Метод минимизации второго порядка для негладких функций, допускающих выпуклую квадратичную аппроксимацию приращения

М.Э. Аббасов
Санкт-Петербургский Государственный Университет, Санкт-Петербург
e-mail: abbasov.majid@gmail.com

Будем рассматривать задачу минимизации функции $f\colon \mathbb{R}^n \to \mathbb{R}$, допускающей разложение (см. [1, 2])

$$f(x+\Delta) = f(x) + \max_{i\in I}\left(a_i(x) + \langle v_i(x), \Delta\rangle + \frac{1}{2}\langle A_i(x)\Delta, \Delta\rangle\right) + o_x(\Delta).$$

где I – конечное множество индексов, все матрицы $A_i(x)$ при $i \in I$ будем считать неотрицательно определенными для любых x из области определения f.

Опишем условие экстремума для изучаемой функции, а далее на его основе построим оптимизационный алгоритм.

Теорема 1. *Пусть функция $f(x)$ допускает разложение*

$$f(x+\Delta) = f(x) + \max_{i\in I}\left(a_i(x) + \langle v_i(x), \Delta\rangle + \frac{1}{2}\langle A_i(x)\Delta, \Delta\rangle\right) + o_x(\Delta),$$

где все матрицы $A_i(x) \in I$ неотрицательно определенные при любых $x \in \mathbb{R}^n$, а $o_x(\Delta)$ удовлетворяет

$$\lim_{\alpha\downarrow 0}\frac{o_x(\alpha\Delta)}{\alpha^2} = 0, \quad \forall\ \Delta \in \mathbb{R}^n.$$

Тогда, если $x_* = \operatorname*{argmin}_{x\in\mathbb{R}^n} f(x)$, то

$$0_n = \operatorname*{argmin}_{\Delta\in\mathbb{R}^n}\max_{i\in I}\left(a_i(x_*) + \langle v_i(x_*), \Delta\rangle + \frac{1}{2}\langle A_i(x_*)\Delta, \Delta\rangle\right). \qquad (1)$$

Перейдем теперь к самому алгоритму. Пусть x_0 – начальная точка. Предположим, мы получили точку x_k. Дальше определяем

$$y_k = \operatorname*{argmin}_{y\in\mathbb{R}^n}\max_{i\in I}\left(a_i(x_k) + \alpha\langle v_i(x_k), y\rangle + \frac{\alpha^2}{2}\langle A_i(x_k)y, y\rangle\right),$$

$$\alpha_k = \operatorname*{argmin}_{\alpha} f\left(x_k + \alpha \frac{y_k}{\|y_k\|}\right),$$

$$x_{k+1} = x_k + \alpha_k \frac{y_k}{\|y_k\|}.$$

В результате строим последовательность такую, что

$$f(x_{k+1}) < f(x_k)$$

Теорема 2. *Пусть множество*

$$P = \{x \in \mathbb{R}^n |\ f(x) \le f(x_0)\}$$

ограничено, x_* – предельная точка последовательности $\{x_k\}$, отображение $C(x) = \Big\{[a_i(x), v_i(x), A_i(x)] \Big|\ i \in I\Big\}$ непрерывно в метрике Хаусдорфа, а функция $o_x(\Delta)$ такова, что $\frac{o_x(\alpha\Delta)}{\alpha^2} \xrightarrow[\alpha\downarrow 0]{} 0$ равномерно по x из некоторой окрестности x_* и по Δ из $\mathbb{S} = \{\Delta \in \mathbb{R}^n \mid \|\Delta\| = 1\}$. Тогда точка x_* является стационарной точкой функции f на $\mathbb{R}^n$, то есть в ней выполнено условие (1).

Работа выполнена при поддержке РФФИ, грант №12-01-00752, а также при поддержке гранта СПбГУ 9.38.205.2014.

Литература

[1] *Демьянов В.Ф., Рубинов А.М.* Основы негладкого анализа и квазидифференциальное исчисление. М.: Наука, 1990. 432 с.

[2] *Демьянов В.Ф., Аббасов М.Э.* Условия экстремума негладкой функции в терминах экзостеров и коэкзостеров // Труды Института математики и механики УрО РАН. 2009, 15, № 4, С. 10–19.