Градиент.

$\mathbb{R}^n = \{(x_1, \ldots, x_n) | \forall i : x_i \in \mathbb{R}\}$ – этом множество наборов из n вещественных чисел. Элементы $\mathbb{R}^n$ мы будем называть векторами и обозначать жирными буквами, а значения отдельных чисел в векторе – той же буквой, только нежирной и с индексом. Например x_i – это i-тое число в векторе $\mathbf{x}$. Вектора можно складывать и умножать на числа. Любую функцию от n переменных можно считать функцией из $\mathbb{R}^n$ в $\mathbb{R}$.

Стандартное скалярное произведение *двух векторов* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *это:*

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i \cdot y_i$$

Для функции $\mathbf{u} : \mathbb{R} \to \mathbb{R}^n$, $\mathbf{u}(t) = (u_1(t), \ldots, u_n(t))$. Будем писать, что $\mathbf{u}(t) = o(t)$, если

$$\forall i \in \{1, \ldots, n\} : u_i(t) = o(t).$$

??? разные нормы и их эквивалентность

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

переделать! *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 $v \in \mathbb{R}^n$*

$$f(\mathbf{x} + \mathbf{v} \cdot t + o(t)) = f(\mathbf{x}) + (v, \nabla \mathbf{x}) \cdot t + o(t)$$

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

Задача 1. 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 2. 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 3. Докажите, что $f(g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$ – дифференцируемая функция из $\mathbb{R}^n$ в $\mathbb{R}$, если $g_i : \mathbb{R}^n \to \mathbb{R}$ и $f : \mathbb{R}^m \to \mathbb{R}$ – дифференцируемы.

Задача 4. Докажите, что следующие функции $\mathbb{R}^n \to \mathbb{R}$ дифференцируемы:

а) $x_1 \cdot \ldots \cdot x_n$

б) $\sin(x_1 + \ldots + x_n)$

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

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

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

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

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 5. Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке $\mathbf{x}$. Докажите что:

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

Задача 6* (ненужная). Докажите, что f дифференцируема в точке $\mathbf{x}$, если ее частные производные определены в некоторой окрестности $\mathbf{x}$ и непрерывны в $\mathbf{x}$.