

# Основные элементы математического анализа, используемые в теории оптимизации

Мальковский Н. В.

Санкт-Петербургский Академический Университет



# Скалярное произведение

## Определение

Пусть  $X$  – линейное пространство над  $\mathbb{R}$  (или  $\mathbb{C}$ ). Скалярным произведением на  $X$  называется функция  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$  (или  $\mathbb{C}$ ), обладающей следующими свойствами

- 1 Симметричность

$$\langle x, y \rangle = \overline{\langle y, x \rangle}, \quad x, y \in X,$$

где  $\bar{y}$  – комплексное сопряжение  $y$ .

- 2 Линейность

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle, \quad x, y \in X,$$

$$\langle \alpha(x + z), y \rangle = \alpha \langle x + z, y \rangle, \quad x \in X, \quad \alpha \in \mathbb{R}(\mathbb{C}).$$

- 3 Положительная определенность,  $\langle x, x \rangle \in \mathbb{R}$  и

$$\langle x, x \rangle \in \mathbb{R} \geq 0$$

$$\langle x, x \rangle \in \mathbb{R} = 0 \Leftrightarrow x = 0.$$

## Скалярное произведение

Стандартное скалярное произведение в  $\mathbb{R}^n$ :  $x = (x_1, \dots, x_n)^T$ ,  
 $y = (y_1, \dots, y_n)^T$ , тогда

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^T y = y^T x.$$

Последнее обозначение – стандартное матричное произведение, в дальнейшем будет использовано в основном только оно.

Скалярное произведение в  $\mathbb{C}^n$ :

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i = y^\dagger x.$$

Скалярное произведение функций из  $C([a, b])$ :

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx$$

## Определение

Пусть  $X$  – линейное пространство над  $\mathbb{R}$  (или  $\mathbb{C}$ ). Нормой на  $X$  называется функция  $\|\cdot\| : X \rightarrow \mathbb{R}$ , для которой выполняются следующие свойства

- 1 Строгая положительная определенность

$$\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0, x \in X.$$

- 2 Гомогенность

$$\|\alpha x\| = |\alpha| \cdot \|x\|, x \in X, \alpha \in \mathbb{R}(\mathbb{C}).$$

- 3 Неравенство треугольника

$$\|x + y\| \leq \|x\| + \|y\|, x, y \in X.$$

# Норма и нормированные пространства

Норма, индуцированная скалярным произведением: если  $\langle \cdot, \cdot \rangle$  – скалярное произведение, то  $\|x\| = \sqrt{\langle x, x \rangle}$  – норма.

$p$ -норма в  $\mathbb{R}^n$ : пусть  $p \geq 1$ , тогда функция

$$g(x) = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$$

является нормой. Принято обозначение  $\|x\|_p := g(x)$ . При  $p = 2$  получаем *евклидову норму*, она же является нормой, индуцированной стандартным скалярным произведением в  $\mathbb{R}^n$ :

$$\|x\|_2^2 = \sum_{i=1}^n x_i^2 = x^T x.$$

В дальнейшем по умолчанию в  $\mathbb{R}^n$  будет использоваться именно эта норма.

# Норма и нормированные пространства

При  $p = 1$  получаем

$$\|x\|_1 = \sum_{i=1}^n |x_i|.$$

Функция  $\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p$  также является нормой:

$$\begin{aligned} \|x\|_\infty &= \lim_{p \rightarrow \infty} \|x\|_p = \lim_{p \rightarrow \infty} \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} = \lim_{p \rightarrow \infty} \max_j |x_j| \left( \underbrace{\sum_{i=1}^n \left( \frac{|x_i|}{\max_j |x_j|} \right)^p}_{1 \leq \dots \leq n} \right)^{1/p} \\ &= \max_j |x_j|. \end{aligned}$$

Метрика, индуцированная нормой: если  $\|\cdot\|$  – норма, то  $\rho(x, y) = \|x - y\|$  – метрика.

## Операторная и двойственная нормы

Если  $\|\cdot\|$  – норма в  $X$ , индуцированная скалярным произведением, то

$$\|A\|_{op} := \sup_{x \in X} \frac{\|Ax\|}{\|x\|}$$

норма в пространстве ограниченных операторов над  $X$ .

Двойственная норма: если  $\|\cdot\|$  – норма в  $X$ , то

$$\|s\|_* := \sup_{x \in X} \frac{|\langle s, x \rangle|}{\|x\|}$$

является нормой в пространстве линейных функционалов над  $X$  (обозначается  $X^*$ ). Двойственная норма является операторной нормой:  $\|s\|_* = \|\langle s, \cdot \rangle\|_{op}$ . Основное свойство этих норм:

$$\|Ax\| \leq \|A\|_{op} \|x\|$$

$$\langle s, x \rangle \leq \|s\|_* \|x\|.$$

# Операторная и двойственная нормы

Пусть  $p, q > 0, p^{-1} + q^{-1} = 1$ . Из неравенства Гёльдера

$$\frac{|\langle s, x \rangle|}{\|x\|_p} \leq \frac{\|s\|_q \|x\|_p}{\|x\|_p} = \|s\|_q,$$

причем равенство достигается при  $s = x$ , а значит  $\|\cdot\|_q$  – норма, двойственная к  $\|\cdot\|_p$ . В частности, евклидова норма двойственна сама себе. Также,  $\|\cdot\|_1$  двойственна к  $\|\cdot\|_\infty$ .

Наконец, двойственность симметрична, т.е.

$$\|x\|_{**} = \|x\|.$$



# Дифференцируемость одномерной функции

## Определение

Пусть  $f : \mathcal{D} \subset \mathbb{R} \rightarrow \mathbb{R}$ ,  $x$  – внутренняя точка  $\mathcal{D}$ . Говорят, что  $f$  дифференцируема в точке  $x$ , если существует такое  $a$ , что выполняется

$$f(x + h) = f(x) + ah + o(h).$$

Число  $a$  принято называть производной  $f$  в точке  $x$ .

*Замечание 1.* Функция дифференцируема на множестве  $A \in \mathcal{D}$ , если она дифференцируема в каждой точке  $A$ , обозначается  $f \in C^{(1)}(A)$ . Производная функции  $f$  – функция, сопоставляющая каждой точке производную  $f$  в этой точке, обозначается  $f'$ ,  $\frac{df}{dx}$  или  $\frac{d}{dx}f$ .

*Замечание 2.* Если  $f'$  также является дифференцируемой функцией, то говорят, что  $f$  – дважды дифференцируема. Вторая производная  $f$  – производная производной  $f$ , обозначается  $f''$ ,  $\frac{d^2f}{dx^2}$  или  $\frac{d^2}{dx^2}f$ . Последующие производные вводятся индуктивно, обозначаются  $f^{(n)}$ ,  $\frac{d^n f}{dx^n}$  или  $\frac{d^n}{dx^n}$ , для дифференцируемости используется обозначение  $f \in C^{(n)}(A)$ .

## Определение

Пусть  $f \in C^{(n)}([a, b])$ ,  $x \in (a, b)$ , рядом Тейлора функции  $f$  в точке  $x$  называется следующее разложение  $f$  в точке  $x$

$$f(x+h) = \sum_{i=0}^n f^{(i)}(x) \frac{h^i}{i!} + R(x, h).$$

*Замечание 1.* Остаток в форме Пеано:

$$f(x+h) = \sum_{i=0}^n f^{(i)}(x) \frac{h^i}{i!} + o(h^n).$$

*Замечание 2.* Остаток в форме Лагранжа ( $f \in C^{(n+1)}([a, b])$ ): Существует точка  $c \in (x, x+h)$ , что выполняется

$$f(x+h) = \sum_{i=0}^n f^{(i)}(x) \frac{h^i}{i!} + f^{(n+1)}(c) \frac{h^{n+1}}{(n+1)!}.$$

*Замечание 3.* Есть и другие формы остатка: интегральная, Коши и т. д.

# Частные производные и градиент

## Определение

Пусть  $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\mathcal{D}_1 = \{z \mid z \in \mathbb{R}, y \in \mathbb{R}^{n-1}, (z, y) \in \mathcal{D}\}$ ,  $g \in \mathcal{D}_1 \rightarrow \mathbb{R}$ ,  $x = (x_1, x_2, \dots, x_n) \in \mathcal{D}$  и при этом

$$g(z) = f(z, x_2, \dots, x_n).$$

Частной производной  $f$  по переменной  $x_1$  в точке  $x$  называется производная функции  $g$  в точке  $x_1$ , обозначается  $\frac{\partial f}{\partial x_1}(x)$ ,  $\frac{\partial}{\partial x_1} f(x)$  или  $\nabla_{x_1} f(x)$ . Аналогичным образом вводятся производные по переменным  $x_2, \dots, x_n$ .

## Определение

Вектор  $\left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^T$  называется градиентом функции  $f$  в точке  $x$ . Обозначается  $\nabla f(x)$ .

*Замечание.* Дифференцируемость в многомерном случае: если существует линейное приближение  $f(x+h) = f(x) + a^T h + o(\|h\|)$ , то  $f$  имеет все частные производные и при этом  $a = \nabla f(x)$ .

## Определение

Пусть  $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f$  дифференцируема в точке  $x$  и при этом существует симметричная матрица  $H$  такая, что

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T H h + o(\|h\|^2),$$

то  $f$  дважды дифференцируема в точке  $x$ , а матрицу  $H$  принято называть матрицей Гессе, гессианом или матрицей вторых производных и обозначать  $\nabla^2 f(x)$ .

*Замечание 1.* В этом определении правая часть неравенства является квадратичным приближением, что есть ни что иное как ряд Тейлора для многомерной функции.

*Замечание 2.* Если  $f$  дважды дифференцируема в точке  $x$  и дифференцируема в окрестности  $x$ , то  $\nabla^2 f(x) = \nabla(\nabla f(x))$ .

## Определение

Подмножество  $\mathcal{D}$  некоторого линейного пространства над  $\mathbb{R}$  называется выпуклым, если  $\forall t \in (0, 1), x, y \in \mathcal{D}$  выполняется

$$tx + (1 - t)y \in \mathcal{D}.$$

# Выпуклость функций

## Определение

Пусть  $\mathcal{D} \subset \mathbb{R}^n$  – выпуклое множество. Функция  $f : \mathcal{D} \rightarrow \mathbb{R}$  называется выпуклой, если  $\forall x, y \in \mathcal{D}, 0 < t < 1$  выполняется

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

## Определение

Пусть  $\mathcal{D} \subset \mathbb{R}^n$  – выпуклое открытое множество. Функция  $f : \mathcal{D} \rightarrow \mathbb{R}$  называется сильно выпуклой с константой  $m > 0$ , если  $\forall x, y \in \mathcal{D}, 0 < t < 1$  выполняется

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - t(1 - t)\frac{m}{2}\|x - y\|_2^2.$$

*Замечание.* Формально можно считать, что выпуклость – это сильная выпуклость с константой  $m = 0$ .

## Условия выпуклости первого порядка

### Теорема (Условия выпуклости первого порядка)

Если  $f$  дифференцируема, то сильная выпуклость  $f$  с параметром  $m$  равносильна

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|^2 \quad \forall x, y \in \mathcal{D}.$$

**Док-во.** Если  $f$  выпукла и дифференцируема, то для  $0 < t < 1$

$$f(x + t(y - x)) \leq f(x) + t(f(y) - f(x)) - t(1 - t)\frac{m}{2}\|x - y\|^2.$$

После деления  $t$  и переноса слагаемых получаем

$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t} + (1 - t)\frac{m}{2}\|x - y\|^2.$$

Устремляя  $t$  к нулю получаем желаемое неравенство.

## Условия выпуклости первого порядка

С другой стороны, если  $x, y \in \mathcal{D}$ ,  $z = tx + (1 - t)y$  и выполняются условия

$$f(x) \geq f(z) + \nabla f(z)^T(x - z) + \frac{m}{2}\|x - z\|^2, \quad f(y) \geq f(z) + \nabla f(z)^T(y - z) + \frac{m}{2}\|y - z\|^2,$$

то складывая первое неравенство умноженное на  $t$  и второе неравенство умноженное на  $1 - t$ , учитывая  $x - z = (1 - t)(x - y)$ ,  $y - z = t(y - x)$  получаем

$$\begin{aligned} tf(x) + (1 - t)f(y) &\geq f(z) + t(1 - t)\nabla f(z)^T(x - y) + (1 - t)t\nabla f(z)^T(y - x) \\ &\quad + t^2(1 - t)\frac{m}{2}\|x - y\|^2 + (1 - t)^2t\frac{m}{2}\|x - y\|^2 \\ &= f(z) + t(1 - t)\frac{m}{2}\|x - y\|^2 \quad \blacksquare \end{aligned}$$



## Условия выпуклости первого порядка

Другое эквивалентное определение можно получить следующим образом:  
сложив неравенства

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2} \|x - y\|^2$$

и

$$f(x) \geq f(y) + \nabla f(y)^T (x - y) + \frac{m}{2} \|x - y\|^2$$

получаем

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq m \|x - y\|^2.$$

Аналогично для выпуклых функций

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

и

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0.$$

## Условия выпуклости второго порядка

### Теорема (Условия выпуклости второго порядка)

Если  $f$  дважды дифференцируема, то сильная выпуклость  $f$  с параметром  $m$  равносильна

$$\nabla^2 f(x) \succeq ml \quad \forall x \in \mathcal{D}.$$

**Док-во.** В силу дифференцируемости  $f$  и сильной выпуклости  $f$  выполняется условие первого порядка

$$(\nabla f(x + t(y - x)) - \nabla f(x))^T t(y - x) \geq mt^2 \|x - y\|^2.$$

Делим на  $t^2$  и устремляем  $t$  к нулю:

$$(y - x)^T \nabla^2 f(x)(y - x) \geq m \|x - y\|^2 = (y - x)^T (ml)(y - x). \quad \blacksquare$$

Как следствие, для выпуклых функций получаем

$$\nabla^2 f(x) \succeq 0.$$

# Связь выпуклости для множеств и для функций

## Теорема (О надграфике выпуклой функции)

Если  $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$  – выпуклая функция,  $\mathcal{D}$  – выпуклое множество, тогда множество  $\mathbb{G} = \{(x, y) \mid x \in \mathcal{D}, y \in \mathbb{R}, y \geq f(x)\}$  выпукло.

**Док-во.** Непосредственным вычислением для любых точек  $(x, y), (u, v) \in \mathbb{G}$  и  $0 \leq t \leq 1$  получаем

$$ty + (1 - t)v \geq tf(x) + (1 - t)f(u) = tf(x) + (1 - t)f(u) \geq f(tx + (1 - t)u).$$

Следовательно,  $t(x, y) + (1 - t)(u, v) \in \mathbb{G}$ , а значит  $\mathbb{G}$  выпукло.

## Связь выпуклости для множеств и для функций

### Теорема (О выпуклости множества, заданном неравенством)

Если  $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$  – выпуклая функция,  $\mathcal{D}$  – выпуклое множество,  $c \in \mathbb{R}$ , тогда множество  $\mathbb{E} = \{x \in \mathcal{D} \mid f(x) \leq c\}$  выпукло.

**Док-во.** Пусть  $x, y \in \mathbb{E}$ , тогда для  $0 \leq t \leq 1$

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) \leq tc + (1 - t)c = c,$$

а значит  $tx + (1 - t)y \in \mathbb{E}$  и, следовательно,  $\mathbb{E}$  выпукло.

# Представление выпуклых множеств пересечением полупространств

## Теорема (О существовании опорной функции)

Пусть  $\mathcal{D} \in \mathbb{R}^n$  – замкнутое выпуклое множество, тогда существует функция  $\phi_{\mathcal{D}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  такая, что

$$\mathcal{D} = \bigcap_{x \in \mathbb{R}^n} \{y \in \mathbb{R}^n \mid \langle x, y \rangle \leq \phi_{\mathcal{D}}(x)\}$$

**Док-во.** Рассмотрим функцию

$$\phi_{\mathcal{D}}(x) := \sup_{y \in \mathcal{D}} \langle x, y \rangle.$$

Если  $x \in \mathcal{D}$ , то для любого  $z \in \mathbb{R}^n$ :  $\langle z, x \rangle \leq \sup_{y \in \mathcal{D}} \langle z, y \rangle = \phi_{\mathcal{D}}(z)$ , т. е.  $x \in \{y \in \mathbb{R}^n \mid \langle z, y \rangle \leq \phi_{\mathcal{D}}(z)\}$ , а значит  $x$  принадлежит соответствующему пересечению.

Пусть  $x \notin \mathcal{D}$ ,  $\mathcal{D}$  – замкнуто, то существуют  $a \in \mathbb{R}^n$ ,  $c \in \mathbb{R}$  (разделяющая гиперплоскость) такие, что

$$\begin{cases} \langle a, y \rangle < c \quad \forall y \in \mathcal{D}, \\ \langle a, x \rangle > c. \end{cases}$$

# Представление выпуклых множеств пересечением полупространств

Таким образом

$$\phi_{\mathcal{D}}(a) = \sup_{y \in \mathcal{D}} \langle a, y \rangle \leq c,$$

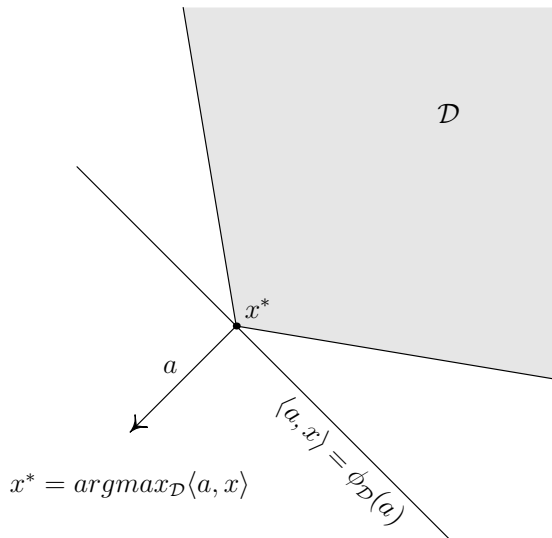
а значит  $x \notin \{z \in \mathbb{R}^n \mid \langle z, a \rangle \leq f(a)\}$  и, следовательно,  $\mathcal{D}$  совпадает с указанным пересечением. ■

*Замечание 1.* Функция  $\phi_{\mathcal{D}}$  называется опорной функцией множества  $\mathcal{D}$ . Величина  $\phi_{\mathcal{D}}(x)$  измеряет насколько “далеко” множество  $\mathcal{D}$  распространяется в направлении  $x$ .

*Замечание 2.* В общем случае указанное пересечение задает выпуклую оболочку замыкания множества  $\mathcal{D}$ .

*Замечание 3.* Если опорные функции двух замкнутых выпуклых множеств совпадают, то совпадают и сами множества.

# Геометрическая интерпретация опорной функции



# Неявно заданные функции

## Теорема (О неявной функции)

Пусть  $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ ,  $f$  непрерывно дифференцируема в окрестности  $(x^*, y^*)$ ,  $f(x^*, y^*) = 0_m$ ,  $|\nabla_y f(x^*, y^*)| \neq 0$ , тогда существуют множества  $V_x \in \mathbb{R}^n$ ,  $x^* \in V_x$ ,  $V_y \in \mathbb{R}^m$ ,  $y^* \in V_y$  и единственная функция  $\varphi : V_x \rightarrow V_y$  такая, что

$$f(x, \varphi(x)) = 0_m \quad \forall x \in V_x.$$

Более того,  $\varphi$  непрерывно дифференцируема на  $V_x$  и имеет место равенство

$$\nabla \varphi(x) = -[\nabla_y f(x, y)]^{-1} \nabla_x f(x, y).$$



## Неявно заданные функции

**Док-во.** Индукция по  $m$ . Для  $m = 1$  имеем  $\nabla_y f(x^*, y^*) \neq 0$ , не умаляя общности можно считать, что  $\nabla_y f(x^*, y^*) > 0$ . Из непрерывности  $\nabla_y f$  следует, что существует  $\epsilon > 0$  такое, что  $\nabla_y f(x^*, y) > 0$  при  $|y - y^*| \leq \epsilon$ . Следовательно  $f(x^*, y - \epsilon) < 0 < f(x^*, y + \epsilon)$ . Пусть  $V_y = \{y \in \mathbb{R} \mid |y - y^*| \leq \epsilon\}$ .

Из непрерывности  $f$  и  $\nabla_y f$  следует, что существует  $\delta > 0$  такое, что  $f(x, y - \epsilon) < 0 < f(x, y + \epsilon)$  и  $\nabla_y f(x, y) > 0$  при  $y \in V_y$ ,  $\|x - x^*\| \leq \delta$ . Пусть  $V_x = \{x \in \mathbb{R}^n \mid \|x - x^*\| \leq \delta\}$ .

Зафиксируем  $x \in V_x$ . Из непрерывности  $f$  и строгой монотонности  $f(x, \cdot)$  при  $x \in V_x$  следует существование единственной точки  $\varphi(x)$  такой, что  $f(x, \varphi(x)) = 0$ .

Пусть  $x, x + h \in V_x$ . По формуле Лагранжа

$$0 = f(x+h, \varphi(x+h)) - f(x, \varphi(x)) = \nabla_x f(x', \varphi(x'))^T h + \nabla_y f(x', \varphi(x'))^T (\varphi(x+h) - \varphi(x))$$

где  $(x', \varphi(x'))$  – точка на отрезке с концами  $(x, \varphi(x))$ ,  $(x + h, \varphi(x + h))$ .

## Неявно заданные функции

Так как  $V_x \times V_y$  – компакт, а  $\nabla_y f$  непрерывна и  $\nabla_y f(x, y) > 0$ ,  $x \in V_x, y \in V_y$ , то  $\exists \alpha > 0$ :  $\nabla_y f(x, y) \geq \alpha$ , а значит

$$|\nabla_y f(x', \varphi(x'))| \cdot |\varphi(x') - \varphi(x)| = \|\nabla_x f(x', \varphi(x'))(x' - x)\|,$$

$$|\varphi(x') - \varphi(x)| \leq \frac{1}{\alpha} \|\nabla_x f(x', \varphi(x'))(x' - x)\| \xrightarrow{x' \rightarrow x} 0,$$

а значит  $\varphi$  непрерывна на  $V_x$ . Далее

$$\begin{aligned} 0 &= f(x + th, \varphi(x + th)) - f(x, \varphi(x)) \\ &= t \nabla_x f(x', \varphi(x'))^T h + \nabla_y f(x', \varphi(x'))^T (\varphi(x + th) - \varphi(x)). \end{aligned}$$

Делим на  $t$  и устремляем  $t$  к нулю

$$\lim_{t \rightarrow 0} \frac{\varphi(x + th) - \varphi(x)}{t} = -\nabla_y f(x, \varphi(x))^{-1} \nabla_x f(x, \varphi(x))^T h.$$

А значит  $\varphi$  непрерывно дифференцируема на  $V_x$ .

## Неявно заданные функции

Индукционный переход: так как  $|\nabla_y f(x^*, y^*)| \neq 0$ , то хотя бы для одного индекса  $i$  выполняется  $\frac{\partial f_i}{\partial y_1} \neq 0$ . Не умаляя общности можно считать, что  $i = 1$ . Из доказанной базы следует, что существует  $V_x \in \mathbb{R}^n$ ,  $V_y \subset \mathbb{R}$  и единственная функция  $\psi : V_x \rightarrow V_y$  такая, что  $f_1(x, \psi(x)) = 0$ .

Обозначим  $g(x, y) = (f_2(x, y), \dots, f_m(x, y))^T$ . Из индукционного предположения следует, что существует  $V'_x \in \mathbb{R}^n$ ,  $V'_y \subset \mathbb{R}$  и единственная функция  $\gamma : V'_x \rightarrow V'_y$  такая, что  $g(x, \gamma(x)) = 0$ . Таким образом, мы получаем необходимую функцию  $\varphi : V_x \cap V'_x \rightarrow V_y \times V'_y$ ,  $\varphi(x) = \begin{bmatrix} \psi(x) \\ \gamma(x) \end{bmatrix}$ .

Наконец, дифференцируя  $f(x, \varphi(x))$  получаем

$$0 = \nabla_x(f(x, \varphi(x))) = \nabla_x f(x, \varphi(x)) + \nabla_y f(x, \varphi(x)) \nabla \varphi(x).$$

$$\nabla \varphi(x) = -[\nabla_y f(x, \varphi(x))]^{-1} \nabla_x f(x, \varphi(x)). \blacksquare$$

## Неявно заданные функции

*Замечание.* Если выполняются условия теоремы, то для любого вектора

$0_{n+m} \neq (x, y) \in \mathbb{R}^{n+m}$ ,  $\nabla f(x^*, y^*) \begin{bmatrix} x \\ y \end{bmatrix} = 0_m$  существует кривая

$\psi : [-a, a] \rightarrow \mathbb{R}^{n+m}$ ,  $a > 0$  такая, что

$$\begin{cases} f(\psi(t)) = 0_m, \\ \psi(0) = \begin{bmatrix} x^* \\ y^* \end{bmatrix}, \\ \frac{d}{dt}\psi(0) = \begin{bmatrix} x \\ y \end{bmatrix}, \end{cases}$$

**Док-во.** Пусть  $h \in \mathbb{R}^n$ , рассмотрим

$$\psi(t) = \begin{bmatrix} x^* + th \\ \varphi(x^* + th) \end{bmatrix}$$

Непосредственной подстановкой получаем  $f(\psi(t)) = f(x^* + th, \varphi(x^* + th)) = 0_m$ . Дифференцируя  $\psi(t)$  получаем

$$\nabla \psi_x(t) = h,$$

$$\nabla \psi_y(t) = \nabla \varphi(x^* + th)h = -[\nabla_y f(x^* + th, \varphi(x^* + th))]^{-1} \nabla_x f(x^* + th, \varphi(x^* + th))h.$$

## Неявно заданные функции

Следовательно нужно взять  $h = x$ . С другой стороны в силу

$$\nabla f(x^*, y^*) \begin{bmatrix} x \\ y \end{bmatrix} = 0_m \text{ получаем}$$

$$y = [\nabla_y f(x^*, y^*)]^{-1} \nabla_x f(x^*, y^*) x,$$

что совпадает с  $\nabla \psi_y(0)$ .

С другой стороны, если такая кривая существует, то дифференцируя по  $t$  равенство  $f(\psi(t)) = 0_m$  получаем

$$0_m = \left. \frac{d}{dt} f(\psi(t)) \right|_{t=0} = \nabla f(\psi(0)) \nabla \psi(0) = \nabla f(x^*, y^*) \begin{bmatrix} x \\ y \end{bmatrix}.$$