Лекция 8:

(8.1) Линейные отображения

Мы знаем, что $\mathbb{R}^m$ образует векторное пространство. В нём можно ввести базис
$$e_1 := (1,0,\dots,0),\ e_2 := (0,1,0,\dots,0),\ \dots,\ e_m := (0,\dots,0,1).$$
Тогда любой вектор $x \in \mathbb{R}^m$ раскладывается по этому базису:
$$x = x^1 e_1 + \dots + x^m e_m \quad \text{или} \quad x = x^i e_i.$$

Рассмотрим отображение $L: \mathbb{R}^m \to \mathbb{R}^n$. Оно называется **линейным**, если для всех $x_1, x_2 \in \mathbb{R}^m$ и всех $\lambda, \mu \in \mathbb{R}$ выполняется
$$L(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 L(x_1) + \lambda_2 L(x_2).$$

Если в $\mathbb{R}^n$ зафиксирован базис $\tilde{e}_1, \dots, \tilde{e}_n$, то образы базисных векторов имеют вид
$$L(e_i) = a_i^1 \tilde{e}_1 + \dots + a_i^n \tilde{e}_n = a_i^j \tilde{e}_j \quad \text{для всех } i = 1, \dots, m$$
Значит для любого вектора $h \in \mathbb{R}^m$ в силу линейности L
$$L(h) = L(h^i e_i) = h^i L(e_i) = h^i a_i^j \tilde{e}_j = a_i^j h^i \tilde{e}_j$$
или в координатной форме
$$L(h) = (a_i^1 h^i, \dots, a_i^n h^i).$$
Таким образом, линейное отображение $L: \mathbb{R}^m \to \mathbb{R}^n$ можно записать как набор $L = (L^1, \dots, L^n)$ линейных функций $L^j: \mathbb{R}^m \to \mathbb{R}$, $j = 1, \dots, n$.
Учитывая правило умножения матриц,
$$L(h) = \begin{pmatrix} L^1(h) \\ \vdots \\ L^n(h) \end{pmatrix} = \begin{pmatrix} a_1^1 & \dots & a_m^1 \\ \dots & \ddots & \dots \\ a_1^n & \dots & a_m^n \end{pmatrix} \begin{pmatrix} h^1 \\ \vdots \\ h^m \end{pmatrix}.$$
Мы показали, что при фиксированных в $\mathbb{R}^m$ и $\mathbb{R}^n$ базисах, всякое линейное отображение $L: \mathbb{R}^m \to \mathbb{R}^n$ можно отождествить с $m \times n$-матрицей.

(8.2) Норма и скалярное произведение в $\mathbb{R}^m$

Нормой вектора $x = (x^1, \dots, x^m) \in \mathbb{R}^m$ называется величина $\|x\| := \sqrt{(x^1)^2 + \dots + (x^m)^2}$.
Она обладает следующими свойствами:

1°. $\|x\| \geq 0$ и $\|x\| = 0 \Leftrightarrow x = 0$.
2°. $\|\lambda x\| = |\lambda| \cdot \|x\|$ для $\lambda \in \mathbb{R}$.
3°. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$.

любая функция $\|\cdot\|: X \to \mathbb{R}$ на векторном пространстве X удовлетворяющая 1°–3°, наз-ся **нормой** на X

Заметим что ранее введённое нами расстояние $d(\cdot,\cdot)$ в $\mathbb{R}^m$ связано с нормой:
$$d(x_1, x_2) = \|x_2 - x_1\|.$$

Пусть $f: X \to \mathbb{R}^m$, $g: X \to \mathbb{R}^n$ – отображения. Будем говорить, что

$$f(x) = o(g(x)) \text{ при } x \to x_0,$$

если $\|f(x)\|_{\mathbb{R}^m} = o(\|g(x)\|_{\mathbb{R}^n})$ при $x \to x_0$. Заметим, что в силу неравенства

$$|f^i(x)| \leq \|f(x)\| = \|f^1(x)e_1 + \ldots + f^m(x)e_m\| \leq \|f^1(x)e_1\| + \ldots + \|f^m(x)e_1\| \leq \sum_{i=1}^{m} |f^i(x)|$$

можно заключить, что

$$f(x) = o(g(x)) \text{ при } x \to x_0 \iff f^i(x) = o(g(x)) \text{ при } x \to x_0, \quad i = 1, \ldots, m.$$

Аналогично, будем говорить, что

$$f(x) = O(g(x)) \text{ при } x \to x_0,$$

если $\|f(x)\|_{\mathbb{R}^m} = O(\|g(x)\|_{\mathbb{R}^n})$ при $x \to x_0$. При этом

$$f(x) = O(g(x)) \text{ при } x \to x_0 \iff f^i(x) = O(g(x)) \text{ при } x \to x_0, \quad i = 1, \ldots, m.$$

Пример 8.1: Пусть $h \doteq (h^1, \ldots, h^m) \in \mathbb{R}^m$ и $L: \mathbb{R}^m \to \mathbb{R}^n$ – линейное отображение. Тогда

$$\|L(h)\|_{\mathbb{R}^n} = \left\| \sum_{i=1}^{m} h^i L(e_i) \right\| \leq \sum_{i=1}^{m} \|L(e_i)\| |h^i| \leq \left(\sum_{i=1}^{m} \|L(e_i)\| \right) \|h\|_{\mathbb{R}^m},$$

а значит

$$\|L(h)\|_{\mathbb{R}^n} = O(\|h\|_{\mathbb{R}^m}) \text{ при } h \to 0, \text{ т.е. } L(h) = O(h) \text{ при } h \to 0.$$

Из этого факта следует, что линейное отображение непрерывно в любой $x_0 \in \mathbb{R}^m$:

$$L(x - x_0) = L(x) - L(x_0) \to 0 \text{ при } x \to x_0.$$

Более того, отображение L равномерно непрерывно на $\mathbb{R}^m$.

Вспомним, что **скалярным произведением** в векторном пространстве X над $\mathbb{R}$ наз-ся функция $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$, удовлетворяющая условиям

1°. $\langle x, x \rangle \geq 0$ и $\langle x, x \rangle = 0 \iff x = 0$.

2°. $\langle x_1, x_2 \rangle = \langle x_2, x_1 \rangle$.

3°. $\langle \lambda x_1, x_2 \rangle = \lambda \langle x_1, x_2 \rangle$ для $\lambda \in \mathbb{R}$.

4°. $\langle x_1 + x_2, x_3 \rangle = \langle x_1, x_3 \rangle + \langle x_2, x_3 \rangle$.

Пространство X со скалярным произведением наз-ся **евклидовым**.

Пусть в X зафиксирован базис $e_1, \ldots, e_m$. Тогда для векторов $x = (x^1, \ldots, x^m)$, $y = (y^1, \ldots, y^m) \in X$ скалярное произведение запишется

$$\langle x, y \rangle = g_{ij} x^i x^j,$$

где числа $g_{ij} = \langle e_i, e_j \rangle$ образуют матрицу Грамма.

В **ортонормированном базисе** $g_{ij} = \delta_{ij}$, где $\delta_{ij} = 1$, если $i = j$, и $\delta_{ij} = 0$ в ином случае; т.е.

$$\langle x, y \rangle = x^1 y^1 + \ldots + x^m \cdot y^m.$$

Алгебраический факт: Пусть $L: \mathbb{R}^m \to \mathbb{R}$ – линейная функция в евклидовом пространстве. Тогда существует единственный вектор $\xi \in \mathbb{R}^m$, т.ч. $L(x) = \langle \xi, x \rangle$.

(8.3) Дифференцируемость и дифференциал функции

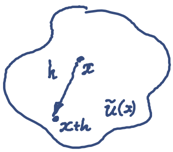
Определение 8.1: Пусть $f: \tilde{U}(x) \to \mathbb{R}^n$ — отображение, где $\tilde{U}(x) \subset \mathbb{R}^m$ — некоторая окрестность точки $x \in \mathbb{R}^m$. Тогда отображение f **дифференцируемо** в т. x, если

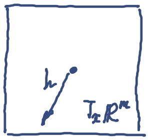
(8.1) $$f(x+h) - f(x) = L(x)h + \alpha(x;h),$$

где $L(x): \mathbb{R}^m \to \mathbb{R}^n$ — линейное относительно h отображение ($L(x)(h) := L(x)h$), $\alpha(x;h) = o(h)$ при $h \to 0$, $x+h \in \tilde{U}(x)$.

Линейная функция $L(x): \mathbb{R}^m \to \mathbb{R}^n$ называется **дифференциалом**, **касательным отображением** или **производным отображением** отображения f в т. x. (Обозначения: $df(x)$, $Df(x)$ или $f'(x)$)

Заметим, что вектор h является вектором-смещением из т. x. Совокупность всех таких векторов будем обозначать $T_x\mathbb{R}^m$ ($\simeq \mathbb{R}^m$) и называть **касательным пространством** к $\mathbb{R}^m$ в т. $x \in \mathbb{R}^m$. В новых обозначениях дифференциал $df(x): T_x\mathbb{R}^m \to T_{f(x)}\mathbb{R}^n$.





Векторное равенство (8.1) равносильно n скалярным равенствам

(8.2) $$f^i(x+h) - f^i(x) = L^i(x)h + \alpha^i(x;h), \quad i = 1,\dots,n,$$

где $L^i(x): \mathbb{R}^m \to \mathbb{R}$ — линейные функции, $\alpha^i(x,h) = o(h)$ при $h \to 0$, а $x+h \in \tilde{U}(x)$.

Предложение 8.1: Отображение $f: \tilde{U}(x) \to \mathbb{R}^n$ дифференцируемо в т. x тогда и только тогда, когда дифференцируемы в x его компоненты $f^i: \tilde{U}(x) \to \mathbb{R}$, $i = 1,\dots,n$.

Если $x = (x^1,\dots,x^m)$ и $h = (h^1,\dots,h^m)$, то в (8.2) линейные функции, составляющие отображение $L(x)$ из (8.1), равны

$$L^i(x)h = a^i_1(x)h^1 + \dots + a^i_m(x)h^m, \quad i = 1,\dots,m,$$

а, следовательно, приращения

(8.3) $$f^i(x^1+h^1,\dots,x^m+h^m) - f^i(x^1,\dots,x^m) = a^i_1(x)h^1 + \dots + a^i_m(x)h^m + o(h), \quad h \to 0,$$

где $a^i_j(x) \in \mathbb{R}$.

Если функция f^i дифференцируема в т. x, то (8.3) справедливо для всех $h \in T_x\mathbb{R}^m$. В частности, выбрав $h = h^k e_k$, мы получим

$$f^i(x^1,\dots,x^{k-1}, x^k+h^k, x^{k+1},\dots,x^m) - f^i(x^1,\dots,x^m) = a^i_k(x)h^k + o(h^k) \text{ при } h^k \to 0$$

Из последнего равенства находим, что
$$a^i_k(x) = \frac{f^i(x^1,\dots,x^{k-1}, x^k+h^k, x^{k+1},\dots,x^m) - f^i(x^1,\dots,x^m)}{h^k} + \frac{o(h^k)}{h^k}.$$
Переходя к пределу по $h^k \to 0$, получаем
$$a^i_k(x) = \lim_{h^k\to 0} \frac{f^i(x^1,\dots,x^{k-1}, x^k+h^k, x^{k+1},\dots,x^m) - f^i(x^1,\dots,x^m)}{h^k}.$$
Этот предел называется **частной производной** функции f^i в т. $x=(x^1,\dots,x^m)$ по переменной x^k. Распространённые обозначения для частных производных:
$$\frac{\partial f}{\partial x^k}(x),\quad \partial_k f(x),\quad D_k f(x),\quad f'_{x^k}(x).$$

Предложение 8.2: Пусть функция $f\colon E\to\mathbb{R}^m$, где $E\subset\mathbb{R}^m$, дифференцируема во внутренней точке $x\in E$. Тогда функция f имеет в т. x частные производные по каждой переменной, а дифференциал однозначно определяется ими:
$$df(x)h = \frac{\partial f}{\partial x^1}(x)h^1 + \dots + \frac{\partial f}{\partial x^m}(x)h^m.$$

Предложение 8.3: Пусть отображение $f\colon E\to\mathbb{R}^n$, где $E\subset\mathbb{R}^m$, дифференцируемо во внутренней точке $x\in E$. Тогда отображение $f(x)=(f^1(x),\dots,f^n(x))$ имеет в т. x единственный дифференциал $df(x)$, причём имеющий координатное представление
$$df(x)(h) = \begin{pmatrix} df^1(x)h \\ \vdots \\ df^n(x)h \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x) & \dots & \frac{\partial f^1}{\partial x^m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial x^1}(x) & \dots & \frac{\partial f^n}{\partial x^m}(x) \end{pmatrix}}_{\text{матрица Якоби или якобиан отображения } f\colon E\to\mathbb{R}^n \text{ в точке } x\in E} \begin{pmatrix} h^1 \\ \vdots \\ h^m \end{pmatrix}.$$

Заметим, что в (8.1) $L(x)h \to 0$ при $h\to 0$, поэтому в случае дифференцируемости в т. x приращение $f(x+h)-f(x) \to 0$ при $h\to 0$, т.е. дифференцируемость в т. x влечёт непрерывность в точке x.

Пример 8.2: Рассмотрим функцию
$$f(x^1,x^2) = \begin{cases} 0, & \text{если } x^1\cdot x^2 = 0, \\ 1, & \text{если } x^1\cdot x^2 \neq 0. \end{cases}$$
Очевидно, что $\frac{\partial f}{\partial x^1}(0,0)=0$, $\frac{\partial f}{\partial x^2}(0,0)=0$. Однако, f не является дифференцируемой в точке $(0,0)$, т.к. она разрывна в $(0,0)$. ◀