УДК 519.8

РЕЛАКСАЦИОННЫЙ МЕТОД ОБОБЩЕННОГО ГРАДИЕНТА

Л.В.Васильев, В.Ф.Демьянов

Пусть на открытом выпуклом множестве $\Omega' \subset E_n$ задана ограниченная выпуклая функция f и пусть Ω - замкнутое выпуклое множество, содержащееся в Ω'. Рассматривается задача минимизации функции f на множестве Ω.

Известно, что f непрерывна на Ω' и в каждой точке $x \in \Omega$ дифференцируема по любому направлению $g \in E_n$, причем [I]

$$\frac{\partial f(x)}{\partial g} = \lim_{\alpha \to +0} \frac{f(x+\alpha g) - f(x)}{\alpha} = \max_{v \in M(x)} (v, g),$$

где $M(x) = \{v \in E_n / f(y) - f(x) \geq (v, y-x), \quad y \in \Omega\}$ - замкнутое выпуклое и ограниченное множество субградиентов (обобщенных градиентов). Точечно-множественное отображение $M(x)$ полунепрерывно сверху.

Пусть $\Gamma(x)$ - конус возможных направлений множества Ω в точке x, а $\Gamma^+(x)$ - сопряженный к нему ($\Gamma^+(x) = \{w / (w, v) \geq 0, \forall v \in \Gamma(x)\}$). Нетрудно установить, что точечно-множественное отображение $\Gamma^+(x)$ полунепрерывно сверху.

Для того чтобы в точке $x^* \in \Omega$ функция f достигала своего минимального на Ω значения, необходимо и достаточно, чтобы

$$M(x^*) \cap \Gamma^+(x^*) \neq \emptyset . \qquad (1)$$

Для наших целей понадобится

ЛЕММА (ср. [2, стр.183]). Если множество Q имеет внутренние точки, то условие (1) эквивалентно условию

$$0 \in L(x^*),$$

где $L(x) = co\{M(x) \cup T_\eta(x)\}$,

$$T_\eta(x) = \begin{cases} M(x), & \text{если } \Gamma^+(x) = \{0\}, \\ \{v \mid -v \in \Gamma^+(x),\ \|v\| = \eta\}, & \text{если } \Gamma^+(x) \neq \{0\}, \end{cases}$$

$\eta > 0$ – произвольное число.

Для минимизации выпуклой функции на всем пространстве в [3,4] разработан метод обобщенного градиента, в [5] этот метод распространен на случай минимизации выпуклой функции на выпуклом множестве. Метод обобщенного градиента не является релаксационным (метод называется релаксационным, если значения минимизируемой функции не возрастают на последовательных приближениях). В [6,7] разработан релаксационный метод для минимизации выпуклой функции на всем пространстве. Ниже описывается релаксационный метод для минимизации выпуклой функции на выпуклом множестве. При отсутствии ограничений предлагаемый метод близок по идее к методу [6,7].

Выберем произвольно $x_0 \in Q$. Предположим, что множество $D(x_0) = \{x \in Q \mid f(x) \le f(x_0)\}$ ограничено.

Пусть $b = \sup\limits_{x \in D(x_0)} \sup\limits_{v \in M(x)} \|v\|$. Ясно, что $b < \infty$.

Зафиксируем положительные числа ε, μ, η и h, причем $\varepsilon < b$, $\eta \le b$, и целое число m такое, что

$$m > \frac{\ln \frac{\varepsilon}{b}}{\ln q}, \qquad \text{где} \qquad q = \frac{1}{\sqrt{1 + (\frac{\varepsilon}{b})^2}} < 1.$$

Пусть уже найдено приближение $x_k \in Q$. Положим $x_{k0} = x_k$. Выберем произвольно $v_{k0} \in L(x_{k0})$. Если $\|v_{k0}\| \le \varepsilon$, то процесс прекращается. В противном случае на луче $x_{k0}(\alpha) = x_{k0} - \alpha v_{k0}$, $\alpha > 0$, найдем точку $x_{k1} = x_{k0}(\alpha_{k0})$, в которой

$$f(x_{k1}) = \min_{\alpha > 0,\ x_{k0}(\alpha) \in Q} f(x_{k0}(\alpha)). \qquad (2)$$

Если $\|x_{k1} - x_{k0}\| > h$ или $f(x_{k0}) - f(x_{k1}) \geq \mu$, то полагаем $x_{k+1} = x_{k1}$. В противном случае находим $\bar{v}_{k0} \in L(x_{k1})$ такой, что

$$(v_{k0}, \bar{v}_{k0}) \leq 0. \qquad (3)$$

В силу (2) такой вектор $\bar{v}_{k0}$ найдется. Действительно, в точке x_{k1} должно быть

либо $\dfrac{\partial f(x_{k1})}{\partial g_{k0}} = \max_{v \in M(x_{k1})} (v, g_{k0}) \geq 0$, где $g_{k0} = -v_{k0}$, и должен найтись такой вектор $\bar{v}_{k0} \in M(x_{k1})$, для которого выполнено (3);

либо $x_{k1} - \beta v_{k0} \notin Q$ при любом $\beta > 0$, т.е.

$$-v_{k0} \in \{v = \lambda(z - x_{k1}) \mid \lambda > 0,\ z \in Q\},$$

а тогда найдется $\bar{v}_{k0} \in \Gamma^{+}(x_{k1})$ такой, что имеет место (3). Без ограничения общности можно считать, что $\|\bar{v}_{k0}\| = \eta$.

Если $\|\bar{v}_{k0}\| \leq \varepsilon$, то полагаем $x_{k+1} = x_{k1}$, и процесс прекращается. Если же $\|\bar{v}_{k0}\| > \varepsilon$, то находим такой вектор v_{k1},

$$v_{k1} = \beta_{k0} v_{k0} + (1 - \beta_{k0}) \bar{v}_{k0}, \qquad 0 < \beta_{k0} < 1, \qquad \text{что}$$

$$\|v_{k1}\| = \min_{0 \leq \beta \leq 1} \|\beta v_{k0} + (1-\beta)\bar{v}_{k0}\|.$$

Построим луч $x_{k1}(\alpha) = x_{k1} - \alpha v_{k1}$, $\alpha > 0$, и найдем на нем точку $x_{k2} = x_{k1}(\alpha_{k1})$, в которой

$$f(x_{k2}) = \min_{\alpha > 0,\ x_{k1}(\alpha) \in Q} f(x_{k1}(\alpha)).$$

Далее продолжаем аналогично до тех пор, пока не найдем точку x_{kt_k} такую, что либо а) $t_k = m+1$, либо б) $\|x_{kt_k} - x_{kt_k-1}\| \geq h$, либо в) $f(x_{kt_k-1}) - f(x_{kt_k}) \geq \mu$, либо г) $\|\bar{v}_{kt_k}\| \leq \varepsilon$.

В случаях а), б) и в) полагаем $x_{k+1} = x_{kt_k}$, а в случае г) процесс прекращается.

В результате построим последовательность $\{x_k\}$.

Положим $\rho(x) = \min\limits_{v \in L(x)} \|v\|$. Если полученная последовательность конечна, то для ее последней точки x_k будет $\rho(x_k) \le \varepsilon$. В противном случае справедлива

ТЕОРЕМА. Если функция f строго выпукла на Ω, то для всякой предельной точки x^* последовательности $\{x_k\}$ выполнено соотношение $\rho(x^*) \le \varepsilon$.

ДОКАЗАТЕЛЬСТВО. Допустим противное. Пусть $x_{k_s} \longrightarrow x^*$ и $\rho(x^*) > \varepsilon$. Нетрудно установить полунепрерывность сверху точечно-множественного отображения $L(x)$. Поэтому найдется такое число $\delta > 0$, что

$$\min_{x \in S_\delta(x^*)} \min_{v \in L(x)} \|v\| > \varepsilon,$$

где $S_\delta(x^*) = \{x \in \Omega \mid \|x - x^*\| \le \delta\}$.

В силу строгой выпуклости f при достаточно больших k_s будет $t_{k_s} = m+1$, $x_{k_s i} \in S_\delta(x^*)$, $\forall i \in 0 : t_{k_s}$, а также

$$\|v_{k_s i}\| > \varepsilon, \quad \forall i \in 0 : m. \tag{4}$$

Так как $\|v_{k_s i}\| \le b$, то из (3), (4) следует

$$\|v_{k_s, i+1}\| \le \|v_{k_s i}\| \frac{1}{\sqrt{1 + (\frac{\varepsilon}{b})^2}} = \|v_{k_s i}\| q.$$

Отсюда получаем $\|v_{k_s i}\| \le \|v_{k_s 0}\| q^i \le b q^i$. Значит, $\|v_{k_s m}\| \le b q^m$. Но поскольку $m > \frac{\ln(\varepsilon/b)}{\ln q}$, то имеем $\|v_{k_s m}\| < \varepsilon$, что противоречит (4). Теорема доказана.

Отметим, что можно построить метод, в котором минимум на луче ищется приближенно, при этом условие (3) нужно заменить условием $(v_{ki}, \bar{v}_{ki}) \le \gamma$, $\gamma > 0$.

Если функция f выпукла, то ее можно аппроксимировать строго выпуклой функцией, например, $f_1(x) = f(x) + \nu x^2$, $\nu > 0$.

Метод был опробован при решении минимаксных задач. При хорошей организации вычислений, он, несмотря на свою простоту, незначительно уступает методу наискорейшего спуска.

Литература

I. ПШЕНИЧНЫЙ Б.Н. Необходимые условия экстремума. М., "Наука", 1969.

2. ДЕМЬЯНОВ В.Ф., МАЛОЗЕМОВ В.Н. Введение в минимакс. М., "Наука", 1972.

3. ЕРМОЛЬЕВ Ю.М., ШОР Н.З. О минимизации недифференцируемых функций. - "Кибернетика", 1967, № I.

4. ШОР Н.З. Обобщенный градиентный спуск. М., Труды I зимней школы по математическому программированию, г. Дрогобыч, 1969, вып. 3.

5. ПОЛЯК Б.Т. Минимизация негладких функционалов. -"Журн.вычислит.математики и мат.физики", 1969, т.9, № 3, с.509-52I.

6. WOLFE Ph. Note on a method of conjugate subgradients for minimizing nondifferentiable functions. - "Mathematical Programming", 1974, v.7, p.380-383.

7. LEMARCHAL C. Note on an extension of Davidon methods to nondifferentiable functions. -"Math Programming", 1974, v.7, p.384 - 387.