МЕТОДИ ОПТИМІЗАЦІЇ, ОПТИМАЛЬНЕ УПРАВЛІННЯ І ТЕОРІЯ ІГОР

УДК 519.8

ОПТИМИЗАЦИЯ И РЕШЕНИЕ ВЫПУКЛЫХ ВАРИАЦИОННЫХ НЕРАВЕНСТВ

Ю.М. ДАНИЛИН, И.А. ШУБЕНКОВА

Рассматривается подход к решению вариационных неравенств, сводящий исходную задачу к задаче оптимизации. Дальнейшее исследование позволяет построить функцию в пространстве исходных переменных, минимизация которой эквивалентна решению исходной задачи.

Решение вариационного неравенства на множестве S n-мерного евклидова пространства E^n заключается в определении точки $x_* \in S$, для которой выполняется условие

$$\langle F(x_*), x - x_* \rangle \geq 0, \quad \forall x \in S, \tag{1}$$

где оператор $F(x): E^n \to E^n$.

Символы $\langle \cdot,\cdot \rangle$ и $|\cdot|$ обозначают скалярное произведение и норму в E^n.

Будем считать, что $F(x)$ — непрерывный оператор; $S = \{x : g_j(x) \leq \leq 0,\ j = 1,\dots,t\}$, причем S ограничено и имеет непустую внутренность, а $g_j(x)$ — выпуклые непрерывно дифференцируемые функции.

Если x_* удовлетворяет неравенству (1), то, очевидно, $x_* = \arg\min f(x)$, $f(x) = \langle F(x_*), x - x_* \rangle$, $x \in S$.

В точке x_* выполняются следующие необходимые и достаточные условия минимума функции $f(x)$ на множестве S:

существуют такие множители $\lambda_*^j \geq 0$, что

$$F(x_*) + \sum_{j=1}^{t} \lambda_*^j g_j'(x_*) = 0, \quad \lambda_*^j g_j(x_*) = 0, \quad j = 1,\dots,t. \tag{2}$$

С учетом (2) оказывается [1, 2], что решение вариационного неравенства (1) может быть сведено к задаче

$$\min \varphi(x,\lambda), \quad (x,\lambda) \in Q, \tag{3}$$

$$Q=\left\{(x,\lambda): x\in S,\ \lambda^j\ge 0,\ j=1,\dots,t\right\},$$

$$\varphi(x,\lambda)=\frac{1}{2}\left|F(x)+\sum_{j=1}^{t}\lambda^j g'_j(x)\right|^2-\sum_{j=1}^{t}\lambda^j g_j(x). \qquad (4)$$

Функция $\varphi(x,\lambda)$ (4) является двойственной к функции

$$\psi(x)=\min_{p\in E^n}\left\{\langle F(x),p\rangle+\frac{1}{2}|p|^2 : g_j(x)+\langle g'_j(x),p\rangle\le 0,\ j=1,\dots,t\right\}, \qquad (5)$$

и, в силу теоремы двойственности (при сделанных предположениях справедлива теорема Куна-Таккера)

$$\psi(x)=\min_{\lambda\ge 0}\varphi(x,\lambda)=\varphi(x,\lambda(x)). \qquad (6)$$

При $x\in S$ выполняется неравенство $\psi(x)\ge 0$.

Следовательно, при $(x,\lambda)\in Q$

$$0\le\varphi(x,\lambda(x))\le\varphi(x,\lambda),$$

и для точки (x_*,λ_*), удовлетворяющей соотношениям (2),

$$0\le\varphi(x_*,\lambda(x_*))\le\varphi(x_*,\lambda_*)=0,$$

т.е.

$$\psi(x_*)=\varphi(x_*,\lambda(x_*))=\varphi(x_*,\lambda_*)=0.$$

Таким образом, в точке x_* реализуется абсолютный минимум функции $\psi(x)$ на множестве S, и ,как вытекает из (4), в точке $(x_*,\lambda(x_*))$ выполняются условия (2). А это означает, что решение вариационного неравенства (1) эквивалентно задаче

$$\min\psi(x),\quad x\in S. \qquad (7)$$

Остановимся теперь на изучении некоторых свойств функции $\psi(x)$ (6).

Теорема 1. На произвольном ограниченном множестве $\Omega\subset E^n$ функция $\psi(x)$ непрерывна.

Доказательство. По предположению множество S имеет непустую внутренность, т.е. существует точка $z\in S$ такая, что $g_j(z)\le-\gamma$, $\gamma>0$, $j=1,\dots,t$. При этом задача (5) при любом x имеет единственное решение $p(x)$, и выполняются следующие необходимые и достаточные условия экстремума:

$$\begin{gathered}F(x)+p(x)+\sum_{j=1}^{t}\lambda^j(x)g'_j(x)=0,\quad \lambda^j(x)\ge 0,\\ \lambda^j(x)\left(g_j(x)+\langle g'_j(x),p(x)\rangle\right)=0,\quad j=1,\dots,t.\end{gathered} \qquad (8)$$

Множители $\lambda^j(x)$, удовлетворяющие (8), одновременно реализуют и решение задачи $\min\limits_{\lambda\geq 0}\varphi(x,\lambda)$, т.е. определяют функцию $\psi(x)$.

В работе [3] доказана оценка

$$\sum_{j=1}^{t}\lambda^j(x)\leq\frac{1}{2\gamma}\left|F(x)+(z-x)\right|^2 .$$

С учетом непрерывности $F(x)$ это означает, что на ограниченном множестве Ω множители $\lambda^j(x)$ также ограничены, т.е. $\lambda(x)\subset\Lambda\subset E^t$, где $\lambda(x)$ — вектор с компонентами $\lambda^j(x)$, а множество Λ — компакт. Таким образом, на ограниченном множестве Ω

$$\psi(x)=\min_{\lambda\in\Lambda}\varphi(x,\lambda).$$

В силу непрерывности $\varphi(x,\lambda)$ (по переменным x и λ) следует и непрерывность функции $\psi(x)$ (см., напр., [4], теорема 7.2). ■

При некоторых дополнительных предположениях можно установить дифференцируемость $\psi(x)$ в локальной окрестности точки x_*.

Предварительно докажем следующий результат.

Лемма. Решение задачи (5) — вектор $p(x)\to 0$ при $x\to x_*$.

Доказательство. На ограниченном множестве Ω величина $|p(x)|$ ограничена — это следует из (8) в силу ограниченности $|F(x)|$, $\lambda^j(x)$, $|g'_j(x)|$. Пусть $\{x_k\}$ — произвольная последовательность, сходящаяся к точке x_*, и $\{x_{k_\xi}\}$ — произвольная бесконечная подпоследовательность из $\{x_k\}$. В силу ограниченности $\{p(x_{k_\xi})\}$, $\{\lambda(x_{k_\xi})\}$ существуют предельные точки этих последовательностей $\tilde{p}$, $\tilde{\lambda}$. Переходя в (8) к пределу при $x_{k_\xi}\to x_*$, получаем

$$F(x_*)+\tilde{p}+\sum_{j=1}^{t}\tilde{\lambda}^j g'_j(x_*)=0,$$

$$\tilde{\lambda}^j\left(g_j(x_*)+\left\langle g'_j(x_*),\tilde{p}\right\rangle\right)=0,\quad \tilde{\lambda}^j\geq 0,\quad j=1,\dots,t. \tag{9}$$

Условия (9) показывают, что $\tilde{p}$ — решение задачи (5) в точке x_*. В силу единственности решения $\tilde{p}=p(x_*)$. В силу непрерывности функций $\psi(x)$ и $\varphi(x,\lambda)$

$$\psi(x_*)=\lim_{k_\xi\to\infty}\psi\left(x_{k_\xi}\right)=\lim_{k_\xi\to\infty}\varphi\left(x_{k_\xi},\lambda\left(x_{k_\xi}\right)\right)=\varphi\left(x_*,\tilde{\lambda}\right).$$

Но $\psi(x_*)=0$, т.е. $\varphi\left(x_*,\tilde{\lambda}\right)=0$. Из этого следует (4), что

$$F(x_*)+\sum_{j=1}^{t}\tilde{\lambda}^j g'_j(x_*)=0 .$$

Поэтому в силу (9) $\widetilde{p} = p(x_*) = 0$.

Итак, предельной точкой любой бесконечной последовательности $\{p(x_{k_\xi})\}$ является нулевой вектор. Из этого следует, что $p(x_k) \to 0$, $k \to \infty$. Поскольку $\{x_k\}$ — произвольная последовательность (сходящаяся к x_*), то $p(x) \to 0$ при $x \to x_*$.■

Далее будем считать, что выполняются условия:

а) оператор $F(x)$ дважды, а функции $g_j(x)$ трижды непрерывно дифференцируемы;

б) в точке x_*, удовлетворяющей (2), градиенты $g'_j(x_*)$, $j \in J(x_*)$ — линейно независимы и $\lambda_*^j > 0$, $j \in J(x_*)$.

Здесь

$$J(x_*) = \{j : g_j(x_*) = 0,\ j = 1,...,t\}.$$

Теорема 2. Пусть выполняются условия *а*, *б*. Тогда в локальной окрестности точки x_* функция $\psi(x)$ дважды непрерывно дифференцируема, и

$$\psi'(x) = \varphi'_x(x, \lambda(x)),$$

$$\psi''(x) = \varphi''_{xx}(x, \lambda(x)) - \varphi''_{x\lambda}(x, \lambda(x))[\varphi''_{\lambda\lambda}(x, \lambda(x))]^{-1}\varphi''_{\lambda x}(x, \lambda(x)). \qquad (10)$$

Доказательство. При выполнении условий *б* существует окрестность $\Omega(x_*, \delta) = \{x : |x - x_*| \le \delta\}$, для всех точек которой выполняется условие

$$J(x) = J(x_*),$$

$$J(x) = \{j : g_j(x) + \langle g'_j(x),\ p(x)\rangle = 0,\quad j = 1, \dots, t\}. \qquad (11)$$

Действительно, для $j \notin J(x_*)$ будет $g_j(x) \to g_j(x_*) = -\beta_*^j < 0$. В силу леммы $p(x) \to 0$ ($x \to x_*$). Следовательно, если окрестность $\Omega(x_*, \delta)$ достаточно мала, в любой точке $x \in \Omega(x_*, \delta)$

$$g_j(x) + \langle g'_j(x), p(x)\rangle < 0,\quad j \notin J(x_*), \qquad (12)$$

и в силу (8) $\lambda^j(x) = 0$.

Пусть $\{x_k\}$ — произвольная последовательность, сходящаяся к x_*. Предположим, что при сколь угодно больших номерах k может оказаться $J(x_k) \ne J(x_*)$. С учетом (12) это будет означать, что $J(x_k) \subset J(x_*)$. Поскольку существует лишь конечное число различных множеств $J(x) \subset J(x_*)$, найдется подпоследовательность $\{x_{k_\xi}\}$ из $\{x_k\}$ такая, что $J(x_{k_\xi}) = J \subset J(x_*)$ ($J \ne J(x_*)$).

При этом, переходя в (8) к пределу ($x_{k_\xi} \to x_*$, $p(x_{k_\xi}) \to 0$, $\lambda(x_{k_\xi}) \to \widetilde{\lambda}$, получаем

$$F(x_*)+\sum_{j\in J}\tilde{\lambda}^j g'_j(x_*)=0\,,$$

что противоречит условию *б* ($\lambda_*^j>0$, $j\in J(x_*)$). Поэтому условие $J(x_k)\neq J(x_*)$ при $k\to\infty$ выполняться не может. В силу произвольного выбора $\{x_k\}$ следует справедливость (11).

При выполнении (11) соотношения (8) представимы в виде

$$F(x)+p(x)+\sum_{j\subset J(x_*)}\lambda^j(x)g'_j(x)=0\,,$$

$$g_j(x)+\langle g'_j(x),p(x)\rangle=0,\quad j\in J(x_*),\tag{13}$$

$$\lambda^j(x)=0,\quad j\notin J(x_*)\,.$$

Полагая окрестность $\Omega(x_*,\delta)$ настолько малой, что градиенты $g'_j(x)$, $j\in J(x_*)$ — линейно независимы при любом $x\in\Omega(x_*,\delta)$, из (13) получаем

$$\begin{gathered}\lambda_J(x)=\left(g'(x)g'^{\mathrm{T}}(x)\right)^{-1}\left(g(x)-g'(x)F(x)\right),\\ p(x)=-F(x)-g'^{\mathrm{T}}(x)\lambda_J(x),\end{gathered}\tag{14}$$

где $\lambda_J(x)$ — вектор-столбец с компонентами $\lambda^j(x)$, $j\in J(x_*)$; $g(x)$ — вектор-столбец с компонентами $g_j(x)$, $j\in J(x_*)$; $g'(x)$ — матрица, строками которой являются векторы $g'_j(x)$, $j\in J(x_*)$.

Поскольку в окрестности $\Omega(x_*,\delta)$ множители $\lambda^j(x)\equiv 0$, $j\notin J(x_*)$, далее будем считать, что в этой окрестности $\lambda=\lambda_J$ и

$$\varphi(x,\lambda)=\frac{1}{2}\left|F(x)+\sum_{j\in J(x_*)}\lambda^j g'_j(x)\right|^2-\sum_{j\in J(x_*)}\lambda^j g_j(x)\,.$$

Из (14), в силу условий *а*, следует, что $\lambda_J(x)$ дважды непрерывно дифференцируемая функция, и поэтому функция $\psi(x)=\varphi(x,\lambda(x))$ также дважды непрерывно дифференцируема в окрестности $\Omega(x_*,\delta)$.

Установим теперь справедливость выражений (10).

Если $\lambda'(x)$ — матрица, строками которой являются векторы $\left(\lambda^j(x)\right)'$, то градиент функции $\psi(x)$

$$\psi'(x)=\varphi'_x(x,\lambda(x))+\lambda'^{\mathrm{T}}(x)\varphi'_\lambda(x,\lambda(x))=$$

$$=\varphi'_x(x,\lambda(x))+\sum_{j\in J(x_*)}\varphi'_{\lambda^j}(x,\lambda(x))\left(\lambda^j(x)\right)'.\tag{15}$$

Будем считать окрестность $\Omega(x_*,\delta)$ настолько малой, что множители $\lambda^j(x)>0$, $j\in J(x_*)$. При этом в силу необходимых условий экстремума (на множестве $\lambda\ge 0$)

$$\frac{\partial\varphi(x,\lambda(x))}{\partial\lambda^j}=\varphi'_{\lambda^j}(x,\lambda(x))=0,$$

$$\forall\ x\in\Omega(x_*,\delta),\quad j\in J(x_*). \tag{16}$$

Следовательно (15), $\psi'(x)=\varphi'_x(x,\lambda(x))$.

Определим теперь $\psi''(x)$.

Обозначим

$$\varphi''_{xx}(x,\lambda)=\left[\frac{\partial^2\varphi(x,\lambda)}{\partial x^i\partial x^j}\right],\quad \varphi''_{\lambda\lambda}(x,\lambda)=\left[\frac{\partial^2\varphi(x,\lambda)}{\partial\lambda^i\partial\lambda^j}\right],$$

i, j — индексы строки и столбца,

$$\varphi''_{x\lambda}(x,\lambda)=\left[\frac{\partial^2\varphi(x,\lambda)}{\partial x^i\partial\lambda^j}\right],\quad \varphi''_{\lambda x}(x,\lambda)=\left[\frac{\partial^2\varphi(x,\lambda)}{\partial\lambda^i\partial x^j}\right].$$

Тогда

$$\psi''(x)=\left[\varphi'_x(x,\lambda(x))\right]'_x=\varphi''_{xx}(x,\lambda(x))+\varphi''_{x\lambda}(x,\lambda(x))\lambda'(x). \tag{17}$$

Далее,

$$\left[\varphi'_\lambda(x,\lambda(x))\right]'_x=\varphi''_{\lambda x}(x,\lambda(x))+\varphi''_{\lambda\lambda}(x,\lambda(x))\lambda'(x). \tag{18}$$

При выполнении (16) очевидно

$$\left[\varphi'_\lambda(x,\lambda(x))\right]'_x=0,\quad \forall x\in\Omega(x_*,\delta). \tag{19}$$

В окрестности $\Omega(x_*,\delta)$ матрица $\varphi''_{\lambda\lambda}(x,\lambda(x))=g'(x)g'^{\mathrm{T}}(x)$ является невырожденной, поэтому из (18), (19) следует

$$\lambda'(x)=-\varphi''_{\lambda\lambda}(x,\lambda(x))^{-1}\varphi''_{\lambda x}(x,\lambda(x)).$$

Таким образом, выражение (17) можно представить в виде

$$\psi''(x)=\varphi''_{xx}(x,\lambda(x))-\varphi''_{x\lambda}(x,\lambda(x))\left[\varphi''_{\lambda\lambda}(x,\lambda(x))\right]^{-1}\varphi''_{\lambda x}(x,\lambda(x)).$$

Итак, равенства (10) справедливы. ■

Теорема 3. Пусть (x_*,λ_*) — точка, удовлетворяющая соотношениям (2), и выполняются условия *а*, *б*.

Тогда

$$\psi'(x_*)+\sum_{j\in J(x_*)}\lambda_*^j g'_j(x_*)=0. \tag{20}$$

Если к тому же оператор $F(x)$ — сильно монотонный, т.е. $\langle F'(x)p, p\rangle \geq$ $\geq m|p|^2$, $\forall p \in E^n$, $m > 0$, то

$$\left\langle \left(\psi''(x_*) + \sum_j \lambda_*^j g_j''(x_*) \right) p, p \right\rangle > 0 \tag{21}$$

для любого вектора $p \neq 0$ такого, что

$$\langle g_j'(x_*), p \rangle = 0, \quad j \in J(x_*). \tag{22}$$

Доказательство. Поскольку x_* — точка минимума функции $\psi(x)$ на множестве S, существуют такие множители $\mu_*^j \geq 0$, $j \in J(x_*)$, что

$$\psi'(x_*) + \sum_{j \in J(x_*)} \mu_*^j g_j'(x_*) = 0\,. \tag{23}$$

Но, поскольку $F(x_*) + \sum_{j \in J(x_*)} \lambda_*^j g_j'(x_*) = 0$, то (10)

$$\psi'(x_*) = \varphi_x'(x_*, \lambda(x_*)) = -\sum_{j \in J(x_*)} \lambda_*^j g_j'(x_*).$$

Следовательно (23),

$$\sum_{j \in J(x_*)} \left(\mu_*^j - \lambda_*^j \right) g_j'(x_*) = 0\,. \tag{24}$$

В силу линейной независимости градиентов $g_j'(x_*)$, $j \in J(x_*)$ из (24) следует, что $\mu_*^j = \lambda_*^j$, $j \in J(x_*)$, т.е. равенство (20) справедливо.

Установим справедливость (21).

Поскольку $\varphi_{\lambda\lambda}''(x_*, \lambda(x_*)) = g'(x_*) g'^{\mathrm{T}}(x_*)$,

$$\varphi_{\lambda x}''(x_*, \lambda(x_*)) = g'(x_*) \left[\left(F'(x_*) + \sum_{j \in J(x_*)} \lambda_*^j g_j''(x_*) \right) - I \right],$$

где I — единичная матрица; $\varphi_{x\lambda}''(x, \lambda(x)) = [\varphi_{\lambda x}''(x, \lambda(x))]^{\mathrm{T}}$, тогда

$$\varphi_{x\lambda}''(x_*, \lambda(x_*)) [\varphi_{\lambda\lambda}''(x_*, \lambda(x_*))]^{-1} \varphi_{\lambda x}''(x_*, \lambda(x_*)) =$$

$$= \left[\left(F'(x_*) + \sum_{j \in J(x_*)} \lambda_*^j g_j''(x_*) \right) - I \right]^{\mathrm{T}} g'^{\mathrm{T}}(x_*) \left[g'(x_*) g'^{\mathrm{T}}(x_*) \right]^{-1} g'(x_*) \times$$

$$\times \left[\left(F'(x_*) + \sum_{j \in J(x_*)} \lambda_*^j g_j''(x_*) \right) - I \right].$$

Обозначим

$$A_* = F'(x_*) + \sum_{j \in J(x_*)} \lambda_*^j g_j''(x_*),$$

$$G_* = g'^{\mathrm{T}}(x_*)\left[g'(x_*) g'^{\mathrm{T}}(x_*)\right]^{-1} g'(x_*).$$

Тогда

$$[A_* - I]^{\mathrm{T}} G_* [A_* - I] = A_*^{\mathrm{T}} G_* A_* - G_* A_* - A_*^{\mathrm{T}} G_* + G_* . \tag{25}$$

Заметим, что G_* — оператор проектирования на подпространство, образованное векторами $g_j'(x_*)$, $j \in J(x_*)$. В силу свойств оператора проектирования, если вектор p такой, что выполняются условия (22), то

$$\begin{gathered} \langle G_* A_* p, p \rangle = \langle A_* p, G_* p \rangle = 0, \\ A_*^{\mathrm{T}} G_* p = 0, \quad G_* p = 0, \end{gathered} \tag{26}$$

$$\begin{aligned} \langle A_*^{\mathrm{T}} G_* A_* p, p \rangle &= \langle G_* A_* p, A_* p \rangle = \langle G_* G_* A_* p, A_* p \rangle = \\ &= \langle G_* A_* p, G_* A_* p \rangle = |G_* A_* p|^2 . \end{aligned} \tag{27}$$

По предположению, оператор $F(x)$ — сильно монотонный, а функции $g_j(x)$ — выпуклы, поэтому матрица A_* положительно определена, т.е. $\langle A_* p, p \rangle > 0$, $\forall p \in E^n$, $p \neq 0$. Если для вектора p выполняются условия (22), то вектор $A_* p$ не может принадлежать подпространству, образованному векторами $g_j'(x_*)$, $j \in J(x_*)$, в противном случае оказалось бы, что $\langle A_* p, p \rangle = 0$.

Из этого следует

$$|G_* A_* p| < |A_* p| . \tag{28}$$

Учитывая, что $\varphi_{xx}''(x_*, \lambda) = A_*^{\mathrm{T}} A_* - \sum_{j \in J(x_*)} \lambda_*^j g''(x_*)$, в силу второго из равенств (10) получаем

$$\psi''(x_*) + \sum_j \lambda_*^j g_j''(x_*) = A_*^{\mathrm{T}} A_* - [A_* - I]^{\mathrm{T}} G_* [A_* - I].$$

Из (25) с учетом (26), (27) следует, что для любого вектора p , удовлетворяющего условиям (22),

$$\left\langle [A_* - I]^{\mathrm{T}} G_* [A_* - I] p, p \right\rangle = |G_* A_* p|^2 .$$

Следовательно,

$$\left\langle \left(\psi''(x_*) + \sum_j \lambda_*^j g_j''(x_*) \right) p, p \right\rangle = |A_* p|^2 - |G_* A_* p|^2 .$$

С учетом (28) это означает справедливость (21). ■

Поскольку при выполнении условий *а*, *б* справедливо равенство (20), а при условии, что $F(x)$ — сильно монотонный оператор, и неравенство (21), то точка x_* реализует невырожденный минимум функции $\psi(x)$.

Итак, решение вариационного неравенства (1) может быть сведено к решению задачи минимизации функции $\psi(x)$ на множестве S (т.е. к оптимизационной задаче в пространстве исходных переменных). Изученные в данной работе свойства функции $\psi(x)$ (теоремы 1, 2, 3) показывают, что, по крайне мере, в локальной окрестности решения может быть использован весь хорошо развитый аппарат численных методов оптимизации. В частности, возможно использование ньютоновских и квазиньютоновских процессов, базирующихся на методах последовательного квадратичного программирования [2, 5].

ЛИТЕРАТУРА

1. *Danilin Yu. M.* On an approach to solution of variational inequalities // II Укр. конф. «Автоматика-95». — Львів, 1995. — **1**. — С. 20.
2. *Данилін Ю.М., Шубенкова І.А.* Про розв'язання варіаційних нерівностей в скінченновимірних просторах // Праці міжнар. конф. «Питання оптимізації обчислень». Київ. — 1997. — С. 86–89.
3. *Пшеничный Б.Н., Калжанов М.У.* Метод решения вариационных неравенств // Кибернетика и системный анализ. — 1992. — №6. — С. 48–55.
4. *Зангвилл У.И.* Нелинейное программирование. — М.: Сов. радио, 1973. — 307 с.
5. *Данилин Ю.М.* Последовательное квадратичное программирование и модифицированные функции Лагранжа // Кибернетика и системный анализ. — 1994. — №5. — С. 51–67.

Поступила 25.07.2003