Лекция 31

Дифференцируемость отображений

Теперь мы будем рассматривать только отображения между пространствами $\mathbb{R}^d$ и $\mathbb{R}^m$ с введённой на каждом из них евклидовой метрикой. Прежде всего отметим, что любое отображение $f : \mathbb{R}^k \to \mathbb{R}^m$ в любой точке $\boldsymbol{x} \in \mathbb{R}^k$, в которой оно определено, можно записать в виде $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_m(\boldsymbol{x}))$, где отображения $f_j : \mathbb{R}^k \to \mathbb{R}$ $(j = 1, 2, ..., m)$ называются *координатными функциями*. Мы увидим, что вопрос о дифференцируемости отображений сводится к вопросу о дифференцируемости координатных функций.

Напомним некоторые понятия из линейной алгебры.

Определение 1. *Отображение* $L : \mathbb{R}^d \to \mathbb{R}^m$ *называется линейным, если* $\forall \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^d$ *и* $\forall \alpha,\ \beta \in \mathbb{R}$ *выполнено равенство*

$$L(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha L(\boldsymbol{x}) + \beta L(\boldsymbol{y}),$$

где $L\boldsymbol{x},\ L\boldsymbol{y} \in \mathbb{R}^m$.

Упражнение 1. Как устроены координатные функции линейного отображения?

Если в пространстве $\mathbb{R}^d$ задан базис $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ ...,\ \boldsymbol{e}_d$, а в пространстве $\mathbb{R}^m$ – базис $\boldsymbol{e}'_1,\ \boldsymbol{e}'_2,\ ...,\ \boldsymbol{e}'_m$, то для любого вектора вектора $\boldsymbol{h} \in \mathbb{R}^d$ в силу определения линейного отображения имеем:

$$L(\boldsymbol{h}) = L\left(\sum_{j=1}^{d} h_j \boldsymbol{e}_j\right) = \sum_{j=1}^{d} h_j L(\boldsymbol{e}_j) = \sum_{j=1}^{d} h_j \sum_{i=1}^{m} a_{ij}\boldsymbol{e}'_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{d} h_j a_{ij}\right)\boldsymbol{e}'_i,$$

где h_j $(j = 1,\ ...,\ d)$ – координаты вектора $\boldsymbol{h}$ в базисе $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ ...,\ \boldsymbol{e}_d$, а a_{ij} $(i = 1,\ ...,\ m)$ координаты вектора $L\boldsymbol{e}_j$ в базисе $\boldsymbol{e}'_1,\ \boldsymbol{e}'_2,\ ...,\ \boldsymbol{e}'_m$. Таким образом, столбец координат образа $L(\boldsymbol{h})$ вектора $\boldsymbol{h}$ при отображении $L : \mathbb{R}^d \to \mathbb{R}^m$ в базисе $\boldsymbol{e}'_1,\ \boldsymbol{e}'_2,\ ...,\ \boldsymbol{e}'_m$ равен

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1d} \\ a_{21} & a_{22} & ... & a_{2d} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{md} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ ... \\ h_d \end{pmatrix}.$$

Значит, любое линейное отображение при фиксированных базисах в пространствах $\mathbb{R}^d$ и $\mathbb{R}^m$ задаётся матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1d} \\ a_{21} & a_{22} & ... & a_{2d} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{md} \end{pmatrix}.$$

По столбцам этой матрицы стоят координаты образов $L(\boldsymbol{e}_j)$ $(j = 1,\ ...,\ d)$ базисных векторов $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ ...,\ \boldsymbol{e}_d$ в базисе $\boldsymbol{e}'_1,\ \boldsymbol{e}'_2,\ ...,\ \boldsymbol{e}'_m$. Эта матрица называется *матрицей отображения* L. Если фиксировать другие базисы, то матрица оператора изменится.

Если заданы базисы, то действие линейного отображения на вектор сводится к умножению матрицы этого линейного отображения на данный вектор, поэтому в дальнейшем для краткости мы будем писать просто $L\boldsymbol{x}$ вместо $L(\boldsymbol{x})$, то есть скобки писать не будем.

Теперь дадим определение дифференцируемого отображения. Напомним, что мы рассматриваем пространства с евклидовыми метриками на них.

Определение 2. *Отображение $f : \mathbb{R}^d \to \mathbb{R}^m$ называется дифференцируемым в точке $\boldsymbol{x} \in \mathbb{R}^d$, если для любого ненулевого вектора $h \in \mathbb{R}^d$ из некоторого шара в $\mathbb{R}^d$ с центром в начале координат выполнено равенство*

$$f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) = L_{\boldsymbol{x}}\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|,$$

где $L_{\boldsymbol{x}} : \mathbb{R}^d \to \mathbb{R}^m$ – линейное отображение, зависящее от $\boldsymbol{x}$, $\alpha(\boldsymbol{h}) \in \mathbb{R}^m$, причём

$$\|\alpha(\boldsymbol{h})\| \to \boldsymbol{0}$$

(здесь $\boldsymbol{0}$ – нейтральный элемент пространства $\mathbb{R}^m$) при $\|\boldsymbol{h}\| \to 0$.

*Линейное отображение $L_{\boldsymbol{x}} : \mathbb{R}^d \to \mathbb{R}^m$ называется **дифференциалом** отображения f в точке $\boldsymbol{x}$ и обозначается $df(\boldsymbol{x})$.*

Если $d = 2,\ m = 1$, то отображение f является функцией двух переменных, которая паре чисел (x, y) ставит в соответствие число z. Визуально можно представить себе график такой функции как поверхность в $\mathbb{R}^3$, а тогда дифференцируемость функции f в точке $\boldsymbol{x}$ с координатами (x, y) равносильна тому, что в этой точке можно провести касательную плоскость к поверхности, являющейся графиком функции.

Пример 1. *Рассмотрим отображение $f : \mathbb{R}^2 \to \mathbb{R}^2$, принимающее на векторе $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ значение $f(\boldsymbol{x}) = \begin{pmatrix} x_1^2 \\ x_2^3 \end{pmatrix}$. Тогда при $\boldsymbol{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ имеем*

$$f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) = \begin{pmatrix} (x_1+h_1)^2 \\ (x_2+h_2)^3 \end{pmatrix} - \begin{pmatrix} x_1^2 \\ x_2^3 \end{pmatrix} = \begin{pmatrix} 2x_1h_1 + h_1^2 \\ 3x_2^2h_2 + 3x_2h_2^2 + h_2^3 \end{pmatrix} = \begin{pmatrix} 2x_1 & 0 \\ 0 & 3x_2^2 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} +$$

$$+ \begin{pmatrix} \frac{h_1^2}{\sqrt{h_1^2+h_2^2}} \\ \frac{3x_2h_2^2+h_2^3}{\sqrt{h_1^2+h_2^2}} \end{pmatrix}\sqrt{h_1^2+h_2^2} = f(\boldsymbol{x}) + L_{\boldsymbol{x}}\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|,$$

где $L_{\boldsymbol{x}}$ задаётся матрицей $\begin{pmatrix} 2x_1 & 0 \\ 0 & 3x_2^2 \end{pmatrix}$, а $\alpha(\boldsymbol{h}) = \begin{pmatrix} \frac{h_1^2}{\sqrt{h_1^2+h_2^2}} \\ \frac{3x_2h_2^2+h_2^3}{\sqrt{h_1^2+h_2^2}} \end{pmatrix}$. Проверим, что при $\|\boldsymbol{h}\| \to 0$ вектор $\alpha(\boldsymbol{h})$ стремится к нейтральному элементу пространства $\mathbb{R}^2$, то есть к вектору $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Действительно, из предыдущей лекции следует, что условие $\|\boldsymbol{h}\| \to 0$ равносильно тому, что $h_1 \to 0$ и $h_2 \to 0$, а тогда

$$\lim_{h_1\to 0} \frac{h_1^2}{\sqrt{h_1^2+h_2^2}} = \lim_{h_1\to 0} \frac{h_1}{\frac{\sqrt{h_1^2+h_2^2}}{h_1}} = \lim_{h_1\to 0} \frac{h_1}{\sqrt{1+h_2^2/h_1^2}} = 0$$

и, аналогично, $\lim\limits_{h_2\to 0} \frac{3x_2h_2^2+h_2^3}{\sqrt{h_1^2+h_2^2}} = 0$.

Таким образом, отображение f удовлетворяет определению дифференцируемого отображения, поэтому f дифференцируемо.

Упражнение 2. Найти матрицу дифференциала и $\alpha(\boldsymbol{h})$ для отображения $f : \mathbb{R}^2 \to \mathbb{R}^2$, принимающего на векторе $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ значение $f(\boldsymbol{x}) = \begin{pmatrix} x_1 + x_2 \\ x_1 \cdot x_2 \end{pmatrix}$.

Отметим, что *в фиксированной точке $\boldsymbol{x}$ отображение $df(\boldsymbol{x})$ задаётся фиксированной матрицей, зависящей только от координат точки $\boldsymbol{x}$.*

Важен вопрос о том, единственно ли отображение $df(\boldsymbol{x})$. Приведённый пример не проясняет этот вопрос, так как может создаться впечатление, что мы могли разбить вектор

$$\begin{pmatrix} 2x_1h_1 + h_1^2 \\ 3x_2^2h_2 + 3x_2h_2^2 + h_2^3 \end{pmatrix}$$

на слагаемые и по-другому. Однако такое впечатление ошибочно и верен следующий факт.

Предложение 1. *Пусть отображение* $f : \mathbb{R}^d \to \mathbb{R}^n$ *дифференцируемо в точке* $\boldsymbol{x}$. *Тогда его дифференциал* $L_{\boldsymbol{x}}$ *определён однозначно.*

Доказательство этого факта относится к необязательному материалу и потому может быть пропущено.

Доказательство. Пусть L_1 и L_2 – линейные отображения, удовлетворяющие определению дифференцируемой функции, то есть

$$f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) = L_1\boldsymbol{h} + \alpha_1(\boldsymbol{h})\|\boldsymbol{h}\|,$$

$\alpha_1(\boldsymbol{h}) \to \boldsymbol{0}$ и

$$f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) = L_2\boldsymbol{h} + \alpha_2(\boldsymbol{h})\|\boldsymbol{h}\|,$$

$\alpha_2(\boldsymbol{h}) \to \boldsymbol{0}$. Пусть $L = L_1 - L_2$, тогда, вычитая из верхнего равенства нижнее, получим

$$L\boldsymbol{h} = (\alpha_2(\boldsymbol{h}) - \alpha_1(\boldsymbol{h}))\|\boldsymbol{h}\|.$$

Так как разность двух линейных отображений снова является линейным отображением, то последнее равенство можно переписать в виде $L\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|} = \alpha_2(\boldsymbol{h}) - \alpha_1(\boldsymbol{h})$. Если вместо вектора $\boldsymbol{h}$ подставить в последнее равенство вектор $\lambda\boldsymbol{h}$ ($\lambda \neq 0$) и взять норму от обеих частей, то получим равенство

$$\left\| L\frac{\lambda\boldsymbol{h}}{\|\lambda\boldsymbol{h}\|} \right\| = \|\alpha_2(\lambda\boldsymbol{h}) - \alpha_1(\lambda\boldsymbol{h})\|.$$

В левой части этого равенства вектор, на который действует линейное отображение L, всегда единичный, а в правой части при $\lambda \to 0$ предел будет равен нулю, откуда мы делаем вывод, что на любом единичном векторе значение отображения L дает нулевой вектор. Но тогда отображение L тождественно нулевое, а значит для любого вектора $\boldsymbol{h}$ имеем равенство

$$L_1\boldsymbol{h} = L_2\boldsymbol{h}.$$

□

В следующем предложении выясняется связь между непрерывностью и дифференцируемостью.

Предложение 2. *Если отображение* $f : \mathbb{R}^d \to \mathbb{R}^m$ *дифференцируемо в точке* $\boldsymbol{x} \in \mathbb{R}^d$, *то оно непрерывно в этой точке.*

Доказательство. Пусть в $\mathbb{R}^k$ задан базис $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \ldots,\ \boldsymbol{e}_k$. Так как отображение f дифференцируемо в точке $\boldsymbol{x}$, то в силу линейности дифференциала и неравенства треугольника получим

$$\|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})\| = \|L_{\boldsymbol{x}}\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|\| = \|df(\boldsymbol{x})\left(\sum_{j=1}^{d} h_j\boldsymbol{e}_j\right) + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|\| \leq$$

$$\leq \sum_{j=1}^{k} \|df(\boldsymbol{x})\boldsymbol{e}_j\| \cdot |h_j| + \|\alpha(\boldsymbol{h})\| \cdot \|\boldsymbol{h}\| \to 0$$

при $\|\boldsymbol{h}\| \to 0$, откуда следует, что и $f(\boldsymbol{x}+\boldsymbol{h}) \to f(\boldsymbol{x})$ при $\|\boldsymbol{h}\| \to 0$, что означает непрерывность f в точке $\boldsymbol{x}$. □

Упражнение 3. Доказать, что обратное неверно. Для этого рассмотреть, например, отображение $f : \mathbb{R}^2 \to \mathbb{R}$, принимающее на векторе $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ значение $f(\boldsymbol{x}) = |x_1| + |x_2|$. Оно непрерывно в точке $(0, 0)$, но не дифференцируемо в этой точке.

На будущее запомним, что *дифференцируемость отображения в точке равносильна дифференцируемости каждой из координатных функций отображения в этой точке.* Этот факт мы не раз будем использовать в дальнейшем.

Производная по направлению и частная производная

Рассмотрим более подробно отображение из $\mathbb{R}^2$ в $\mathbb{R}$. На рисунке 1 ниже нарисован пример графика такого отображения. Зафиксируем точку $\boldsymbol{x} = (x_1, x_2)$ и единичный вектор $\boldsymbol{v} \in \mathbb{R}$. Множество всех точек $z = f(\boldsymbol{x} + t\boldsymbol{v})$ при переменном $t \in \mathbb{R}$ представляет собой кривую на поверхности, являющейся графиком функции. Действительно, умножая на различные t вектор $\boldsymbol{v}$ и откладывая его от точки $\boldsymbol{x}$, мы получаем прямую l на плоскости Oxy, а множество значений отображения f в точках этой прямой на графике задаётся пересечением поверхности с плоскостью Π, проходящей через прямую l и перпендикулярной Oxy. Мы можем ставить вопрос о наличии у кривой, образованной пересечением плоскости Π и графиком отображения, касательной в точке $\boldsymbol{x}$. Этот вопрос важен, если, например, мы хотим выяснить, в направлении какого вектора $\boldsymbol{v}$ наклон поверхности максимальный, то есть угол между прямой l и касательной максимален. В этом случае вопрос, по сути, сводится к существованию производной функции одной переменной $\varphi(t) = f(\boldsymbol{x} + t\boldsymbol{v})$, график которой лежит в плоскости Π. Такая производная будет называться *производной по направлению* отображения f. Ниже будет дано точное определение.

Напомним, что дифференцируемость в точке $\boldsymbol{x}$ означает наличие касательной плоскости в этой точке. В этом случае прямая, образованная пересечением плоскости Π и касательной плоскости как раз и будет касательной в точке $\boldsymbol{x}$ к нашей кривой.

Отметим, что $\mathbb{R}^2$ и $\mathbb{R}$ были выбраны только для того, чтобы мы имели возможность визуализировать понятие производной по направлению, но в общем случае, когда отображение f действует из $\mathbb{R}^d$ в $\mathbb{R}^m$ все обозначения и объяснения имеют ту же суть, что и в рассмотренном нами просто случае.

Сейчас будет доказано утверждение, в котором формализуются введенные понятия.

Предложение 3. *Пусть отображение $f : \mathbb{R}^d \to \mathbb{R}^m$ дифференцируемо в точке $\boldsymbol{x} \in \mathbb{R}^d$. Тогда для любого вектора $\boldsymbol{v} \in \mathbb{R}^d$ единичной длины функция $\varphi(t) = f(\boldsymbol{x} + t\boldsymbol{v})$ дифференцируема в точке $t = 0$ и*

$$\lim_{t \to 0} \frac{f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x})}{t} = df(\boldsymbol{x})\boldsymbol{v}.$$

Напомним, символ $df(\boldsymbol{x})\boldsymbol{v}$ означает действие линейного отображения $df(\boldsymbol{x})$ на вектор $\boldsymbol{v}$.

Доказательство. Положим в определении дифференцируемости $\boldsymbol{h} = t\boldsymbol{v}$. Тогда, в силу линейности дифференциала и определения нормы, получим:

$$f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x}) = df(\boldsymbol{x})t\boldsymbol{v} + \alpha(t\boldsymbol{v})\|t\boldsymbol{v}\| = tdf(\boldsymbol{x})\boldsymbol{v} + |t|\alpha(t\boldsymbol{v})\|\boldsymbol{v}\|.$$

Деля на t, имеем

$$\frac{f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x})}{t} = df(\boldsymbol{x})\boldsymbol{v} + \operatorname{sgn} t \cdot \alpha(t\boldsymbol{v})\|\boldsymbol{v}\|,$$

откуда, переходя к пределу при $t \to 0$, получаем справедливость нашего утверждения. □

Определение 3. *Предел из доказанного предложения обозначается $f'_t(\boldsymbol{x}+t\boldsymbol{v})|_{t=0}$ и называется* ***производной** f **по направлению** $\boldsymbol{v}$ **в точке** $\boldsymbol{x}$. Производная по направлению в точке $\boldsymbol{x}$ обозначается ещё символом $\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x})$.*

В случае отображения $f:\mathbb{R}^2\to\mathbb{R}$ ниже приведена иллюстрация, проясняющая геометрический смысл производной по направлению. В этом случае плоскость Π, перпендикулярная координатной плоскости Oxy, пересекает поверхность, являющуюся графиком функции f, по кривой, к которой в точке с координатами $(x, y, f(x,y))$ проведена касательная прямая.

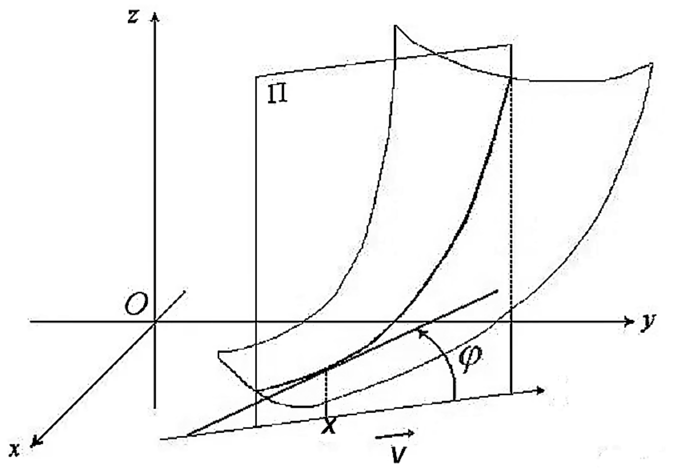

Рис. 1: На рисунке $f'_t(\mathbf{x}+t\mathbf{v})|_{t=0}=\operatorname{tg}\varphi$

Отметим, что дифференцируемость отображения f позволяет найти производную отображения f по любому направлению, что видно из доказательства последнего предложения, однако обратное неверно, то есть из существования производной по любому направлению ещё не следует дифференцируемость отображения. В качестве упражнения можно попробовать построить соответствующий пример.

Производные по некоторым направлениям играют особую роль, так как через них можно выразить дифференциалы отображений. К таким направлениям относятся те, которые параллельны осям координат, то есть это направления, совпадающие с базисными векторами. О таких производных речь идёт в определении ниже. Это определение верно для отображений со значениями на вещественной прямой. Такие отображениями чаще называют просто функциями многих переменных, поэтому мы далее будем называть отображения со значениями в $\mathbb{R}$ функциями.

Определение 4. ***Частной производной** функции $f:\mathbb{R}^d\to\mathbb{R}$ по переменной x_j в точке $\boldsymbol{x}=(x_1,x_2,...,x_j,...,x_d)$ называется производная этой функции в точке $\boldsymbol{x}$ вдоль базисного вектора $\boldsymbol{e}_j=(0,0,...,0,\underbrace{1}_{j-\text{е место}},0,...,0)$ пространства $\mathbb{R}^d$. Частная производная функции f по переменной x_j в точке $\boldsymbol{x}$ обозначается $f'_{x_j}(\boldsymbol{x})$ или $\frac{\partial f}{\partial x_j}(\boldsymbol{x})$. В терминах пределов это определение записывается так:*

$$\frac{\partial f}{\partial x_j}(\boldsymbol{x})=\lim_{h_j\to 0}\frac{f(x_1,...,x_{j-1},x_j+h_j,x_{j+1},...,x_d)-f(x_1,...,x_j,...,x_d)}{h_j}.$$

Для поиска частной производной функции f необходимо продифференцировать её просто как функцию одной переменной x_j, считая остальные переменные фиксированными константами, то есть

$$\frac{\partial f}{\partial x_j}(\boldsymbol{x}) = \frac{df(x_1, ..., x_{j-1}, t, x_{j+1}, ..., x_k)}{dt}|_{t=x_j}.$$

Пример 2. *Пусть функция трёх переменных $f : \mathbb{R}^3 \to \mathbb{R}$ задаётся формулой*

$$f(x, y, z) = e^{xy} \sin z.$$

Найдём частные производные этой функции по всем трём переменным в точке $A(1, 2, 3)$. Для нахождения f'_x отметим, что меняется только переменная x, а переменные y и z фиксированы и равны 2 и 3., поэтому

$$f'_x(1, 2, 3) = 2e^{2x} \sin 3|_{x=1} = 2e^2 \sin 3.$$

Мы взяли производную по x, считая y и z просто константами. Отметим, что так как нам требуется найти частные производные функции f при указанных числовых значениях переменных, то можно как сначала подставить в функцию числа, соответствующие тем переменным, которые мы фиксируем, так и сначала продифференцировать по переменной, считая все остальные константами, а потом уже подставить числовые значения. Найдём частную производную по y таким способом:

$$f'_y(1, 2, 3) = xe^{xy} \sin z|_{x=1,y=2,z=3} = e^2 \sin 3.$$

Наконец,

$$f'_z(1, 2, 3) = (e^2 \sin z)' = e^2 \cos z.$$