ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2004 МАТЕМАТИКА № 1 (500)

УДК 517.929

М.Ю. АНДРАМОНОВ

МЕТОД ДОВЕРИТЕЛЬНЫХ ОКРЕСТНОСТЕЙ ДЛЯ МИНИМИЗАЦИИ КОДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ

Введение

В данной работе предлагается алгоритм нахождения локального минимума негладких невыпуклых функций. Предполагается, что целевая функция кодифференцируема [1], [2]. Такая задача имеет многочисленные приложения, но известно мало алгоритмов для ее решения.

Алгоритм принадлежит к классу методов доверительных окрестностей [3], [4], которые были впервые разработаны в [5] и [6] для минимизации суммы квадратов. Одно из преимуществ методов доверительных окрестностей — их глобальная сходимость. Известно, что использование доверительных окрестностей в ряде случаев лучше минимизации функции по направлению.

Идея таких методов проста. Строится функция, аппроксимирующая целевую, и минимизируется с ограничениями на шаг, при которых следующая итерационная точка должна принадлежать доверительной окрестности. Если получается достаточное убывание целевой функции, доверительная окрестность расширяется, и производится переход к новой итерационной точке. В противном случае итерационная точка остается прежней, и доверительная окрестность сокращается. Вспомогательная задача минимизации аппроксимации целевой функции может решаться приближенно с помощью двойственных оценок.

1. Постановка задачи

В статье решается задача безусловной минимизации

$$\min\{f(x) \mid x \in \mathbb{R}^n\}.$$

Предполагается, что целевая функция $f(x)$ ограничена снизу. Пусть функция $m(s)$ аппроксимирует $f(x+s)-f(x)$ локально в окрестности x, где s — вектор малой нормы. Будем называть эту функцию моделью.

В дифференцируемом случае берут

$$m(s) = \langle f'(x), s\rangle.$$

Для дважды дифференцируемых функций можно использовать квадратичную модель [3]

$$m(s) = \langle f'(x), s\rangle + \langle f''(x)s, s\rangle.$$

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант № 01-01-00070.

2. Метод доверительных окрестностей

Полагаем $k := 0$. Выбираются начальный радиус Δ_0 и начальная точка x_0. Выбираются параметры $0 < \mu < \eta < 1$ и $0 < \alpha_1 < 1 < \alpha_2$. Пусть найдена точка x_k и выбран радиус Δ_k, $k \geq 0$.

В точке x_k решается задача

$$\min\{m(s) \mid \|s\| \leq \Delta_k\},$$

пусть s_k — ее решение. Вычисляется $l_k = \frac{f(x_k+s_k)-f(x_k)}{m(s_k)}$.

Если $l_k \geq \mu$, то $x_{k+1} = x_k + s_k$; иначе $x_{k+1} = x_k$.

Если при этом $l_k \geq \eta$, то выбираем $\Delta_{k+1} \in [\Delta_k; \alpha_2\Delta_k]$.

Если $\eta > l_k \geq \mu$, то выбираем $\Delta_{k+1} \in [\alpha_1\Delta_k; \Delta_k]$.

Если $l_k < \mu$, то выбираем $\Delta_{k+1} \in (0; \alpha_1\Delta_k]$. Вычисления повторяются при замене k на $k+1$.

Имеется много результатов сходимости для линейной или квадратичной модели ([3], [4], [7]). Если целевая функция недифференцируема, нами предлагается использовать негладкую модель.

Предположим, что целевая функция $f(x)$ дважды кодифференцируема на $\mathbb{R}^n$, т. е.

$$f(x+s) = f(x) + \max_{(A,b,c)\in\underline{d}^2(x)}(\langle As,s\rangle + \langle b,s\rangle + c) + \min_{(P,q,r)\in\overline{d}^2(x)}(\langle Ps,s\rangle + \langle q,s\rangle + r) + o(\|s\|^2),$$

где A, P — матрицы размерности $n \times n$; $b, q \in \mathbb{R}^n$; $c, r \in \mathbb{R}$. Пара $[\underline{d}^2(x); \overline{d}^2(x)]$ называется вторым кодифференциалом функции f в точке x ([1], [2]).

Предлагается использовать в методе доверительных окрестностей модель вида

$$z_x(s) = \max_{(A,b,c)\in\underline{d}^2(x)}(\langle As,s\rangle + \langle b,s\rangle + c) + \min_{(P,q,r)\in\overline{d}^2(x)}(\langle Ps,s\rangle + \langle q,s\rangle + r).$$

В описанном методе доверительных окрестностей модель $m(s)$ заменяется на $z_x(s)$, но для доказательства сходимости нужна другая методика.

3. Сходимость алгоритма

Назовем $\Omega = \{x : \min_{\|s\|\leq\beta} z_x(s) \geq 0\}$ множеством стационарных точек второго порядка. Здесь $\beta > 0$ — малый положительный параметр. Пусть Ω_ε — ε-окрестность множества Ω.

Обозначим также $L_0 = \{x : f(x) \leq f(x_0)\}$, $\overline{\Omega_\varepsilon} = L_0 \setminus \Omega_\varepsilon$. Пусть L_0 ограничено и выполнены три предположения.

Предположение 1. Для каждого $y \in \overline{\Omega_\varepsilon} \cap L_0$ и каждого $\varepsilon > 0$ существуют такие $\overline{\Delta} > 0$ и убывающая на $(0; \overline{\Delta}]$ функция $c(\Delta)$, не зависящие от y, что

$$c(\Delta) < 0, \quad \Delta \in (0, \overline{\Delta}], \quad \min_{s\in U_\Delta} z_y(s) \leq c(\Delta), \quad \Delta \in (0; \overline{\Delta}], \quad \text{где } U_\Delta = \{s \in \mathbb{R}^n : \|s\| \leq \Delta\}.$$

Предположение 2. Для каждого $x \in \overline{\Omega_\varepsilon} \cap L_0$ и каждых $\varepsilon > 0$, $0 < \alpha < 1$ существует такое $\widetilde{\Delta}$, не зависящее от x, что

$$|f(x+s) - f(x) - z_x(s)| \leq \alpha\|s\|^2, \quad \|s\| \in (0; \widetilde{\Delta}].$$

Предположение 3. $\lim\limits_{\|s\|\to 0} \dfrac{\|s\|^2}{c(\|s\|)} = C > -\infty.$

Теорема 3.1. *При предположениях 1, 2, 3 для каждого $\varepsilon > 0$ среди точек последовательности, строящейся по методу доверительных окрестностей, найдется $x_k \in \Omega_\varepsilon$.*

Доказательство. Пусть α настолько мало, что

$$1 + \frac{\alpha\|s\|^2}{c(\|s\|)} \geq \eta, \quad \|s\| \in (0, \widetilde{\Delta}].$$

Предположим, что $\Delta_k \in (0, \min(\overline{\Delta}, \widetilde{\Delta}))$. Имеем

$$\frac{f(x_k + s_k) - f(x_k)}{z_{x_k}(s_k)} \geq 1 + \frac{\alpha \|s_k\|^2}{c(\|s_k\|)} \geq \eta,$$

поскольку $\|s_k\| \leq \widetilde{\Delta}$. Тогда $f(x_k+s_k) - f(x_k) \leq \eta c(\|s_k\|)$, и итерационная точка пересчитывается. Выполнено неравенство $f(x_k + s_k) - f(x_k) \leq \eta c(\Delta_k)$, т. к. $z_{x_k}(s_k) \leq c(\Delta_k)$ по предположению 1. Поэтому имеем убывание функции $f(x)$ на константу.

Если Δ_k остается больше, чем $\min(\overline{\Delta}, \widetilde{\Delta})$, то на некоторых итерациях радиус доверительной окрестности растет, и целевая функция убывает по крайней мере на $\eta \min(\overline{\Delta}, \widetilde{\Delta})$. Тогда после некоторого числа итераций также получим убывание функции $f(x)$ на константу. Поскольку f ограничена снизу, итерационный процесс остановится, и найдется точка из Ω_ε. $\square$

Пусть $\overline{s}$ — направление наискорейшего спуска в произвольной точке $y \notin \overline{\Omega_\varepsilon}$. Предположим, что для некоторого $\overline{\Delta} > 0$ выполнено

$$f(y + \Delta\overline{s}) < f(y), \quad \Delta \in (0; \overline{\Delta}], \quad f(y + \overline{\Delta}\overline{s}) < \nu(\varepsilon) < 0.$$

Тогда можно взять

$$c(\Delta) = \max_{(A,b,c)\in \underline{d}^2(x)} (\langle A\Delta\overline{s}, \Delta\overline{s}\rangle + \langle b, \Delta\overline{s}\rangle + c) + \min_{(P,q,r)\in \overline{d}^2(x)} (\langle P\Delta\overline{s}, \Delta\overline{s}\rangle + \langle q, \Delta\overline{s}\rangle + r).$$

Предположим, что максимум и минимум берутся по конечному числу индексов. Тогда для выполнения условий теоремы сходимости достаточно условия

$$\lim_{\Delta\to 0} \frac{\Delta^2}{\Delta^2\langle (A_j + P_i)\overline{s}, \overline{s}\rangle + \Delta(\langle (b_j + q_i), \overline{s}\rangle + (c_j + r_i))} = C > -\infty.$$

Это условие означает, что $c_j + r_i$, или $\langle b_j + q_i, \overline{s}\rangle$, или $\langle (A_j + P_i)\overline{s}, \overline{s}\rangle$ не равны нулю для всех индексов j, i и всех точек $y \notin \Omega_\varepsilon$.

Назовем

$$Z = \{x : \min_{\|s\|\leq\beta} (\max_{(a,b)\in\underline{d}(x)} (\langle a, s\rangle + b) + \min_{(q,r)\in\overline{d}(x)} (\langle q, s\rangle + r)) \geq 0\}$$

множеством стационарных точек первого порядка. Здесь пара $[\underline{d}(x); \overline{d}(x)]$ — кодифференциал функции f в x (см. [1]), $\beta > 0$ — малый параметр. Пусть Z_ε — ε-окрестность множества Z. Обозначим $\overline{Z_\varepsilon} = L_0 \setminus Z_\varepsilon$.

Вместо модели второго порядка можно использовать модель первого порядка, предполагая, что целевая функция только кодифференцируема. В этом случае модель имеет вид

$$z_x(s) = \max_{(a,b)\in\underline{d}(x)} (\langle a, s\rangle + b) + \min_{(q,r)\in\overline{d}(x)} (\langle q, s\rangle + r).$$

Введем три условия.

Условие 1. Для каждого $x \in \overline{Z}_\varepsilon \cap L_0$ и каждого $\varepsilon > 0$ существуют такие $\overline{\Delta} > 0$ и убывающая на $(0; \overline{\Delta}]$ функция $c(\Delta)$, не зависящие от x, что

$$c(\Delta) < 0, \quad \Delta \in (0, \overline{\Delta}], \quad \min_{s\in U_\Delta} z_x(s) \leq c(\Delta), \quad \Delta \in (0; \overline{\Delta}],$$

где $U_\Delta = \{s \in \mathbb{R}^n : \|s\| \leq \Delta\}$.

Условие 2. Для каждого $x \in \overline{Z}_\varepsilon \cap L_0$ и каждых $\varepsilon > 0$, $0 < \alpha < 1$ найдется такое $\widetilde{\Delta}$, не зависящее от x, что

$$|f(x + s) - f(x) - z_x(s)| \leq \alpha\|s\|, \quad \|s\| \in (0; \widetilde{\Delta}].$$

Условие 3. $\lim\limits_{\|s\|\to 0} \dfrac{\|s\|}{c(\|s\|)} = C > -\infty.$

Теорема 3.2. *При выполнении условий 1, 2, 3 для каждого $\varepsilon > 0$ найдется $x_k \in Z_\varepsilon$.*

Доказательство проводится дословным повторением доказательства теоремы 3.1 с заменой ссылки на предположение 1 ссылкой на условие 1. □

Сходимость в этом случае медленнее, но легче решать вспомогательную задачу минимизации модели.

Пусть $\overline{s}$ — направление наискорейшего спуска в произвольной точке $y \notin Z_\varepsilon$. Предположим, что для некоторого $\overline{\Delta} > 0$ выполняется

$$f(y + \Delta\overline{s}) < f(y), \quad \Delta \in (0; \overline{\Delta}], \quad f(y + \overline{\Delta}\overline{s}) < \omega(\varepsilon) < 0.$$

Тогда можно взять

$$c(\Delta) = \max_{(a,b)\in\underline{d}^2(x)} (\langle a, \Delta\overline{s}\rangle + b) + \min_{(q,r)\in\overline{d}^2(x)} (\langle q, \Delta\overline{s}\rangle + r).$$

Предположим, что максимум и минимум берутся по конечному числу индексов. Тогда для выполнения условий теоремы сходимости достаточно, чтобы

$$\lim_{\Delta\to 0} \frac{\Delta}{\Delta\langle(a_j + q_i), \overline{s}\rangle + (b_j + r_i)} = C > -\infty.$$

Это условие означает, что $b_j + r_i$ или $\langle(a_j + q_i), \overline{s}\rangle$ не равны нулю для всех индексов j, i.

Для минимизации модели можно использовать методы имитации замораживания (simulated annealing, см. [8]), т. к. число локальных минимумов в малой окрестности обычно невелико.

4. Решение задачи минимизации модели

В следующих случаях задачу минимизации модели решить легко.

1) $z_x(s) = \min_{i\in I}(\langle Ps, s\rangle + \langle q, s\rangle + r)$, где I — конечное число индексов, P_i — произвольные матрицы. В этом случае требуется решить ряд задач минимизации квадратичной функции с ограничением по норме (их можно решать методами из [3], [7]).

Такие методы базируются на следующих теоремах из [7].

Теорема 4.1. *Если x — решение задачи*

$$\min\{f + \langle g^T, w\rangle + \tfrac{1}{2}w^T Bw, \ \|w\| \le \Delta\}, \tag{1}$$

где B — матрица размерности $n \times n$, g — вектор, то x — решение уравнения $(B + \lambda I)x = -g$, где $\lambda > 0$, $\lambda(w^T w - \Delta^2) = 0$, $B + \lambda I$ — положительно определенная, а I — единичная матрицы.

Теорема 4.2. *Пусть $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$ удовлетворяют условию $(B + \lambda I)x = -g$ и $B + \lambda I$ положительно полуопределена. В случаях, когда $\lambda = 0$, $\|x\| \le \Delta$ и $\lambda \ge 0$, $\|x\| = \Delta$, x — решение задачи* (1)*. Если же $\lambda = 0$, $\|x\| = \Delta$, то x — решение задачи*

$$\min\{f + \langle g^T, \omega\rangle + \tfrac{1}{2}\omega^T B\omega, \ \|\omega\| = \Delta\}.$$

Если матрица $B + \lambda I$ положительно определена, то решение x единственно в каждом случае.

2) $z_x(s) = \max_{i\in I}(\langle A_i s, s\rangle + \langle b_i, s\rangle + c_i) + \min_{j\in J}(\langle P_j s, s\rangle + \langle q_j, s\rangle + r_j)$, где матрицы $A_i + P_j$ положительно определены. Тогда нужно применять много раз методы минимизации выпуклых недифференцируемых функций.
3) Для достаточно малого Δ функция $z_x(s)$ имеет единственную точку глобального минимума в окрестности нуля для каждого $x \in \Omega_\varepsilon$.

Если модель имеет много локальных минимумов, следует модифицировать алгоритм. Вычисляются направление наискорейшего спуска v_k, $\|v_k\| = 1$, в x_k и локальный минимум s_k функции $z_{x_k}(s)$. Делается сравнение

$$(f(x_k + s_k) - f(x_k))/z_{x_k}(s_k) \ge \mu,$$

и вычисляется точка

$$x_k + \gamma_k v_k = \arg\min_{\Delta\in(0;\Delta_k]} f(x_k + \Delta v_k).$$

Тогда полагается $x_{k+1} = \arg\min(f(x_k + s_k), f(x_k + \gamma_k v_k))$. В этом случае можно увеличивать и уменьшать Δ_k, но оно должно быть ограничено снизу.

Задача

$$\min \max_{j\in J}\{([A_j s, s] + [b_j, s] + c_j),\ \|s\| \le \Delta_k\}$$

эквивалентна последовательности задач

$$\min K_m(s) = [A_m s, s] + [b_m, s] + c_m,$$
$$K_i(s) = [A_i s, s] + [b_i, s] + c_i - [A_m s, s] - [b_m, s] - c_m \le 0, \quad i \ne m,$$
$$K_0(s) = \|s\| \le \Delta_k,$$

у которых целевая функция и ограничения квадратичны. Такие задачи могут решаться методами из [8]. Если решение ищется в подпространстве, то задачи имеют меньшую размерность и могут быть решены эффективно.

Такие задачи могут решаться методами ветвей и границ с двойственными оценками из [9], которые строятся следующим образом. Пусть $N = |J|$ и $\lambda \in \mathbb{R}^N_+$. Пусть

$$L(s,\lambda) = K_m(s) + \sum_{i\ne m} K_i(s)$$

— функция Лагранжа, определенная на $\mathbb{R}^n \times \mathbb{R}^N_+$. Пусть T — множество допустимых решений задачи. Если $\overline{s} \in T$, $\overline{\lambda} \in \mathbb{R}^N_+$, то $\overline{\lambda} K_i(\overline{s}) \le 0$ для $i \ne m$ и $L(\overline{s}, \overline{\lambda}) \le K_m(\overline{s})$. Поэтому

$$\varphi(\overline{\lambda}) = \inf_{s\in\mathbb{R}^n} L(s,\overline{\lambda}) \le \inf_{s\in T} L(s,\overline{\lambda}) \le \inf_{s\in T} K_m(s) = K^*,$$

где K^* — оптимальное решение задачи для всех $\overline{\lambda} \in \mathbb{R}^N_+$. Тогда $\psi(\lambda) = \inf_{s\in\mathbb{R}^n} L(x,\lambda)$ и $\psi^* = \sup_{\lambda\in\mathbb{R}^N} \psi(\lambda)$ — нижние оценки для K^* при $\lambda \in \mathbb{R}^N_+$. Функцию Лагранжа можно представить как

$$L(s,\lambda) = [A(\lambda)s, s] + [b(\lambda), s] + c(\lambda),$$

где

$$A(\lambda) = A_m + \sum_{i\ne m} \lambda_i (A_i - A_m), \quad b(\lambda) = b_m + \sum_{i\ne m} \lambda_i (b_i - b_m), \quad c(\lambda) = c_m + \sum_{i\ne m} \lambda_i (c_i - c_m).$$

Пусть D и $\overline{D}$ — подмножества $\mathbb{R}^N$, на которых $A(\lambda)$ положительно определена или положительно полуопределена соответственно. Функция $\psi(\lambda)$ конечна и вогнута на D ([9]). Если $\lambda \in D$, то задача вычисления $\psi(\lambda)$ — задача выпуклого квадратичного программирования. В противном случае она многоэкстремальна и сложна.

Ввиду этого вместо ψ^* рассмотрим оценку $\overline{\psi}^* = \sup\limits_{\lambda\in\overline{D}\cap\mathbb{R}^N_+} \psi(\lambda)$. Для $\lambda \in D$ ([10]) $\psi(\lambda) = -\frac{1}{4}[A^{-1}(\lambda)b(\lambda), b(\lambda)] + c(\lambda)$. Функция $\psi(\lambda)$ непрерывно дифференцируема на $D \cap \mathbb{R}^N_+$. Для ее максимизации можно применить метод эллипсоидов, штрафные функции или метод обобщенного градиентного спуска. Если точная верхняя грань на $\overline{D} \cap \mathbb{R}^N_+$ достигается на D, то $\overline{\psi}^* = K^*$ и оценка точная.

При использовании модели первого порядка решать вспомогательную задачу легко (можно использовать $\|\cdot\|_\infty$). Тогда для конечных множеств индексов получаем задачу

$$\min\,(\min_{i\in I}([a_i, s] + b_i) + \max_{j\in J}([p_j, s] + r_j), \quad \|s\|_\infty \le \Delta_k),$$

или

$$\min_{i\in I} \max_{j\in J}([a_i, s] + b_i + [p_j, s] + r_j), \quad \|s\|_\infty \le \Delta_k.$$

Достаточно решить для каждого i задачу

$$\min \max_{j\in J}([a_i,s]+b_i+[p_j,s]+r_j), \quad \|s\|_\infty \le \Delta_k,$$

которая сводится к задаче линейного программирования.

5. Метод доверительных окрестностей с масштабированием

В данном параграфе методы доверительных окрестностей обобщаются введением масштабирования. В них подзадача имеет вид

$$\min\{z_{x_k}(s),\ \|D_k s\| \le \gamma_k\},$$

где D_k — матрицы масштаба. Легко доказать сходимость, если матрицы D_k диагональны и неотрицательны.

Полагается $k := 0$. Выбираются начальный радиус γ_0 и начальная матрица масштаба D_0. Выбирается начальная точка x_0. Задаются параметры $0<\mu<\eta<1$ и $0<\alpha_1<1<\alpha_2$.

Для каждого $k=0,1,\dots$ решается задача

$$\min\{m(s),\ \|D_k s\| \le \gamma_k\}.$$

Пусть s_k — ее решение. Вычисляется $l_k = \frac{f(x_k+s_k)-f(x_k)}{m(s_k)}$.

Если $l_k \ge \mu$, то $x_{k+1}=x_k+s_k$; иначе $x_{k+1}=x_k$.

Если $l_k \ge \eta$, то выбирается $\gamma_{k+1}\in[\gamma_k;\alpha_2\gamma_k]$.

Если $\eta > l_k \ge \mu$, то выбирается $\gamma_{k+1}\in[\alpha_1\gamma_k;\gamma_k]$.

Если $l_k<\mu$, то выбирается $\gamma_{k+1}\in(0;\alpha_1\gamma_k]$. Пересчитывается матрица масштаба и вычисляется D_{k+1}.

Приведем теорему сходимости для модели второго порядка.

Теорема 5.1. *Пусть выполняются шесть следующих предположений.*

1) *Для каждого $y \in \overline{\Omega_\varepsilon}\cap L_0$ и каждого $\varepsilon>0$ найдутся такие $\overline{\Delta}>0$ и убывающая на $(0;\overline{\Delta}]$ функция $c(\Delta)$, не зависящие от y, что*

$$c(\Delta)<0, \quad \Delta\in(0,\overline{\Delta}], \quad \min_{s\in U_\Delta} z_y(s)\le c(\Delta), \quad \Delta\in(0;\overline{\Delta}],$$

где $U_\Delta=\{s\in\mathbb{R}^n : \|s\|\le\Delta\}$.

2) *Для каждого $x\in\overline{\Omega_\varepsilon}\cap L_0$ и каждых $\varepsilon>0, 0<\alpha<1$ найдется такое $\widetilde{\Delta}$, не зависящее от x, что*

$$|f(x+s)-f(x)-z_x(s)| \le \alpha\|s\|^2, \quad \|s\|\in(0;\widetilde{\Delta}].$$

3) *Для каждого k и каждого $\Delta>0$ найдется $\gamma>0$ такое, что*

$$\|D_k s\|\le\gamma \Rightarrow \|s\|\le\Delta.$$

4) *Множество $\|D_k s\|\le\gamma$ для всех k включает шар $\|s\|\le c_1\gamma$, где $c_1>0$.*

5)

$$\lim_{\Delta\to 0}\frac{\Delta^2}{c(\lambda\Delta)} = C(\lambda) > -\infty \quad \forall\lambda>0.$$

6) *γ_k и α такие, что $\|s\|\le\widehat{\Delta}(\alpha)\le\min\{\overline{\Delta},\widetilde{\Delta}\}$.*

Тогда $\gamma_k\ge\beta\widehat{\Delta}(\alpha)$ для некоторого $\beta>0$ и всех k.

Доказательство. Выполняется

$$\frac{f(x_k+s_k)-f(x_k)}{m(s_k)} - 1 \ge \frac{\alpha\|s_k\|^2}{c(c_1\gamma_k)} \ge \frac{\widehat{\Delta}^2(\alpha)}{c(c_1\gamma_k)} \ge \frac{\widehat{\Delta}^2(\alpha)}{c(c_1\beta\widehat{\Delta}(\alpha))} \ge \eta-1$$

для достаточно малого α по нашим предположениям. □

Можно решать следующую вспомогательную задачу, требуя, чтобы направление s принадлежало подпространству S малой размерности:

$$\min\{z_{x_k}(s),\ \|s\| \le \Delta_k,\ s \in S\}.$$

Например, можно взять $S = \{x = \alpha s_N + \beta v\}$, где v — направление наискорейшего спуска, s_N — направление, найденное методом Ньютона для дважды кодифференцируемых функций ([1]).

Пусть выполняются следующие предположения.

a) Для каждого $y \in \overline{\Omega_\varepsilon} \cap L_0$ и каждого $\varepsilon > 0$ найдутся такие $\overline{\Delta} > 0$ и убывающая на $(0;\overline{\Delta}]$ функция $c(\Delta)$, не зависящие от y, что

$$c(\Delta) < 0,\quad \Delta \in (0,\overline{\Delta}],\quad \min_{s\in U_\Delta, s\in S} z_y(s) \le c(\Delta),\quad \Delta \in (0;\overline{\Delta}],$$

где $U_\Delta = \{s \in \mathbb{R}^n : \|s\| \le \Delta\}$.

b) Для каждого $x \in \overline{\Omega_\varepsilon} \cap L_0$ и каждых $\varepsilon > 0$, $0 < \alpha < 1$ найдется такое $\widetilde{\Delta}$, не зависящее от x, что

$$|f(x+s) - f(x) - z_x(s)| \le \alpha\|s\|^2,\quad \|s\| \in (0;\widetilde{\Delta}].$$

c) $\lim\limits_{\|s\|\to 0} \dfrac{\|s\|^2}{c(\|s\|)} = C > -\infty$.

Теорема 5.2. *При предположениях* a), b), c) *для каждого* $\varepsilon > 0$ *найдется* $x_k \in \Omega_\varepsilon$.

Доказательство аналогично доказательству теоремы 5.1.

Автор выражает благодарность В.Ф. Демьянову и Я.И. Заботину за ряд ценных советов и замечаний.

Литература

1. Демьянов В.Ф., Рубинов А.М. *Основы негладкого анализа и квазидифференциальное исчисление.* – М.: Наука, 1990. – 431 с.
2. Demyanov V.F., Rubinov A.M. *Constructive non-smooth analysis.* – Frankfurt: Peter Lang, 1995. – 590 p.
3. Moré J.J. *Recent developments in algorithms and software for trust region methods* // Math. Progr.: State of the Art. – 1982. – P. 259–287.
4. Powell M.J.D. *Convergence properties of a class of minimization algorithms* // Non. Progr. 2. – Academ. Press, 1975. – P. 1–27.
5. Levenberg K. *A method for the solution of certain nonlinear problems in least squares* // Quart. of Appl. Math. – 1944. – V. 2. – P. 164–168.
6. Marquardt D.W. *An algorithm for least squares estimation of nonlinear parameters* // SIAM J. on Appl. Math. – 1963. – V. 11. – P. 431–441.
7. Sorensen D.C. *Newton's method with a model trust region modification* // SIAM J. on Num. Anal. – 1982. – V. 19. – P. 409–426.
8. Horst R., Pardalos P. (Eds.) *Handbook on Global Optimization.* – Dordrecht: Kluwer Academ. Publ., 1996. – 900 p.
9. Shor N.Z., Stetsenko S.I. *Quadratic Extremal Problems and Nondifferentiable Optimization.* – Kiev, Nauk. Dumka, 1989. – 220 c.
10. Shor N.Z. *Dual estimates in multiextremal problems* // J. of Global Optim. – 1995. – V. 7. – P. 75–91.

Казанский государственный университет

Поступила 03.09.2002