in Q$. В частности, $(\varphi'(x) + \varphi''(z)\xi)(\tilde{x}^* - \tilde{x}) \ge 0$. Далее, учитывая условие 1^0 невырожденности и условие 2^0 постоянства φ вдоль направлений из H, получим

$$\gamma\|\tilde{x} - \tilde{x}^*\|^2 \le \varphi'(x^*)(\tilde{x} - \tilde{x}^*) + \varphi''(x^*)(\tilde{x} - \tilde{x}^*)(\tilde{x} - \tilde{x}^*) =$$

$$= \varphi'(x)(\tilde{x}-\tilde{x}^*) + \varphi''(z)\xi(\tilde{x}-\hat{x}^*) + \int_0^1 \varphi''(x+\lambda(x^*-x))(x^*-x)(\tilde{x}-\tilde{x}^*)\,d\lambda -$$

$$- \varphi''(z)\xi(\tilde{x}-\tilde{x}^*) + \varphi''(z^*)(\tilde{x}-\tilde{x}^*)(\tilde{x}-\tilde{x}^*) \leq$$

$$\leq \int_0^1 \varphi''(x+\lambda(x^*-x))(x^*-x)(\tilde{x}-\tilde{x}^*)\,d\lambda -$$

$$- \varphi''(z)(x^*-x)(\tilde{x}-\tilde{x}^*) + L\|z-z^*\|^{\alpha}\cdot\|\tilde{x}-\tilde{x}^*\|^2.$$

Если $z = x$, то

$$\int_0^1 [\varphi''(x+\lambda(x^*-x)) - \varphi''(z)](x^*-x)(\tilde{x}-\tilde{x}^*)\,d\lambda \leq \frac{L}{1+\alpha}\|x-x^*\|^{1+\alpha}\|\tilde{x}-\tilde{x}^*\|$$

и получаем (14). В противном случае

$$\int_0^1 \varphi''(x+\lambda(x^*-x))(x^*-x)(\tilde{x}-\tilde{x}^*)\,d\lambda - \varphi''(z)(x^*-x)(\tilde{x}-\tilde{x}^*) =$$

$$= \int_0^1 [\varphi''(x+\lambda(x^*-x)) - \varphi''(x^*)](x^*-x)(\tilde{x}-\tilde{x}^*)\,d\lambda + [\varphi''(z^*)-\varphi''(z)](x^*-x)(\tilde{x}-\tilde{x}^*) \leq$$

$$\leq \frac{L}{1+\alpha}\|x-x^*\|^{1+\alpha}\|\tilde{x}-\tilde{x}^*\| + L\|z-z^*\|^{\alpha}\|x-x^*\|\cdot\|\tilde{x}-\tilde{x}^*\|,$$

и получаем (13). Лемма доказана.

ЛЕММА 2. Если $\|x-x^*\| \leq \varepsilon$, $\|z-z^*\| \leq \varepsilon$ и $\xi \in B(x,z)$, то при $q_2 < \frac{1}{2}$ и достаточно малых положительных ε окажется

$$\varphi(x) - \varphi(x+\xi) + q_2\,\varphi'(x)\xi \geq 0.$$

ДОКАЗАТЕЛЬСТВО. Положим $x + \xi = \tilde{x}$. Так как $\varphi'(x)\xi + \varphi''(z)\xi\xi \leq 0$, то

$$\varphi(x) - \varphi(x+\xi) + \frac{1}{2}\varphi'(x)\xi \geq \frac{1}{2}[\varphi''(z) - \varphi''(x+\theta\xi)]\xi\xi =$$

$$\frac{1}{2}[\varphi''(z) - \varphi''(z^*) + \varphi''(x^*+\theta(\tilde{x}^*-x^*)) - \varphi''(x+\theta(\tilde{x}-x))](\tilde{x}-\tilde{x}^*+x^*-x)(\tilde{x}-\tilde{x}^*+x^*-x) \geq$$

$$\geq -\frac{L}{2}(\|z-z^*\|^{\alpha} + (\|x-x^*\| + \|\tilde{x}-\tilde{x}^*\|)^{\alpha})(\|\tilde{x}-\tilde{x}^*\| + \|x-x^*\|)^2.$$

Согласно лемме I при достаточно малом ε окажется

$$\varphi(x) - \varphi(x+\xi) + \frac{1}{2}\varphi'(x)\xi \geq -\frac{\gamma}{4}\left(\frac{1}{2} - q_2\right)\|x-x^*\|^2.$$

С другой стороны,

$$\varphi'(x)\xi \leq \varphi'(x)\xi + \frac{1}{2}\varphi''(z)\xi\xi \leq \varphi'(x)(x^*-x) + \frac{1}{2}\varphi''(z)(x^*-x)(x^*-x) =$$

$$= -\varphi'(x^*)(x-x^*) - \frac{1}{2}\varphi''(x^*)(x-x^*)(x-x^*) +$$

$$+\int_0^1[(\tfrac{1}{2}\varphi''(z)-\tfrac{1}{2}\varphi''(z^*)+\varphi''(x^*)-\varphi''(x^*+\lambda(x-x^*))](x-x^*)(x-x^*)d\lambda\le$$

$$\le-\tfrac{\gamma}{2}\|x-x^*\|^2+(\tfrac{L}{2}\|z-z^*\|^{\alpha}+\tfrac{L}{1+\alpha}\|x-x^*\|^{\alpha})\|x-x^*\|^2.$$

Можно считать ε настолько малым, что $\varphi'(x)\xi\le-\frac{\gamma}{4}\|x-x^*\|^2$, откуда и следует утверждение леммы.

Для решения задачи (II) построим сначала модифицированный метод Ньютона. Для этого положим $B_0(x)=B(x,x)$, $B(x)=\bigcup_z B(x,z)$ и $\Psi(x,\xi)=-q_2\varphi'(x)\xi$, где $q_2\in(0,\frac{1}{2})$. Начав с произвольной точки $x_0\in Q$, построим последовательности $\{x_n\}$ и $\{z_n\}$ следующим образом. Выберем $\xi\in B(x_n,z_n)$. Если $\varphi(x)-\varphi(x+\xi)\ge\Psi(x,\xi)$, то положим $z_{n+1}=z_n$, $x_{n+1}=x_n+\xi$. В противном случае положим $z_{n+1}=x_n$, возьмем $\xi_n\in B_0(x_n)$ и используем способы I, II и III для выбора длины шага.

Проверим выполнение всех требований первого параграфа. Прежде всего, $M(a)\le d$, где d – диаметр компакта Q. Если $x\in Q$, $\xi\in B(x,z)$ при некотором $z\in Q$, то

$$\frac{1}{q_2}\Psi(x,\xi)=-\varphi'(x)\xi\ge-\varphi'(x)\xi-\tfrac{1}{2}\varphi''(z)\xi\xi\ge-\min_{\mu\in[0,1]}\{\mu\varphi'(x)(x^*-x)+$$
$$+\frac{\mu^2}{2}\varphi''(z)(x^*-x)(x^*-x)\}=\mathcal{K}(x,z).$$

Если $\varphi'(x)(x^*-x)+\varphi''(z)(x^*-x)(x^*-x)>0$, то

$$\mathcal{K}(x,z)=\frac{1}{2}\frac{[\varphi'(x)(x^*-x)]^2}{\varphi''(z)(x^*-x)(x^*-x)}\ge\frac{[\varphi(x)-\varphi_{min}]^2}{2M_2\|x-x^*\|^2},$$

где M_2 – максимум $\|f''(x)\|$ на Q. Если же $\varphi'(x)(x^*-x)+\varphi''(z)(x^*-x)(x^*-x)\le0$, то

$$\mathcal{K}(x,z)=-\varphi'(x)(x^*-x)-\tfrac{1}{2}\varphi''(z)(x^*-x)(x^*-x)\ge-\tfrac{1}{2}\varphi'(x)(x^*-x)\ge\frac{\varphi(x)-\varphi_{min}}{2}$$

Таким образом,

$$\delta(a)\ge q_2\min\left\{\frac{(a-\varphi_{min})^2}{2M_2d^2},\ \frac{a-\varphi_{min}}{2}\right\}$$

Далее, $\varphi(x+\mu\xi)=\varphi(x)+\mu\varphi'(x)\xi+\frac{\mu^2}{2}\varphi''(x+\theta\mu\xi)\xi\xi\le\varphi(x)+\mu\varphi'(x)\xi+\frac{\mu^2}{2}M_2d^2$. Следовательно, $\varphi(x)-\varphi(x+\mu\xi)\ge\mu\Psi(x,\xi)+$

$+\mu\left(\frac{1}{q_2}-1\right)\delta(\varphi(x))-\frac{\mu^2}{2}M_2 d^2$. Поэтому

$$\nu(x)\geq \min\left\{1;\frac{2(1-q_2)\,\delta(\varphi(x))}{q_2 M_2 d^2}\right\}.$$

По теореме 2 получаем, что $\lim \varphi(x_n)=\varphi_{min}$, и ввиду компактности Q и непрерывности φ оказывается, что $\lim \|x_n-x_n^*\|=0$. Покажем, что последовательность $\{z_n\}$ стабилизируется. Если бы это было не так, то $\|z_n-z_n^*\|$ стремилась бы к нулю и по лемме 2, начиная с некоторого шага, $\varphi(x_n)-\varphi(x_n+\xi)\geq \psi(x_n,\xi_n)$ при $\xi\in B(x_n,z_n)$, то есть $z_{n+1}=z_n$. Так что последовательность $\{z_n\}$ действительно стабилизируется.

Оценим скорость сходимости. При $x\in Q$

$$\varphi(x)=\varphi(x^*)+\varphi'(x^*)(x-x^*)+\frac{1}{2}\varphi''(x^*+\theta(x-x^*))(x-x^*)(x-x^*)\geq$$

$$\geq \varphi(x^*)+\frac{1}{2}[\varphi''(x^*+\theta(x-x^*))-\varphi''(x^*)](x-x^*)(x-x^*)+\frac{\gamma}{2}\|x-x^*\|^2\geq$$

$$\geq \varphi_{min}+\left(\frac{\gamma}{2}-\frac{L}{2}\|x-x^*\|^\alpha\right)\|x-x^*\|^2.$$

Так что $\pi(x,z)\geq\dfrac{\varphi(x)-\varphi_{min}}{2}\cdot\min\left\{1;\ \dfrac{\gamma-L\|x-x^*\|^\alpha}{2\|\varphi''(z)\|}\right\}$.

Если при $n\geq n_0$ оказалось $z_n=z_{n_0}$ и $\lambda_n=1$, то при этих n

$$\varphi(x_n)-\varphi(x_{n+1})\geq -q_2\varphi'(x_n)\xi_n\geq q_2\pi(x_n,z_{n_0})\geq\bar\lambda_n(\varphi(x_n)-\varphi_{min}),$$

где

$$\bar\lambda_n=\frac{q_2}{2}\min\left\{1;\ \frac{\gamma-L\|x_n-x_n^*\|^\alpha}{2\|\varphi''(z_{n_0})\|}\right\}.$$

В пределе получим

$$\overline{\lim}\ \frac{\varphi(x_{n+1})-\varphi_{min}}{\varphi(x_n)-\varphi_{min}}\leq(1-\bar\lambda). \tag{15}$$

$$\bar\lambda=\frac{q_2}{2}\min\left\{1;\frac{\gamma}{2\|\varphi''(z_{n_0})\|}\right\}. \tag{16}$$

Таким образом, доказана следующая

ТЕОРЕМА 5. Если выполнены предположения о невырожденности, то в модифицированном методе Ньютона $\lim \varphi(x_n)=\varphi_{min}$,

$\lim \|x_n - x_n^*\| = 0$ и, начиная с некоторого шага, $\lambda_n = 1$, а x_n не меняется (то есть не нужно вычислять $\varphi''(x_n)$). При этом сходимость имеет скорость геометрической прогрессии, и эта скорость определяется формулами (I5) и (I6).

ЗАМЕЧАНИЕ 3. Стабилизация последовательности $\{x_n\}$ еще, вообще говоря, не свидетельствует о том, что установившееся x_n достаточно близко к x_n^*. Формула (I3), однако, показывает, что при достаточно малой $\|x_n - x_n^*\|$ можно гарантировать скорость геометрической прогрессии со сколь угодно малым знаменателем. Поэтому при практическом применении метода в зависимости от поведения разностей $\varphi(x_n) - \varphi(x_{n+1})$ можно на некоторых шагах принудительно положить $x_{n+1} = x_n$, перевычислив вторую производную.

Построим теперь метод Ньютона, положив для этого $B(x) = B_0(x)$. Как и раньше, по теореме 2 найдем, что $\lim \varphi(x_n) = \varphi_{min}$ и $\lim \|x_n - x_n^*\| = 0$. По лемме 2, начиная с некоторого шага, $\lambda_n = 1$, а по лемме I

$$\overline{\lim} \frac{\|x_{n+1} - x_{n+1}^*\|}{\|x_n - x_n^*\|^{1+\alpha}} \leqslant \frac{L}{\gamma(1+\alpha)} . \tag{17}$$

Получена следующая

ТЕОРЕМА 6. Если выполнены предположения о невырожденности, то последовательность $\{x_n\}$, построенная по методу Ньютона, приближается к множеству решений задачи (II), причем скорость сходимости имеет порядок $1+\alpha$ и определяется формулой (I7).

Литература

1. ЛЕВИТИН Е.С. и ПОЛЯК Б.Т. Методы минимизации при наличии ограничений. – "Ж. выч. мат. и матем. физики", I966, т.6, № 5, с. 787-783.
2. ПШЕНИЧНЫЙ Б.Н. Метод Ньютона для решения систем равенств и неравенств. – "Матем. заметки", I970, т.8, № 5, с. 635-640.
3. ДАНИЛИН Ю.М. Минимизация нелинейных функционалов в задачах с ограничениями. – "Кибернетика", I970, № 3, с. II0-II7.
4. БУЛАВСКИЙ В.А. О расширении области сходимости итеративных методов повышенной точности. – "Докл. АН СССР", I972, т.205, № 2, с. 274-276.

5. БУЛАВСКИЙ В.А. Один специальный алгорифм квадратичного программирования. - "Оптимизация", Новосибирск, 1972, вып. 5(22), с. 23-36.

6. FLETCHER R. An algorithm for solving linearly constrained optimization problems.-"Math. Program.", 1972, v. 2, N2, p. 133-165.

Поступила в ред.-изд. отд.
29. XI. 1972 г.