МЕТОД ЛИНЕАРИЗАЦИИ
В НЕЛИНЕЙНОМ ПРОГРАММИРОВАНИИ
(ПРИНЦИПИАЛЬНАЯ СХЕМА)

В. Н. Малозёмов

v.malozemov@spbu.ru

14 января 2019 г.

1°. Рассмотрим задачу нелинейного программирования с нелинейными ограничениями-неравенствами:

$$
\begin{gathered}
f_0(x) \to \inf, \\
f_i(x) \leqslant 0,\ i \in 1:m.
\end{gathered}
\tag{1}
$$

Здесь $f_0(x), f_1(x), \dots, f_m(x)$ — функции из класса $C^1(\mathbb{R}^n)$. Множество векторов, удовлетворяющих ограничениям задачи (1), обозначим Ω.

Положим

$$
F(x) = \max\bigl\{0, f_1(x), \dots, f_m(x)\bigr\}.
$$

Очевидно, что $F(x) \geqslant 0$ при всех $x \in \mathbb{R}^n$, и $F(x) = 0$ тогда и только тогда, когда $x \in \Omega$.

Возьмём достаточно большое число $N > 0$ и введём функцию

$$
\Phi_N(x) = f_0(x) + NF(x).
$$

Для решения задачи (1) можно использовать метод линеаризации [1, с. 45–51], к описанию которого мы и переходим.

Параметрами метода являются $\varepsilon \in (0, 1)$ и штрафной коэффициент $N > 0$.

В качестве начального приближения можно взять произвольную точку $x_0 \in \mathbb{R}^n$. Пусть имеется k-е приближение x_k.

Рассмотрим вспомогательную задачу квадратичного программирования:

$$
\begin{gathered}
\bigl\langle f_0'(x_k), p \bigr\rangle + \tfrac{1}{2}\|p\|^2 \to \inf, \\
f_i(x_k) + \bigl\langle f_i'(x_k), p \bigr\rangle \leqslant 0,\ i \in 1:m.
\end{gathered}
\tag{2}
$$

Эта задача представляет собой линеаризованный в окрестности точки x_k вариант задачи (1). Пусть p_k — решение задачи (2).

Если $p_k = \mathbb{O}$, то вычисления заканчиваются. При $p_k \neq \mathbb{O}$ полагаем

$$x_{k+1} = x_k + \alpha_k p_k,$$

где α_k — первое число в последовательности $1, \frac{1}{2}, \frac{1}{4}, \ldots$, при котором выполняется неравенство

$$\Phi_N(x_k + \alpha_k p_k) \leqslant \Phi_N(x_k) - \varepsilon\alpha_k\|p_k\|^2. \tag{3}$$

После этого вычисления повторяются.

2°. Напомним, что точка $x \in \mathbb{R}^n$ называется *стационарной точкой* для задачи (1), если $x \in \Omega$ и существуют числа u_i, $i \in 1:m$, такие, что

$$\begin{gathered} f_0'(x) + \sum_{i=1}^{m} u_i f_i'(x) = \mathbb{O}, \\ u_i f_i(x) = 0 \quad \text{и} \quad u_i \geqslant 0 \quad \text{при} \quad i \in 1:m. \end{gathered} \tag{4}$$

Условия (4) называются *условиями Куна-Таккера.*

ЛЕММА 1. *Для того чтобы точка $x \in \mathbb{R}^n$ была стационарной для задачи* (1)*, необходимо и достаточно, чтобы задача квадратичного программирования*

$$\begin{gathered} Q(p) := \langle f_0'(x), p\rangle + \tfrac{1}{2}\|p\|^2 \to \inf, \\ f_i(x) + \langle f_i'(x), p\rangle \leqslant 0,\ i \in 1:m, \end{gathered} \tag{5}$$

имела решение $p = p(x)$ и чтобы выполнялось равенство $p(x) = \mathbb{O}$.

Доказательство. Необходимость. Пусть x — стационарная точка. В частности, $x \in \Omega$. Отсюда следует, что множество планов задачи (5) непусто — оно содержит вектор $p = \mathbb{O}$. Целевая функция задачи (5) сильно выпуклая. Значит, она ограничена снизу на $\mathbb{R}^n$. По теореме существования решения для задачи квадратичного программирования [2, с. 111–112] задача (5) имеет решение $p = p(x)$. Сильная выпуклость целевой функции $Q(p)$ гарантирует единственность этого решения.

Покажем, что $p(x) = \mathbb{O}$. По критерию оптимальности для задачи квадратичного программирования план p задачи (5) будет оптимальным, тогда и только тогда, когда существуют числа $v_i(x)$, $i \in 1:m$, такие, что

$$\begin{gathered} f_0'(x) + p + \sum_{i=1}^{m} v_i(x) f_i'(x) = \mathbb{O}, \\ v_i(x)\Big(f_i(x) + \langle f_i'(x), p\rangle\Big) = 0 \quad \text{и} \quad v_i(x) \geqslant 0 \quad \text{при} \quad i \in 1:m. \end{gathered} \tag{6}$$

В силу (4) условия (6) выполняются при $p = \mathbb{O}$ и $v_i(x) = u_i$, $i \in 1 : m$. Значит, нулевой вектор является решением задачи (5). На основании единственности решения заключаем, что $p(x) = \mathbb{O}$.
Достаточность. Пусть при некотором $x \in \mathbb{R}^n$ задача (5) имеет решение $p(x) = \mathbb{O}$. Отсюда, в частности, следует, что $x \in \Omega$. Критерий оптимальности (6) принимает вид

$$f_0'(x) + \sum_{i=1}^{m} v_i(x) f_i'(x) = \mathbb{O},$$
$$v_i(x) f_i(x) = 0 \quad \text{и} \quad v_i(x) \geqslant 0 \quad \text{при} \quad i \in 1 : m.$$

При $u_i = v_i(x)$ получаем условия Куна-Таккера (4). Значит, x — стационарная точка для задачи (1).

Лемма доказана. □

Напомним, что метод линеаризации прекращает свою работу, когда $p_k = \mathbb{O}$. В этом случае, по лемме, x_k — стационарная точка для задчи (1). В дальнейшем считаем, что $p_k \neq \mathbb{O}$ при всех $k = 0, 1, 2, \ldots$.

3°. Анализ сходимости метода линеаризации будем проводить при следующих предположениях:

(a) все градиенты $f_i'(x)$ удовлетворяют условию Липшица, то есть существует константа L, такая, что
$$\big\| f_i'(x) - f_i'(y) \big\| \leqslant L \|x - y\|$$
при любых x, y из $\mathbb{R}^n$ и всех $i \in 0 : m$;

(b) существует констата $N > 0$, такая, что множество
$$\Omega_N = \big\{ x \in \mathbb{R}^n \mid \Phi_N(x) \leqslant \Phi_N(x_0) \big\}$$
ограничено;

(c) задача квадратичного программирования (5) имеет решение $p(x)$ при любом $x \in \Omega_N$, и существуют такие множители Лагранжа $v_i(x)$, $i \in 1 : m$, что
$$v_i(x) \geqslant 0 \quad \text{при} \quad i \in 1 : m \quad \text{и} \quad \sum_{i=1}^{m} v_i(x) \leqslant N.$$

ЛЕММА 2. *Пусть выполнены условия* (a)-(c). *Тогда все точки x_k, построенные методом линеаризации, принадлежат множеству Ω_N. Кроме того, справедливо предельное соотношение*

$$\lim_{k \to \infty} p_k = \mathbb{O}. \tag{7}$$

Д о к а з а т е л ь с т в о. По определению Ω_N точка x_0 принадлежит Ω_N.

Пусть $x_k \in \Omega_N$. В силу условия (c) имеем

$$f_i(x_k) + \big\langle f_i'(x_k), p_k \big\rangle \leqslant 0,\ \ i \in 1:m. \tag{8}$$

При тех же i и $\alpha > 0$ запишем разложение

$$f_i(x_k + \alpha p_k) = f_i(x_k) + \alpha\big\langle f_i'(x_k), p_k\big\rangle + \alpha\big\langle f_i'(x_k + \theta_k^i \alpha p_k) - f_i'(x_k), p_k\big\rangle,$$

где $\theta_k^i \in (0,1)$. На основании условия (a) и (8) получаем

$$f_i(x_k + \alpha p_k) \leqslant f_i(x_k) - \alpha f_i(x_k) + L\alpha^2 \|p_k\|^2.$$

Если при этом $\alpha \leqslant 1$, то

$$f_i(x_k + \alpha p_k) \leqslant (1-\alpha)F(x_k) + L\alpha^2\|p_k\|^2,\ \ i \in 1:m.$$

Отсюда следует, что при $\alpha \in (0,1]$ справедливо неравенство

$$F(x_k + \alpha p_k) \leqslant (1-\alpha)F(x_k) + L\alpha^2\|p_k\|^2. \tag{9}$$

Теперь оценим целевую функцию. Согласно (6) (включая условие дополнительности) имеем

$$\big\langle f_0'(x_k), p_k\big\rangle = -\|p_k\|^2 - \sum_{i=1}^m v_i(x_k)\big\langle f_i'(x_k), p_k\big\rangle =$$
$$= \sum_{i=1}^m v_i(x_k) f_i(x_k) - \|p_k\|^2.$$

Поэтому

$$\begin{gathered} f_0(x_k + \alpha p_k) = f_0(x_k) + \alpha\big\langle f_0'(x_k), p_k\big\rangle + \\ + \alpha\big\langle f_0'(x_k + \theta_k^0\alpha p_k) - f_0'(x_k), p_k\big\rangle \leqslant \\ \leqslant f_0(x_k) + \alpha\Big(\sum_{i=1}^m v_i(x_k) f_i(x_k) - \|p_k\|^2\Big) + L\alpha^2\|p_k\|^2. \end{gathered}$$

Окончательно с учётом (9) при $\alpha \in (0,1]$ получаем

$$\begin{gathered} \Phi_N(x_k + \alpha p_k) = f_0(x_k + \alpha p_k) + NF(x_k + \alpha p_k) \leqslant f_0(x_k) + NF(x_k) + \\ + \alpha\Big(\sum_{i=1}^m v_i(x_k) f_i(x_k) - NF(x_k)\Big) - \alpha\|p_k\|^2 + L\alpha^2(N+1)\|p_k\|^2. \end{gathered}$$

В силу условия (c) и неотрицательности множителей Лагранжа

$$\sum_{i=1}^{m} v_i(x_k) f_i(x_k) - NF(x_k) \leqslant 0,$$

так что

$$\Phi_N(x_k + \alpha p_k) \leqslant \Phi_N(x_k) - \alpha \|p_k\|^2 \bigl(1 - L\alpha(N+1)\bigr).$$

Более того, при $\alpha \in (0, \hat{\alpha}]$, где

$$\hat{\alpha} = \min\Bigl\{1, \frac{1-\varepsilon}{L(N+1)}\Bigr\},$$

выполняется неравенство

$$\Phi_N(x_k + \alpha p_k) \leqslant \Phi_N(x_k) - \varepsilon\alpha \|p_k\|^2. \qquad (10)$$

Пусть $\frac{1}{2^s} < \hat{\alpha} \leqslant \frac{1}{2^{s-1}}$. Тогда по определению α_k имеем

$$\alpha_k \geqslant \frac{1}{2^s} = \frac{1}{2 \cdot 2^{s-1}} \geqslant \frac{1}{2}\hat{\alpha}. \qquad (11)$$

Значит, при всех k шаг α_k будет найден не более чем за $s+1$ испытаний неравенства (10).

Неравенство (3) перепишем в виде

$$\Phi_N(x_{k+1}) \leqslant \Phi_N(x_k) - \varepsilon\alpha_k \|p_k\|^2.$$

Отсюда следует, что $x_{k+1} \in \Omega_N$. Кроме того, в силу (11) получаем

$$\Phi_N(x_k) - \Phi_N(x_{k+1}) \geqslant \tfrac{1}{2}\varepsilon\hat{\alpha}\|p_k\|^2. \qquad (12)$$

По условию (b) монотонно убывающая последовательность $\Phi_N(x_k)$, $k = 0$, 1, $2, \ldots$, ограничена снизу. Значит, она имеет предел. В частности, левая часть неравенства (12) стремится к нулю при $k \to \infty$. Учитывая, что правая часть неравенства (12) неотрицательна, заключаем, что $p_k \to \mathbb{O}$.

Лемма доказана. □

4°. Переходим к основному результату.

ТЕОРЕМА (о сходимости метода линеаризации)**.** *Пусть выполнены условия* (a)-(c)*. Тогда* $F(x_k) \to 0$ *при* $k \to \infty$*, и любая предельная точка последовательности* $\{x_k\}$ *является стационарной точкой для задачи* (1)*.*

Доказательство. Мы находимся в условиях леммы 2. Поэтому все x_k принадлежат множеству Ω_N, и справедливы формулы (7) и (8). Обозначим

$$C = \max_{i\in 1:m} \max_{x\in\Omega_N} \big\| f_i'(x) \big\|.$$

Согласно (8) имеем

$$f_i(x_k) \leqslant -\big\langle f_i'(x_k), p_k \big\rangle \leqslant C\|p_k\|, \; i \in 1:m.$$

Отсда следует, что

$$0 \leqslant F(x_k) \leqslant C\|p_k\|.$$

Учитывая предельное соотношение (7), заключаем, что $F(x_k) \to 0$ при $k \to \infty$. Это значит, что последовательность $\{x_k\}$ приближается к множеству Ω.

Далее, на основании критерия оптимальности (6) запишем

$$\begin{gathered} f_0'(x_k) + p_k + \sum_{i=1}^m v_i^k f_i'(x_k) = \mathbb{O}, \\ v_i^k \Big(f_i(x_k) + \big\langle f_i'(x_k), p_k \big\rangle \Big) = 0 \quad \text{и} \quad v_i^k \geqslant 0 \quad \text{при} \quad i \in 1:m. \end{gathered} \tag{13}$$

Пусть x_* — предельная точка последовательности $\{x_k\}$. В силу компактности множества Ω_N хотя бы одна такая точка существует. Так как $F(x_k) \to 0$, то $F(x_*) = 0$, т. е. $x_* \in \Omega$. Условие (c) гарантирует, что существуют предельные точки и у всех последовательностей $\{v_i^k\}$, $i \in 1:m$. Обозначим их через v_i^*. Переходя к пределу в (13), получаем

$$\begin{gathered} f_0'(x_*) + \sum_{i=1}^m v_i^* f_i'(x_*) = \mathbb{O}; \\ v_i^* f_i(x_*) = 0 \quad \text{и} \quad v_i^* \geqslant 0 \quad \text{при} \quad i \in 1:m. \end{gathered}$$

Вместе с условием $x_* \in \Omega$ это означает, что x_* — стационарная точка для задачи (1).

Теорема доказана. □

5°. В методе линеаризации на каждом шаге приходится решать задачу квадратичного программирования (2). Запишем двойственную к ней задачу.

Обозначим через A матрицу со строками $-f_i'(x_k)$, $i = 1, \ldots, m$, и пусть

$$c = f_0'(x_k); \quad b[i] = f_i(x_k), \quad i = 1, \ldots, m.$$

Тогда задача (2) примет вид

$$\begin{gathered} \frac{1}{2}\langle Ep, p\rangle + \langle c, p\rangle \to \inf, \\ Ap \geqslant b, \end{gathered} \tag{14}$$

где E — единичная матрица порядка n. Экстремальная задача

$$\begin{gathered}
-\frac{1}{2}\langle Ez, z\rangle + \langle b, v\rangle \to \sup, \\
-z + A^T v = c, \\
v[i] \geqslant 0,\ i \in 1:m,
\end{gathered} \tag{15}$$

является двойственной к (14) задачей [2, с. 114].

В двойственной задаче (15) вектор z можно исключить и перейти к задаче

$$\begin{gathered}
\frac{1}{2}\langle A^T v - c, A^T v - c\rangle - \langle b, v\rangle \to \inf, \\
v[i] \geqslant 0,\ i \in 1:m,
\end{gathered}$$

которую можно переписать так:

$$\begin{gathered}
\frac{1}{2}\langle AA^T v, v\rangle - \langle Ac + b, v\rangle \to \inf, \\
v[i] \geqslant 0,\ i \in 1:m.
\end{gathered} \tag{16}$$

Допустим, что мы нашли решение v^k задачи (16). Тогда пара (z^k, v^k), где $z^k = A^T v^k - c$, будет решением задачи (15). По первой теореме двойственности [2, с. 115] существует решение p^k и у задачи (14). По второй теореме двойственности [2, с. 116] $E(p^k - z^k) = \mathbb{O}$, так что $p^k = z^k$. Значит, вектор

$$p^k = A^T v^k - c$$

является решением задачи (14). Мы нашли и решение p^k задачи (2), и соответствующий двойственный вектор v^k.

6°. Выше была описана принципиальная схема метода линеаризации, в которой используется штрафной коэффициент N. Он служит равномерной оценкой для суммы множителей Лагранжа (условие (c)). Заранее такой параметр указать трудно. В связи с этим приведём адаптированный вариант метода линеаризации для решения задачи (1) [1, с. 51].

Возьмём произвольное начальное приближение x_0 и достаточно большое число $N_0 > 0$.

***k*-й шаг**. Пусть имеются x_k и N_k. Решаем вспомогательную задачу (2). Пусть p_k — её решение и v^k — соответствующий двойственный вектор. Если $p_k = \mathbb{O}$, то x_k — стационарная точка. Вычисления заканчиваются. В противном случае полагаем

$$x_{k+1} = x_k + \alpha_k p_k,$$

где α_k — первое число в последовательности $1, \frac{1}{2}, \frac{1}{4}, \ldots,$ при котором

$$\Phi_{N_k}(x_k + \alpha_k p_k) \leqslant \Phi_{N_k}(x_k) - \varepsilon\alpha_k \|p_k\|^2.$$

Если при этом

$$\sigma_k := \sum_{i=1}^{m} v_i^k > N_k,$$

то полагаем $N_{k+1} = 2\sigma_k$. Иначе $N_{k+1} = N_k$.

ЛИТЕРАТУРА

1. Пшеничный Б. Н. *Метод линеаризации.* М.: Наука, 1983. 136 с.

2. Гавурин М. К., Малозёмов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984, 176 с.