

Линейные функции

Пусть X и Y - нормированные пространства. $L: X \rightarrow Y$ - линейное отображение.

Опр: 1. Отображение $L: X \rightarrow Y$ называется линейным (или еще говорят линейным оператором), если выполнено следующее равенство:

$$L(\alpha_1 x + \alpha_2 y) = \alpha_1 L(x) + \alpha_2 L(y), \forall \alpha, \beta \in \mathbb{R}, x, y \in X$$

Непрерывность линейных отображений

Конечномерный случай: Линейные отображения $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ задаются следующим образом:

$$L(x) = \underset{m \times n}{A} x$$

У этого отображения k -ый столбец матрицы A равен $L(e_k)$, где $e_k = (0 \dots 0 \ 1 \ 0 \dots 0)^T$. В частном случае $m = 1$ это видно сразу, поскольку оператор будет иметь следующий вид

$$L(x) = a_1 x_1 + \dots + a_n x_n$$

Если координаты к чему-то стремятся, то это выражение будет стремиться к сумме произведения коэффициентов на координатные пределы. Следовательно, в конечномерных случаях имеем непрерывные функции, как комбинацию непрерывных функций.

Бесконечномерный случай: В этом случае непрерывности может не быть:

$$L: C^1[0, 1] \rightarrow \mathbb{R}, L(f) = f'(0)$$

L - не является непрерывным отображением, но это линейное отображение.

Теорема 1. Пусть X и Y - нормированные пространства. $L: X \rightarrow Y$ - линейная функция, тогда следующие утверждения равносильны:

- (1) L - непрерывна;
- (2) L - непрерывна в 0;
- (3) $\exists C > 0: \|L(x)\|_Y \leq C\|x\|_X, \forall x$;

□

(1) \Rightarrow (2) Очевидно: если функция непрерывна, то она непрерывна в 0.

(2) \Rightarrow (3) Функция L непрерывна в 0 \Rightarrow ограничена в окрестности 0, тогда:

$$\exists C_1 > 0, \exists B(0, \delta): \|L(x)\| \leq C_1, \forall x \in B(0, \delta)$$

Пусть $x \neq 0 \Rightarrow$ рассмотрим вместо x следующий вектор: $\frac{\delta}{2} \cdot \frac{x}{\|x\|}$, длина этого вектора равна $\frac{\delta}{2}$. Тогда:

$$\frac{\delta}{2} \cdot \frac{x}{\|x\|} \in B(0, \delta) \Rightarrow \left\| L\left(\frac{\delta}{2} \cdot \frac{x}{\|x\|}\right) \right\| = \left\| \frac{\delta}{2} \cdot \frac{1}{\|x\|} L(x) \right\| = \frac{\delta}{2} \cdot \frac{1}{\|x\|} \cdot \|L(x)\| \leq C_1 \Rightarrow \|L(x)\| \leq \frac{2C_1}{\delta} \|x\| = C\|x\|$$

(3) \Rightarrow (1) $\|L(x - y)\| = \|L(x) - L(y)\| \leq C\|x - y\| \Rightarrow$ по определению непрерывности L - непрерывна. ■

Линейное пространство линейных непрерывных операторов

$\mathcal{L}(X, Y)$ - линейное пространство линейных непрерывных (ограниченных) операторов (функций, отображений). Пусть $L \in \mathcal{L}(X, Y)$, определим норму

$$\|L\| = \sup_{x \neq 0} \frac{\|L(x)\|}{\|x\|}$$

Данный супремум конечен из-за свойства (3) теоремы выше.

Упр. 1. Проверить, что пространство \mathcal{L} с указанной выше нормой: $(\mathcal{L}, \|\cdot\|)$ - нормированное.

□ Поскольку пространство линейное, то нам необходимо проверить свойства нормы:

- (0) Выполнено по определению функции $\|\cdot\|_X$, так как $x \neq 0 \Rightarrow \|x\|_X > 0 \Rightarrow \sup_{x \neq 0} \frac{\|L(x)\|_Y}{\|x\|_X} \geq 0$;
- (1) $\|L\| = 0 \Leftrightarrow \forall x \neq 0, \|L(x)\|_Y = 0 \Leftrightarrow \forall x \neq 0, L(x) = 0$, а поскольку $L(0) = L(0) + L(0) \Rightarrow L(0) = 0$, то получим, что $L(x) = 0, \forall x \in X \Leftrightarrow L = 0$;
- (2) $\|\alpha L\| = \sup_{x \neq 0} \frac{\|\alpha L(x)\|}{\|x\|} = \sup_{x \neq 0} \frac{|\alpha| \cdot \|L(x)\|}{\|x\|} = |\alpha| \cdot \sup_{x \neq 0} \frac{\|L(x)\|}{\|x\|} = |\alpha| \cdot \|L\|, \forall \alpha \in \mathbb{R}$;
- (3) $\frac{\|L(x) + M(x)\|}{\|x\|} \leq \frac{\|L(x)\|}{\|x\|} + \frac{\|M(x)\|}{\|x\|} \leq \|L\| + \|M\| \Rightarrow \|L + M\| = \sup_{x \neq 0} \frac{\|L(x) + M(x)\|}{\|x\|} \leq \|L\| + \|M\|$;

Таким образом, все свойства нормы выполнены и пространство является нормированным. ■

Rm: 1. Сделаем ряд замечаний, относительно конечномерных случаев:

- 1) Пусть $L: \mathbb{R}^n \rightarrow \mathbb{R}$, тогда $L(x) = a_1x_1 + \dots + a_nx_n = \langle a, x \rangle \Rightarrow \|L(x)\| = |L(x)|$, поскольку находимся в пространстве \mathbb{R} . Тогда, в силу неравенства Коши-Буняковского $|L(x)| \leq \|a\| \cdot \|x\|$. Отсюда следует, что L - непрерывное линейное отображение;

- 2) Пусть $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$, тогда $L(x) = \begin{pmatrix} L_1(x) \\ L_2(x) \\ \vdots \\ L_m(x) \end{pmatrix}$ - координаты вектора, который получается отображе-

нием $x \in \mathbb{R}^n$. Очевидно, что $L_k: \mathbb{R}^n \rightarrow \mathbb{R}$ - линейно (следует из покоординатного сложения)

$$L(\alpha x + \beta y) = \begin{pmatrix} L_1(\alpha x + \beta y) \\ L_2(\alpha x + \beta y) \\ \vdots \\ L_m(\alpha x + \beta y) \end{pmatrix} = \alpha L(x) + \beta L(y) = \alpha \begin{pmatrix} L_1(x) \\ L_2(x) \\ \vdots \\ L_m(x) \end{pmatrix} + \beta \begin{pmatrix} L_1(y) \\ L_2(y) \\ \vdots \\ L_m(y) \end{pmatrix}$$

тогда, в силу первого пункта замечания:

$$|L_k(x)| \leq C_k \|x\|$$

Отсюда мы получаем следующее:

$$\|L(x)\| \leq \|C\| \cdot \|x\| = \sqrt{\sum_k C_k^2} \cdot \|x\|$$

Таким образом, L - непрерывный линейный оператор.

Дифференцируемые отображения

Пусть X и Y - нормированные пространства и $a \in X$. Предположим, что $f: U \rightarrow Y$, где U - открытое множество, содержащее a , то есть f определена в окрестности точки a .

Опр: 2. Функция f дифференцируема в точке a , если существует непрерывное линейное отображение $L: X \rightarrow Y$ и функция $\alpha: X \rightarrow Y$ такие, что:

$$f(a+h) - f(a) = L(h) + \alpha(h) \cdot \|h\|, \quad \alpha(h): \lim_{h \rightarrow 0} \alpha(h) = 0$$

Пример: Пусть $L: X \rightarrow Y$ - линейная непрерывная функция, тогда L - дифференцируема в каждой точке. Действительно:

$$L(a+h) - L(a) = L(a+h-a) = L(h) + 0 \cdot \|h\|$$

Утв. 1. Если f дифференцируема в точке a , то f непрерывна в точке a .

□ $f(a+h) - f(a) = L(h) + \alpha(h) \cdot \|h\| \xrightarrow{h \rightarrow 0} L(0) + 0 = 0 \Rightarrow \lim_{x \rightarrow a} f(x) = f(a)$, поскольку $L(h)$ - непрерывная линейная функция и $x = a+h$, $h \rightarrow 0 \Leftrightarrow x \rightarrow a$. ■

Rm: 2. Дифференцируемые функции всегда непрерывны, поэтому в качестве недифференцируемых функций подойдут линейные разрывные.

Теорема 2. f дифференцируема в точке $a \Leftrightarrow$ существует линейное непрерывное отображение L такое, что:

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0$$

В частности, если f дифференцируема в точке a и L - линейная часть приращения f , то:

$$\forall v, \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t} = L(v)$$

Rm: 3. В прошлом семестре, дифференцируемость давалась через: $f(a+h) - f(a) = Ah + \alpha(h) \cdot h$. После чего, давалось определение производной: $f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$. Затем давалось, что дифференцируемость и существование производной равносильны и если f - дифференцируема, то $A = f'(a)$.

Rm: 4. Если не брать предел по h , который подходит произвольно к нулю, а брать как предел подходящий к нулю вдоль вектора v , то получим ситуацию аналогичную одномерному случаю.

□ Функция f - дифференцируема в точке $a \Leftrightarrow f(a+h) - f(a) = L(h) + \alpha(h) \cdot \|h\|$, $\lim_{h \rightarrow 0} \alpha(h) = 0$, но это то же самое, что:

$$\frac{f(a+h) - f(a) - L(h)}{\|h\|} = \alpha(h) \rightarrow 0 \Leftrightarrow \lim_{h \rightarrow 0} \|\alpha(h)\| = 0 \Leftrightarrow \lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0$$

Рассмотрим частный случай и возьмем $h = tv$, $v \neq 0$, тогда:

$$\frac{f(a+tv) - f(a) - L(tv)}{\|tv\|} \xrightarrow{t \rightarrow 0} 0 \Leftrightarrow \frac{1}{\|v\|} \cdot \frac{t}{|t|} \cdot \left(\frac{f(a+tv) - f(a)}{t} - L(v) \right) \xrightarrow{t \rightarrow 0} 0$$

Поскольку $\left\| \frac{1}{\|v\|} \cdot \frac{t}{|t|} \right\| = \left| \frac{1}{\|v\|} \cdot \frac{t}{|t|} \right| = \frac{1}{\|v\|} \neq 0$, то получаем, что:

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} - L(v) = 0 \Rightarrow \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = L(v)$$

■

Опр: 3. Предел $\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$ называют производной функции f по вектору v и обозначают следующим образом:

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$$

Рм: 5. Если $\|v\| = 1$, то эту производную $\frac{\partial f}{\partial v}(a)$ еще называют производной по направлению. Иногда требуют, чтобы $t \rightarrow 0+$.

Следствие 1. $L(v) = \frac{\partial f}{\partial v}(a)$, $\forall v$ и L определено однозначно.

□ Следует из доказательства для частного случая теоремы выше, поскольку значение этого отображения на векторе v однозначно определено функцией f . ■

Опр: 4. Линейный оператор L из определения дифференцируемости f называется дифференциалом f . Обозначается, как $df(a, h)$ или $df(h)$, если понимаем в какой точке все происходит.

Тогда определение дифференцируемости записывается следующим образом:

Опр: 5. Функция f дифференцируема в точке a , если выполнено следующее равенство:

$$f(a + h) - f(a) = df(a, h) + \alpha(h) \cdot \|h\|, \lim_{h \rightarrow 0} \alpha(h) = 0$$

где $h \mapsto df(a, h)$ - непрерывное линейное отображение, а также $df(a, v) = df(v) = \frac{\partial f}{\partial v}(a)$.

Частные случаи дифференцируемости функций

(I) Отображения из \mathbb{R}^n в \mathbb{R}

Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируема в точке $a = (a_1, \dots, a_n)$. Запишем определение дифференцируемости:

$$f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n) = df(\vec{h}) + \alpha(\vec{h}) \cdot \|\vec{h}\|$$

где $df(\vec{h})$ - линейное отображение из \mathbb{R}^n в \mathbb{R} , тогда:

$$df(h) = h_1 \cdot df(e_1) + \dots + h_n \cdot df(e_n), e_k = (0 \dots 0 \ 1 \ 0 \dots 0)^T$$

По определению $df(e_k) = \frac{\partial f}{\partial e_k}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_k) - f(a)}{t}$, то есть фиксируем все координаты значениями a_j , кроме k -ой координаты:

$$\lim_{t \rightarrow 0} \frac{f(a + te_k) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_{k-1}, a_k + t, a_{k+1}, \dots, a_n) - f(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n)}{t}$$

Таким образом, мы рассматриваем функцию $x_k \mapsto f(a_1, \dots, x_k, \dots, a_n)$ одной переменной x_k и дифференцируем её по этой переменной.

Опр: 6. Производная вдоль вектора e_k : $\frac{\partial f}{\partial e_k}(a)$ называется частной производной функции f по переменной x_k в точке a и обозначается следующим образом:

$$\frac{\partial f}{\partial x_k}(a) = \frac{\partial f}{\partial e_k}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_k) - f(a)}{t}$$

Частная производная означает дифференцирование вдоль координат. Таким образом, мы можем записать дифференциал в следующем виде:

$$df(h) = \frac{\partial f}{\partial x_1}(a) \cdot h_1 + \dots + \frac{\partial f}{\partial x_n}(a) \cdot h_n$$

Опр: 7. Вектор состоящий из частных производных $\left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\right)$ в точке a называется градиентом функции f и обозначается следующим образом:

$$\text{grad } f(a) = \nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\right)$$

Рассмотрим следующее выражение:

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = df(v) = \frac{\partial f}{\partial x_1}(a) \cdot v_1 + \dots + \frac{\partial f}{\partial x_n}(a) \cdot v_n = \langle \nabla f(a), v \rangle$$

Таким образом, производная по направлению это скалярное произведение градиента с вектором v . Возникает вопрос, а что можно говорить о поведении функции, глядя на градиент?

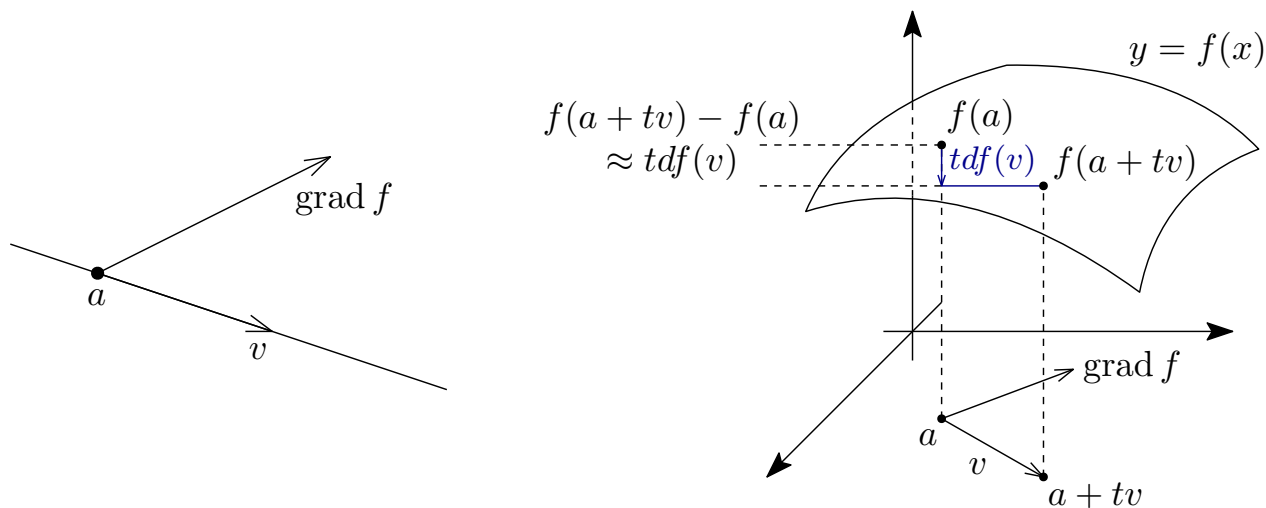


Рис. 1: Геометрический смысл производной по вектору v .

Геометрический смысл производной по направлению: При достаточно маленьких значениях t , приращение функции f в точке a будет примерно равно $tdf(v) \approx f(a + tv) - f(a)$. Таким образом, производная по вектору v указывает, как изменяется функция, когда мы сдвигаемся по вектору v .

Но мы знаем, что $df(v) = \langle \nabla f(a), v \rangle$. Тогда если возьмем v перпендикулярно градиенту, то скалярное произведение будет равно 0 \Rightarrow приращение функции примерно нулевое. То есть двигаясь вдоль градиента, значения функции остаются постоянными. Если будем двигаться в сторону градиента, то это скалярное произведение будет максимальным, в противоположную сторону - минимальным.

Утв. 2. Предположим, что $f: \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируема в точке a и $\nabla f(a) \neq 0$. Тогда максимум производной функции f по направлению v : $\max_{\|v\|=1} \frac{\partial f}{\partial v}(a)$ достигается на $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$ и соответственно минимум $\min_{\|v\|=1} \frac{\partial f}{\partial v}(a)$ достигается на $v = -\frac{\nabla f(a)}{\|\nabla f(a)\|}$.

□ Рассмотрим производную, как скалярное произведение: $\frac{\partial f}{\partial v}(a) = \langle \nabla f(a), v \rangle$. Поскольку $\|v\| = 1$, используем неравенство Коши-Буняковского и получим следующий результат:

$$-\|\nabla f(a)\| \cdot \|v\| = -\|\nabla f(a)\| \leq \frac{\partial f}{\partial v}(a) = \langle \nabla f(a), v \rangle \leq \|\nabla f(a)\| \cdot \|v\| = \|\nabla f(a)\|$$

Равенство возможно только тогда, когда $\nabla f(a)$ и v - линейно зависимы, а поскольку $\|v\| = 1$, то для равенства в правой части $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$ и $v = -\frac{\nabla f(a)}{\|\nabla f(a)\|}$ для равенства в левой части. ■

Пример: $f(x_1, x_2) = x_1^2 + x_2^2$, возьмем точку (a_1, a_2) и найдем градиент функции в этой точке:

$$\nabla f(a_1, a_2) = \left(\frac{\partial f}{\partial x_1}(a_1, a_2), \frac{\partial f}{\partial x_2}(a_1, a_2) \right) = (2a_1, 2a_2)$$

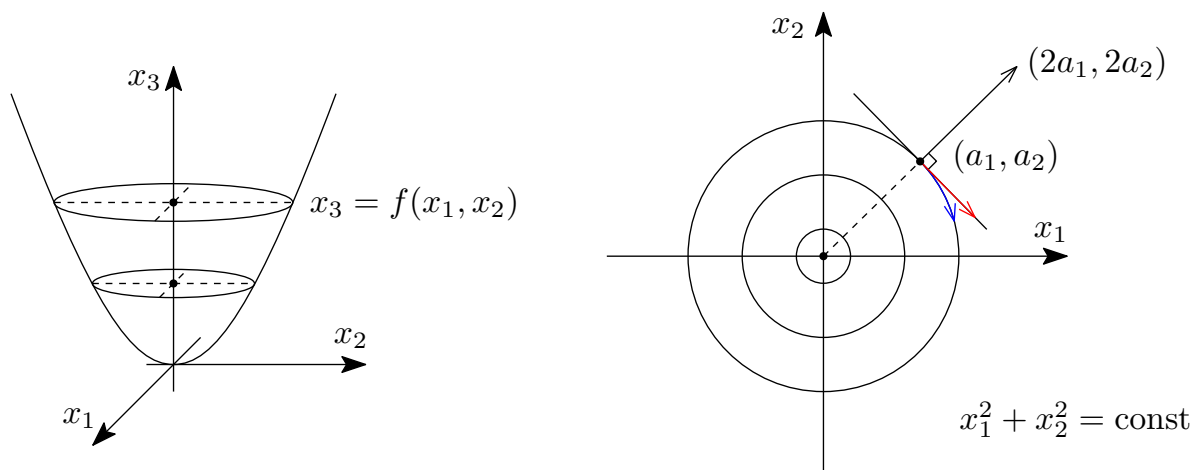


Рис. 2: Поиск максимального возрастания функции f с помощью линий уровня и градиента $(2a_1, 2a_2)$.

Видно, что вектор градиента перпендикулярен линии уровня и это отражает наше наблюдение, что если двигаться перпендикулярно, то значения функции не должны изменяться. Как раз если будем сдвигаться перпендикулярно градиенту (красная линия), то будем двигаться вдоль линии уровня (синяя линия) и тем самым значение функции будет постоянным.

Если будем двигаться в сторону градиента, то функция будет максимально увеличиваться. Минус градиент будет направлен к нулю и если будем двигаться вдоль него, то функция будет максимально уменьшаться.

Rm: 6. На этой идее построен метод градиентного спуска.