Доклады Академии наук СССР 1984. Том 278, No.1, стр. 24–27

УДК 519.8 МАТЕМАТИКА

Ю.Г. ЕВТУШЕНКО, В.А. ПУРТОВ

ДОСТАТОЧНЫЕ УСЛОВИЯ МИНИМУМА ДЛЯ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

(Представлено академиком А.А. Дородницыным 3 X 1983)
Пересмотрено 18 VI 2002

$\boxed{\mathbf{1.}}$ Рассмотрим общую задачу нелинейного программирования

$$\min_{x\in X} f(x); \tag{1}$$

здесь $X \subset X_0 \subset \mathbb{E}^n$; $\mathbb{E}^i$ есть i-мерное евклидово пространство; функция $f(x)$ определена на X_0 и принимает значения в действительной области. Обозначим через X_* множество всех решений задачи (1). Всюду в дальнейшем считаем, что X_* непусто.

Введем функцию $H(x,y) = f(x) + \xi(x,y)$, где $y \in \mathbb{E}^m$, $\xi(x,y) : X_0 \times \mathbb{E}^m \to \mathbb{E}^1$, и определим точечно-множественные отображения

$$\begin{aligned} X(y) &= \operatorname*{Argmin}_{x\in X_0} H(x,y); \\ Y(x) &= \{y \in \mathbb{E}^m : \xi(z,y) \le \xi(x,y),\ \ \forall\, z \in X\}. \end{aligned}$$

Как правило, множество X_0 выпуклое и имеет сравнительно простую структуру, поэтому задача нахождения точек из $X(y)$ проще задачи (1). В частности, X_0 может совпадать с $\mathbb{E}^n$. В ряде случаев можно подобрать H так, чтобы для некоторых $y \in \mathbb{E}^m$ было

$$X(y) = X_*. \tag{2}$$

В качестве примеров укажем

$$\begin{aligned} H(x,y) &= |f(x) - y|^p + S(x), \\ H(x,y) &= (f(x) - y)_+^p + S(x), \end{aligned} \tag{3}$$

где $S(x) : \mathbb{E}^n \to \mathbb{E}^1$, $S(x) = 0$ для всех $x \in X$ и $S(x) > 0$ для любого $x \notin X$, $\varphi_+ = \max[0,\varphi]$, $p > 0$. Если $y = f(x_*)$, $x_* \in X_*$, то (2) имеет место. Однако значение $f(x_*)$ обычно бывает неизвестно, что усложняет использование (3) для численных расчетов. Дадим другие достаточные условия минимума в задаче (1).

Определение 1. *Точку* $(x,y) \in \mathbb{E}^n \times \mathbb{E}^m$ *назовем* **особой точкой функции** $H(x,y)$, *если* $x \in X \cap X(y)$, $y \in Y(x)$.

Теорема 1. *Пусть существует особая точка* (x,y) *функции* $H(x,y)$. *Тогда* $x \in X_* \subset X(y)$.

Теорема 2. *Пусть существует такой вектор* $y \in \mathbb{E}^m$, *что множество* $X(y)$ *непусто,* $X(y) \subset X$ *и функция* $\xi(x,y)$ *принимает постоянное значение при всех* $x \in X$. *Тогда имеет место* (2).

Теорема 3. *Пусть* $f(x)$ *непрерывна на* X_0 *и существует такая последовательность* $\{x_k, y_k\}$, *что* $x_k \in X(y_k)$, $y_k \in Y(x_k)$, $\lim_{k\to\infty} x_k = x_* \in X$. *Тогда* $x_* \in X_*$.

Теорема 1 представляет собой новую формулировку теоремы 1.7.4 из [1]. Вектор y может не входить в функцию H; тогда, обозначив

$$\bar{X} = \operatorname*{Argmin}_{x \in X_0} H(x), \qquad H(x) = f(x) + \xi(x), \qquad \xi(x) : \mathbb{E}^n \to \mathbb{E}^1,$$

из теоремы 1 получим

Следствие 1. *Пусть $\bar{x} \in X \cap \bar{X}$ и функция $\xi(x)$ такова, что $\xi(x) \le \xi(\bar{x})$ для всех $x \in X$. Тогда $\bar{x} \in X_* \subset \bar{X}$.*

Введем функции

$$R(x, y) = f(x) + \tilde{\xi}(x, y), \qquad \tilde{\xi}(x, y) = \xi(x, y) - \gamma(y), \qquad \gamma(y) = \sup_{x \in X} \xi(x, y).$$

Область определения $\gamma(y)$ обозначим Y. Очевидно, $Y \supset \bigcup\limits_{x \in X_0} Y(x)$. Рассмотрим задачу отыскания

$$\sup_{y \in Y} \inf_{x \in X_0} R(x, y) \tag{4}$$

и наложим на R условие

$(\mathbf{A}_0)$. Для каждого $x \in X$ существует такая точка $y \in Y$, что $\tilde{\xi}(x, y) = 0$; для каждого $x \in X_0 \backslash X$ существует такая последовательность $\{y_k\}$, что все $y_k \in Y$, $k = 1, 2, \ldots$, и $\lim\limits_{k \to \infty} \tilde{\xi}(x, y_k) = \infty$.

Теорема 4. *Для того чтобы точка (x, y) была особой точкой $R(x, y)$, необходимо, а если имеет место условие* $(\mathbf{A}_0)$*, то и достаточно, чтобы она была седловой в задаче* (4).

$\boxed{\mathbf{2.}}$ Рассмотрим редукцию исходной задачи нелинейного программирования к задаче отыскания точек (x, y), удовлетворяющих условиям

$$G(x, y) = 0, \qquad x \in X(y); \tag{5}$$

здесь $G(x, y) : X_0 \times \mathbb{E}^m \to \mathbb{E}^t$. Напомним определение из [2].

Определение 2. *Пара $\{H, G\}$* **согласована с задачей** (1)*, если множество решений* (5) *непусто и любая точка (x, y), удовлетворяющая* (5)*, такова, что $x \in X_*$.*

Наложим условие на функции H и G.

$(\mathbf{A}_1)$. Если существуют x и y, удовлетворяющие уравнению $G(x, y) = 0$, то $x \in X$, $y \in \in Y(x)$.

Теорема 5. *Для всякой задачи нелинейного программирования* (1) *с непустым множеством решений существует согласованная с ней пара. Если для пары $\{H, G\}$ выполнено условие* $(\mathbf{A}_1)$ *и множество решений* (5) *непусто, то эта пара согласована с задачей* (1).

Введем вспомогательную мажорирующую функцию $\eta(y)$, определенную на некотором множестве $Y_H \subset Y$ и удовлетворяющую на нем неравенству $\gamma(y) \le \eta(y)$. Определим отображение

$$W(x) = \{y \in Y_H : \eta(y) \le \xi(x, y)\}.$$

Очевидно, $W(x) \subset Y(x)$ для любого $x \in X_0$. Тогда в формулировке теоремы 5 условие $(\mathbf{A}_1)$ можно заменить условием

$(\mathbf{A}_2)$. Если точка (x, y) такова, что $G(x, y) = 0$, то $x \in X$, $y \in W(x)$.

При дополнительных предположениях относительно функции ξ условие $(\mathbf{A}_2)$ можно ослабить. Пусть, например, $\xi(x, y) = \eta(y)$ для всех $(x, y) \in X \times Y_H$; тогда вместо $(\mathbf{A}_2)$ используется условие

$(\mathbf{A}_3)$. Если точка (x, y) такова, что $G(x, y) = 0$, то $x \in X$, $y \in Y_H$.

Если к тому же $X(y) \subset X$ для любого $y \in Y_H$, то $(\mathbf{A}_2)$ заменится на условие

$(\mathbf{A}_4)$. Если точка (x, y) такова, что $G(x, y) = 0$, то $y \in Y_H$.

$\boxed{\mathbf{3.}}$ Дадим примеры построения согласованных пар. Конкретизируем множество X. Пусть

$$X = \{x \in \mathbb{E}^n : g(x) = 0, \;\; h(x) \le 0\}; \tag{6}$$

здесь $g(x) : \mathbb{E}^n \to \mathbb{E}^\ell$, $h(x) : \mathbb{E}^n \to \mathbb{E}^c$. Обозначим через $\mathbb{E}^i_+$ неотрицательный ортант пространства $\mathbb{E}^i$. Определим множество $Y_L = \mathbb{E}^\ell \times \mathbb{E}^c_+$. В качестве H возьмем функцию Лагранжа

$$L(x, y) = f(x) + \xi(x, y), \qquad \xi(x, y) = \sum_{i=1}^{\ell} u^i g^i(x) + \sum_{j=1}^{c} v^j h^j(x), \qquad y = (u, v) \in \mathbb{E}^{\ell+c}.$$

Для любой задачи (1) с допустимым множеством (6) $Y_L \subset Y$ и $\gamma(y) \le 0$ на Y_L, т.е. функция $\eta(y) \equiv 0$ мажорирует $\gamma(y)$ на Y_L. Функцию G строим в виде

$$G(x, y) = \{\alpha(g^1(x), u^1), \ldots, \alpha(g^\ell(x), u^\ell), \; \beta(h^1(x), v^1), \ldots, \beta(h^c(x), v^c)\}. \tag{7}$$

На $\alpha(a, b)$, $\beta(a, b)$, отображающие $\mathbb{E}^2$ в $\mathbb{E}^1$, наложим следующие условия:

$(\mathbf{B}_1)$. Если уравнение $\alpha(a, b) = 0$ имеет решение, то $a = 0$.

$(\mathbf{B}_2)$. Если уравнение $\beta(a, b) = 0$ имеет решение, то $a \le 0$, $b \ge 0$, $ab = 0$.

Легко построить функции $\alpha(a, b)$, удовлетворяющие $(\mathbf{B}_1)$: a, $\sinh a$, $\arctan a$ и т.п. Приведем примеры функций $\beta(a, b)$, удовлетворяющих $(\mathbf{B}_2)$: $(a + b)_+ - b$, $ab + a_+^2 + b_-^2$, $ab + (a - b)_+^2$, $b \arctan a + a_+^2 + b_-^2$, $a \arctan b + a_+^2 + b_-^2$, $(a + b)_+^3 - b^3 - (a_-^2 b)/(1 + a^2)$.

Если функция G определяется из (7), выполнены условия $(\mathbf{B}_1)$, $(\mathbf{B}_2)$, то имеет место $(\mathbf{A}_2)$. Если к тому же в задаче (1) есть седловая точка функции Лагранжа, то пара $\{L, G\}$ согласована с задачей (1).

Выберем теперь в качестве H функцию

$$L_2(x, y) = f(x) + \sum_{i=1}^{\ell} u^i g^i(x) + \sum_{j=1}^{c} (v^j)^2 h^j(x).$$

Для любой задачи (1) $Y = \mathbb{E}^{\ell+c}$, $\gamma(y) \le 0$, поэтому $\eta(y) \equiv 0$ мажорирует $\gamma(y)$ на $\mathbb{E}^{\ell+c}$. Функцию G будем по-прежнему строить в виде (7). Наложим на $\beta(a, b)$ условие

$(\mathbf{B}_3)$. Если уравнение $\beta(a, b) = 0$ имеет решение, то $a \le 0$, $ab = 0$.

Очевидно, что если для $G(x, y)$ выполнены условия $(\mathbf{B}_1)$, $(\mathbf{B}_3)$, то имеет место $(\mathbf{A}_2)$. Если же в задаче (1) есть седловая точка функции L_2, то пара $\{L_2, G\}$ согласована с этой задачей. Примерами функций $\beta(a, b)$, удовлетворяющих условию $(\mathbf{B}_3)$, могут служить $ae^{-b} - a_-$, $a_+^p + ab^2$, $a_+^p + a_-^p b$; здесь $p > 0$ — натуральное число. Отметим, что $Y_L \subset \mathbb{E}^{\ell+c}$, поэтому при построении согласованной пары $\{L_2, G\}$ можно использовать функции β, удовлетворяющие $(\mathbf{B}_2)$.

$\boxed{4.}$ Широкий класс согласованных пар получается, если в качестве H выбрать

$$H(x,y) = f(x) + \sum_{i=1}^{\ell} \varphi(g^i(x), u^i) + \sum_{j=1}^{c} \psi(h^j(x), v^j), \tag{8}$$

а G по-прежнему определять по формуле (7). Здесь $\varphi(a,b) : \mathbb{E}^2 \to \mathbb{E}^1$, $\psi(a,b) : \mathbb{E}^2 \to \to \mathbb{E}^1$. В общем случае имеет смысл подобрать функции φ и ψ так, чтобы можно было выделить достаточно представительное подмножество $Y_H \subset Y$ и мажоранту $\eta(y)$ для всех задач (1) с допустимым множеством в виде (6). Затем определить отображение $W(x)$ и построить такие функции α, β, чтобы для G выполнялось одно из условий $\mathbf{(A_2)}$–$\mathbf{(A_4)}$ (или непосредственно условие $\mathbf{(A_1)}$). Воспользуемся построенными в п. 3 функциями α, β. Для этого наложим условия на $\psi(a,b)$. Обозначим $\delta(b) = \sup\limits_{a \le 0} \psi(a,b)$.

$\mathbf{(C_1)}$. $\delta(b)$ принимает на $\mathbb{E}^1_+$ конечные значения, и если $a \le 0$, $b \ge 0$, $ab = 0$, то $\psi(a,b) = \delta(b)$.

$\mathbf{(C_2)}$. $\delta(b)$ определена всюду на $\mathbb{E}^1$, и если $a \le 0$, $ab = 0$, то $\psi(a,b) = \delta(b)$.

Пусть $H(x,y)$ определяется по формуле (8) и $\psi(a,b)$ удовлетворяет условию $\mathbf{(C_1)}$, ($\mathbf{(C_2)}$). Тогда $\eta(y) = \sum\limits_{i=1}^{\ell} \varphi(0, u^i) + \sum\limits_{j=1}^{c} \delta(v^j)$ мажорирует $\gamma(y)$ на $Y_L(\mathbb{E}^{\ell+c})$ для любой задачи (1). Если к тому же G определяется из (7), для α, β выполнены условия $\mathbf{(B_1)}$, $\mathbf{(B_2)}$ ($\mathbf{(B_1)}$, $\mathbf{(B_3)}$), то пара $\{H, G\}$ удовлетворяет условию $\mathbf{(A_2)}$. Согласно теореме 5 эта пара согласована с задачей (1) при условии, что множество решений (5) непусто.

Приведем примеры $\psi(a,b)$, для которых выполняется условие $\mathbf{(C_1)}$: $r(a+b)_+^p$, $ab + ra_+^p$, $be^a + ra_+^p$. Следующие функции $\psi(a,b)$ удовлетворяют условию $\mathbf{(C_2)}$: $ab^2 + ra_+^p$, $a_+^{2p-1}b + 1/2a_-^{2p-1}b^2$. Здесь $p > 0$ — натуральное число, $r > 0$ — действительное число; p и r подбираются таким образом, чтобы обеспечить решение задачи минимизации $H(x,y)$ на X_0, если это возможно. Приведем примеры функций $\varphi(a,b)$: $ab + ra^p$, $ab + r \cosh a$, $a(b-1) + re^a$.

$\boxed{5.}$ В заключение отметим, что переход от задачи (1) к задаче (5) позволяет применить для решения исходной задачи нелинейного программирования разнообразные численные методы решения систем нелинейных уравнений в сочетании с методами поиска минимума на множествах простой структуры. Например, если размерности G и y совпадают, то для решения (5) можно воспользоваться методом простой итерации

$$y_{k+1} = y_k + \alpha G(x_k, y_k), \qquad x_k \in X(y_k),$$

где α — коэффициент, обеспечивающий сходимость процесса. При этом известные достаточные условия сходимости метода простой итерации переформулируются в терминах функций H, G. Аналогично можно поступить, используя другие методы решения систем уравнений.

ЛИТЕРАТУРА

1. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982, 432 с.

2. *Евтушенко Ю.Г., Жадан В.Г.* — Техническая кибернетика, 1983, No.1, с. 47–59.