ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

А.В.СТЕПАНОВ

Аннотация. Мы будем рассматривать в основном функции на *конечномерных* линейных пространствах. Однако, бо́льшая часть формулировок (но не доказательств!) будет верна и для бесконечномерных пространств с необходимыми оговорками. Это одна из причин, почему я стараюсь давать формулировки в *инвариантной* форме (т.е. не в координатной форме – в форме, независимой от выбора базиса линейного пространсва). Дисциплины, которые занимаются случаем бесконечномерных пространств называются функциональным анализом и вариационным исчислением.

Содержание

1. Нормированные линейные пространства.

Определение 1.1. Пусть V – линейное пространство. Функция $\nu : V \to \mathbb{R}$ называется **нормой** на V, если она удовлетворяет следующим условиям. Если $u, v \in V$, $\lambda \in \mathbb{R}$, то:

(1) $\nu(\lambda u) = |\lambda|\nu(u)$ (линейность);
(2) $\nu(u + v) \leqslant \nu(u) + \nu(v)$ (неравенство треугольника);
(3) $\nu(u) \geqslant 0$ (положительная определенность);
(4) $\nu(u) = 0 \iff u = 0$ (невырожденность).

Вместо $\nu(u)$ мы будем обычно писать $\|u\|$.

Примеры 1.2. 1. Норма в $\mathbb{R}^n$, заданная равенством $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$, называется **стандартной** нормой. Она индуцирована стандартным скалярным произведением. Формулы $\|x\| = \max(|x_1|, \ldots, |x_n|)$ и $\|x\| = |x_1| + \cdots + |x_n|$ также задают нормы в $\mathbb{R}^n$.

2. Если V – пространство непрерывных функций $[a, b] \to \mathbb{R}$, то можно задать норму формулой $\|f\| = \sup\limits_{x\in[a,b]} |f(x)|$.

3. Пусть U, V – конечномерные нормированные линейные пространства, а $\mathcal{L}(U, V)$ – линейное пространство всех линейных отображений $U \to V$. Тогда

$$\|L\| = \sup_{x\neq 0} \frac{\|L(x)\|}{\|x\|} = \sup_{\|x\|=1} \|L(x)\|.$$

Вскоре (предложение 1.6) мы докажем, что супремум в этом определении всегда конечен в случае конечномерных пространств U и V.

Определение 1.3. Пусть μ и ν – нормы на линейном пространстве U. Они называются эквивалентными, если существуют такие положительные константы C_1 и C_2, что для любого $u \in U$ выполнено неравенство $C_1\mu(x) \leqslant \nu(x) \leqslant C_2\mu(x)$.

Теорема 1.4. *Все нормы на данном конечномерном линейном пространстве эквивалентны между собой.*

Без доказательства.

Определение 1.5. Пусть U, V – нормированные линейные пространства (возможно бесконечномерные), а L – линейное отображение $U \to V$. L называется **ограниченным**, если существует константа C такая, что $\|L(x)\| \leqslant C\|x\|$ для любого $x \in U$, или, что то же самое, $\|L(x)\| \leqslant C$ для любого $x \in U$ с $\|x\| = 1$. Множество всех ограниченных линейных отображений из U в V обозначается через $\mathcal{L}(U, V)$.

Предложение 1.6. *Пусть L – линейное отбражение конечномерных нормированных линейных пространств $U \to V$. Тогда L ограничен. Таким образом, если U и V конечномерны, то $\mathcal{L}(U, V)$ является множеством всех линейных отображений $U \to V$.*

Доказательство. Пусть $u_1, \ldots, u_m$ – базис U, а $v_1, \ldots, v_n$ – базис v. Так как все нормы в конечномерном пространстве эквивалентны, то можно считать, что $\|x\| = \|x_u\|$ и $\|y\| = \|y_v\|$, где $x \in U$, $y \in V$, а нормы в пространствах $\mathbb{R}^m$ и $\mathbb{R}^n$ – стандартные. Пусть $A = L_{u,v}$ – матрица отображения L. Обозначим через $a^{(1)}, \ldots, a^{(m)}$ столбцы матрицы A. Тогда, если $x \in U$ – вектор с координатами $x_1, \ldots, x_m$ с нормой 1, то $|x_i| \leqslant 1$, и значит

$$\|L(x)\| = \|Ax_u\| = \|\sum_{i=1}^{m} a^{(i)}x_i\| \leqslant \sum_{i=1}^{m} \|a^{(i)}\| \cdot |x_i| \leqslant \sum_{i=1}^{m} \|a^{(i)}\| = M$$

Так как последнее выражение не зависит от x, то оно и есть верхняя граница (не обязательно наименьшая) для интересующего нас множества. □

1.7. В дальнейшем нам понадобяться простейшие сведения о сопряженном пространстве. Пусть U – нормированное линейное пространство. **Линейным функционалом** на U называется ограниченный линейный оператор из U в $\mathbb{R}$. Множество всех линейных функционалов обозначается через V^* (оно же $\mathcal{L}(U, \mathbb{R})$ – см.1.5) и называется **пространством, сопряженным с** U. В случае $U = \mathbb{R}^m$, U^* естественно отождествляется с множеством строк длины m. Следовательно, для любого m-мерного пространства U пространство U^* также имеет размерносlargeь m. Поэтому, U^* изоморфно U для любого конечномерного векторного пространства U. Более того, если на U зафиксировано скалярное произведение, то можно выбрать канонический изоморфизм, а именно, вектору $v \in V$ сопоставляется функционал L_v, действующий по правилу $L_v(u) = (v, u)$. В случае стандартного скалярного произведения на $U = \mathbb{R}^m$ канонический изоморфизм – это просто транспонирование (проверьте).

2. Предел функции нескольких переменных.

Определение 2.1. Пусть U, V – нормированные линейные пространства, $f : U \to V$, и $u \in U$. $\lim\limits_{x \to u} f(x) = v \in V$, если $\lim\limits_{\|x-u\| \to 0} \|f(x) - v\| = 0$. Другими словами, $\lim\limits_{x \to u} f(x) = v$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\|x - u\| < \delta$ выполнено $\|f(x) - v\| < \varepsilon$.

Сравним понятие пределов для функций $\mathbb{R} \to \mathbb{R}$ и $\mathbb{R}^2 \to \mathbb{R}$.

Предложение 2.2. *Пусть $f : \mathbb{R}^2 \to \mathbb{R}$, и существует $\lim\limits_{x \to u} f(x)$ равный v.*
Тогда $\lim\limits_{x_1 \to u_1} \big(\lim\limits_{x_2 \to u_2} f(x_1, x_2)\big) = v$

Замечание 2.3. Обратное неверно! Например, $f(x, y) = \frac{2xy}{x^2+y^2}$. Тогда повторный предел в точке $(0, 0)$ существует и равен 0, в то время как на прямой $x = y$ функция тождественно равна 1, и, следовательно, $\lim\limits_{(x,y) \to (0,0)} f(x, y)$ не существует.

Теорема 2.4. *Пусть функция $f : \mathbb{R}^n \to V$ непрерывна в некоторой проколотой окрестности точки u. Тогда предел $f(x)$ при $x \to u$ существует и равен v тогда и только тогда, когда $\lim\limits_{t\to 0} f(u + tw) = v$ при любом $w \in U$.*

Без доказательства.

Определение предела функции $f : U \to V$ apriori зависит от норм, выбранных в пространствах U и V. Однако сейчас мы докажем, что если нормы эквивалентны, то пределы по этим нормам совпадают. Так как в дальнейшем мы будем рассматривать в основном конечномерные пространства, в которых все нормы эквивалентны между собой (теорема 1.4), то понятия предела, непрерывности, дифференцируемости и т.п. не будут зависеть от выбранных нами норм. Пусть $\lim^{\nu,\tau}$ обозначает предел по нормам ν и τ, т.е. предел, в определении которого задействована норма ν на пространстве U и норма τ на пространстве V.

Предложение 2.5. *Пусть μ и ν – эквивалентные нормы на пространстве U, σ и τ – эквивалентные нормы на пространстве V, $u \in U$, а f – функция $U \to V$. Тогда $\lim^{\mu,\sigma} f(x) = \lim^{\nu,\tau} f(x)$ при $x \to u$, или оба этих предела не существуют.*

Доказательство. Пусть $\lim^{\mu,\sigma} f(x) = v$. По определению 1.3 существуют $C_1, C_2 > 0$ такие, что $C_1\mu(x) \leqslant \nu(x)$ и $C_2\tau(y) \leqslant \sigma(y)$ для любых $x \in U$ и $y \in V$. Возьмем $\varepsilon > 0$. По определению предела существует $\delta > 0$ такое, что $\sigma(f(x)-v) < C_2\varepsilon$ при $\mu(x-v) < \delta$. Если $\nu(x-v) < C_1\delta$, то $C_1\mu(x-v) \leqslant \nu(x-v) < C_1\delta$, следовательно, $\mu(x-v) < \delta$. Отсюда $C_2\tau(f(x)-v) \leqslant \sigma(f(x)-v) < C_2\varepsilon$, т.е. $\tau(f(x)-v) < \varepsilon$. Это означает, что $\lim^{\nu,\tau} f(x) = v$ при $x \to u$. Таким образом, если один из пределов существует, то второй тоже существует и равен первому. $\square$

Определение непрерывности, свойства пределов и непрерывных функций аналогичны тому, что известно для функций $\mathbb{R} \to \mathbb{R}$ (см. также приложение).

Предложение 2.6. *Линейное отображение $L : U \to V$ непрерывно тогда и только тогда, когда оно ограничено.*

Доказательство. Если L ограничен, то $0 \leqslant \|L(x) - L(u)\| = \|L(x-u)\| \leqslant \|L\| \cdot \|x-u\| \to 0$ при $x \to u$, что и означает, что L непрерывен.

Обратно, если L непрерывен, то существует $\delta > 0$ такое, что $\|L(u)\| < 1$ при $\|u\| < \delta$. Пусть $\|x\| = 1$. Положим $u = \frac{\delta}{2}x$ так, что $\|u\| < \delta$. Имеем: $\|L(x)\| = \frac{2}{\delta}\|L(u)\| < \frac{2}{\delta}$, откуда $\sup\limits_{\|x\|=1} \|L(x)\| \leqslant \frac{2}{\delta}$, т.е. L ограничен. $\square$

3. Определение и свойства дифференциала и производной.

3.1. Пусть U и V – конечномерные линейные пространства. Из курса линейной алгебры известно, что любое линейное отображение в фиксированных базисах пространств U и V является оператором умножения на матрицу. Сопоставление линейному отображению его матрицы в выбранных базисах задает изоморфизм $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathrm{M}(m \times n, \mathbb{R})$. В частности, $\mathcal{L}(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$.

Пусть $g : U \to V$ – отображение нормированных линейных пространств. Предположим, что $g(x) \to 0$ при $x \to u$. Напомним, что $o\big(g(x)\big)$ (при $x \to u$) обозначает *любую* функцию, обладающую свойством $\frac{o\left(g(x)\right)}{\|g(x)\|} \to 0$ при $x \to u$. Во избежание недоразумений, будем писать $o_1\big(g(x)\big)$, $o_2\big(g(x)\big), \dots$, если в формуле задействованы несколько функций, обладающих указанным свойством.

Определение 3.2. Пусть U, V – нормированные линейные пространства, $u \in U$, и $f : U \to V$. Предположим, что существует ограниченное линейное отображение $L : U \to V$ такое, что

$$f(x) = f(u) + L(x-u) + o(x-u)$$

Тогда говорят, что функция f дифференцируема в точке u. При этом L называется **(полным) дифференциалом** функции f в точке u и обозначается через $d_u f$.

Предложение 3.3. *Если f дифференцируема в точке U, то дифференциал $d_u f$ определяется единственным образом.*

Доказательство. Пусть L_1, L_2 – различные линейные отображения, удовлетворяющие условию последнего определения, и $L = L_1 - L_2$. Тогда $\lim\limits_{x\to u} L\left(\frac{x-u}{\|x-u\|}\right) = 0$. Другими словами, $\forall\varepsilon > 0\ \exists\delta > 0 : \|x-u\| < \delta \Rightarrow \|L\left(\frac{x-u}{\|x-u\|}\right)\| < \varepsilon$. Предположим, что $L(y) \neq 0$ для некоторого $y \in U$. Положим $\varepsilon = \frac{\|L(y)\|}{2\|y\|}$ и выберем δ по этому ε. Положим $x = u + \frac{\delta}{2\|y\|}y$. Тогда $\|x-u\| = \frac{\delta}{2} < \delta$, но $\|L\left(\frac{x-u}{\|x-u\|}\right)\| = \frac{\|L(y)\|}{\|y\|} = 2\varepsilon > \varepsilon$, что противоречит выбору δ. □

3.4. В случае $U = V = \mathbb{R}$ любое линейное отображение (например, $d_u f$) – это умножение на число. Это число и называется **производной функции** f **в точке** u. Оно обозначается через $f'(u)$. Связь дифференциала функции и ее производной может быть выражена формулой $d_u f(\Delta x) = f'(u)\cdot\Delta x$. В физике, технике и во многих математических учебниках символ $d_x f$ или даже просто df по умолчанию означает $d_x f(dx)$, где dx – “приращение аргумента”. Отсюда формула $df = f'(x)dx$.

Предложение 3.5. *Если $f : \mathbb{R} \to \mathbb{R}$, то $f'(u) = \lim\limits_{x\to u} \frac{f(x)-f(u)}{x-u}$.*

3.6. Рассмотрим теперь случай функции $f : U = \mathbb{R}^m \to V = \mathbb{R}^n$. **Полной производной** такой функции в точке $u \in \mathbb{R}^m$ называется матрица отображения $d_u f$ в стандартных базисах пространств $\mathbb{R}^m$ и $\mathbb{R}^n$.

3.7. Если $f : U = \mathbb{R}^m \to V = \mathbb{R}^n$ и $x \in U$, то значение $f(x)$ – это столбец $(f_1(x), \dots, f_m(x))^T$. В этом случае функции $f_k : \mathbb{R}^m \to \mathbb{R}$ называются **координатными функциями** функции f. Производная k-ой координатной функции в точке $u \in \mathbb{R}^m$ – это строка, которая, как легко видеть, совпадает с k-ой строкой матрицы $f'(u)$. Следующее определение говорит о том, из чего обычно состоит эта строка (см. предложение 3.9 ниже).

Определение 3.8. Пусть $f : \mathbb{R}^m \to \mathbb{R}$ – скалярная функция n переменных. **Частной производной** по i-ой переменной в точке $u = (u_1, \dots, u_m)^T$ функции f (обозначение: $D_i f(u)$ или $\frac{\partial f}{\partial x_i}(u)$) называется производная функции $g : \mathbb{R} \to \mathbb{R}$ в точке u_i, где g задана равенством $g(x) = f(u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_m)$. Другими словами,

$$D_i f(u) = \frac{\partial f}{\partial x_i}(u) = \lim_{\Delta x\to 0} \frac{f(u_1, \dots, u_{i-1}, u_i + \Delta x, u_{x+1}, \dots, u_m) - f(u)}{\Delta x}$$

Предложение 3.9. *Пусть $f = (f_1, \dots, f_n)^T : \mathbb{R}^m \to \mathbb{R}^n$, $u \in \mathbb{R}^m$. Предположим, что все частные производные координатных функций $D_i f_j(u)$ существуют и непрерывны в некоторой окрестности точки u. Тогда матрица $f'(u)$ состоит из частных производных $D_i f_j(u)$ (т.е. $f'(u)_{ij} = D_i f_j(u)$).*

Доказательство. Как было отмечено в пункте 3.7, полная производная координатной функции f_k является k-ой строкой матрицы f'. Поэтому достаточно провести доказательство для $n = 1$.

Нам надо доказать, что

$$(\star) \qquad \lim_{x\to u} \frac{\left|f(x) - f(u) - \sum\limits_{k=1}^{m} \frac{\partial f}{\partial x_k}(u)(x_k - u_k)\right|}{\|x-u\|} = 0$$

Заметим, что $|x_k - u_k| \leqslant \|x-u\|$, и, следовательно, $x_k \to u_k$ при $x \to u$. Преобразуем левую часть формулы $(\star)$. Она равна

$$\lim_{x\to u} \sum_{k=1}^{m} \frac{\left|f(u_1, \dots, u_{k-1}, x_k, \dots, x_m) - f(u_1, \dots, u_k, x_{k+1}, \dots, x_m) - \frac{\partial f}{\partial x_k}(u)(x_k - u_k)\right|}{\|x-u\|} \leqslant$$

$$\leqslant \lim_{x\to u} \sum_{k=1}^{m} \frac{\left| f(u_1,\ldots,u_{k-1},x_k,\ldots,x_m) - f(u_1,\ldots,u_k,x_{k+1},\ldots,x_m) - \frac{\partial f}{\partial x_k}(u)(x_k-u_k)\right|}{|x_k-u_k|} \leqslant$$
$$\leqslant \lim_{x\to u} \sum_{k=1}^{m} \left| \frac{\partial f}{\partial x_k}(u_1,\ldots,u_k,x_{k+1},\ldots,x_m) - \frac{\partial f}{\partial x_k}(u)\right| = 0\,,$$

что и требовалось. □

Свойства дифференциала и производной.

3.10. *Непрерывность.* Если функция $f : U \to V$ дифференцируема в точке u, то она непрерывна в этой точке.

Доказательство. По определению дифференциала $f(x)-f(u) = d_u f(x-u) + o(x-u)$. Поэтому

$$0 \leqslant \|f(x)-f(u)\| = \big(\|d_u f(x-u) + o(x-u)\|\big) \leqslant \|d_u f\|\cdot\|x-u\| + \|o(x-u)\| \xrightarrow[x\to u]{} 0,$$

откуда $f(x) \to f(u)$ при $x \to u$. □

3.11. *Линейность.* $d(f+g) = df + dg$, $d(\lambda f) = \lambda df$. Аналогичные свойства выполнены для производных.

3.12. *Дифференциал постоянной.* Пусть $u \in U$, $v \in V$, а $f : U \to V$ задана формулой $f(x) = v$. Тогда $d_u f = 0$.

3.13. *Дифференциал линейного отображения.* Если $L : U \to V$ – ограниченное линейное отображение, то $d_u L = L$ в любой точке $u \in U$.

3.14. *Дифференциал композиции.* Пусть $g : U \to V$, $f : V \to W$, и $u \in U$. Тогда $d_u(f\circ g) = d_{g(u)} f \circ d_u g$. Если $U = \mathbb{R}^m$, $V = \mathbb{R}^n$, а $W = \mathbb{R}^k$, то последнее равенство можно переписать в виде $(f\circ g)'(u) = f'\big(g(u)\big)\cdot g'(u)$ (здесь в левой части стоит матрица размера $k\times m$, а в правой – произведение матриц размеров $k \times n$ и $n\times m$).

При $k = 1$ в координатной форме это равенство будет выглядеть так:

$$\frac{\partial (f\circ g)}{\partial x_i}(u) = \sum_{j=1}^{n} \frac{\partial f}{\partial y_j}(g(u)) \cdot \frac{\partial g_j}{\partial x_i}(u)$$

для любого $i = 1,\ldots,m$ (здесь $x = (x_1,\ldots,x_m)^T$ – столбец переменных в пространстве $U = \mathbb{R}^m$, а $y = (y_1,\ldots,y_n)^T$ – в пространстве $V = \mathbb{R}^n$.

Доказательство. Обозначим через A линейное отображение $d_{g(u)} f : V \to W$, а через B – отображение $d_u g : U \to V$. Обозначим через Δx приращение аргумента, а через $\Delta g = g(u + \Delta x) - g(u)$ – приращение функции g. По определению дифференциала $\Delta g = B(\Delta x) + o_2(x)$, а по свойству 3.10 $\Delta g \to 0$ при $\Delta x \to 0$. Рассмотрим приращение функции $f \circ g$.

$$f\big(g(u+\Delta x)\big) - f\big(g(u)\big) = f\big(g(u) + \Delta g\big) - f\big(g(u)\big) = A(\Delta g) + o_1(\Delta g) =$$
$$= A\big(B(\Delta x) + o_2(\Delta x)\big) + o_1(\Delta g) = A\big(B(\Delta x)\big) + A\big(o_2(\Delta x)\big) + o_1(\Delta g)$$

Далее,

$$\frac{\|A\big(o_2(\Delta x)\big)\|}{\|\Delta x\|} \leqslant \|A\|\cdot\frac{\|o_2(\Delta x)\|}{\|\Delta x\|} \to 0 \qquad \text{и}$$
$$\frac{\|o_1(\Delta g)\|}{\|\Delta x\|} = \frac{\|o_1(\Delta g)\|}{\|\Delta g\|}\cdot\frac{\|B(\Delta x) + o_2(\Delta x)\|}{\|\Delta x\|} \leqslant \frac{\|o_1(\Delta g)\|}{\|\Delta g\|}\left(\|B\| + \frac{\|o_2(\Delta x)\|}{\|\Delta x\|}\right) \to 0$$

Следовательно, $A\big(o_2(\Delta x)\big) + o_1(\Delta g) = o(\Delta x)$, откуда $f\big(g(u+\Delta x)\big) = f\big(g(u)\big) + A\circ B(\Delta x) + o(\Delta x)$, а это и означает, что $d_u(f\circ g) = A\circ B$. □

3.15. *Дифференциал произведения.* Пусть $f, g : U \to \mathbb{R}$, и $u, \Delta x \in U$. Тогда $d_u(fg)(\Delta x) = d_u f(\Delta x)\cdot g(u) + f(u)\cdot d_u g(\Delta x)$. Если $U = \mathbb{R}^m$, то $(fg)' = gf' + fg'$.

Доказательство. Представим функцию fg в виде композиции

$$U \xrightarrow{\ \delta\ } U\times U \xrightarrow{f\times g} \mathbb{R}\times\mathbb{R} \xrightarrow{\ M\ } \mathbb{R},$$

где $\delta(u) = (u,u)$, $(f\times g)(u,v) = \big(f(u), g(v)\big)$, а $M(x,y) = xy$. Функция δ очевидно является ограниченным линейным отображением. Легко показать, что $d(f\times g) = df\times dg$, а $d_{(x,y)}M(\Delta x, \Delta y) = x\Delta y + y\Delta x$. Теперь доказываемое утверждение следует из свойств 3.13 и 3.14. $\square$

Аналогичным образом можно получить формулы дифференцирования частного, степени (т.е. функции f^g) и т.п.

4. Теоремы об обратной и неявной функции.

Теорема 4.1 (об обратной функции)**.** *Пусть U и V – конечномерные нормированные линейные пространства, функция $f : U \to V$ непрерывно дифференцируема в некоторой окрестности точки $u \in U$. Если $d_u f$ является обратимым линейным отображением, то существуют окрестности Ω_u точки u и $\Omega_{f(u)}$ точки $f(u)$ и непрерывно дифференцируемая функция $g : \Omega_{f(u)} \to \Omega_u$, являющаяся обратной к сужению f на Ω_u. Кроме того, $d_u f = (d_{f(u)}g)^{-1}$.*

Доказательство. Заметим, что, если существует обратимое линейное отображение $U \to V$, то размерности пространств U и V равны. Как обычно, выбирая базисы в пространствах U и V, можно считать, что $U = V = \mathbb{R}^m$, а по теореме 1.4 можно считать, что на этих пространствах заданы стандартные нормы.

Докажем сначала существование окрестности Ω_u, на которой f инъективна. Предположим противное. Тогда в каждой окрестности $\Omega_u^{1/n} = \{x \in U \mid \|x-u\| < \frac{1}{n}\}$ найдутся точки $x_n \neq y_n$ такие, что $f(x_n) = f(y_n)$. Для каждой пары n, k рассмотрим функцию $g : \mathbb{R} \to \mathbb{R}$, заданную равенством $g(t) = f_k\big(x_n + t(y_n - x_n)\big)$, где f_k – координатная функция. Так как $g(0) = g(1)$, то по теореме Ролля существует точка $\theta \in (0,1)$ такая, что $g'(\theta) = 0$. Заметим, что $c_{n,k} = x_n + \theta(y_n - x_n)$ лежит в окрестности $\Omega_u^{1/n}$. Далее, для каждого k имеем $0 = g'(\theta) = f_k'(c_{n,k})\cdot(y_n - x_n)$. Составим матрицу A_n из строк $f_k'(c_{n,k})$. Так как $A_n(y_n - x_n) = 0$, а $y_n - x_n \neq 0$, то $\det A_n = 0$. По построению, $\lim\limits_{n\to\infty} c_{n,k} = u$, а так как частные производные функции f непрерывны, то $\lim\limits_{n\to\infty} A_n = f'(u)$. Так как определитель – непрерывная функция элементов матрицы, то $\det f'(u) = \lim\limits_{n\to\infty} \det A_n = 0$, что противоречит условию.

Таким образом, при некотором n функция f инъективна на $\Omega_u = \Omega_u^{1/n}$. По теореме 10.19 (см. приложение) ее образ $\Omega_{f(u)} = f(\Omega_u)$ открыт, а обратная функция $g : \Omega_{f(u)} \to \Omega_u$ непрерывна.

Пусть теперь $\Delta x \in \Omega_u$, а $\Delta y = g(u+\Delta x) - g(u)$. Заметим, что $\Delta y \to 0$ при $\Delta x \to 0$ по непрерывности функции g. Обозначим $f'\big(g(u)\big)$ через A. Так как A обратима, то $\|x\| = \|A^{-1}Ax\| \leqslant \|A^{-1}\|\cdot\|Ax\|$, откуда $\frac{\|x\|}{\|A^{-1}\|} \leqslant \|Ax\| \leqslant \|A\|\cdot\|x\|$ для любого $x \in U$ (здесь норма матрицы – это норма оператора умножения на эту матрицу). Далее,

$$u + \Delta x = f\circ g(u+\Delta x) = f\big(g(u) + \Delta y\big) = f\big(g(u)\big) + A\Delta y + o(\Delta y)$$

откуда $\Delta x = A\Delta y + o(\Delta y)$, а $\Delta y = A^{-1}\Delta x + A^{-1}o(\Delta y)$. Сейчас мы покажем, что $A^{-1}o(\Delta y)$ есть $o(\Delta x)$. Как только это будет сделано, подставляя два последних равенства в определение Δy, мы получим

$$g(u+\Delta x) - g(u) = A^{-1}\Delta x + o(\Delta x),$$

что и будет означать, что $g'(u)$ существует и равно $f'\big(g(u)\big)^{-1}$.

Итак, осталось доказать, что

$$\tfrac{\|A^{-1}o(\Delta y)\|}{\|\Delta x\|} = \tfrac{\|A^{-1}o(\Delta y)\|}{\|A\Delta y + o(\Delta y)\|} \xrightarrow[\Delta x\to 0]{} 0.$$

Числитель последней дроби не превосходит $\|A^{-1}\|\|o(\Delta y)\|$, а знаменатель не меньше, чем $\big|\|A\Delta y\| - \|o(\Delta y)\|\big| \geqslant \Big|\frac{\|\Delta y\|}{\|A^{-1}\|} - \|o(\Delta y)\|\Big|$. Таким образом,

$$\frac{\|A^{-1}o(\Delta y)\|}{\|\Delta x\|} \leqslant \frac{\|A^{-1}\|\|o(\Delta y)\|}{\Big|\frac{\|\Delta y\|}{\|A^{-1}\|} - \|o(\Delta y)\|\Big|} = \frac{\|A^{-1}\|\frac{\|o(\Delta y)\|}{\|\Delta y\|}}{\Big|\frac{1}{\|A^{-1}\|} - \frac{\|o(\Delta y)\|}{\|\Delta y\|}\Big|} \xrightarrow[\Delta y \to 0]{} 0,$$

что завершает доказательство. □

Теорема 4.2 (о неявной функции)**.** *Пусть $a \in \mathbb{R}^m$, $n < m$, функция $f : \mathbb{R}^m \to \mathbb{R}^n$ непрерывно дифференцируема в некоторой окрестности точки a, $f(a) = 0$ и* $\begin{vmatrix} D_1 f_1(a) & \dots & D_n f_1(a) \\ \dots & \dots & \dots \\ D_1 f_n(a) & \dots & D_n f_n(a) \end{vmatrix} \neq 0.$ *Тогда существует окрестность $V_0 \subset \mathbb{R}^{m-n}$ точки $(a_{n+1}, \dots, a_m)^T$ и функция $g : V_0 \to \mathbb{R}^n$ такая, что при $y \in V_0$ уравнение $f(x, y) = 0$ равносильно уравнению $x = g(y)$ (здесь $x = (x_1, \dots, x_n)^T$, а $y = (y_1, \dots, y_{m-n})^T$).*

Более того, функция g дифференцируема, и $g'(u) = -f'_x\big(g(u), u\big)^{-1} f'_y\big(g(u), u\big)$, (здесь f'_x – матрица, составленная из первых n столбцов матрицы f', f'_y – матрица, составленная из последних $m - n$ столбцов матрицы f', а $u \in V_0$).

Доказательство. Рассмотрим функцию $w : \mathbb{R}^m \to \mathbb{R}^m$, заданную равенствами

$$w_k(x_1, \dots, x_n, y_1, \dots, y_{m-n}) = \begin{cases} f_k(x_1, \dots, x_n, y_1, \dots, y_{m-n}), & \text{при } k \leqslant n \\ y_{k-n}, & \text{при } k > n \end{cases}$$

Легко видеть, что

$$\det\big(w'(a)\big) = \begin{vmatrix} f'_x(a) & * \\ 0 & E \end{vmatrix} = \det\big(f'_x(a)\big) \neq 0.$$

По теореме об обратной функции существует окрестность U точки a, окрестность V точки $w(a) = (0, \dots, 0, a_{n+1}, \dots, a_m)^T$ и функция $h : V \to U$ обратная для w. Другими словами, $w(x, y) = \begin{pmatrix} z \\ t \end{pmatrix} \iff h(z, t) = \begin{pmatrix} x \\ y \end{pmatrix}$ при $\begin{pmatrix} x \\ y \end{pmatrix} \in U$, $\begin{pmatrix} z \\ t \end{pmatrix} \in V$ (здесь $z \in \mathbb{R}^n$, а $t \in \mathbb{R}^{m-n}$). Из определения функции w следует тогда, что $t = y$, и что $f(x, y) = z \iff x = h(z, y)$ при $\begin{pmatrix} x \\ y \end{pmatrix} \in U$, $\begin{pmatrix} z \\ t \end{pmatrix} \in V$. Пусть $V_0 = \{y \in \mathbb{R}^{m-n} \,|\, (0, y) \in V\}$, а $g(y) = h(0, y)$. Очевидно, что V_0 – окрестность точки $(a_{n+1}, \dots, a_m)^T$, а g – непрерывная функция, удовлетворяющие условию теоремы.

Функция g дифференцируема, так как по теореме об обратной функции h дифференцируема. По определению g при всех $y \in V_0$ выполнено равенство $f\big(g(y), y\big) = 0$. Дифференцируя это равенство в точке $u \in V_0$ получим $f'\big(g(u), u\big) \begin{pmatrix} g'(u) \\ E \end{pmatrix} = 0$, откуда $f'_x\big(g(u), u\big) g'(u) + f'_y\big(g(u), u\big) = 0$, т.е. $g'(u) = -f'_x\big(g(u), u\big)^{-1} f'_y\big(g(u), u\big)$. □

5. Дифференциалы высших порядков, формула Тэйлора.

5.1. В соответствии с предложением 3.3, для фиксированной функции $f : U \to V$ отображение $u \mapsto d_u f$ определяет функцию $df : U \to \mathcal{L}(U, V)$. Норма оператора (см. 1.6) превращает $\mathcal{L}(U, V)$ в нормированное линейное пространство. Если функция df является дифференцируемой, то мы получаем функцию $d(df) : U \to \mathcal{L}\big(U, \mathcal{L}(U, V)\big)$. Обозначим через $\mathcal{L}^2(U, V)$ множество всех билинейных отображений $U \times U \to V$. Используя изоморфизм

$$\varphi : \mathcal{L}\big(U, \mathcal{L}(U, V)\big) \xrightarrow{\ \sim\ } \mathcal{L}^2(U, V), \text{ заданный формулой } \varphi(f)(u, w) = f(u)(w),$$

получаем функцию $d^2 f : U \to \mathcal{L}^2(U, V)$, где $d^2 f = \varphi \circ d(df)$, причем $\mathcal{L}^2(U, V)$ снова является нормированным линейным пространством.

По индукции, множество $\mathcal{L}^k(U,V)$ всех полилинейных отображений $\underbrace{U\times\dots\times U}_{k}\to V$ является нормированным линейным пространством, и, если $d^{k-1}f$ является дифференцируемой функцией $U\to\mathcal{L}^{k-1}(U,V)$, то можно определить $d^kf:U\to\mathcal{L}^k(U,V)$, как композицию дифференциала $d(d^{k-1}f):U\to\mathcal{L}\big(U,\mathcal{L}^{k-1}(U,V)\big)$ с естественным изоморфизмом $\mathcal{L}\big(U,\mathcal{L}^{k-1}(U,V)\big)\cong\mathcal{L}^k(U,V)$. В этом случае говорят, что f является k раз дифференцируемой функцией. Если к тому же d^kf непрерывна, то говорят что f является k раз непрерывно дифференцируемой.

5.2. Предположим, что U и V – конечномерные пространства с базисами $u_1,\dots,u_m$ и $v_1,\dots,v_m$. Ясно, что любое полилинейное отображение $P\in\mathcal{L}^k(U,V)$ полностью определяется своими значениями $P(u_{i_1},\dots,u_{i_k})\in V$ на всевозможных наборах базисных векторов пространства U. Обозначим через $P^j_{i_1,\dots,i_k}$ координаты вектора $P(u_{i_1},\dots,u_{i_k})$ в базисе v. Тогда набор этих чисел полностью определяет отображение P:

$$P(x^{(1)},\dots,x^{(k)})=\sum_{j=1}^{n}\sum P^j_{i_1,\dots,i_k}x^{(1)}_{i_1}\cdot\dots\cdot x^{(k)}_{i_k}v_j,$$

где $(x^{(\ell)}_1,\dots,x^{(\ell)}_m)^T$ – столбец координат вектора $x^{(\ell)}$ в базисе u, а внутренняя сумма берется по всем наборам $(i_1,\dots,i_k)\in\{1,\dots,m\}^k$.

Другими словами, набор отображений $E^j_{i_1,\dots,i_k}\in\mathcal{L}^k(U,V)$, $j\in\{1,\dots,n\}$, $i_\ell\in\{1,\dots,m\}$, заданных равенствами

$$E^j_{i_1,\dots,i_k}(x^{(1)},\dots,x^{(k)})=x^{(1)}_{i_1}\cdot\dots\cdot x^{(k)}_{i_k}v_j,$$

является базисом в $\mathcal{L}^k(U,V)$. При этом числа $P^j_{i_1,\dots,i_k}$ являются координатами отображения P в этом базисе. В случае, когда u и v – стандартные базисы пространств $\mathbb{R}^m$ и $\mathbb{R}^n$, набор E отображений $E^j_{i_1,\dots,i_k}$ называется стандартным базисом пространства $\mathcal{L}^k(\mathbb{R}^m,\mathbb{R}^n)$.

5.3. Пусть теперь $U=\mathbb{R}^m$, $V=\mathbb{R}$, в этих пространствах выбраны стандартные базисы, а $f:\mathbb{R}^m\to\mathbb{R}$ является k раз непрерывно дифференцируемой. Определим k-ую частную производную $D^k_{i_1,\dots,i_k}f=\frac{\partial^kf}{\partial x_{i_1}\dots\partial x i_k}$, соответствующую набору $(i_1,\dots,i_k)\in\{1,\dots,m\}^k$ индукцией по k. При $k=1$ положим $D^1_if=D_if$. При $k>1$: $D^k_{i_1,\dots,i_k}f=D_{i_k}\big(D^{k-1}_{i_1,\dots,i_{k-1}}f\big)$. При $k=0$ будем считать, что $D^0f=f$.

Теперь из теоремы 3.9 нетрудно вывести, что значения частных производных $D^k_{i_1,\dots,i_k}f(a)$ являются координатами полилинейного отображения $d^k_af\in\mathcal{L}^k(\mathbb{R}^m,\mathbb{R})$ в стандартном базисе E (здесь $a\in\mathbb{R}^m$). В случае $V=\mathbb{R}^n$ все проведенные рассуждения можно применить к координатным функциям функции f.

5.4. Рассмотрим теперь более внимательно второй дифференциал скалярной функции f от m переменных. В этом случае d^2_af является билинейной формой на $\mathbb{R}^m$. Как мы только что показали, матрица этой билинейной формы состоит из вторых частных производных $D^2_{ij}f(a)=\frac{\partial^2f}{\partial x_i\partial x_j}(a)$, где $i,j=1,\dots,m$. Эта матрица называется полной второй производной функции f в точке a и обозначаетя через $f''(a)$. В соответствии с определением матрицы билинейной формы мы получаем

$$d^2_af(y,z)=\sum_{i,j=1}^{m}\frac{\partial^2f}{\partial x_i\partial x_j}(a)y_jz_i=z^Tf''(u)y.$$

Теорема 5.5. *Пусть $a,y,z\in U$, и $f:U\to V$ дважды непрерывно дифференцируема в окрестности точки a. Тогда $d^2_af(y,z)=d^2_af(z,y)$. В частности, вторые частные производные не зависят от порядка дифференцирования, т.е. $\frac{\partial^2f}{\partial x_i\partial x_j}=\frac{\partial^2f}{\partial x_j\partial x_i}$.*

Доказательство этой теоремы будет дано после формулы Тэйлора.

Определение 5.6. Подмножество Ω линейного пространства U называется **выпуклым**, если для любых $a,x\in\Omega$ отрезок $[a,x]=\{a+t(x-a)\,|\,t\in[0,1]\,\}$ целиком содержится в Ω.

Теорема 5.7 (Формула Тэйлора с остаточным членом в форме Лагранжа). *Пусть $f : U \to \mathbb{R}$ дифференцируема $n+1$ раз в выпуклой окрестности Ω точки $a \in U$, и $x \in \Omega$. Тогда существует $\theta \in \Omega$ (зависящее от a, x, n) такое, что*

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} d_a^k f(x-a, \dots, x-a) + \frac{1}{(n+1)!} d_\theta^{(n+1)} f(x-a, \dots, x-a).$$

Доказательство. Обозначим $\Delta x = x - a$, и рассмотрим вспомогательную функцию $g(t) = f(a + \Delta x t) : \mathbb{R} \to \mathbb{R}$. Заметим, что по правилу дифференцирования сложной функции

$$g'(t) = d_{a+\Delta x t} f(\Delta x), \dots, g^{(k)}(t) = d^k_{a+\Delta x t} f(\Delta x, \dots, \Delta x), \dots$$

Теперь утверждение теоремы следует из формулы Тэйлора для функции g в окрестности нуля. □

Теорема 5.8 (Формула Тэйлора с остаточным членом в форме Пеано). *Пусть U, V – конечномерные пространства, а $f : U \to \mathbb{R}^m$ является n раз непрерывно дифференцируемой в окрестности Ω точки $a \in U$, и $x \in \Omega$. Тогда*

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} d^k f(a)(x-a, \dots, x-a) + o(\|x-a\|^n)$$

Доказательство. Мы докажем это утверждение только для $(n+1)$ раз дифференцируемых функций. Так как нас интересует только поведение функции в некоторой, сколь угодно малой, окрестности точки a, то можно считать, что Ω – выпуклая окрестность, например, открытый шар. Как всегда можно считать, что $V = \mathbb{R}^m$ со стандартной нормой. Если $m = 1$, то учитывая предыдущую теорему нам достаточно доказать, что

$$\frac{\frac{1}{(n+1)!} d_\theta^{(n+1)} f(x-a, \dots, x-a)}{\|x-a\|^n} \xrightarrow[x \to a]{} 0$$

Но, по определению нормы оператора, это выражение не превосходит $\frac{1}{(n+1)!} \|d_\theta^{(n+1)} f\| \cdot \|x - a\|$, что очевидно является бесконечно малой.

Утверждение в общем случае следует теперь из утверждения для координатных функций и из того, что норма вектора не превосходит суммы модулей его координат. □

теоремы 5.5. Зафиксируем $x, y \in U$, и пусть $\varepsilon, \delta \in \mathbb{R}$. Рассмотрим функцию $g : \mathbb{R}^2 \to V$:

$$g(\varepsilon, \delta) = \big(f(a + \varepsilon x + \delta y) - f(a + \varepsilon x)\big) - \big(f(a + \delta y) - f(a)\big)$$

По формуле Тэйлора для функций f и df в окрестностях точек a и $a + \varepsilon x$, при $\varepsilon, \delta \to 0$ имеет место равенство

$$\begin{aligned} g(\varepsilon, \delta) &= d_{a+\varepsilon x} f(\delta y) - d_a f(\delta y) + \tfrac{1}{2} d^2_{a+\varepsilon x} f(\delta y, \delta y) - \tfrac{1}{2} d^2_a f(\delta y, \delta y) + o_1(\|\delta y\|^2) - o_2(\|\delta y\|^2) \\ &= \delta\big(d_{a+\varepsilon x} f - d_a f\big)(y) + \frac{\delta^2}{2}\big(d^2_{a+\varepsilon x} f - d^2_a f\big)(y, y) + o_3(\delta^2) \\ &= \delta\big(d^2_a f(\varepsilon x, y) + o_4(\|\varepsilon x\|)\big) + \frac{\delta^2}{2}\big(d^2_{a+\varepsilon x} f - d^2_a f\big)(y, y) + o_3(\delta^2) \end{aligned}$$

Аналогично, меняя местами слагаемые в определении g, получим

$$g(\varepsilon, \delta) = \varepsilon\big(d^2_a f(\delta y, x) + o_5(\|\delta y\|)\big) + \frac{\varepsilon^2}{2}\big(d^2_{a+\delta y} f - d^2_a f\big)(x, x) + o_6(\varepsilon^2)$$

Положим $\delta = \varepsilon$, приравняем полученные выражения для $g(\varepsilon, \varepsilon)$ и сократим на ε^2. Имеем:

$$d^2_a f(x, y) + \frac{1}{2}\big(d^2_{a+\varepsilon x} f - d^2_a f\big)(y, y) + o_7(1) = d^2_a f(y, x) + \frac{1}{2}\big(d^2_{a+\varepsilon y} f - d^2_a f\big)(x, x) + o_8(1),$$

где $o_7(1), o_8(1) \to 0$ при $\varepsilon \to 0$. Учитывая, что $d^2 f$ непрерывен, при $\varepsilon \to 0$ получаем $d^2_a f(x, y) = d^2_a f(y, x)$.

Если $f : \mathbb{R}^m \to \mathbb{R}$, то доказанное равенство можно переписать так: $y^T f''(a)x = x^T f''(a)y$. Полагая $x = e_i$, $y = e_j$, получаем $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$, что завершает доказательство. □

6. Экстремумы.

Пусть $f : U \to \mathbb{R}$.

Определение 6.1. Точка $a \in U$ называется точкой **локального минимума (максимума)** функции f, если существует такая окрестность этой точки, что при любом x из этой окрестности $f(a) \leqslant f(x)$ (соответственно, $f(a) \geqslant f(x)$). **Экстремум** функции – это точка локального минимума или максимума.

Теорема 6.2 (Необходимое условие экстремума)**.** *Пусть $f : U \to \mathbb{R}$ дифференцируема в точке $\theta \in U$ и имеет в этой точке локальный экстремум. Тогда $d_\theta f = 0$.*

Доказательство. Проведем доказательство в случае локального максимума в точке θ. По определению производной в точке θ

$$f(x) - f(\theta) = d_\theta f(x - \theta) + o(x - \theta)$$

Положим $x = \theta + \lambda v$ и заметим, что по предположению $f(x) - f(\theta) \leqslant 0$ при любом x из некоторой окрестности Ω точки θ. Получаем: $d_\theta f(\lambda v) + o(\lambda v) \leqslant 0$ при *любых* достаточно маленьких λ. Поделив на $|\lambda|$ и устремив λ к нулю получим $\pm d_\theta f(v) \leqslant 0$, где плюс или минус выбирается в зависимости от знака λ. Отсюда следует, что $df(\theta)(v) = 0$ для любого $v \in U$, а это и означает, что $df(\theta) = 0$. □

Замечание 6.3. В случае $U = \mathbb{R}^m$ заключение теоремы означает, что все частные производные $\frac{\partial f}{\partial x_i}$ равны нулю.

6.4. Симметричные билинейные формы $U \times U \to \mathbb{R}$ взаимно однозначно соответствуют квадратичным формам $U \to \mathbb{R}$, причем их матрицы в фиксированном базисе совпадают. Квадратичная форма, соответствующая билинейной форме $d_a^2 f$, называется **квадратичной формой второго дифференциала** в точке $a \in U$. Как мы уже знаем, если $U = \mathbb{R}^m$, то ее матрица в стандартном базисе состоит из частных производных

Напомним, что квадратичная форма Q называется положительно определенной, если $Q(x) > 0$ при любом $x \neq 0$; отрицательно определенной, если $Q(x) < 0$ при любом $x \neq 0$; и вырожденной, если $Q(x) = 0$ при некотором $x \neq 0$.

Лемма 6.5. *Если Q положительно или отрицательно определенная квадратичная форма на конечномерном пространстве U, то $\inf\limits_{x \neq 0} \frac{|Q(x)|}{\|x\|^2} > 0$.*

Доказательство. Из курса линейной алгебры известно, что любая квадратичная форма приводится к виду $Q(x) = \lambda_1 x_1^2 + \dots + \lambda_m x_m^2$, где $x_1, \dots, x_n$ – координаты x в некотором базисе. По теореме об эквивалентности норм можно считать, что норма на U задана равенством $\|x\| = \sqrt{x_1^2 + \dots + x_m^2}$. Очевидно, что форма положительно или отрицательно определена, если все λ_i имеют один знак. Так как в условии стоит модуль Q, можно считать, что все $\lambda_i > 0$. Пусть λ_k – наименьшее из λ_i. Тогда

$$\frac{|Q(x)|}{\|x\|^2} = \frac{\lambda_1 x_1^2 + \dots + \lambda_m x_m^2}{x_1^2 + \dots + x_m^2} \geqslant \lambda_k$$

Если x – k-ый базисный вектор (т.е. $x_k = 1$, $x_i = 0$ при $i \neq k$), то последнее неравенство превращается в равенство. Таким образом, $\inf\limits_{x \neq 0} \frac{|Q(x)|}{\|x\|^2} = \lambda_k > 0$. □

Теорема 6.6 (Достаточное условие экстремума)**.** *Пусть $f : U \to \mathbb{R}$ дважды непрерывно дифференцируема в окрестности точки a, и $d_a f = 0$. Если квадратичная форма второго дифференциала в точке a положительна определена, то a – точка локального минимума; если она отрицательна определена, то a – точка локального максимума.*

Доказательство. Предположим, что квадратичная форма $Q(\Delta x) = d_a^2 f(\Delta x, \Delta x)$ положительно определена. Пусть $\lambda = \inf\limits_{x \neq 0} \frac{d^2 f(a)(x-a, x-a)}{\|x\|^2} > 0$. По формуле Тэйлора с остаточным членом в форме Пеано 5.8

$$f(x) - f(a) = \frac{1}{2} d_a^2 f(x-a, x-a) + o(\|x-a\|^2) \geqslant \left(\frac{\lambda}{2} + o(1)\right) \|x-a\|^2$$

Так как $o(1) \to 0$ при $x \to a$, то существует $\delta > 0$ такое, что $\frac{\lambda}{2} + o(1) > 0$ при $\|x-a\| < \delta$. Таким образом, при $\|x-a\| < \delta$ мы доказали, что $f(x) > f(a)$, что и означает, что a – локальный минимум. $\square$

7. Условные экстремумы.

Рассмотрим функции $f : U \to \mathbb{R}$ и $g : U \to V$, где U и V – нормированные линейные пространства.

Определение 7.1. Будем говорить, что $a \in U$ является точкой **условного минимума (максимума)** функции f при условии $g = 0$, если $g(a) = 0$, и существует окрестность Ω точки a такая, что $f(z) \geqslant f(a)$ (соотв. $f(z) \leqslant f(a)$) при всех $z \in \Omega$, удовлетворяющих уравнению $g(z) = 0$.

Предложение 7.2. *Если при некотором $\lambda \in V^*$ функция $F(z) = f(z) - \lambda g(z)$ имеет локальный экстремум в точке a такой, что $g(a) = 0$, то f имеет условный экстремум при условии $g = 0$ в точке a.*

Доказательство. Пусть a – локальный минимум функции F. По определению существует окрестность Ω точки a, в которой $F(z) \geqslant F(a)$. Тогда для любого $z \in \Omega$, удовлетворяющего условию $g(z) = 0$, имеем: $f(z) = F(z) \geqslant F(a) = f(a)$. $\square$

7.3. Функция F называется **функцией Лагранжа**. Далее в этом параграфе мы рассмотрим только конечномерные пространства U и V. Как всегда, без ограничения общности можно считать, что $U = \mathbb{R}^{n+m}$, а $V = \mathbb{R}^m$ (задача на нахождение условного экстремума является интересной, только если $\dim U > \dim V$, т.е. $n > 0$). В этом случае V^* естественно отождествляется с множеством строк длины m, см.1.7. Элементы строки λ называются тогда **множителями Лагранжа**.

Элементы пространства U (т.е. аргументы функций f и g) будем обозначать через $z = (x^T, y^T)^T$, где $x \in \mathbb{R}^n$, а $y \in \mathbb{R}^m$, т.е. $z = (x_1, \ldots, x_n, y_1, \ldots, y_m)$ (эти обозначения соответствуют разложению пространства $\mathbb{R}^{n+m}$ в прямую сумму $\mathbb{R}^n \oplus \mathbb{R}^m$).

Пусть $(a, b) \in \mathbb{R}^{n+m}$ и $g(a, b) = 0$. Предположим, что существует функция $h : \mathbb{R}^n \to \mathbb{R}^m$ такая, что в некоторой окрестности Ω точки (a, b) уравнение $g(x, y) = 0$ равносильно уравнению $y = h(x)$. Например, это выполнено, если g непрерывно дифференцируема в окрестности точки a и $\det g'_y(a, b) \neq 0$ (см. теорему о неявной функции 4.2). Пусть $\varphi : \mathbb{R}^n \to \mathbb{R}$ задана формулой $\varphi(x) = f\big(x, h(x)\big)$. Тогда очевидно, что f имеет условный экстремум при условии $g = 0$ в точке (a, b) тогда и только тогда, когда φ имеет экстремум в точке a.

Чаще всего мы не сможем выразить y через x, т.е. найти выражение для функции h через *элементарные* функции. Так что существование функции φ и связь экстремумов этой функции с условными экстремумами f имеет скорее теоретическое значение. Несмотря на это, нам удастся написать систему уравнений для нахождения стационарных точек функции φ.

Теорема 7.4. *Пусть $(a, b) \in \mathbb{R}^{n+m}$ и $g(a, b) = 0$. Предположим, что f и g непрерывно дифференцируемы в некоторой окрестности точки (a, b) и $\det g'_y(a, b) \neq 0$. Тогда точка a является стационарной точкой функции φ тогда и только тогда, когда существует строка λ длины m такая, что точка (a, b) является стационарной точкой функции $F(x, y) = f(x, y) - \lambda g(x, y)$.*

Доказательство. По теореме о неявной функции 4.2, производная функции h вычисляется по формуле $h'(x) = -g'_y\big(x, h(x)\big)^{-1} g'_x\big(x, h(x)\big)$. По формуле для производной сложной функции 3.14,

$$\varphi'(x) = f'_x\big(x, h(x)\big) + f'_y\big(x, h(x)\big) h'(x) = f'_x\big(x, h(x)\big) - f'_y\big(x, h(x)\big) g'_y\big(x, h(x)\big)^{-1} g'_x\big(x, h(x)\big).$$

Заметим, что, так как $g(a,b)=0$, то $h(a)=b$ по определению функции h. Положим $\lambda = f'_y(a,b)g'_y(a,b)^{-1}$. Тогда $\varphi'(a)=F'(a,b)$ так, что, если a – стационарная точка φ, то $F'(a,b)=0$.

Обратно, если $F'(a,b)=0$ при каком-то значении λ, то, в частности, $f'_y(a,b)-\lambda g'_y(a,b)=0$, откуда $\lambda = f'_y(a,b)g'_y(a,b)^{-1}$. А при этом значении λ мы уже видели,что $\varphi'(a)=F'(a,b)$, т.е. $\varphi'(a)=0$. □

7.5. Таким образом, точки, подозрительные на условный экстремум функции f при условии $g=0$ (они же – стационарные точки функции φ), находятся из системы $m+n+1$ уравнения с неизвестными $x_1,\dots,x_n,y_1,\dots,y_m,\lambda$:

$$\begin{cases} D_1 f(x,y) - \lambda D_1 g(x,y) = 0 \\ \dots\dots\dots\dots\dots\dots \\ D_{m+n} f(x,y) - \lambda D_{m+n} g(x,y) = 0 \\ g(x,y) = 0 \end{cases}$$

Напомним 7.2, что точки экстремума функции F обязательно являются точками условного экстремума. Обратное в общем случае неверно. Сформулируем без доказательства более точное достаточное условие существования условного экстремума.

Теорема 7.6. *В обозначениях пункта 7.3 предположим, что функции f и g дважды непрерывно дифференцируемы в окрестности точки (a,b) такой, что $g(a,b)=0$. Пусть $\lambda = f'_y(a,b)g'_y(a,b)^{-1}$, а $A=h'(a)=-g'_y(a,b)^{-1}g'_x(a,b)$ – матрица $m\times n$. Получим:*

$$\varphi''(a) = \begin{pmatrix} E & A^T \end{pmatrix} \cdot F''(a,b) \cdot \begin{pmatrix} E \\ A \end{pmatrix}.$$

Предположим, далее, что $F'(a,b)=0$. Тогда, если $\varphi''(a)$ положительно определена, то f имеет условный минимум при условии $g=0$, отрицательно определена – максимум, знакопеременная – условного экстремума нет.

Другими словами, для нахождения условного экстремума надо сделать следующее:

- решить систему 7.5;
- в каждой стационарной точке (a,b) найти матрицу $F''(a,b)$ (значение λ уже будет к этому моменту известно из решения системы);
- выразить dy через dx из условия $d_{(a,b)}g = g'_x(a,b)dx + g'_y(a,b)dy = 0$;
- в квадратичную форму второго дифференциала $d^2_{(a,b)}F$ подставить dy из предыдущего пункта, получив квадратичную форму второго дифференциала $d^2_a\varphi$;
- выяснить, является ли эта квадратичная форма положительно или отрицательно определенной;
- записать ответ, используя достаточное условие экстремума 6.6.

8. Многообразия, касательные пространства.

Понятие многообразия обобщает понятия кривой и поверхности. Интуитивно, кривая – это "изогнутая" прямая, а поверхность – "изогнутая" плоскость. Однако, даже на примере окружности видно, что нельзя определить кривую, как топологическое пространство, гомеоморфное прямой (см.10.18). Еще хуже определить кривую, как образ прямой при непрерывном отображении, так как существует (см.10.17) непрерывное *сюръективное* отображение прямой на квадрат с двумя выколотыми граничными точками (очевидно, не хочется, чтобы квадрат был бы примером кривой). На самом деле, кривая определяется, как топологическое пространство "склеенное" из кусочков, гомеоморфных прямой, а поверхрость – пространство, "склеенное" из кусочков, гомеоморфных плоскости. В общем случае, n-мерное многообразие "склеено" из кусочков, гомеоморфных $\mathbb{R}^n$.

Определение 8.1. Подмножество $M\subseteq\mathbb{R}^m$ называется **n-мерным многообразием** (без края), если для любой точки $a\in M$ существует окрестность Ω_a и гомеоморфизм $\varphi_a:\mathbb{R}^n\to\Omega_a$.

Функция φ_a называется **локальной параметризацией** окрестности Ω_a.

M называется n**-мерным многообразием с краем**, если для любой точки $a \in M$ существует окрестность Ω_a, гомеоморфная либо $\mathbb{R}^n$, либо замкнутому полупространству $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_n \geqslant 0\}$.

Множество всех точек, у которых не существует окрестности, гомеоморфной $\mathbb{R}^n$, называется **краем** многообразия M. Оно обозначается через ∂M (не путайте с границей топологического подпространства, которое обозначается тем же символом).

Одномерное многообразие называется **кривой**, а двумерное – **поверхностью**.

Замечание 8.2. Так как $\mathbb{R}^n$ не гомеоморфно $\mathbb{R}^m$ при $n \neq m$ (см. 10.16), то размерность многообразия определена единственным образом.

Так как $\mathbb{R}^n$ гомеоморфно открытому n-мерному шару (более того, гомеоморфизм может быть выбран сколько угодно раз дифференцируемым, см.10.15), то можно считать, что локальные параметризации действуют не из $\mathbb{R}^n$ в Ω_a, а из Φ_a в Ω_a, где Φ_a – некоторое открытое подмножество $\mathbb{R}^n$ (ср. со следующим примером).

Примеры 8.3. (i) Любое открытое подмножество Φ в $\mathbb{R}^n$ является n-мерным многообразием. Действительно, для любой точки $a \in \Phi$ существует открытый шар $\Omega_a \subseteq \Phi$ с центром в a. Тогда локальная параметризация – гомеоморфизм $\mathbb{R}^n \to \Omega_a$.
(ii) Окружность является одномерным многообразием. Рассмотрим окружность C в $\mathbb{R}^2$, заданную уравнением $x^2 + y^2 = 1$. Пусть $\varphi : \mathbb{R} \to \mathbb{R}^2$ задана формулой $\varphi(t) = (\cos t, \sin t)^T$. Эта функция не является параметризацией, так как она не взаимно однозначна (как уже было сказано в начале параграфа, окружность нельзя параметризовать целиком). Однако для каждой точки $a \in C$ существует открытый отрезок $\Phi_a \subset \mathbb{R}$, который гомеоморфно отображается в некоторую окрестность $\Omega_a \subset C$ точки a под действием функции φ. В этой и аналогичных ситуациях будем говорить, что функция φ **индуцирует** набор локальных параметризаций.
(iii) Более сложный пример двумерного многообразия – тор (поверхность бублика). Набор локальных параметризаций индуцируется функцией $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ с координатными функциями: $\varphi_1(u, v) = (2 + \cos v) \cos u$, $\varphi_2(u, v) = (2 + \cos v) \sin u$ и $\varphi_3(u, v) = \sin v$.

8.4. Многообразия M_1 и M_2 называются **гомеоморфными**, если они гомеоморфны, как топологические пространства, т.е. существуют взамно-обратные непрерывные функции $f : M_1 \to M_2$ и $g : M_2 \to M_1$.

Определение 8.5. n-Мерное многообразие M называется **гладким**, если все локальные параметризации φ_a могут быть выбраны непрерывно дифференцируемыми, причем $\operatorname{rank} \varphi_a'(b) = n$, где $b = \varphi_a^{-1}(a)$.

В дальнейшем слово многообразие будет означать гладкое многообразие без края.

Пусть $M_1 \subseteq \mathbb{R}^{m_1}$ и $M_2 \subseteq \mathbb{R}^{m_2}$ – многообразия, а f – функция $M_1 \to M_2$. Можно было бы дать такое же определение дифференцируемости f, как в 3.2. Однако, дифференциал f таким образом определить *нельзя*, так как может существовать много линейных отображений L, удовлетворяющих указанному в этом определении свойству. Интуитивно это видно из того, что значения L на векторах, не представляющихся в виде $x - a$ при $a, x \in M_1$ не играют никакой роли. Поэтому, мы сначала определим касательное пространство к многообразию, а потом дифференциал f, как линейное отображение из касательного пространства к M_1 в точке a в касательное пространство к M_2 в точке $f(a)$. Для этого нам понадобиться следующая лемма.

Лемма 8.6. *Пусть Φ_1 и Φ_2 – открытые подмножества в $\mathbb{R}^n$, а $\varphi_i : \Phi_i \to \Omega$ – непрерывно дифференцируемые гомеоморфизмы Φ_i на некоторое подмножество $\Omega \subseteq \mathbb{R}^m$, причем $\operatorname{rank} \varphi_i'(b) = n$ для всех $b \in \Phi_i$ (здесь $i = 1$ или 2). Тогда $\varphi = \varphi_2^{-1} \circ \varphi_1$ – непрерывно дифференцируемый гомеоморфизм $\Phi_1 \to \Phi_2$ и $\det \varphi'(b) \neq 0$ при любом $b \in \Phi_1$.*

Доказательство. Ясно, что φ является гомеоморфизмом. Докажем, что она непрерывна дифференцируема в некоторой окрестности произвольной точки $a \in \Phi_1$. Пусть $b = \varphi(a)$. Так как $\operatorname{rank} \varphi_2'(b) = n$, то существует подматрица A в $\varphi_2'(b)$ размера $n \times n$, определитель которой не равен нулю. Пусть эта подматрица состоит из строк с номерами $k_1, \ldots, k_n$. Обозначим через ψ_i функции $\Phi_i \to \mathbb{R}^n$ с координатными функциями $(\psi_i)_j = (\varphi_i)_{k_j}$ (здесь $i = 1$ или 2, а

$j = 1, \ldots, n$). Очевидно, что $\psi_2'(b) = A$, причем ψ_i непрерывно дифференцируемы. По теореме об обратной функции 4.1 в некоторой окрестности Ω' точки $\psi_2(b)$ существует непрерывно дифференцируемая функция ξ, обратная к ψ_2 (точнее к сужению ψ_2 на некоторую окрестность точки b). Теперь легко проверить, что сужение функции φ на $\psi_1^{-1}(\Omega')$ равно $\xi \circ \psi_1$, которая непрерывно дифференцируема, как композиция непрерывно дифференцируемых функций.

Аналогично докажем, что $\varphi^{-1} = \varphi_1^{-1} \circ \varphi_2$ непрерывно дифференцируема. Тогда, по формуле для производной обратной функции, матрица $\varphi'(b)$ должна быть обратима при любом $b \in \Phi_1$.

Интуитивно, смысл этого доказательства заключается в том, что образ $\varphi_i(b)$ любой точки $b \in \Phi_i$ однозначно определяется своими координатами с номерами $k_1, \ldots, k_n$. □

Определение 8.7. Пусть M – n-мерное многообразие, φ_a - локальная параметризация M в точке a, а $b = \varphi_a^{-1}(a) \in \mathbb{R}^n$ прообраз точки a. **Касательным пространством** к M в точке a называется образ отображения $d_b\varphi_a$, Оно обозначается через D_aM. Другими словами, D_aM – это линейная оболочка столбцов матрицы $\varphi_a'(b)$. По определению 8.5, эти столбцы линейно независимы и, следовательно, являются базисом D_aM. Поэтому размерность касательного пространства равна размерности многообразия.

Множество $T_aM = a + D_aM \subseteq \mathbb{R}^m$ называется **касательным многообразием** к M в точке a.

Лемма 8.8. *Касательное пространство не зависит от выбора параметризации.*

Доказательство. Пусть Ω_1 и Ω_2 – окрестности точки a, а $\varphi_1 : \mathbb{R}^n \to \Omega_1$ и $\varphi_2 : \mathbb{R}^n \to \Omega_2$ – локальные параметризации из определения 8.5. Обозначим $\Omega = \Omega_1 \cap \Omega_2$. Тогда $\Phi_i = \varphi_i^{-1}(\Omega)$ – окрестности точек $a_i = \varphi_i^{-1}(a)$ (здесь $i = 1$ или 2). По лемме 8.6 существует непрерывно дифференцируемая функция $\varphi : \Phi_1 \to \Phi_2$ с невырожденной производной такая, что $\varphi_1 = \varphi_2 \circ \varphi$ на Φ_1. По формуле дифференциала композиции 3.14 $d_{a_1}\varphi_1 = d_{a_2}\varphi_2 \circ d_{a_1}\varphi$. Так как матрица оператора $d_{a_1}\varphi$ невырождена, он отображает $\mathbb{R}^n$ на себя биективно. Поэтому образы операторов $d_{a_1}\varphi_1$ и $d_{a_2}\varphi_2$ совпадают. □

Одно из важнейших применений дифференциального исчисления – возможность заменить "маленькую" окрестность точки на многообразии на "кусочек" касательного многообразия, что происходит, например, при определении интеграла от скалярной функции по многообразию (в частности, при определении площади поверхности).

Определение 8.9. Пусть M_1 – n_1-мерное, а M_2 – n_2-мерное многообразие. Рассмотрим непрерывную функцию $f : M_1 \to M_2$, точку $a \in M_1$ и обозначим $b = f(a)$. Пусть $\Omega_b \subseteq M_2$ – окрестность точки b, а $\Omega_a = f^{-1}(\Omega_b) \subseteq M_1$ – окрестность точки a такие, что существуют локальные параметризации $\varphi_a : \mathbb{R}^{n_1} \to \Omega_a$ и $\varphi_b : \mathbb{R}^{n_2} \to \Omega_b$. Положим $\bar{a} = \varphi_a^{-1}(a)$, и $\bar{b} = \varphi_b^{-1}(b)$. Обозначим через ψ композицию функций

$$\mathbb{R}^{n_1} \xrightarrow{\varphi_a} \Omega_a \xrightarrow{f} \Omega_b \xrightarrow{\varphi_b^{-1}} \mathbb{R}^{n_2}.$$

Если функция ψ непрерывно дифференцируема, то f называется непрерывно дифференцируемой. При этом **дифференциалом** f в точке a называется линейное отображение d_af являющееся композицией линейных отображений

$$D_aM_1 \xrightarrow{(d_{\bar{a}}\varphi_a)^{-1}} \mathbb{R}^{n_1} \xrightarrow{d_{\bar{a}}\psi} \mathbb{R}^{n_2} \xrightarrow{d_{\bar{b}}\varphi_b} D_bM_2.$$

Доказательство того, что d_af не зависит от выбора параметризаций φ_a и φ_b, оставим в качестве упражнения.

Определение 8.10. Многообразия M_1 и M_2 называются **диффеоморфными**, если если существуют взаимно-обратные непрерывно дифференцируемые функции $f : M_1 \to M_2$ и $g : M_2 \to M_1$. Сами функции f и g называются **диффеоморфизмами**.

Очевидно, что, если f – диффеоморфизм, то $d_af : D_aM_1 \to D_{f(a)}M_2$ – изоморфизм линейных пространств (обратное отображение $d_{f(a)}g$, сравните с теоремой 4.1). Этот факт, в частности, говорит о том, что касательное пространство независимо от вложения многообразия в $\mathbb{R}^m$ с точностью до *канонического* изоморфизма.

Следствие 8.11. *Пусть M_1 – подмногообразие в M, а $a \in M_1$. Тогда $D_aM_1 \subseteq D_aM$ и $T_aM_1 \subseteq T_aM$. В частности, касательная к любой кривой, лежащей на поверхности, содержится в касательной плоскости к этой поверхности.*

Доказательство. Применить определение 8.9 к отображению вложения $M_1 \to M$. □

Довольно часто подмножества в $\mathbb{R}^m$ заданы не параметрически, как в определении многообразия, а уравнением, системой уравнений или, как график функции. Сначала мы покажем, что график непрерывно дифференцируемой функции является многообразием.

Предложение 8.12. *Пусть Φ – открытое подмножество в $\mathbb{R}^n$. График непрерывно дифференцируемой функции $g : \Phi \to \mathbb{R}^m$ является n-мерным многообразием в $\mathbb{R}^{n+m}$.*

Доказательство. Функция $\varphi : \Phi \to \mathbb{R}^{m+n}$, с координатными функциями

$$\varphi_k(x_1, \ldots, x_n) = \begin{cases} g_k(x_1, \ldots, x_n), & \text{при } k \leqslant n \\ x_{k-n}, & \text{при } k > n \end{cases}$$

и будет локальной параметризацией в любой точке a. Действительно, уравнение $y = g(x_1, \ldots, x_n)$, задающее график функции g очевидно равносильно уравнению $(y_1, \ldots, y_m, x_1, \ldots, x_n)^T = \varphi(x)$, а функция φ инъективна. Кроме того, как указано в замечании 8.2, локальная параметризация может быть задана на открытом подмножестве $\mathbb{R}^n$. Наконец, из непрерывной дифференцируемости g сразу вытекает непрерывная дифференцируемость φ. □

Следующая теорема, вытекающая из теоремы о неявной функции 4.2, дает достаточное условие того, чтобы система уравнений задавала многообразие.

Теорема 8.13. *Пусть Φ – открытое подмножество в $\mathbb{R}^m$, а $f : \Phi \to \mathbb{R}^n$ – непрерывно дифференцируемая функция, причем $\operatorname{rank} f'(a) = n$ в любой точке $a \in \Phi$. Тогда множество точек $z \in \Phi$, удовлетворяющих уравнению $f(z) = 0$, является $(m-n)$-мерным многообразием в $\mathbb{R}^m$.*

Доказательство. Предположим для определенности, что подматрица, состоящая из первых n столбцов матрицы $f'(a)$, невырождена. Обозначим $z = (x^T, y^T)^T$, где $x \in \mathbb{R}^n$, а $y \in \mathbb{R}^{m-n}$. По теореме о неявной функции 4.2 существует окрестность $\Phi \subset \mathbb{R}^{m-n}$ точки $(a_{n+1}, \ldots, a_m)^T$ и непрерывно дифференцируемая функция $g : \Phi \to \mathbb{R}^n$ такая, что при $y \in \Phi$ уравнение $f(x, y) = 0$ равносильно уравнению $x = g(y)$, т.е. наше множество является графиком функции g. По предыдущему предложению, оно является многообразием. □

9. Кривые и поверхности. Градиент, производная по направлению.

В настоящем параграфе мы рассмотрим только что введенные понятия в маленьких размерностях так, чтобы все можно было бы легко представить наглядно. Вообще, видимо бóльшая часть людей, говоря про n-мерные многообразия в $\mathbb{R}^m$, представляют себе кривые и поверхности в $\mathbb{R}^3$. Буквы x, y и z будут теперь обозначать обычные декартовы координаты в трехмерном пространстве. Все функции, по умолчанию, будут считаться непрерывно дифференцируемыми.

9.1. *Способы задания кривых и поверхностей в $\mathbb{R}^2$ и $\mathbb{R}^3$.*

(i) Параметрическое задание – это задание из определения многообразия. Т.е. кривая (локально) задается, как образ открытого отрезка или прямой, а поверхность – как образ плоскости, открытого круга, прямоугольника, или другого открытого множества из $\mathbb{R}^2$, под действием биективного непрерывно дифференцируемого отображения.

(ii) Явное задание – это задание кривой (поверхности) как графика функции f. Для кривой в $\mathbb{R}^2$ f действует из $\mathbb{R}$ в $\mathbb{R}$ (это известно еще со школы !), для кривой в $\mathbb{R}^3$ – $f : \mathbb{R} \to \mathbb{R}^2$, а для поверхности в $\mathbb{R}^3$ – $f : \mathbb{R}^2 \to \mathbb{R}$. Явное задание очевидным образом переводится в параметрическое (см. доказательство предложения 8.12).

(iii) Неявное задание – это задание уравнением или системой уравнений $f = 0$. Для кривой в $\mathbb{R}^2$ f действует из $\mathbb{R}^2$ в $\mathbb{R}$, для кривой в $\mathbb{R}^3$ – $f : \mathbb{R}^3 \to \mathbb{R}^2$, а для поверхности в $\mathbb{R}^3$ – $f : \mathbb{R}^3 \to \mathbb{R}$.

В случае кривой в $\mathbb{R}^3$ мы имеем систему уравнений $\begin{cases} f_1(x, y, z) = 0 \\ f_2(x, y, z) = 0 \end{cases}$. Каждое из уравнений этой системы задает поверхность. Так что в этом случае можно считать, что кривая задана, как пересечение двух поверхностей.

9.2. В этом пункте мы напишем в явном виде уравнения касательных многообразий к кривым и поверхностям в точке $a = (x_0, y_0, z_0)$ при различных способах задания. Буквы t, u и v будут обозначать параметры, а t_0, u_0 и v_0 – те значения параметров, при которых мы попадаем в точку a. Номера (i)-(iii) будут соответствовать нумерации из 9.1, а буквы a, b, c – кривой на плоскости, кривой в пространстве и поверхности в пространстве, соответственно.

(i.a) Аналогично (i.b), но без координаты z.

(i.b) Кривая: $x = X(t)$, $y = Y(t)$ и $z = Z(t)$.
Касательная: $x = x_0 + X'(t_0)t$, $y = y_0 + Y'(t_0)t$, $z = z_0 + Z'(t_0)t$.

(i.c) Поверхность: $x = X(u, v)$, $y = Y(u, v)$ и $z = Z(u, v)$. Касательная плоскость: $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$, где вектор $(A, B, C)^T$ равен векторному произведению векторов $\big(\frac{\partial X}{\partial u}(u_0, v_0), \frac{\partial Y}{\partial u}(u_0, v_0), \frac{\partial Z}{\partial u}(u_0, v_0)\big)^T$ и $\big(\frac{\partial X}{\partial v}(u_0, v_0), \frac{\partial Y}{\partial v}(u_0, v_0), \frac{\partial Z}{\partial v}(u_0, v_0)\big)^T$.

(ii.a) Кривая является графиком функции $Y(x)$.
Касательная: $y = y_0 + Y'(x_0)(x - x_0)$.

(ii.b) Кривая: $y = Y(x)$ и $z = Z(x)$.
Касательная: $y = y_0 + Y'(x_0)(x - x_0)$, $z = z_0 + Z'(x_0)(x - x_0)$.

(ii.c) Поверхность: $z = Z(x, y)$.
Касательная плоскость: $z = z_0 + \frac{\partial Z}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial Z}{\partial y}(x_0, y_0)(y - y_0)$.

(iii.a) Кривая: $f(x, y) = 0$. Касательная: $\frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0) = 0$.

(iii.b) Кривая: $f(x, y, z) = 0$ и $g(x, y, z) = 0$. Касательная: $\frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0) + \frac{\partial f}{\partial z}(z - z_0) = 0$ и $\frac{\partial g}{\partial x}(a)(x - x_0) + \frac{\partial g}{\partial y}(a)(y - y_0) + \frac{\partial g}{\partial z}(z - z_0) = 0$.

(iii.c) Поверхность: $f(x, y, z) = 0$.
Касательная плоскость: $\frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0) + \frac{\partial f}{\partial z}(z - z_0) = 0$.

Сейчас мы выясним, как связаны между собой нормальные вектора к касательной плоскости из (i.c) и (iii.c), а в следующем определении обобщим это. Пусть поверхность M задана уравнением $f(x, y, z) = 0$. По теореме 8.13 в окрестности точки a ее можно задать параметрически, как в (i.c). Тогда вектора из (i.c), состоящие из частных производных, по определению 8.7 являются базисом в D_aM (а значит, параллельны касательной плоскости). Так как $f\big(X(u, v), Y(u, v), Z(u, v)\big) = 0$, то из свойства дифференцирования композиции 3.14 следует, что вектор $f'(a)^T$ ортогонален этим векторам, а следовательно, и всей касательной плоскости. Это доказывает формулу для касательной плоскости из (iii.c).

Определение 9.3. Пусть $f : \mathbb{R}^m \to \mathbb{R}$. **Градиентом** этой функции в точке $a \in \mathbb{R}^m$ называется вектор $\operatorname{grad}_a f = f'(a)^T$.

Замечание 9.4. Производная $f'(a)$ – это матрица линейного функционала $\mathbb{R}^m \to \mathbb{R}$. Сопоставление функционалу d_af вектора $\operatorname{grad}_a f$ – это отображение, обратное к каноническому изоморфизму, указанному в 1.7.

Предложение 9.5. *Градиент* $\operatorname{grad}_a f$ *ортогонален касательному подпространству* D_aM *к многообразию* M*, заданному уравнением* $f = 0$.

Это утверждение уже фактически было доказано перед определением градиента. Для того, чтобы сформулировать еще одно замечательное свойство градиента, введем понятие производной по направлению для функции $f : \mathbb{R}^m \to \mathbb{R}$, которая показывает скорость возрастания функции f, если аргумент “движется” по направлению вектора v.

Определение 9.6. Пусть $a, v \in \mathbb{R}^m$. **Производной** функции f **по направлению** вектора v в точке a называется число $\frac{\partial f}{\partial v}(a)$ равное производной функции $g(t) = f\big(a + t\frac{v}{\|v\|}\big)$ в нуле. Другими словами,

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f\big(a + t\frac{v}{\|v\|}\big) - f(a)}{t}.$$

Предложение 9.7. $\frac{\partial f}{\partial v}(a) = \frac{(v, \operatorname{grad}_a f)}{\|v\|}$. *Если f и a фиксированы, то производная по направлению принимает наибольшее значение при $v = \operatorname{grad}_a f$.*

Формула сразу следует из определения и свойства дифференцирования композиции 3.14. Второе утверждение вытекает из неравенства Коши–Буняковского.

10. *Приложение.* Элементарные сведения из топологии.

В отличие от основного текста, почти все утверждения в приложении даются без доказательств (в некоторых местах даже не акцентируется внимание, что там нужно что-то доказывать). Аккуратное изложение большей части приложения можно найти на первых страницах любой книги по общей топологии. Однако несколько утверждений являются *сложными* теоремами, для доказательства которых используется теория гомологий и/или гомотопий топологических пространств (для чего эти теории, в частности, и нужны). На эти утверждения в тексте обращается особое внимание.

Определение 10.1. Пусть X – множество, а $\mathcal{U}$ – набор подмножеств в X. $\mathcal{U}$ называется **топологией** на X, если выполнены следующие условия.

(i) $\varnothing, X \in \mathcal{U}$.
(ii) Если $U_1, U_2 \in \mathcal{U}$, то $U_1 \cap U_2 \in \mathcal{U}$.
(iii) Если $\mathcal{T} \subseteq \mathcal{U}$, то $\bigcup\limits_{U \in \mathcal{T}} U \in \mathcal{U}$.

Множество X с заданной на нем топологией называется **топологическим пространством** и обозначается $(X, \mathcal{U})$. Там, где это не вызывает недоразумений (т.е. почти всегда), будем писать X вместо $(X, \mathcal{U})$. Множества из $\mathcal{U}$ называются **открытыми** множествами. **Замкнутым** множеством называется множество C, дополнение к которому $X \setminus C$ открыто. **Открытая окрестность** точки $x \in X$ – это любое открытое множество Ω, содержащее точку x. **Окрестность** точки $x \in X$ – это любое множество, содержащее открытую окрестность точки x.

В настоящем тексте все окрестности по умолчанию считаются открытыми.

Из определения следует, что пересечение *конечного* числа открытых множеств открыто, объединение *любого* количества открытых открыто, пересечение *любого* количества замкнутых множеств замкнуто, и объединение *конечного* числа замкнутых замкнуто.

Пусть Y – подмножество топологического пространства X. **Замыканием** Y называется пересечение всех замкнутых множеств, содержащих Y. Двойственно, **внутренностью** Y называется объединение всех открытых, содержащихся в Y. Замыкание Y обозначается через $\operatorname{Cl} Y$ или иногда $\bar{Y}$, а внутренность – через $\operatorname{Int} Y$. **Границей** Y называется разность $\partial Y = \operatorname{Cl} Y \setminus \operatorname{Int} Y$.

Для нас будут важны примеры метрических топологических пространств.

Определение 10.2. Пусть X – множество, а $\rho : X \times X \to [0, +\infty)$ – функция, удовлетворяющая следующим свойствам.

(i) $\rho(x, y) = 0 \iff x = y$.
(ii) $\rho(x, y) = \rho(y, x)$.
(iii) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$.

Тогда ρ называется **метрикой** на X. Множество X, наделенное метрикой, называется **метрическим пространством**.

10.3. В метрическом пространстве X топология задается с помощью **открытых шаров** $B_r(a) = \{x \in X \mid \rho(x, a) < r\}$, где $a \in X$, а $r > 0$. А именно, подмножество U называется открытым, если оно является объединением некоторого количества открытых шаров. Эквивалентно, U открыто в X, если $\forall u \in U \ \exists r > 0 : \ B_r(u) \subseteq U$.

10.4. Если Y – подмножество метрического пространства (X, ρ), то та же функция ρ (точнее, ее сужение на $Y \times Y$) задает метрику на Y. Мы всегда будем считать, что подмножество метрического пространства будет метрическим пространством с такой метрикой.

Заметим, что открытые шары в Y – это пересечение Y с открытыми шарами из X. Следовательно, то же самое выполнено для открытых множеств. Такая топология на Y называется топологией, индуцированной с X.

Точнее, пусть $(X, \mathfrak{U}_X)$ – произвольное топологическое пространство, а $Y \subseteq X$. Пусть $\mathfrak{U}_Y = \{U \cap Y \mid U \in \mathfrak{U}_X\}$. Тогда $\mathfrak{U}_Y$ является топологией на Y, которая называется **индуцированной топологией**. В этом случае $(Y, \mathfrak{U}_Y)$ называется (топологическим) подпространством в $(X, \mathfrak{U}_X)$. Очевидно, что индуцированная с X метрика задает индуцированную с X топологию.

10.5. Основными примерами метрических пространств будут у нас подмножества $\mathbb{R}^n$, или, более общо, подмножества нормированных линейных пространств с индуцированной метрикой.

Если X – нормированное линейное пространство, то метрика на X задается формулой $\rho(x, y) = \|x - y\|$. Таким образом X, а следовательно и любое его подмножество, превращается в метрическое пространство и, значит, в топологическое пространство. Мы *всегда* будем рассматривать на подмножествах нормированных линейных пространств именно такую топологию.

10.6. В метрическом пространстве обычным образом определяется понятие предела последовательности или функции из другого метрического пространства. На языке пределов можно переформулировать определение замкнутости следующим образом. Подмножество C метрического пространства X является замкнутым тогда и только тогда, когда для любой сходящейся в X последовательности $x_n \to x$, из того что $x_n \in C$ при всех $n \in \mathbb{N}$ следует, что $x \in C$.

На языке пределов также легко сформулировать определение непрерывности функции $f : X \to Y$ для метрических пространств X и Y: функция f называется непрерывной в точке $a \in X$, если $\forall \varepsilon > 0\ \exists \delta > 0$ такое, что $\rho(x, a) < \delta \Longrightarrow \rho\big(f(x), f(a)\big) < \varepsilon$.

Более удобным однако является топологическое определение непрерывности, которое в случае метрических пространств эквивалентно определению приведенному выше.

Определение 10.7. Пусть $(X, \mathfrak{U}_X)$ и $(Y, \mathfrak{U}_Y)$ – топологические пространства. Функция $f : X \to Y$ называется **непрерывной**, если прообраз любого открытого в X множества является открытым в Y.

Функция f называется **гомеоморфизмом**, если она непрерывна, биективна, и обратная функция $f^{-1} : Y \to X$ также непрерывна. Другими словами, f – гомеоморфизм, если она биективно отображает $X \to Y$, а $\mathfrak{U}_X \to \mathfrak{U}_Y$. В этом случае пространства X и Y называются **гомеоморфными**. Таким образом, гомеоморфные пространства с точки зрения топологии отличаются только обозначениями. Отсюда следующее определение.

Свойство топологического пространства называется **топологическим свойством**, если оно сохраняется при гомеоморфизмах.

Любой изоморфизм *конечномерных* нормированных линейных пространств является гомеоморфизмом (это утверждение очевидно равносильно теореме об эквивалентности норм 1.4). Поэтому все топологические свойства подмножеств $\mathbb{R}^n$ сохраняются для подмножеств произвольных конечномерных нормированных линейных пространств.

Сейчас мы определим несколько топологических свойств, которые играют важную роль в математическом анализе.

Определение 10.8. Топологическое пространство X называется **несвязным**, если оно разбивается в объединение двух непустых непересекающихся открытых множеств. В противном случае X называется **связным**.

X называется **линейно связным**, если для любых точек $x, y \in X$ существует непрерывная функция $f : [0, 1] \to X$ такая, что $f(0) = x$, а $f(1) = y$.

Предложение 10.9. *Образ связного множества под действием непрерывного отображения связен. Образ линейно связного множества под действием непрерывного отображения линейно связен. Следовательно, связность и линейная связность являются топологическими свойствами.*

Линейно связное множество связно.

Определение 10.10. Пусть $(X, \mathcal{U})$ – топологическое пространство, а $\mathcal{C} \subseteq \mathcal{U}$ – набор открытых множеств. $\mathcal{C}$ называется **открытым покрытием** X, если $X = \bigcup_{C \in \mathcal{C}} C$. Покрытие называется конечным, если оно состоит из конечного числа открытых множеств. **Подпокрытие** покрытия $\mathcal{C}$ – это подмножество $\mathcal{B} \subseteq \mathcal{C}$, которое само является покрытием.

Топологическое пространство X называется **компактным**, если из любого открытого покрытия можно выделить конечное подпокрытие.

Предложение 10.11. *Образ компактного множества при непрерывном отображении компактен. Поэтому компактность является топологическим свойством.*

Замкнутое подмножество компакта – компакт.

Определение компактности для подмножеств в $\mathbb{R}^n$ можно переформулировать на гораздо более понятном языке.

Теорема 10.12. *Если X – топологическое подпространство в $\mathbb{R}^n$, то X компактно тогда и только тогда, когда оно замкнуто в $\mathbb{R}^n$ и ограничено (т.е. существует константа $\mu \in \mathbb{R}$ такая, что $\|x - y\| < \mu$ для всех $x, y \in X$).*

Компактные метрические пространства обладают рядом замечательных свойств. Для формулировки одного из них нам понадобится следующее определение.

Определение 10.13. Пусть X и Y – метрические пространства. Напомним, что функция $f : X \to Y$ называется **равномерно непрерывной**, если $\forall \varepsilon > 0$ $\exists \delta > 0$ такое, что $\rho(x, y) < \delta \Longrightarrow \rho\big(f(x), f(y)\big) < \varepsilon$.

От определения непрерывности из 10.6 это отличается тем, что δ может быть выбрана независимо от x и y.

Теорема 10.14. *Компактное метрическое пространство X обладает следующими свойствами.*

(i) *Из любой последовательности $\{x_n\}_{n \in \mathbb{N}}$ элементов из X можно выбрать подпоследовательность x_{n_k}, сходящуюся при $k \to \infty$ к некоторому элементу из X.*

(ii) *Для любой непрерывной функции $f : X \to \mathbb{R}$ существуют точки $\underline{x}$ и $\overline{x}$ такие, что $f(\underline{x}) \leqslant f(x) \leqslant f(\overline{x})$ для любого $x \in X$.*

(iii) *Любая непрерывная функция $f : X \to Y$ является равномерно непрерывной.*

(iv) *Если X – подпространство метрического пространства Y, то X замкнуто в Y.*

Свойство быть замкнутым в некотором объемлющем пространстве очевидно не является топологическим свойством (например, $[0, 1)$ незамкнуто в $\mathbb{R}$, но замкнуто в интервале $(-1, 1)$, который, как обычно, рассматривается как подпространство в $\mathbb{R}$ с индуцированной топологией). Тем более удивительно, что топологическое свойство компактности влечет нетопологическое свойство быть замкнутым в объемлющем метрическом пространстве.

Наконец, приведем несколько важных примеров гомеоморфных и не гомеоморфных топологических пространств.

10.15. Открытый шар в $\mathbb{R}^n$ гомеоморфен $\mathbb{R}^n$.

Например, для единичного шара B с центром в нуле гомеоморфизм $f : \mathbb{R}^n \to B$ можно задать формулой $f(v) = \frac{2 \operatorname{arctg} \|v\|}{\pi \|v\|} \cdot v$ (каждый ненулевой вектор сохраняет направление, а его норма загоняется в промежуток $(0, 1)$; нулевой вектор отображается в себя).

10.16. $\mathbb{R}^n$ не гомеоморфно $\mathbb{R}^m$ при $m \neq n$.

В общем случае это *сложная* теорема. Докажем ее в случае $n = 1$. Пусть $f : \mathbb{R} \to \mathbb{R}^m$ – гомеоморфизм. Выколем точку 0 из $\mathbb{R}$ и ее образ $f(0)$ из $\mathbb{R}^m$. Тогда сужение f будет гомеоморфизмом $\mathbb{R} \setminus \{0\} \to \mathbb{R}^m \setminus \{f(0)\}$. Но $\mathbb{R} \setminus \{0\}$ несвязно (оно равно объединению $(-\infty, 0) \cup (0, +\infty)$ открытых множеств, в то время, как $\mathbb{R}^m \setminus \{f(0)\}$ очевидно линейно связно и, следовательно, связно. Это противоречит тому, что связность является топологическим свойством.

10.17. Не существует непрерывного биективного отображения замкнутого отрезка I на квадрат (или n-мерный куб при $n > 1$).

Отрезок I компактен, поэтому на незамкнутый квадрат он отображаться не может (см.10.11). Пусть f – непрерывное биективное отображения I на замкнутый квадрат. Замкнутое подмножество в I компактно, поэтому оно отображается в компактное множество, которое по 10.14 является замкнутым. Раз образ замкнутого – замкнут, а отображение биективно, то образ открытого – открыт. Поэтому обратное к f отображение непрерывно, следовательно, f – гомеоморфизм. То, что этого не может быть, доказывается так же, как и в предыдущем примере, выкалывая внутреннюю точку отрезка.

Забавно, что *сюръективное* непрерывное отображение отрезка на квадрат существует. Действительно, любая точка отрезка представляется в виде $x = \sum\limits_{n=1}^{\infty} \frac{\varepsilon_n}{2^n}$, где все ε_n – нули или единицы (двоичная запись числа). Отобразим x в пару чисел $(\sum\limits_{n=1}^{\infty} \frac{\varepsilon_{2n}}{2^{2n}}, \sum\limits_{n=1}^{\infty} \frac{\varepsilon_{2n-1}}{2^{2n-1}})$. Легко показать, что эта пара не зависит от представления x (двузначность представления возникает, когда все ε_n, начиная с некоторого номера, равны 1), и что такое отображение непрерывно. Из-за той же самой двузначности представления точек квадрата, это отображение не инъективно, и, как мы только что видели, нет никакой возможности подправить его, чтобы оно стало биективным и осталось непрерывным.

10.18. Окружность не гомеоморфна прямой.

Доказательство такое же, как в примере 10.16: выкалывая точку из прямой, получим несвязное пространство, а из окружности – связное; поэтому они не могут быть гомеоморфны.

И наконец, *сложная* теорема из той же серии (это значит, что в доказательстве существенную роль играет сохранение топологических свойств).

Теорема 10.19. *Пусть $f : \mathbb{R}^n \to \mathbb{R}^n$ – непрерывное инъективное отображение. Тогда f – открытое отображение, т.е. образ открытого множества открыт.*