

Градиент.

Любую функцию от n переменных можно считать функцией из \mathbb{R}^n в \mathbb{R} . Пусть у нас есть величина $q(\mathbf{u})$, зависящая от вектора $\mathbf{u} \in \mathbb{R}^n$. Будем говорить, что $q(\mathbf{u}) = o(\mathbf{u})$ (читается "q(u) есть о маленькое от u"), если для любой последовательности векторов $\mathbf{u}_1, \mathbf{u}_2, \dots \in \mathbb{R}^n$ такой что $\|\mathbf{u}_i\| \rightarrow 0$ последовательность $q(\mathbf{u}_i)$ стремится к нулю.

Функция f от n переменных **дифференцируема** в точке \mathbf{x} , если существует такой вектор $\nabla f(\mathbf{x}) \in \mathbb{R}^n$, что для любого $\mathbf{v} \in \mathbb{R}^n$

$$f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + \langle \mathbf{v}, \nabla f(\mathbf{x}) \rangle + o(\mathbf{v})$$

Вектор $\nabla f(\mathbf{x})$ называется **градиентом** функции f в точке \mathbf{x} .

Задача 1. эквивалентность обычным определениям для $n=1$

Задача 2. Пусть f и g – дифференцируемые функции из \mathbb{R} в \mathbb{R} и из \mathbb{R}^n в \mathbb{R} соответственно. Докажите, что $f(g(\mathbf{x}))$ – дифференцируемая функция из \mathbb{R}^n в \mathbb{R} .

Задача 3. Пусть f и g – дифференцируемые функции из \mathbb{R}^n в \mathbb{R} . Докажите, что следующие функции дифференцируемы (и выразите их градиенты, через градиенты f и g):

а) $f + g$

б) $f \cdot g$

в) f/g (в точках \mathbf{x} , где $g(\mathbf{x}) \neq 0$)

Задача 4. Докажите, что $f(g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))$ – всюду дифференцируемая функция из \mathbb{R}^n в \mathbb{R} , если $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ и $f : \mathbb{R}^m \rightarrow \mathbb{R}$ всюду дифференцируемы.

Задача 5. Докажите, что следующие функции $\mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируемы:

а) $x_1 \cdot \dots \cdot x_n$

б) $\sin(x_1 + \dots + x_n)$

в) $\log\left(\frac{1}{1+\exp(-(\mathbf{x}, \mathbf{w}))}\right)$, где \mathbf{w} – какой-то вектор из \mathbb{R}^n

г) $\text{softmax}(\mathbf{x}) = \frac{(e^{x_1}, \dots, e^{x_n})}{e^{x_1} + \dots + e^{x_n}}$

Частная производная функции $f : \mathbb{R}^n \rightarrow \mathbb{R}$ по i -той переменной в точке \mathbf{x} это:

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \lim_{\varepsilon \rightarrow 0} \frac{f(x_1, \dots, x_i + \varepsilon, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{\varepsilon}$$

Иными словами мы фиксируем все переменные кроме i -той, рассматриваем f как функцию от одной переменной и берем ее производную в точке x_i .

Задача 6. Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируема в точке \mathbf{x} . Докажите что:

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right).$$

Задача 7* (ненужная). Докажите, что f дифференцируема в точке \mathbf{x} , если ее частные производные определены в некоторой окрестности \mathbf{x} и непрерывны в \mathbf{x} .

Задача 8. Градиент направление самого быстрого....

??? дифференцируемость $\mathbb{R}^n \rightarrow \mathbb{R}^m$