МОМО-16. Домашняя работа 1

Срок сдачи: 12 сентября 2016, 10:30

1 Пусть $x^* \in \mathbb{R}^n$ — точка локального минимума функции $f : \mathbb{R}^n \to \mathbb{R}$, т. е. найдется $r > 0$, такое что для всех $x \in \mathbb{R}^n$, $\|x - x^*\| < r$, выполняется $f(x) \geq f(x^*)$. Докажите, что если функция f дифференцируема в точке x^*, то необходимо выполняется *условие оптимальности первого порядка*: $\nabla f(x^*) = 0$.

2 Рассмотрим непрерывно дифференцируемую функцию $f : \mathbb{R}^n \to \mathbb{R}$, градиент которой удовлетворяет условию Липшица с константой $L > 0$, т. е. для всех $x, y \in \mathbb{R}^n$ выполняется

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|.$$

Докажите, что в этом случае справедлива следующая оценка на погрешность линейной аппроксимации: для всех $x, y \in \mathbb{R}^n$ верно

$$|f(y) - f(x) - \nabla f(x)^\top (y - x)| \leq \frac{L}{2}\|y - x\|^2.$$

Указание: Воспользуйтесь *формулой Ньютона-Лейбница*: для всех $x, y \in \mathbb{R}^n$ справедливо

$$f(y) - f(x) = \int_0^1 \nabla f(x + t(y - x))^\top (y - x) dt.$$

3 Определите скорость сходимости следующих последовательностей ($k \geq 1$):

(a) $r_k = (0.5)^{k^2}$

(b) $r_k = 1/\sqrt{k}$

(c) $r_k = 1/k^k$

(d) $r_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k}, & k \text{ четное}, \\ \frac{r_{k-1}}{k}, & k \text{ нечетное}. \end{cases}$

(e) $(r_k) = (1, \frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{4}, \frac{1}{16}, \frac{1}{8}, \dots)$

Примечание: Для суперлинейно-сходящихся последовательностей необходимо дополнительно выяснить, имеет ли место квадратичная сходимость.

4 Рассмотрим последовательность $(r_k)_{k \geq 0}$ из положительных чисел ($r_k > 0$). Известно, что $r_{k+1} \leq C r_k^2$ для всех $k \geq 0$ и некоторой константы $0 < C < \infty$. При каких условиях на C и r_0 можно утверждать, что $r_k \to 0$? Какова при этом скорость сходимости?

5 Вычислить производные $Df(x)[\Delta x]$ (а также градиенты $\nabla f(x)$ для случаев с $f(x) \in \mathbb{R}$):

(a) $f(x) = \|x\|_2^3 \equiv (x^\top x)^{3/2}, \quad x \in \mathbb{R}^n$

(b) $f(X) = \mathrm{Tr}(AX^{-1}B), \quad X \in \mathbb{R}^{n \times n}$

(c) $f(X) = \mathrm{Det}(X)\,\mathrm{Tr}(AX^{-1}B), \quad X \in \mathbb{R}^{n \times n}$

(d) $f(x) = xx^\top, \quad x \in \mathbb{R}^n$

(e) $f(x) = \mathrm{Det}(2I + xx^\top), \quad x \in \mathbb{R}^n$

Указание: Используйте следующие формулы:

$$D(X^{-1})[\Delta X] = -X^{-1}(\Delta X)X^{-1}$$

$$D(\mathrm{Det}(X))[\Delta X] = \mathrm{Det}(X)\,\mathrm{Tr}(X^{-1}(\Delta X))$$