СХОДИМОСТЬ МЕТОДА ЛИНЕАРИЗАЦИИ ПО КОЭФФИЦИЕНТУ НЕСТАЦИОНАРНОСТИ*

В. Н. Малозёмов

v.malozemov@spbu.ru

3 ноября 2016 г.

Аннотация. В докладе рассматривается принципиальная схема метода линеаризации Пшеничного для решения дискретных задач чебышёвского приближения. Приводится доказательство сходимости метода по коэффициенту нестационарности.

1°. Рассмотрим дискретную задачу чебышёвского приближения:

$$\varphi(x) := \max_{i\in I}\bigl|f_i(x)\bigr| \to \inf_{x\in\mathbb{R}^n}. \tag{1}$$

Здесь $f_i(x)$ — непрерывно дифференцируемые на $\mathbb{R}^n$ функции. Предполагается, что

$$\mu := \inf_{x\in\mathbb{R}^n} \varphi(x) > 0.$$

Наша цель — найти стационарную точку функции $\varphi(x)$.

2°. Обозначим через $R(x)$ множество индексов максимального уклонения в точке x,

$$R(x) = \Bigl\{ i \in I \mid \bigl|f_i(x)\bigr| = \varphi(x) \Bigr\}.$$

Введём вектор-функции

$$V_i(x) = \bigl(\operatorname{sign} f_i(x)\bigr) f_i'(x), \quad i \in R(x).$$

Напомним, что точка $x^* \in \mathbb{R}^n$ называется *стационарной точкой* функции $\varphi(x)$, если найдутся неотрицательные числа α_i, $i \in R(x^*)$, в сумме равные единице, такие что

$$\sum_{i\in R(x^*)} \alpha_i V_i(x^*) = \mathbb{O}. \tag{2}$$

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

3°. С задачей (1) связывают линеаризованную в точке x задачу:

$$\max_{i\in I}\Big|f_i(x)+\big\langle f_i'(x),p\big\rangle\Big|+\tfrac{1}{2}\|p\|^2\to\inf_{p\in\mathbb{R}^n}. \tag{3}$$

Здесь и далее используется евклидова норма векторов.

ЛЕММА 1. *При любом x задача* (3) *имеет решение и это решение единственно.*

Доказательство. При фиксированном x задача (3) эквивалентна задаче квадратичного программирования

$$\begin{aligned}\tfrac{1}{2}\|p\|^2+\eta&\to\inf\\ -\big\langle f_i'(x),p\big\rangle+\eta&\geqslant\ \ f_i(x),\ i\in I,\\ \big\langle f_i'(x),p\big\rangle+\eta&\geqslant -f_i(x),\ i\in I.\end{aligned} \tag{4}$$

Множество планов задачи (4) непусто (планом является пара $p=\mathbb{O}$, $\eta=\varphi(x)$) и целевая функция ограничена снизу нулём. Значит, задача (4) имеет решение $\big(p(x),\eta(x)\big)$. По эквивалентности, $p(x)$ — решение задачи (3).

Ясно, что функция

$$F(x,p)=\max_{i\in I}\Big|f_i(x)+\big\langle f_i'(x),p\big\rangle\Big| \tag{5}$$

при фиксированном x выпукла по p. Слагаемое $\frac{1}{2}\|p\|^2$ делает целевую функцию задачи (3) строго выпуклой. Из строгой выпуклости следует единственность решения $p(x)$. □

4°. С помощью решения $p(x)$ линеаризованной задачи (3) можно получить критерий стационарности для задачи (1).

ЛЕММА 2. *Для того чтобы точка $x^*\in\mathbb{R}^n$ была стационарной точкой функции $\varphi(x)$, необходимо и достаточно, чтобы $p(x^*)=\mathbb{O}$.*

Доказательство. Воспользуемся критерием оптимальности для задачи квадратичного программирования (4) (см. [1, с. 91]): план (p,η) будет оптимальным тогда и только тогда, когда найдутся неотрицательные числа u_i, v_i, $i\in I$, со свойствами

$$\begin{aligned}\sum_{i\in I}(u_i-v_i)f_i'(x)&=-p,\\ \sum_{i\in I}(u_i+v_i)&=1,\\ u_i\Big(-f_i(x)-\big\langle f_i'(x),p\big\rangle+\eta\Big)&=0,\quad i\in I,\\ v_i\Big(f_i(x)+\big\langle f_i'(x),p\big\rangle+\eta\Big)&=0,\quad i\in I.\end{aligned} \tag{6}$$

Пусть x^* — стационарная точка, то есть выполнено условие (2). Покажем, что пара $p^* = \mathbb{O}$, $\eta^* = \varphi(x^*)$ является оптимальным планом задачи (4) при $x = x^*$. Отсюда будет следовать равенство $p(x^*) = \mathbb{O}$.

Вектор (p^*, η^*) удовлетворяет ограничениям задачи (4) при $x = x^*$. Условия оптимальности (6) для него принимают вид

$$\sum_{i\in I}(u_i - v_i)f_i'(x^*) = \mathbb{O}, \quad \sum_{i\in I}(u_i + v_i) = 1, \tag{7}$$

$$u_i\big(-f_i(x^*) + \varphi(x^*)\big) = 0, \quad v_i\big(f_i(x^*) + \varphi(x^*)\big) = 0,\ i \in I. \tag{8}$$

Нетрудно проверить, что эти условия выполняются при

$$u_i = \tfrac{1}{2}\alpha_i\big(1 + \operatorname{sign} f_i(x^*)\big),\ v_i = \tfrac{1}{2}\alpha_i\big(1 - \operatorname{sign} f_i(x^*)\big),\ \text{когда}\ \ i \in R(x^*),$$
$$u_i = v_i = 0,\ \text{когда}\ \ i \in I \setminus R(x^*).$$

Здесь α_i — коэффициенты из условия (2). Значит, $p(x^*) = p^* = \mathbb{O}$.

Наоборот, пусть $p(x^*) = \mathbb{O}$. Тогда пара $p^* = \mathbb{O}$, $\eta^* = \varphi(x^*)$ будет оптимальным планом задачи (4) при $x = x^*$. Критерий оптимальности для этого плана имеет вид (7), (8). Положим $\alpha_i = u_i + v_i$, $i \in I$. Числа α_i неотрицательны и в сумме равны единице. Более того, согласно (8), $u_i = v_i = 0$ при $i \in I \setminus R(x^*)$. При $i \in R(x^*)$ сложим равенства из (8). Получим

$$(u_i - v_i)f_i'(x^*) = \alpha_i\,\varphi(x^*).$$

Поделив на $\big|f_i(x^*)\big| = \varphi(x^*)$, придём к представлению

$$u_i - v_i = \alpha_i \operatorname{sign} f_i(x^*), \quad i \in R(x^*).$$

Теперь первая формула в (7) принимает вид

$$\sum_{i\in R(x^*)} \alpha_i\big(\operatorname{sign} f_i(x^*)\big)f_i'(x^*) = \mathbb{O},$$

что равносильно (2). Установлено, что x^* — стационарная точка.

Лемма доказана. □

5°. Введём величину

$$\beta(x) = \varphi(x) - F\big(x, p(x)\big),$$

где функция F определена формулой (5) и $p(x)$ — решение задачи (3). Учитывая неравенство

$$F\big(x, p(x)\big) + \tfrac{1}{2}\big\|p(x)\big\|^2 \leqslant F(x, \mathbb{O}) = \varphi(x),$$

заключаем, что

$$\beta(x) \geqslant \tfrac{1}{2}\big\|p(x)\big\|^2 \geqslant 0. \tag{9}$$

ЛЕММА 3. *Точка x^* будет стационарной точкой функции $\varphi(x)$ тогда и только тогда, когда $\beta(x^*) = 0$.*

Доказательство. Если $\beta(x^*) = 0$, то, в силу (9), $p(x^*) = \mathbb{O}$. Согласно лемме 2, x^* — стационарная точка.

Наоборот, пусть x^* — стационарная точка. По лемме 2, $p(x^*) = \mathbb{O}$. Согласно определению, $\beta(x^*) = \varphi(x^*) - F(x^*, \mathbb{O}) = 0$. $\square$

Величину $\beta(x)$ назовём *коэффициентом нестационарности* точки x.

6°. Обратимся к методу линеаризации Пшеничного нахождения стационарной точки функции $\varphi(x)$ (см. [2, глава 3]). Опишем его принципиальную схему.

В качестве начального приближения возьмём произвольную точку $x_0 \in \mathbb{R}^n$.

Пусть имеется k-е приближение x_k. Проверим x_k на стационарность. Для этого найдём решение $p_k = p(x_k)$ задачи (3) при $x = x_k$ и вычислим коэффициент нестационарности $\beta(x_k) = \varphi(x_k) - F(x_k, p_k)$. Если $\beta(x_k) = 0$, то x_k — стационарная точка. Процесс закончен. В противном случае (при $\beta(x_k) > 0$) вычисляем очередное приближение по формуле

$$x_{k+1} = x_k + t_k p_k,$$

где t_k — первый элемент последовательности $t = 1, \frac{1}{2}, \frac{1}{4}, \ldots$, при котором выполняется неравенство

$$\varphi(x_k + tp_k) \leqslant \varphi(x_k) - \delta t \beta(x_k). \tag{10}$$

Здесь $\delta \in (0, 1)$ — параметр метода. Условие (10) можно переписать в виде

$$\frac{\varphi(x_k) - \varphi(x_k + tp_k)}{t\beta(x_k)} \geqslant \delta.$$

Описание метода завершено.

Если последовательность $\{x_k\}$ конечна, то последней её элемент является стационарной точкой функции $\varphi(x)$. Рассмотрим основной случай, когда последовательность $\{x_k\}$ бесконечна, то есть когда все коэффициенты $\beta(x_k)$ положительны.

По построению последовательность $\big\{\varphi(x_k)\big\}$ строго убывает. При этом она ограничена снизу числом $\mu > 0$ (см. начало п. 1°). Значит, последовательность $\big\{\varphi(x_k)\big\}$ имеет предел. В частности,

$$\lim_{k \to \infty} \big(\varphi(x_k) - \varphi(x_{k+1})\big) = 0. \tag{11}$$

Мы покажем, что при некотором естественном предположении

$$\lim_{k\to\infty} \beta(x_k) = 0. \tag{12}$$

Именно это мы имеем в виду, когда говорим о сходимости метода линеаризации по коэффициенту нестационарности.

Из (12) следует, что за конечное число шагов будет получена ε-стационарная точка x_k, в которой $\beta(x_k) < \varepsilon$ (при любом $\varepsilon > 0$).

7°. Введём обозначения

$$M_0 = \bigl\{x \in \mathbb{R}^n \mid \varphi(x) \leqslant \varphi(x_0)\bigr\}, \quad M_c = M_0 + B_c,$$

где x_0 —начальное приближение в методе линеаризации и $B_c = \bigl\{x \in \mathbb{R}^n \mid \|x\| \leqslant c\bigr\}$. Предположим, что на множестве M_c при некотором $c > 0$ все градиенты $f_i'(x)$, $i \in I$, удовлетворяют условию Липшица с константой L, то есть

$$\bigl\|f_i'(x) - f_i'(y)\bigr\| \leqslant L\|x - y\| \quad \forall x, y \in M_c. \tag{13}$$

Напомним, что $p(x)$ — это решение вспомогательной задачи (3).

ЛЕММА 4. *Пусть выполнено условие* (13). *Тогда для любой нестационарной точки $x \in M_0$ при $t \in (0, 1]$, $\bigl\|tp(x)\bigr\| \leqslant c$ справедливо неравенство*

$$\Delta(x, t) := \frac{\varphi(x) - \varphi\bigl(x + tp(x)\bigr)}{t\beta(x)} \geqslant 1 - 2Lt. \tag{14}$$

Доказательство. По теореме о среднем

$$f_i\bigl(x + tp(x)\bigr) = f_i(x) + t\Bigl\langle f_i'\bigl(x + \theta_i tp(x)\bigr), p(x)\Bigr\rangle =$$
$$= (1 - t)f_i(x) + t\Bigl[f_i(x) + \bigl\langle f_i'(x), p(x)\bigr\rangle\Bigr] + t\Bigl\langle f_i'\bigl(x + \theta_i tp(x)\bigr) - f_i'(x), p(x)\Bigr\rangle.$$

В силу условий леммы относительно x и t и (13) имеем

$$\Bigl|f_i\bigl(x + tp(x)\bigr)\Bigr| \leqslant (1 - t)\bigl|f_i(x)\bigr| + t\Bigl|f_i(x) + \bigl\langle f_i'(x), p(x)\bigr\rangle\Bigr| + L\,t^2\bigl\|p(x)\bigr\|^2 \leqslant$$
$$\leqslant (1 - t)\varphi(x) + t\,F\bigl(x, p(x)\bigr) + L\,t^2\bigl\|p(x)\bigr\|^2.$$

Отсюда и из (9) следует, что

$$\varphi\bigl(x + tp(x)\bigr) \leqslant (1 - t)\varphi(x) + t\,F\bigl(x, p(x)\bigr) + L\,t^2\bigl\|p(x)\bigr\|^2 \leqslant$$
$$\leqslant \varphi(x) - t\beta(x) + 2L\,t^2\beta(x).$$

и

$$\varphi(x) - \varphi\bigl(x + tp(x)\bigr) \geqslant t\beta(x)[1 - 2L\,t].$$

Поделив на $t\beta(x)$, получим (14).

Лемма доказана. □

8°. Вернёмся к методу линеаризации. Отметим, что в силу неравенства (9) и неравенств $\beta(x_k) \leqslant \varphi(x_k) \leqslant \varphi(x_0)$ справедлива равномерная оценка

$$\|p(x_k)\| \leqslant \sqrt{2\beta(x_k)} \leqslant \sqrt{2\varphi(x_0)}.$$

Положим

$$\tau = \min\Big\{1, \frac{c}{\sqrt{2\varphi(x_0)}}, \frac{1-\delta}{2L}\Big\}.$$

Ясно, что $\tau \in (0,1]$. При $t \in (0,\tau]$ имеем

$$\|tp(x_k)\| \leqslant \frac{c}{\sqrt{2\varphi(x_0)}}\|p(x_k)\| \leqslant c.$$

Поэтому по лемме 4

$$\Delta(x_k, t) \geqslant 1 - 2L\,t \quad \forall t \in (0,\tau].$$

Более того, при тех же t

$$1 - 2L\,t \geqslant 1 - 2L\,\tau \geqslant 1 - 2L\frac{1-\delta}{2L} = \delta,$$

так что при всех k

$$\Delta(x_k, t) \geqslant \delta \quad \forall t \in (0,\tau]. \tag{15}$$

ТЕОРЕМА. *При выполнении условия* (13) *справедливо предельное соотношение*

$$\lim_{k\to\infty} \beta(x_k) = 0.$$

Д о к а з а т е л ь с т в о. Перепишем неравенство (15) в развёрнутом виде

$$\varphi(x_k + tp_k) \leqslant \varphi(x_k) - \delta t\beta(x_k) \quad \forall t \in (0,\tau]. \tag{16}$$

Выберем натуральное число s_0 из условия

$$\frac{1}{2^{s_0}} < \tau \leqslant \frac{1}{2^{s_0-1}}.$$

При $t = \frac{1}{2^{s_0}}$ неравенство (16) выполняется. По определению, t_k первое число в последовательности $t = 1, \frac{1}{2}, \frac{1}{4}, \ldots$ с таким свойством. Значит, $t_k \geqslant \frac{1}{2^{s_0}}$.

Теперь имеем

$$\varphi(x_k) - \varphi(x_{k+1}) \geqslant \delta t_k\beta(x_k) \geqslant \frac{\delta}{2^{s_0}}\beta(x_k) \geqslant 0.$$

Принимая во внимание (11), приходим к заключению теоремы. $\square$

ЛИТЕРАТУРА

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

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