УДК 519.8

Н.Г. ЖУРБЕНКО, Б.М. ЧУМАКОВ

АЛГОРИТМ МИНИМИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ МОДИФИКАЦИИ МЕТОДА ЭЛЛИПСОИДОВ

Для решения задачи минимизации выпуклой функции в конечномерном евклидовом пространстве предлагается ε–субградиентный алгоритм с преобразованием пространства. Алгоритм является модификацией метода эллипсоидов, основан на процедуре одномерного спуска и является, в некотором смысле, монотонным. Приводится оценка трудоемкости алгоритма при решении задачи ε-оптимизации.

Рассматривается задача минимизации ограниченной снизу выпуклой функции $f(x)$ в n-мерном евклидовом пространстве R^n: $f^* = \min\{f(x)|x\in R^n\}$.

Для заданного числа $\bar{\varepsilon} \geq 0$ введем $\bar{\varepsilon}$-оптимальное множество $X^*(\bar{\varepsilon})$:

$$X^*(\bar{\varepsilon}) = \{x\in R^n \,|\, f(x) \leq f^* + \bar{\varepsilon}\}. \qquad (1)$$

Произвольную точку $x \in X^*(\bar{\varepsilon})$ и значение функции $\tilde{f} = f(\tilde{x})$ будем называть решением задачи $\bar{\varepsilon}$-оптимизации ($\bar{\varepsilon}$-решением). Будем использовать несколько отличное от классического [1] определение ε-субградиента. Вектор $g \in R^n$ называется $(\varepsilon, \tilde{f})-$субградиентом функции $f(x)$ в точке z, если для $\forall x \in R^n$ выполняется неравенство

$$f(x) \geq \tilde{f} + (g, x - z) - \varepsilon, \qquad (2)$$

где $\tilde{f} \geq f^*; \varepsilon \in R^1$. Обычное классическое определение ε–субградиента [2] соответствует определению $(\varepsilon, \tilde{f})-$субградиента для $\tilde{f} = f(z)$ и $\varepsilon \geq 0$. Обобщенный градиент функции $f(x)$ в точке z в классическом смысле [3] является $(0, f(z))$–субградиентом. $G(\varepsilon, \tilde{f}, z), \partial f(z)$ будет обозначать множество $(\varepsilon, \tilde{f})-$субградиентов и множество обобщенных градиентов функции

$f(x)$ в точке z соответственно. В дальнейшем предполагается, что имеются алгоритмы вычисления $f(x)$ и $g(x) \in \partial f(x)$ в произвольной точке x.

Пусть $g \in G(\varepsilon_1, \tilde{f}_1, z_1)$; $f^* \le f_2$. Тогда $g \in G(\varepsilon_2, \tilde{f}_2, z_2)$, где

$$\varepsilon_2 = \varepsilon_1 + \tilde{f}_2 - \tilde{f}_1 - (g, z_2 - z_1). \tag{3}$$

Таким образом, $(\varepsilon, \tilde{f})$-субградиент в некоторой точке z_1 является $(\varepsilon, \tilde{f})$-субградиентом в любой другой точке z_2. Формула (3) определяет правило пересчета параметров $(\varepsilon, \tilde{f})$–субградиента. $G(\varepsilon, \tilde{f}, z)$ по своим свойствам аналогично множеству $\partial f(z)$ (выпуклость, замкнутость и т.д.). Например, $\{0 \in G(\varepsilon, \tilde{f}, z), \varepsilon \le \bar{\varepsilon}\} \Leftrightarrow \{\tilde{f} \le f^* + \bar{\varepsilon}\}$).

Пусть $\tilde{X}(\bar{\varepsilon}, \tilde{f}) = \{x \in R^n \mid f(x) \le \tilde{f} - \bar{\varepsilon}\}$.

Отметим, что

$\tilde{X}(\bar{\varepsilon}, \tilde{f}) = \varnothing \Rightarrow \{\tilde{f}$ является $\bar{\varepsilon}$ – решением$\}$;

$$\tilde{f} \ge f^* + 2\bar{\varepsilon} \Rightarrow X^*(\bar{\varepsilon}) \subset \tilde{X}(\bar{\varepsilon}, \tilde{f}). \tag{4}$$

Обозначения: $P(z,\eta) = \{x \in R^n \mid (x - z, \eta) = 0\}$ – плоскость, проходящая через z с нормалью $\eta \in R^n$, $|\eta| > 0$; $P^+(z,\eta) = \{x \in R^n \mid (x - z, \eta) \ge 0\}$ – полупространство, определяемое плоскостью $P(z,\eta)$.

Пусть

$$g \in G(\varepsilon, \tilde{f}, z);\ h = (\bar{\varepsilon} - \varepsilon) / |g|;\ \tilde{z} = z - hg / |g|. \tag{5}$$

Тогда $\tilde{X}(\bar{\varepsilon}, \tilde{f}) \subset P^+(z - \tilde{z}, -g)$. Этот факт соответствует обычному построению отсекающих плоскостей "со сдвигом".

Утверждение 1.

Для заданной точки z, направления спуска η $(|\eta| = 1)$ и $\bar{\varepsilon} > 0$ можно гарантировать вычисление такого ε–субградиента $g \in G(\varepsilon, \tilde{f}, z)$, что $\varepsilon \le \bar{\varepsilon}$ и $(g, \eta) \ge 0$.

Доказательство.

Пусть z_η^* – точка минимума функции на луче $x = z + t\eta$. Возможны два случая.

1. $z_\eta^* = z$. Тогда $f(z + t\eta) \ge f(z) \ge \tilde{f}$, $(g(z + t\eta), \eta) \ge 0$ для $\forall t > 0$, где $g(z + t\eta) \in \partial f(z + t\eta)$. Поэтому $f(x) \ge f(z + t\eta) + (g(z + t\eta), x - (z + t\eta)) \ge$ $\ge \tilde{f} + (g(z + t\eta), x - z) - t(g(x + t\eta), \eta)$. Пусть точность одномерного спуска обеспечивает неравенство $t(g(x + t\eta), \eta) \le \bar{\varepsilon}$. Тогда вектор $g = g(z + t\eta)$

удовлетворяет условиям утверждения. При этом найденное рекордное значение $\tilde{f}$ не изменяется.

2. $z_\eta^* = z + t^*\eta$, $t^* > 0$. Процедурой одномерного поиска определяем две точки $z_i = z + t_i\eta$, $i = 1,2$, для которых выполнены неравенства: $t_2 = t_1 + \delta$; $t_1 > 0$; $\delta > 0$; $(g_1, \eta) \le 0$; $(g_2, \eta) \ge 0$, где $g_i \in \partial f(z_i)$. Заметим, что $\delta = | z_2 - z_1 |$ – точность одномерного поиска минимума. Тогда из определения обобщенного градиента следует неравенство

$$f(x) \ge \lambda_1 f(z_1) + \lambda_2 f(z_2) + (\lambda_1 g_1 + \lambda_2 g_2, x - z) - t_1(\lambda_1 g_1 + \lambda_2 g_2, \eta) - \delta\lambda_2(g_2, \eta), \quad (6)$$

где $g = \lambda_1 g_1 + \lambda_2 g_2; \lambda_i \ge 0;\ \lambda_1 + \lambda_2 = 1$. Выберем λ_i из дополнительного условия $(g, \eta) = 0$. Тогда из (6), с учетом $\lambda_2 \le 1$, следует

$$f(x) \ge \lambda_1 f(z_1) + \lambda_2 f(z_2) + (g, x - z) - \delta(g_2, \eta). \quad (7)$$

Определим новое рекордное значение $\tilde{f} := \min\{\tilde{f}, f_1, f_2\}$. Тогда $\lambda_1 f(z_1) + \lambda_2 f(z_2) \ge \tilde{f}$. Пусть одномерный поиск минимума осуществляется с точностью, обеспечивающей выполнение неравенства $\delta(g_2, \eta) \le \varepsilon$. Тогда из (7) получаем $f(x) \ge \tilde{f} + (g, x - z) - \varepsilon$. Следовательно, вектор g удовлетворяет условию леммы.

Приведенное доказательство утверждения 1 содержит алгоритм вычисления указанного ε–субградиента. Этот алгоритм во многом аналогичен известным процедурам генерации субградиентов множеств [2,4].

Следующее утверждение касается построения описанного эллипсоида минимального объема, содержащего сегмент шара.

Обозначения: $D(z_0, R)$ – шар радиусом $R > 0$ и центром $z_0 \in R^n$; $W = D(z_0, R) \cap P^+(z_1, \eta)$ – часть шара $D(z_0, R)$, определяемая секущей плоскостью $P(z_1, \eta)$, где $z_1 = z_0 + \rho\eta;\ 0 \le \rho < R, |\eta| = 1$; $\sigma = {}^{\rho}/_{R}$; W^* – эллипсоид минимального объема, содержащий множество W; z_0^* -центр эллипсоида W^*; $R_\alpha(\eta) = (\alpha - 1)\eta^T\eta + I$ – оператор растяжения пространства [3] по направлению η с коэффициентом α; $\tilde{D}^*(\bar{z}_0^*, \tilde{R})$ – шар в преобразованном пространстве $\bar{z} = R_\alpha(\eta)z$, где

$$\bar{z}_0^* = R_\alpha(\eta) z_0^*;\ z_0^* = z_0 + h\eta;\ h = \frac{1}{n+1}(1 + n\sigma)R;\ \tilde{R} = \frac{n}{\sqrt{n^2 - 1}}\sqrt{1 - \sigma^2}R;$$

$$\alpha = \sqrt{(n+1)/(n-1)}\sqrt{(1+\sigma)/(1-\sigma)}.$$

Утверждение 2 [5].

$\tilde{D}^*(z_0^*, \tilde{R}) = R_\alpha(\eta)W^*$; $V(W^*) \le qV(D(z_1, R)$; ($V(.)$ – объем тела);

$$q = V(W^*)/V(D(z,\tilde{R}) = \sqrt{\frac{n-1}{n+1}}\left(\frac{n}{\sqrt{n^2-1}}\right)^n \sqrt{\frac{1-\sigma}{1+\sigma}}(1-\sigma^2)^{n/2}.$$

Пусть задан шар начальной локализации некоторого решения x^*задачи (1) $D(z_0, R)$ ($z_0 \in R^n$ – центр шара, $R > 0$ – его радиус): $x^* \in D(z_0, R) \bigcap X^* \neq \varnothing$. $C = \max\{| g || g \in \partial f(x), x \in D(z_0, R)\}$ (C – максимум нормы обобщенного градиента в шаре $D(z_0, R)$). Рассмотрим шар $\tilde{D}(\bar{\varepsilon}) = D(x^*, \bar{\varepsilon}/C)$. Предполагаем, что шар начальной локализации $D(z_0, R)$ наряду с x^*содержит $\tilde{D}(\bar{\varepsilon})$. Пусть $x \in \tilde{D}(\bar{\varepsilon})$. Тогда $f^* \ge f(x) + (g(x), x - x^*) \ge f(x) - |(g(x)|| x - x^*| \ge$
$\ge f(x) - C(\bar{\varepsilon}/C) = f(x) - \bar{\varepsilon}$. $\ge f(x) - C(\bar{\varepsilon}/C) = f(x) - \bar{\varepsilon}$.
Поэтому все точки из $\tilde{D}(\bar{\varepsilon})$. являются решением задачи $\bar{\varepsilon}$-оптимизации (1): $\tilde{D}(\bar{\varepsilon}) \subset X^*(\bar{\varepsilon})$. Отсюда, учитывая (4),

$$\tilde{f} \ge f^* + 2\bar{\varepsilon} \Rightarrow \tilde{D}(\bar{\varepsilon}) \subset \tilde{X}(\bar{\varepsilon}, \tilde{f}). \qquad (8)$$

Пусть $g_0 \in \partial f(z_0)$, $z' = z_0 - \bar{\varepsilon}/|g_0|$. Тогда (см. формулу (3))

$$\tilde{X}(\bar{\varepsilon}, f(z_0)) \subset D^+ = D(z_0 \cap P^+(z_0', -g_0/|g_0|).$$

Введем обозначение: $K(z_0, S)$ –конус, порожденный множеством $S \in R^n$, с полюсом в точке z_0: $K(z_0, S) = \{x \mid x = z_0 + tz; z \in S, t \in R^1, t \ge 0\}$.

Рассмотрим конус $K(z_0, D^+)$. Он является круговым конусом с осью симметрии $-g_0$ и углом 2θ, $\theta = \pi - \varphi$, где $\sin(\theta) = \bar{\varepsilon}/(|g|R) \ge \bar{\varepsilon}/(CR)$. Величину $\gamma = \bar{\varepsilon}/(CR)$ естественно определить как относительную точность решения задачи $\bar{\varepsilon}$-оптимизации (1) (с учетом начального шара-локализатора $D(z_0, R)$). Пусть $u_0^K = z_0 - (R/\sin(\theta))(g_0/|g_0|, R_0^K = R/\cos(\theta)$. Тогда нетрудно видеть, что $K(z_0, D^+) = K(D(u_0^K, R_0^K))$ и $D^+ \in D(u_0^K, R_0^K)$. Таким образом шар $D^K(u_0^K, R_0^K)$ порождает конус $K(z_0, D^+)$. Так как $\tilde{X}(\bar{\varepsilon}, f(z_0)) \subset D^+ \subset K(z_0 \cap D^+)$, то этот конус можно интерпретировать как конус направлений на решение задачи $\bar{\varepsilon}$-оптимизации.

Для описания предлагаемого алгоритма решения задачи $\bar{\varepsilon}$-оптимизации, достаточно рассмотреть одну его итерацию.

Для точки z_0, направления спуска $\eta=(u_0-z_0)/|(u_0-z_0)|$ и $\bar{\varepsilon}>0$, согласно утверждения 2, вычислим такой ε–субградиент $g\in G(\varepsilon,\tilde{f},z_0)$, что $\varepsilon\le\bar{\varepsilon}$ и $(g,\eta)\ge 0$. Тогда множество $\tilde{X}(\bar{\varepsilon},\tilde{f})$ содержится в конусе $\tilde{K}_1=$ $=P^+(z_0,\xi)\cap K(u_0^k,R_0^k)$, где $\xi=-g/|g|$. Однако конус $\tilde{K}_1$ не является круговым. Круговой конус K_1, содержащий конус $\tilde{K}_1$ (следовательно и множество $\tilde{X}(\bar{\varepsilon},\tilde{f})$), определим следующим образом: $W=D(u_0^K,R_0^K)\cap P^+(z_0,\xi)$ – часть шара $D(u_0^K,R_0^K)$, определяемая секущей плоскостью $P(z_1,\xi)$. W^* – эллипсоид минимального объема, содержащий множество W. Параметры эллипсоида W^* определяются в соответствии с утверждением 2. Пусть $K_1=K(W^*)$, тогда нетрудно видеть, что $K_1\subset\tilde{K}_1$. Причем в преобразованном оператором растяжения $R_\alpha(\eta)$ (утверждение 2) пространстве конус $R_\alpha(\eta)K_1$ является круговым: образ эллипсоида W^* в преобразованном пространстве является образующим шаром этого конуса.

Таким образом, на основании процедуры одномерной минимизации по направлению η, получаем новый круговой конус направлений в преобразованном пространстве.

На этом итерация алгоритма закончена. Следующая итерация осуществляется в точности так же, но в преобразованном пространстве.

На основании [6] можно доказать следующее утверждение об эффективности приведенного алгоритма.

Утверждение 3

Для числа итераций k алгоритма, за которые он обеспечивает решение $2\bar{\varepsilon}$-оптимизации, справедлива следующая асимптотическая оценка по $\gamma<<1,\ n>>1$: $k\le 2n^2\ln(1/\gamma)$, где: γ –относительная точность решения задачи, $\gamma=\bar{\varepsilon}/(RC)$; C – оценка сверху норм субградиентов в исходном шаре $D(z,R)$.

Приведенный алгоритм можно использовать непосредственно для решения задачи $\bar{\varepsilon}$-оптимизации. Однако основная цель этого алгоритма состоит в генерации "узких " конусов–локализаторов направлений на решение задачи $\bar{\varepsilon}$-оптимизации. Два таких алгоритма рассматривались в работах [7,8]. В отличие от алгоритмов [7,8], предложенный в данной работе алгоритм использует операцию преобразования пространства.

М.Г. Журбенко, Б.М. Чумаков

АЛГОРИТМ МІНІМІЗАЦІЇ З ВИКОРИСТАННЯМ МОДИФІКАЦІЇ МЕТОДУ ЕЛІПСОЇДІВ

Для розв'язання задачі мінімізації випуклої функції в скінченновимірному евклідовому просторі пропонується ε-субградієнтний алгоритм з перетворенням простору. Алгоритм є модифікацією методу еліпсоїдів, базується на процедурі одномірного спуску і в деякому розумінні монотонний. Наводиться оцінка трудомісткості алгоритму при розв'язанні задачі ε - оптимізації.

N.G. Zhurbenko, B.M. Chumakov

MINIMIZATION ALGORITHM USING THE MODIFIED ELLIPSOID METHOD

An ε-subgradient algorithm for minimization of a convex function in a finite-dimensional Euclidean space is proposed. The algorithm is updating of the ellipsoid method, it is based on a dimensional minimization procedure and it is somewhat monotonous. Algorithm's efficiency evaluation for ε-optimizations problem is given.

1. *Журбенко Н.Г.* Об одном классе методов минимизации с преобразованием пространства // Методы решения экстремальных задач. – Киев: Ин-т кибернетики им.В.М. Глушкова НАН Украины, 1996. – С. 68 – 80.
2. *Lemarechal C., Mifflin K.* Nonsmooth Optimization. – Oxford: Pergamon Press, 1978. – 180 p.
3. *Шор Н.З.* Методы минимизации недифференцируемых функций и их применение. Киев: Наук. думка, 1979. – 200 с.
4. *Ржевский С.В.* Монотонные методы выпуклого программирования. – Киев: Наук. думка, 1993. – 319 с.
5. *Шор Н.З., Гершович В.И.* Об одном семействе алгоритмов для решения задач выпуклого программирования // Кибернетика. – 1979. – №4. – С.62 – 67.
6. *Журбенко Н.Г.* Оценка эффективности одного класса ε–субградиентных методов минимизации с преобразованием пространства // Оптимизация и ее приложения. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1997. – С. 49 – 54.
7. *Журбенко Н.Г.* Об одном ε-субградиентном алгоритме минимизации // Теория оптимальных решений. – Киев: Ин-т кибернетики им.В.М.Глушкова НАН Украины, 2002. – С. 111 – 118.
8. *Журбенко Н.Г.*, *Ненахов Э.И.* Об одном алгоритме ε-субградиентного типа минимизации выпуклой функции // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2004. – С. 27 – 33.

Получено 22.03.2005