5 Элементы функционального анализа

5.1 Линейные, нормированные и банаховы пространства

5.1.1 Определение пространств

Непустое множество X элементов $x, y, z, \ldots$ называется *линейным* (векторным) пространством, если

a) для любых элементов $x, y \in X$ однозначно определен элемент из X, называемый их суммой и обозначаемый $x + y$;

b) для любого числа $\lambda \in \mathbb{R}$ и элемента x однозначно определен элемент, называемый произведением числа λ на элемент x и обозначаемый λx.

Эти операции должны удовлетворять следующим условиям:

I. 1°. $x + y = y + x$ — коммутативность.

2°. $(x + y) + z = x + (y + z)$ — ассоциативность.

3°. Существует элемент 0 такой, что $x + 0 = x \ \forall\, x \in X$. Элемент 0 называется *нулевым* элементом.

4°. Для любого элемента x существует элемент, обозначаемый через $-x$, такой, что $x + (-x) = 0$.

II. 1°. $1 \cdot x = x$.

2°. $\alpha(\beta x) = \alpha\beta(x)$.

III. 1°. $(\alpha + \beta)x = \alpha x + \beta x$.

2°. $\alpha(x + y) = \alpha x + \alpha y$.

Линейное пространство X называется *нормированным*, если на X определен функционал $\|\cdot\|\colon X \to \mathbb{R}$, называемый *нормой* и удовлетворяющий условиям:

$$
\begin{aligned}
&a) \quad \|x\| \geq 0 \quad \forall\, x \in X \text{ и } \|x\| = 0 \Leftrightarrow x = 0;\\
&b) \quad \|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall\, \alpha \in \mathbb{R},\ \forall\, x \in X;\\
&c) \quad \|x_1 + x_2\| \leq \|x_1\| + \|x_2\| \quad \forall\, x_1, x_2 \in X.
\end{aligned}
$$

Линейное нормированное пространство иногда будем называть для краткости нормированным пространством. Иногда, чтобы подчеркнуть, что норма задана именно на пространстве X, мы пишем $\|x\|_X$. Две нормы $\|x\|_1$ и $\|x\|_2$ на X называются *эквивалентными*, если существуют положительные константы C_1 и C_2 такие, что

$$C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1.$$

Всякое нормированное пространство становится *метрическим*, если в нем ввести расстояние $d(x_1, x_2) = \|x_1 - x_2\|$. В метрическом пространстве с помощью открытых и замкнутых шаров естественным образом вводятся понятия открытых и замкнутых множеств, сходимость.

Последовательность точек $\{x_n\}_{n=1}^{\infty}$ метрического пространства называется *фундаментальной*, если она удовлетворяет *критерию Коши*, т. е. если для любого $\varepsilon > 0$ существует N_ε такое, что $d(x_{n_1}, x_{n_2}) < \varepsilon$ для всех $n_1, n_2 > N_\varepsilon$.

Метрическое пространство называется *полным*, если любая фундаментальная последовательность сходится.

Полное относительно введенного расстояния метрическое пространство называется *банаховым* пространством.

Отметим, что всякое конечномерное нормированное пространство является банаховым. Бесконечномерное нормированное пространство не обязано быть банаховым (приведите пример!).

5.1.2 Произведение пространств

Пусть X и Y — линейные нормированные пространства. Декартово произведение $X \times Y$ можно превратить в линейное нормированное пространство, введя норму

$$\|(x, y)\|_{X \times Y} = \max\{\|x\|_X, \|y\|_Y\}.$$

Легко проверить, что все аксиомы нормы выполняются. Возможны и другие эквивалентные нормировки.

Отметим очевидное утверждение: *декартово произведение банаховых пространств банахово.*

5.1.3 Примеры банаховых пространств

1. Конечномерное пространство $\mathbb{R}^n$, состоящее из векторов $x = (x_1, \ldots, x_n)$ с нормой

$$\|x\| = |x| = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}}.$$

Эту норму иногда называют *евклидовой нормой*, а расстояние, вводимое с помощью этой нормы, называют *евклидовым расстоянием.*

2. Конечномерное пространство l_p, $1 \le p \le \infty$, состоящее из векторов $x = (x_1, \ldots, x_n)$ с нормой

$$\|x\|_{l_p} = \begin{cases} \left(\sum\limits_{j=1}^{n} x_j^p\right)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \max\{|x_1|, \ldots, |x_n|\}, & p = \infty. \end{cases}$$

Отметим, что в конечномерном пространстве все нормы эквивалентны.

3. Бесконечномерное пространство l_2, состоящее из последовательностей $x = \{x_n\}_{n=1}^{\infty}$ (иногда пишем, $x = (x_1, \ldots, x_n, \ldots)$), для которых $\sum\limits_{n=1}^{\infty} x_n^2 < \infty$, с нормой

$$\|x\|_{l_2} = \left(\sum_{n=1}^{\infty} x_n^2\right)^{\frac{1}{2}}.$$

4. Пространство $C([t_0, t_1]) := C([t_0, t_1], \mathbb{R})$ непрерывных на отрезке $[t_0, t_1]$ функций $x(\cdot)$ с нормой

$$\|x(\cdot)\|_{C([t_0, t_1])} = \max_{t \in [t_0, t_1]} |x(t)|.$$

Обобщением этого пространства является пространство $C(K, \mathbb{R}^n)$ непрерывных вектор-функций $x(\cdot)\colon K \to \mathbb{R}^n$, заданных на компакте K, с нормой

$$\|x(\cdot)\|_{C(K)} = \max_{t \in K} |x(t)|.$$

5. Пространство $C^1([t_0, t_1]) := C^1([t_0, t_1], \mathbb{R})$ непрерывно дифференцируемых на отрезке $[t_0, t_1]$ функций $x(\cdot)$ с нормой

$$\|x(\cdot)\|_{C^1([t_0, t_1])} = \max\left\{\|x(\cdot)\|_{C([t_0, t_1])}, \|\dot{x}(\cdot)\|_{C([t_0, t_1])}\right\}.$$

Обобщением этого пространства является пространство $C^r(K, \mathbb{R}^n)$ r раз непрерывно дифференцируемых вектор-функций $x(\cdot)\colon K \to \mathbb{R}^n$, заданных на компакте K с нормой

$$\|x(\cdot)\|_{C^r(K)} = \max\left\{\|x(\cdot)\|_{C(K)}, \|\dot{x}(\cdot)\|_{C(K)}, \ldots, \|x^{(r)}(\cdot)\|_{C(K)}\right\}.$$

5.1.4 Сопряженное пространство, сопряженный оператор

Совокупность X^* всех линейных непрерывных функционалов на нормированном пространстве X образует *сопряженное* к X пространство. Оно является банаховым пространством относительно нормы

$$\|x^*\|_{X^*} := \max_{\|x\|_X \le 1} \langle x^*, x\rangle,$$

где $\langle x^*, x\rangle$ означает действие функционала x^* на элемент x.

Пусть X и Y — нормированные пространства. Через $L(X, Y)$ обозначим пространство линейных непрерывных операторов из X в Y. Для оператора $\Lambda \in L(X, Y)$ можно определить *сопряженный оператор* $\Lambda^* \colon Y^* \to X^*$, действующий по формуле

$$\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle \quad \forall\ x \in X.$$

Для линейного непрерывного функционала на произведении пространств имеет место следующая очевидная

Лемма. *Всякий функционал $\Lambda \in (X \times Y)^*$ однозначно представляется в виде*

$$\langle \Lambda, (x, y)\rangle = \langle x^*, x\rangle + \langle y^*, y\rangle,$$

где $x^* \in X^*$, $y^* \in Y^*$.

5.2 Определения производных

Для вещественнозначных функций одного вещественного переменного два определения — существование конечного предела

$$\lim_{h\to 0}\frac{f(\hat{x}+h)-f(\hat{x})}{h} \tag{1}$$

и возможность асимптотического разложения при $h\to 0$

$$f(\hat{x}+h)=f(\hat{x})+f'(\hat{x})h+o(h) \tag{2}$$

— приводят к одному и тому же понятию дифференцируемости. Но уже для функций двух и большего числа переменных существует несколько различных подходов к понятию дифференцируемости (гладкости). Определение (1) ведет к понятиям производной по направлению, вариации по Лагранжу и производной Гато. Определение (2) ведет к понятиям производной Гато, Фреше и строгой дифференцируемости.

Пусть далее в этом пункте X, Y, Z — линейные нормированные пространства. Как правило (если это не оговорено иначе), $f\colon X\to Y$ — отображение пространства X или некоторой окрестности точки $\hat{x}\in X$ в пространство Y.[1]

5.2.1 Производная по направлению

Будем говорить, что отображение f имеет в точке $\hat{x}$ *производную по направлению* h, если существует предел справа

$$\lim_{\lambda\to +0}\frac{f(\hat{x}+\lambda h)-f(\hat{x})}{\lambda} =: \delta_+ f(\hat{x},h),$$

который обозначается $\delta_+ f(\hat{x},h)$. Плюс в индексе здесь указывает на то, что берется именно предел справа.

5.2.2 Вариация по Лагранжу

Если для отображения f в точке $\hat{x}$ существует предел

$$\delta f(\hat{x},h) := \lim_{\lambda\to 0}\frac{f(\hat{x}+\lambda h)-f(\hat{x})}{\lambda} \quad \forall\ h\in X,$$

[1] Но вполне содержателен пример, когда $X=\mathbb{R}^n$, $Y=\mathbb{R}^m$. Элемент из $L(X,Y)$ определяется в этом случае матрицей размера $n\times m$.

то говорят, что отображение f *имеет в точке* $\hat{x}$ *вариацию по Лагранжу*. При этом отображение $h \to \delta f(\hat{x}, h)$ называют *вариацией по Лагранжу*, $\delta f(\hat{x}, \cdot)\colon X \to Y$.

Таким образом, для существования вариации по Лагранжу необходимо существование в точке $\hat{x}$ производной по всем направлениям h и выполнения равенств $\delta_+ f(\hat{x}, h) = -\delta_+ f(\hat{x}, -h)\ \forall\ h \in X$.

5.2.3 Производная Гато

Если оператор вариации по Лагранжу $\delta f(\hat{x}, \cdot)\colon X \to Y$ линеен и непрерывен по h ($\delta f(\hat{x}, \cdot) \in L(X, Y)$), то говорят, что отображение f *дифференцируемо по Гато в точке* $\hat{x}$, а оператор $\delta f(\hat{x}, \cdot)$ называется *производной Гато отображения* f *в точке* $\hat{x}$ и обозначается $f'_G(\hat{x})$.

Отсюда следует, что если отображение f дифференцируемо по Гато в точке $\hat{x}$, то для любого фиксированного вектора h имеет место разложение

$$f(\hat{x} + \lambda h) = f(\hat{x}) + \lambda f'_G(\hat{x})[h] + r(h, \lambda),$$

где $\|r(h, \lambda)\|_Y = o(|\lambda|)$ при $\lambda \to 0$.

Отметим, что отображение, дифференцируемое по Гато в точке $\hat{x}$, не обязано быть непрерывным в этой точке (см. Пример 4, п. 5.2.8).

5.2.4 Производная Фреше

Отображение f называют *дифференцируемым по Фреше в точке* $\hat{x}$ и пишут $f \in D(\hat{x})$, если существуют линейный непрерывный оператор $f'(\hat{x})\colon X \to Y$ и отображение $r\colon X \to Y$ такие, что

$$f(\hat{x} + h) = f(\hat{x}) + f'(\hat{x})[h] + r(h), \qquad (*)$$

где $\|r(h)\|_Y = o(\|h\|_X)$ при $\|h\|_X \to 0$. Оператор $f'(\hat{x})$ называется *производной Фреше*. Это разложение можно кратко записать так:

$$f(\hat{x} + h) = f(\hat{x}) + f'(\hat{x})[h] + o(h),$$

понимая $o(h)$ как элемент пространства Y, для которого $\|o(h)\|_Y = o(\|h\|)$ при $\|h\| \to 0$. Через $f'(\hat{x})[h]$ обозначено значение отображения $f'(\hat{x})$ на элементе h.

Из разложения $(*)$ следует, что функция, дифференцируемая по Фреше, непрерывна в точке дифференцируемости, а также дифференцируема в этой точке по Гато. Уже в двумерном случае эти два понятия различаются: из дифференцируемости функции по Гато не следует ее дифференцируемость по Фреше (см. Пример 4, п. 5.2.8). Аналогично из дифференцируемости по Гато по определению вытекает существование вариации по Лагранжу. И снова (уже в двумерном случае) эти понятия различны. Производная Фреше линейного оператора совпадает с самим оператором.

Если в каждой точке x открытого множества U отображение $f \in D(x)$, а отображение $x \to f'(x)$ непрерывно, то пишем $f \in C^1(U)$.

На языке ε–δ определение дифференцируемости по Фреше отображения f в точке $\hat{x}$ формулируется так: существует оператор $f'(\hat{x}) \in L(X, Y)$ такой, что для любого $\varepsilon > 0$ найдется $\delta > 0$, при котором для любого $\|h\| < \delta$ выполняется неравенство

$$\|f(\hat{x} + h) - f(\hat{x}) - f'(\hat{x})[h]\|_Y \leq \varepsilon \|h\|_X.$$

Из разложения $(*)$ следует, что производная Фреше определена однозначно, ибо равенство $\Lambda_1 h - \Lambda_2 h = o(h)$ для линейных непрерывных операторов Λ_1 и Λ_2 возможно лишь при $\Lambda_1 = \Lambda_2$.

Для функций одной переменной производная по Фреше совпадает с обычной производной, которая изучается в курсе математического анализа.

5.2.5 Строгая дифференцируемость

Во многих задачах конечномерного и бесконечномерного анализа дифференцируемости по Фреше в точке недостаточно для получения содержательного результата. Это побуждает к следующему усилению дифференцируемости в точке.

Пусть отображение f дифференцируемо по Фреше в точке $\hat{x}$. Оно называется *строго дифференцируемым в точке* $\hat{x}$ (при этом пишут $f \in SD(\hat{x})$), если для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех x_1 и x_2, удовлетворяющих неравенствам $\|x_1 - \hat{x}\| < \delta$, $\|x_2 - \hat{x}\| < \delta$, выполнено неравенство

$$\|f(x_1) - f(x_2) - f'(\hat{x})[x_1 - x_2]\|_Y \leq \varepsilon \|x_1 - x_2\|_X.$$

Из определения следует, что функция, строго дифференцируемая по Фреше в точке, непрерывна в некоторой окрестности этой точки.

Ниже будет показано, что если производная Фреше (даже производная Гато) отображения непрерывна в точке, то отображение будет строго дифференцируемо в этой точке.

5.2.6 Частные производные

Пусть X, Y, Z — нормированные пространства. Рассмотрим отображение $f\colon X \times Y \to Z$, $(\hat{x}, \hat{y}) \in X \times Y$. Если отображение $x \to f(x, \hat{y})$ дифференцируемо в точке $\hat{x}$ по Фреше, то его производная называется *частной производной по x отображения f в точке* $(\hat{x}, \hat{y})$ и обозначается $f'_x(\hat{x}, \hat{y})$ или $\dfrac{\partial f(\hat{x}, \hat{y})}{\partial x}$. Аналогично определяется частная производная по y $\quad f'_y(\hat{x}, \hat{y}) = \dfrac{\partial f(\hat{x}, \hat{y})}{\partial y}$.

5.2.7 Производные высших порядков

Дадим теперь определение второй производной Фреше. Если отображение $f\colon X \to Y$ дифференцируемо в каждой точке $x \in X$, то определено отображение $x \to f'(x)$ пространства X в пространство $L(X, Y)$. Поскольку $L(X, Y)$ также является линейным нормированным пространством, то можно определять производную Фреше этого отображения $f'\colon X \to L(X, Y)$, которая и называется *второй производной* отображения f:

$$f''(\hat{x}) = (f')'(\hat{x}) \in L\big(X, L(X, Y)\big).$$

Для вектора $h_1 \in X$ определен оператор $f''(\hat{x})[h_1] \in L(X, Y)$. Для $h_2 \in X$ определен элемент $f''(\hat{x})[h_1][h_2] \in Y$. Тем самым определено отображение $f''(\hat{x})[h_1, h_2] := f''(\hat{x})[h_1][h_2]$. Значит, определено билинейное (линейное по каждому аргументу) отображение $f''(\hat{x})\colon X \times X \to Y$.

Аналогично определяются производные высших порядков.

Теорема (о смешанных производных). [АТФ, с. 156] *Если отображение f дважды дифференцируемо в точке $\hat{x}$ $\big(f \in D^2(\hat{x})\big)$, то*

$$f''(\hat{x})[h_1, h_2] = f''(\hat{x})[h_2, h_1] \quad \forall\ h_1, h_2 \in X.$$

5.2.8 Контрпримеры на дифференцируемость

Приведем несколько контрпримеров, показывающих, что введенные выше понятия дифференцируемости действительно различны.

Пример 1. *Непрерывная функция не имеет в фиксированной точке производной ни по какому направлению:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \begin{cases} x \sin \dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases} \quad \hat{x} = 0.$$

На прямой $\mathbb{R}$ имеются с точностью до умножения на положительную константу два направления: $h_1 = 1$ и $h_2 = -1$. Однако производных ни по какому направлению не существует. Действительно, предел

$$\delta_+ f(\hat{x}, h)\Big|_{\substack{\hat{x}=0\\ h=1}} = \lim_{\lambda \to +0} \frac{f(\hat{x} + \lambda h) - f(\hat{x})}{\lambda}\Big|_{\substack{\hat{x}=0\\ h=1}} = \lim_{\lambda \to +0} \frac{\lambda \sin \frac{1}{\lambda}}{\lambda} = \lim_{\lambda \to +0} \sin \frac{1}{\lambda}$$

не существует. В силу четности функции f не существует также предел по направлению $h = -1$.

Пример 2. *Непрерывная функция имеет в фиксированной точке производную по всем направлениям, но не имеет в этой точке вариации по Лагранжу:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = |x|,\ \hat{x} = 0.$$

Как и в примере 1 на прямой $\mathbb{R}$ имеются с точностью до умножения на положительную константу два направления: $h_1 = 1$ и $h_2 = -1$. Пределы по обоим направлениям существуют:

$$\delta_+ f(\hat{x}, h)\Big|_{\substack{\hat{x}=0\\ h=h_i}} = \lim_{\lambda \to +0} \frac{f(\hat{x} + \lambda h) - f(\hat{x})}{\lambda}\Big|_{\substack{\hat{x}=0\\ h=h_i}} = \lim_{\lambda \to +0} \frac{|\lambda h_i|}{\lambda} = \lim_{\lambda \to +0} \frac{|\lambda|}{\lambda} = 1,$$

$i = 1, 2$. Но предел

$$\delta f(\hat{x}, h)\Big|_{\substack{\hat{x}=0\\ h=1}} = \lim_{\lambda \to 0} \frac{f(\hat{x} + \lambda h) - f(\hat{x})}{\lambda}\Big|_{\substack{\hat{x}=0\\ h=1}} = \lim_{\lambda \to 0} \frac{|\lambda|}{\lambda}$$

не существует. Следовательно, отображение f не имеет в точке $\hat{x} = 0$ вариации по Лагранжу.

Отметим, что если $\hat{x} \neq 0$, то $\delta f(\hat{x}, h) = f'(\hat{x})[h] = sign\, \hat{x} \cdot h$.

Пример 3. *Отображение имеет в фиксированной точке вариацию по Лагранжу, непрерывно в этой точке, но не имеет в этой точке производной Гато.*

Определим отображение в полярных координатах $x = (x_1, x_2) = (r\cos\varphi,\ r\sin\varphi)$ по формуле: $f\colon \mathbb{R}^2 \to \mathbb{R},\ f(x) = r\cos 3\varphi,\ \hat{x} = 0.$

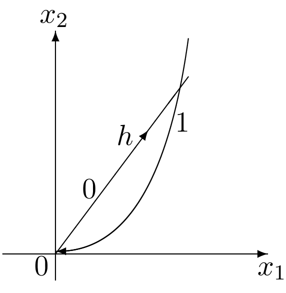
Вычислим вариацию по Лагранжу данного отображения в точке $\hat{x} = 0$. Возьмем произвольное направление $h = (t\cos\alpha, t\sin\alpha)$. Тогда

$$\delta f(\hat{x}, h) = \lim_{\lambda\to 0} \frac{f(\hat{x}+\lambda h) - f(\hat{x})}{\lambda}\bigg|_{\hat{x}=0} = \lim_{\lambda\to 0}\frac{\lambda t\cos 3\alpha}{\lambda} = t\cos 3\alpha.$$

Покажем, что вариация по Лагранжу $\delta f(\hat{x}, h)$ не является линейным оператором по h. Действительно, возьмем два вектора $h_1 = (1,\,0) = (\cos 0, \sin 0)$ $(t = 1, \alpha = 0)$ и $h_2 = (0,\,1) = \left(\cos\dfrac{\pi}{2}, \sin\dfrac{\pi}{2}\right)$ $\left(t = 1, \alpha = \dfrac{\pi}{2}\right)$. Тогда $h_1 + h_2 = (1,1) = \left(\sqrt{2}\cos\dfrac{\pi}{4}, \sqrt{2}\sin\dfrac{\pi}{4}\right)$. Однако $\delta f(\hat{x}, h_1 + h_2) = \sqrt{2}\cos\dfrac{3\pi}{4} = -1 \neq \delta f(\hat{x}, h_1) + \delta f(\hat{x}, h_2) = \cos 0 + \cos\dfrac{3\pi}{2} = 1.$

Пример 4. *Функция f имеет в фиксированной точке производную Гато, но не имеет в этой точке производной Фреше:*

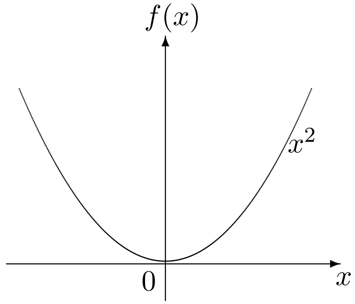
$$f\colon \mathbb{R}^2 \to \mathbb{R},\ f(x) = \begin{cases} 1, & x_2 = x_1^2,\ x_1 > 0, \\ 0, & \text{в остальных случаях,}\end{cases} \quad \hat{x} = (0,0).$$



Поскольку $\delta f(\hat{x}, h) = \lim\limits_{\lambda\to 0}\dfrac{f(\hat{x}+\lambda h) - f(\hat{x})}{\lambda} = \lim\limits_{\lambda\to 0}\dfrac{f(\lambda h)}{\lambda} = 0$ для любого h, то производная Гато существует и $f'_G(\hat{x}) = 0$. С другой стороны, функция f разрывна в точке $\hat{x} = (0,0)$, а функция, дифференцируемая по Фреше, должна быть непрерывна в точке дифференцируемости.

Пример 5. *Функция имеет в фиксированной точке производную Фреше, но не строго дифференцируема в этой точке:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \begin{cases} x^2, & x \text{ рационально}, \\ 0, & x \text{ иррационально},\end{cases} \quad \hat{x} = 0.$$



Выписанная функция одной (!) переменной дифференцируема в точке $\hat{x} = 0$. Значит она дифференцируема по Фреше в этой точке. С другой стороны, функция имеет разрывы в любой окрестности нуля, а строго дифференцируемая функция должна быть непрерывна в некоторой окрестности $\hat{x}$.

5.3 Некоторые теоремы дифференциального исчисления в нормированных пространствах.

Приведем несколько теорем, наиболее часто используемых для решения экстремальных задач.

5.3.1 Теорема о суперпозиции

Теорема. *Пусть X, Y, Z — линейные нормированные пространства, $\varphi\colon X \to Y$, $\psi\colon Y \to Z$, $\varphi(\hat{x}) = \hat{y}$, $f = \psi \circ \varphi\colon X \to Z$ — суперпозиция отображений φ и ψ. Тогда, если отображение ψ дифференцируемо по Фреше в точке $\hat{y}$, а отображение φ в точке $\hat{x}$ имеет вариацию по Лагранжу (дифференцируемо по Гато, дифференцируемо по Фреше), то отображение f обладает в точке $\hat{x}$ тем же свойством, что и отображение φ, и при этом соответственно*

A) $\delta f(\hat{x}, h) = \psi'(\hat{y})[\delta\varphi(\hat{x}, h)] \quad \forall\ h \in X$,

B) $f'_G(\hat{x}) = \psi'(\hat{y}) \circ \varphi'_G(\hat{x})$,

C) $f'(\hat{x}) = \psi'(\hat{y}) \circ \varphi'(\hat{x}) \quad (\Leftrightarrow f'(\hat{x})[h] = \psi'(\hat{y})[\varphi'(\hat{x})[h]])$,

D) Если отображение $\psi \in SD(\hat{y})$ строго дифференцируемо в точке $\hat{y}$, а отображение $\varphi \in SD(\hat{x})$ строго дифференцируемо в точке $\hat{x}$, то отображение $f \in SD(\hat{x})$ строго дифференцируемо в точке $\hat{x}$.

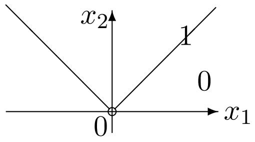
Замечание. *Теорема о суперпозиции не имеет места для производной Гато.*

Доказательство. Пусть $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\varphi(x_1, x_2) = (x_1^2, x_2)$, $\hat{x} = 0$, $\varphi(\hat{x}) = \hat{y} = 0$, $\psi\colon \mathbb{R}^2 \to \mathbb{R}$,

$$\psi(y_1, y_2) = \begin{cases} 1, & y_1 = y_2^2,\ y_2 > 0, \\ 0, & \text{в остальных случаях.} \end{cases}$$

Функция φ дифференцируема по Фреше в точке $\hat{x}$ и даже строго дифференцируема (проверьте!), функция ψ дифференцируема по Гато в точке $\hat{y}$ (см. Пример 4, п. 5.2.8). С другой стороны, функция

$$f(x) = \psi\circ\varphi(x) = \psi(x_1^2, x_2) = \begin{cases} 1, & x_1^2 = x_2^2,\ x_2 > 0 \Leftrightarrow x_2 = |x_1|,\ x_2 > 0, \\ 0, & \text{в остальных случаях,} \end{cases}$$



являющаяся суперпозицией отображений φ и ψ, не дифференцируема по Гато в нуле (и даже не имеет в этой точке производных по направлениям $h = (1, 1)$ и $h = (-1, 1)$). $\triangleright$

5.3.2 Формула Тейлора

Теорема. [АТФ, с. 159] *Пусть X, Y — нормированные пространства, отображение $f : X \to Y$, причем отображение f n раз дифференцируемо по Фреше в точке $\hat{x}$ ($f \in D^n(\hat{x})$). Тогда имеет место разложение в ряд Тейлора*

$$f(\hat{x}+h) = f(\hat{x})+f'(\hat{x})[h]+\frac{1}{2}f''(\hat{x})[h,h]+\dots+\frac{1}{n!}f^{(n)}(\hat{x})[h,\dots,h]+r(h),$$

где $\|r(h)\| = o(\|h\|^n)$ при $h \to 0$.

5.3.3 Теорема о среднем

Хорошо известно, что для числовых функций одного переменного справедлива следующая теорема Лагранжа, называемая иногда также теоремой о среднем значении или формулой конечных приращений:

Теорема Лагранжа. *Если функция $f\colon [a, b] \to \mathbb{R}$ непрерывна на отрезке $[a, b]$ и дифференцируема в интервале (a, b), то существует точка $c \in (a, b)$ такая, что*

$$f(b) - f(a) = f'(c)(b - a). \qquad (*)$$

Замечание 1. *Формула $(*)$ остается справедливой и для числовых функций $f(x)$, аргумент которых принадлежит произвольному линейному нормированному пространству. Дифференцируе-*

мость понимается в смысле Гато. ($f\colon X \to \mathbb{R}$, X — л.н.п., $f \in C[a,\, b]^2, D_G(a,\, b)^3 \Rightarrow (*)$)

Доказательство. Полагая $\varphi(t) := f(a + t(b-a))$, мы сводим доказательство к случаю функции одной переменной:

$$f(b)-f(a) = \varphi(1)-\varphi(0) = \varphi'(\theta) = f'(a+\theta(b-a))(b-a) = f'(c)(b-a),$$

где $\theta \in (0,\, 1)$, $c = a + \theta(b-a) \in (a,\, b)$. $\rhd$

Замечание 2. *Для векторнозначных функций теорема Лагранжа неверна.*

Доказательство. Пусть $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(x) = (\sin x, -\cos x)$. Тогда $f'(x)[h] = (\cos x, \sin x)[h] = (h\cos x, h\sin x)$, $h \in \mathbb{R}$. В то же время

$f(2\pi) - f(0) = (\sin 2\pi, -\cos 2\pi) - (\sin 0, -\cos 0) = (0, 0)$,

$f'(c)[2\pi - 0] = 2\pi(\cos c, \sin c)$.

Равенства $f(2\pi) - f(0) = f'(c)[2\pi - 0] \Leftrightarrow (0,0) = 2\pi(\cos c, \sin c)$ не может быть ни для какого c, так как $\cos c$ и $\sin c$ одновременно в ноль не обращаются. Значит, формула $(*)$ для функции f не имеет места. $\rhd$

Отметим, что в анализе, как правило, используется не сама формула $(*)$, а вытекающая из нее оценка

$$|f(b) - f(a)| \leq \sup_{c \in (a,\, b)} |f'(c)| \cdot |b-a|$$

Покажем, что в этом более слабом виде утверждение распространяется уже на случай произвольных нормированных пространств. По традиции оно сохраняет название “теорема о среднем”, хотя, конечно, его следовало бы назвать “теоремой об оценке конечного приращения”.

Теорема о среднем. *Пусть X, Y — линейные нормированные пространства, отображение $f\colon X \to Y$ дифференцируемо по Гато в каждой точке отрезка $[a,\, b]$. Тогда*

$$\|f(b) - f(a)\| \leq \sup_{c \in (a,\, b)} \|f'_G(c)\| \cdot \|b-a\|.$$

[2] Отрезок $[a,\, b] = \{x \in X \mid x = a + t(b-a),\ 0 \leq t \leq 1\}$.

[3] Интервал $(a,\, b) = \{x \in X \mid x = a + t(b-a),\ 0 < t < 1\}$.

Доказательство. По лемме Банаха (см. п. 5.4) для любого $y \in Y$, а значит, и для $y = f(b) - f(a)$ найдется элемент $y^* \in Y^*$ такой, что $\|y^*\| = 1$ и $\langle y^*, y\rangle = \|y\|$, т. е. $\langle y^*, f(b) - f(a)\rangle = \|f(b) - f(a)\|$.

Обозначим $\varphi(t) = \langle y^*, f(a + t(b - a))\rangle$. Поскольку y^* — линейный непрерывный функционал, а отображение f в каждой точке отрезка $[a,\, b]$ имеет производную Гато, то по теореме о суперпозиции, пользуясь тем, что производная Фреше линейного непрерывного функционала совпадает с ним самим, получим

$$\varphi'(t) = \langle y^*, f'_G(a + t(b - a))[b - a]\rangle \quad \forall\, t \in [0,\, 1].$$

Из дифференцируемости функции φ следует ее непрерывность на отрезке $[0,\, 1]$ и, следовательно, к ней можно применить формулу Лагранжа: $\varphi(1) - \varphi(0) = \varphi'(\theta),\ \theta \in (0,\, 1)$. Поэтому

$$\|f(b) - f(a)\| = \langle y^*, f(b) - f(a)\rangle = \varphi(1) - \varphi(0) = \varphi'(\theta) =$$

$$= \langle y^*, f'_G(a + \theta(b - a))[b - a]\rangle \le \|y^*\| \cdot \|f'_G(a + \theta(b - a))[b - a]\| =$$

$$\overset{\|y^*\|=1}{=} \|f'_G(a + \theta(b - a))[b - a]\| \le$$

$$\le \|f'_G(a + \theta(b - a))\| \cdot \|b - a\| \le \sup_{c \in (a,\, b)} \|f'_G(c)\| \cdot \|b - a\|.$$

Приведем несколько следствий из теоремы о среднем.

Следствие 1. *Пусть выполнены все условия теоремы о среднем* $(f \in D_G[a,\, b])$ *и оператор* $\Lambda \in L(X, Y)$. *Тогда*

$$\|f(b) - f(a) - \Lambda(b - a)\| \le \sup_{c \in (a,\, b)} \|f'_G(c) - \Lambda\| \cdot \|b - a\|.$$

Доказательство. Надо применить теорему о среднем к отображению $g = f - \Lambda$.

Следствие 2. *Пусть* X, Y *— линейные нормированные пространства, отображение* $f\colon X \to Y$ *дифференцируемо по Гато в некоторой окрестности точки* $\hat{x}$, *отображение* $x \to f'_G(x)$ *непрерывно в точке* $\hat{x}$. *Тогда отображение* f *строго дифференцируемо в точке* $\hat{x}$ (а следовательно, и дифференцируемо по Фреше в той же точке).

Доказательство. В силу непрерывности отображения $x \to f'_G(x)$ в точке $\hat{x}$ для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что $\|f'_G(x) - f'_G(\hat{x})\| < \varepsilon$ при $\|x - \hat{x}\| < \delta$.

В силу выпуклости шара $B := \overset{\circ}{B}(\hat{x}, \delta) = \{x \mid \|x - \hat{x}\| < \delta\}$ из условия $x_1, x_2 \in B$ следует, что отрезок $[x_1,\, x_2] \in B$. По следствию 1 теоремы о среднем с $\Lambda = f'_G(\hat{x})$ получаем

$$\|f(x_1) - f(x_2) - f'_G(\hat{x})[x_1 - x_2]\| \leq$$

$$\leq \sup_{x \in (x_1,\, x_2)} \|f'_G(x) - f'_G(\hat{x})\| \cdot \|x_1 - x_2\| \leq \varepsilon \|x_1 - x_2\|,$$

что означает строгую дифференцируемость отображения f в точке $\hat{x}$. $\triangleright$

Следствие 2 показывает, что при проверке дифференцируемости конкретного функционала достаточно доказать существование производной Гато в некоторой окрестности точки $\hat{x}$ и проверить ее непрерывность в точке $\hat{x}$. Это гарантирует строгую дифференцируемость в точке $\hat{x}$ и тем более существование производной Фреше.

Следствие 3. *Существование второй производной Фреше в точке гарантирует строгую дифференцируемость отображения в этой точке* ($f \in D^2(\hat{x}) \Rightarrow f \in SD(\hat{x})$).

Доказательство. Покажем, что все условия для выполнения Следствия 2 имеются. Действительно, из существования второй производной Фреше в точке $\hat{x}$ следует существование производной Фреше (и, следовательно, существование производной Гато) в некоторой окрестности этой точки, а также непрерывность производной Фреше (и, следовательно, непрерывность производной Гато) в точке $\hat{x}$. Тогда по Следствию 2 вытекает строгая дифференцируемость отображения в этой точке. $\triangleright$

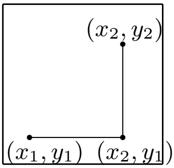
5.3.4 Теорема о полном дифференциале

Теорема. *Пусть X, Y, Z — линейные нормированные пространства, отображение $F\colon X \times Y \to Z$ имеет в каждой точке (x, y) из некоторой окрестности точки $(\hat{x}, \hat{y}) \in X \times Y$ частные производные $F_x(x, y)$ и $F_y(x, y)$ в смысле Гато, являющиеся непрерывными по совокупности переменных в точке $(\hat{x}, \hat{y})$. Тогда отображение $F \in SD(\hat{x}, \hat{y})$ строго дифференцируемо в той же точке и при этом*

$$F'(\hat{x}, \hat{y})[(\xi, \eta)] = F_x(\hat{x}, \hat{y})[\xi] + F_y(\hat{x}, \hat{y})[\eta].$$

Доказательство. В силу непрерывности отображений $F_x(x,y)$ и $F_y(x,y)$ в точке $(\hat{x},\hat{y})$ для любого $\varepsilon > 0$ можно найти $\delta > 0$ такое, что для любой точки (x,y) из "прямоугольной" окрестности $V := \overset{\circ}{B}(\hat{x},\delta) \times \overset{\circ}{B}(\hat{y},\delta)$ точки $(\hat{x},\hat{y})$ существуют частные производные $F_x(x,y)$ и $F_y(x,y)$ в смысле Гато и выполняются неравенства

$$\|F_x(x,y) - F_x(\hat{x},\hat{y})\| < \varepsilon, \quad \|F_y(x,y) - F_y(\hat{x},\hat{y})\| < \varepsilon. \qquad (*)$$



Если точки (x_1,y_1), $(x_2,y_2) \in V$, то и точка $(x_2,y_1) \in V$ и оба отрезка $[(x_1,y_1),(x_2,y_1)]$, $[(x_2,y_1),(x_2,y_2)]$ $\in V$ в силу выпуклости множества V. Поэтому отображения $x \to F(x,y_1)$ и $y \to F(x_2,y)$ дифференцируемы и имеют производные $F_x(x,y_1)$ на отрезке $[x_1,x_2]$, $F_y(x_2,y)$ на отрезке $[y_1,y_2]$.

Рассмотрим норму разности

$$\|F(x_1,y_1) - F(x_2,y_2) - F_x(\hat{x},\hat{y})[x_1-x_2] - F_y(\hat{x},\hat{y})[y_1-y_2]\| =$$

(вычтем и добавим $F(x_2,y_1)$)

$$= \|F(x_1,y_1) - F(x_2,y_1) - F_x(\hat{x},\hat{y})[x_1-x_2] +$$

$$+ F(x_2,y_1) - F(x_2,y_2) - F_y(\hat{x},\hat{y})[y_1-y_2]\| \leq$$

(по неравенству треугольника для норм и Следствию 1 Теоремы о среднем)

$$\leq \sup_{x\in(x_1,x_2)} \|F_x(x,y_1) - F_x(\hat{x},\hat{y})\| \cdot \|x_1-x_2\| +$$

$$+ \sup_{y\in(y_1,y_2)} \|F_y(x_2,y) - F_y(\hat{x},\hat{y})\| \cdot \|y_1-y_2\| \leq \varepsilon\|x_1-x_2\| + \varepsilon\|y_1-y_2\|$$

для любых $(x_1,y_1),(x_2,y_2) \in V$, что и означает строгую дифференцируемость отображения F в точке $(\hat{x},\hat{y})$ и дает явный вид производной Фреше (и полного дифференциала) отображения от двух переменных. $\triangleright$

5.4 Дополнительные сведения из алгебры и функционального анализа

В этом пункте приводятся дополнительные сведения из алгебры и функционального анализа, которые понадобятся для доказательства теорем об условиях экстремума в гладких экстремальных задачах в нормированных пространствах.

Определение. *Аннулятором* $A^{\perp}$ множества A линейного нормированного пространства X называется множество линейных непрерывных функционалов x^*, для которых $\langle x^*, x\rangle = 0$ для любого $x \in A$:

$$A^{\perp} := \{x^* \in X^* \mid \langle x^*, x\rangle = 0 \quad \forall\, x \in A\}.$$

Отметим, что аннулятор $A^{\perp}$ всегда содержит нулевой элемент ($x^* = 0$) сопряженного пространства X^*.

Лемма о нетривиальности аннулятора. *Пусть* L *— замкнутое собственное* (т. е. $L \neq X$) подпространство линейного нормированного пространства X. Тогда аннулятор $L^{\perp}$ содержит ненулевой элемент $x^* \in X^*$.

Доказательство. Поскольку $L \neq X$, то существует точка $\hat{x} \notin L$. По теореме о строгой отделимости точки, не принадлежащей выпуклому замкнутому множеству, от множества существует линейный непрерывный функционал $x^* \in X^*$, строго разделяющий точку $\hat{x}$ и подпространство L (L — подпространство линейного пространства и, следовательно, выпукло)

$$\langle x^*, \hat{x}\rangle > \sup_{x \in L} \langle x^*, x\rangle. \qquad (*)$$

Из этого неравенства следует, что $x^* \neq 0$. Покажем, что $x^* \in L^{\perp}$, т. е. $\langle x^*, x\rangle = 0\ \forall\ x \in L$. Действительно, если бы существовало $x_0 \in L$, для которого $\langle x^*, x_0\rangle \neq 0$, то поскольку $\alpha x_0 \in L$ для любого $\alpha \in \mathbb{R}$, было бы

$$\sup_{x \in L} \langle x^*, x\rangle \geq \sup_{\alpha \in \mathbb{R}} \langle x^*, \alpha x_0\rangle = +\infty.$$

Это не так в силу неравенства $(*)$. Следовательно, $\langle x^*, x\rangle = 0\ \forall\ x \in L$ и, поэтому $x^* \in L^{\perp}$. $\triangleright$

Далее нам понадобятся следующие две теоремы из функционального анализа и некоторые леммы.

Теорема Банаха об открытости. [КФ, с. 262; ГТ, с. 109] *Пусть X, Y — банаховы пространства, Λ — непрерывный линейный оператор из X в Y, являющийся эпиморфизмом $(\Lambda\colon X \overset{на}{\to} Y)$. Тогда образ каждого открытого множества в X открыт в Y.*

Теорема Банаха об обратном операторе. [КФ, с. 213] *Пусть X, Y — банаховы пространства, Λ — непрерывный линейный оператор из X в Y, являющийся эпиморфизмом и $Ker\,\Lambda = 0$. Тогда существуют обратный оператор $\Lambda^{-1}\colon Y \to X$, так же линейный и непрерывный.*

Лемма Банаха. [ГТ, с. 109] *Пусть X — линейное нормированное пространство, $x_0 \in X$. Тогда существует линейный непрерывный функционал $x^* \in X^*$ такой, что $\|x^*\| = 1$, $\langle x^*, x_0\rangle = \|x_0\|$.*

Лемма Банаха является следствием из известной теоремы Хана–Банаха о продолжении линейного функционала.

Лемма о правом обратном отображении. *Пусть X, Y — банаховы пространства, Λ — непрерывный линейный оператор из X в Y, являющийся эпиморфизмом. Тогда существуют отображение $M\colon Y \to X$* (необязательно непрерывное и необязательно линейное) *и константа $K > 0$ такие, что*

$$\Lambda \circ M = I_Y{}^4, \quad \|My\| \le K\|y\| \quad \forall\, y \in Y.$$

Доказательство. Обозначим $B_X := \{x \in X \mid \|x\| < 1\}$ — открытый шар радиуса 1 в пространстве X. По теореме Банаха об открытости образ открытого множества ΛB_X открыт. Поскольку при линейном отображении ноль переходит в ноль, то множество ΛB_X содержит внутренней точкой ноль, и, следовательно, содержит какой-то некоторый открытый шар с центром в нуле $\delta B_Y := \{y \in Y \mid \|y\| < \delta\}$. Значит, для любого $y \in \delta B_Y$ найдется $x(y)$ такой, что $\Lambda x(y) = y$, $\|x(y)\| < 1$. Для любого $y \in Y$, $y \neq 0$, обозначим $My := \dfrac{2\|y\|}{\delta} x\Big(\dfrac{\delta y}{2\|y\|}\Big)$. Тогда $\Big\|\dfrac{\delta y}{2\|y\|}\Big\| = \dfrac{\delta}{2} < \delta$ и из определения отображения M имеем:

$$\Lambda \circ My = \Lambda\bigg(\frac{2\|y\|}{\delta} x\Big(\frac{\delta y}{2\|y\|}\Big)\bigg) = \frac{2\|y\|}{\delta}\Lambda\bigg(x\Big(\frac{\delta y}{2\|y\|}\Big)\bigg) = \frac{2\|y\|}{\delta}\cdot\frac{\delta y}{2\|y\|} = y,$$

[4] Через I_Y обозначается тождественный оператор на пространстве Y.

$$\|My\| = \left\|\frac{2\|y\|}{\delta}x\Big(\frac{\delta y}{2\|y\|}\Big)\right\| = \frac{2\|y\|}{\delta}\cdot\left\|x\Big(\frac{\delta y}{2\|y\|}\Big)\right\| \le \frac{2}{\delta}\|y\|. \ \triangleright$$

Лемма о замкнутости образа. *Пусть X, Y, Z — банаховы пространства, $A\colon X \to Y$, $B\colon X \to Z$ — линейные непрерывные операторы, подпространство $\mathit{Im}\, A$ замкнуто в Y, подпространство $B\mathit{Ker}\, A$ замкнуто в Z, $C\colon X \to Y\times Z$, $Cx := (Ax, Bx)$. Тогда C — линейный непрерывный оператор и подпространство $\mathit{Im}\, C$ замкнуто в $Y \times Z$.*

Доказательство. Очевидно, что оператор C линеен и непрерывен. Докажем замкнутость его образа. Замкнутое подпространство $\widetilde{Y} = \operatorname{Im} A$ банахового пространства Y само банахово и, значит, отображение $A\colon X \to \widetilde{Y}$ — эпиморфизм. По лемме о правом обратном отображении существуют оператор $M\colon \widetilde{Y} \to X$ и константа $K > 0$ такие, что

$$A \circ M = I_{\widetilde{Y}}, \quad \|My\| \le K\|y\| \quad \forall\, y \in \widetilde{Y}.$$

Пусть точка $(y, z) \in \overline{\operatorname{Im} C}$ принадлежит замыканию образа оператора C. Это означает, что найдется последовательность $\{x_n\}$ такая, что последовательность $\{(Ax_n, Bx_n)\}$ сходится к (y, z). В силу замкнутости подпространства $\operatorname{Im} A$ предельная точка $y = \lim Ax_n \in \operatorname{Im} A = \widetilde{Y}$. Значит, определены точки $h_n := M(Ax_n - y)$. По свойству оператора M

$$A(x_n - h_n) = Ax_n - A(M(Ax_n - y)) \overset{A\circ M = I_{\widetilde{Y}}}{=} Ax_n - (Ax_n - y) = y,$$

$\|h_n\| = \|M(Ax_n - y)\| \le K\|Ax_n - y\| \to 0$. Поэтому $Bh_n \to 0$ в силу непрерывности оператора B и

$$z := \lim Bx_n = \lim B(x_n - h_n),$$

т. е. z принадлежит замыканию множества $\Sigma = \{z = Bx \mid Ax = y\}$. Это множество, как легко видеть, является сдвигом подпространства $B\operatorname{Ker} A$, следовательно, замкнуто. Итак, $z \in \overline{\Sigma} = \Sigma$. Это означает, что существует $x \in X \ :\ Ax = y$, $Bx = z$, т. е. $(y, z) \in \operatorname{Im} C$. $\triangleright$

Лемма об аннуляторе ядра регулярного оператора. *Пусть X, Y — банаховы пространства, $A\colon X \to Y$ — линейный непрерывный эпиморфизм. Тогда $(\mathit{Ker}\, A)^{\perp} = \mathit{Im}\, A^*$.*

Доказательство. A) Докажем, что $(\operatorname{Ker} A)^{\perp} \supset \operatorname{Im} A^*$. Возьмем $x^* \in \operatorname{Im} A^* \Rightarrow x^* = A^*y^*$. Следовательно,

$$\langle x^*, x\rangle = \langle A^*y^*, x\rangle = \langle y^*, Ax\rangle = 0 \quad \forall\, x \in \operatorname{Ker} A.$$

Значит, $x^* \in (\operatorname{Ker} A)^{\perp}$. Вложение доказано.

B) Докажем, что $(\operatorname{Ker} A)^{\perp} \subset \operatorname{Im} A^*$. Возьмем $x^* \in (\operatorname{Ker} A)^{\perp}$, т. е. $\langle x^*, x\rangle = 0\ \forall\, x \in \operatorname{Ker} A$. Применим лемму о замкнутости образа для пространств $X, Y, Z = \mathbb{R}$ и отображений A, $Bx := \langle x^*, x\rangle$. Условия леммы выполняются: подпространство $\operatorname{Im} A = Y$ замкнуто в Y, подпространство $B\operatorname{Ker} A = \langle x^*, \operatorname{Ker} A\rangle = 0$, состоящее из одного нуля, замкнуто в $Z = \mathbb{R}$. По лемме о замкнутости образа подпространство $\operatorname{Im} C = \operatorname{Im}\,(A, x^*)$ замкнуто в $Y \times Z = Y \times \mathbb{R}$. Подпространство $\operatorname{Im}\,(A, x^*)$ является собственным, так как точка $(0, 1) \notin \operatorname{Im}\,(A, x^*)$ (если $Ax = 0$, то $\langle x^*, x\rangle = 0 \neq 1$).

По лемме о нетривиальности аннулятора замкнутого собственного подпространства существует ненулевой линейный непрерывный функционал $(y^*, \lambda) \in (\operatorname{Im}\,(A, x^*))^{\perp} \subset (Y \times \mathbb{R})^* = Y^* \times \mathbb{R}^* = Y^* \times \mathbb{R}$ такой, что

$$\langle (y^*, \lambda), (Ax, \langle x^*, x\rangle)\rangle = 0 \iff \langle y^*, Ax\rangle + \lambda\langle x^*, x\rangle = 0 \iff$$

$$\langle A^*y^*, x\rangle + \lambda\langle x^*, x\rangle = 0 \iff \langle A^*y^* + \lambda x^*, x\rangle = 0 \quad \forall\, x \in X$$

$$\iff A^*y^* + \lambda x^* = 0.$$

Причем $\lambda \neq 0$ (ибо иначе $\langle y^*, Ax\rangle = 0\ \forall\, x \in X \overset{AX=Y}{\Rightarrow} y^* = 0$ — противоречие). Тогда $x^* = A^*\big(-\frac{y^*}{\lambda}\big) \in \operatorname{Im} A^*$, т. е. $(\operatorname{Ker} A)^{\perp} \subset \operatorname{Im} A^*$.

Теорема об обратном отображении. *Пусть X, Z — банаховы пространства, отображение $F\colon X \to Z$, $F(\hat{x}) = \hat{z}$, $F \in SD(\hat{x})$, оператор $F'(\hat{x})$ является эпиморфизмом. Тогда существуют обратное отображение $F^{-1}\colon W \subset Z \to X$ некоторой окрестности W точки $\hat{z}$ и константа $K > 0$ такие, что $F^{-1}(\hat{z}) = \hat{x}$ и*

$$F(F^{-1}(z)) = z, \quad \|F^{-1}(z) - F^{-1}(\hat{z})\| \leq K\|z - \hat{z}\| \quad \forall\, z \in W.$$

Теорема Люстерника. *Пусть X, Z — банаховы пространства, отображение $F\colon X \to Z$, $F \in SD(\hat{x})$, оператор $F'(\hat{x})$ является эпиморфизмом. Тогда существуют отображение $\varphi\colon U \subset X \to X$ некоторой окрестности U точки $\hat{x}$ и число $K > 0$ такие, что*

$$F\big(x + \varphi(x)\big) = F(\hat{x}), \quad \|\varphi(x)\| \leq K\|F(x) - F(\hat{x})\| \quad \forall\, x \in U.$$

Доказательство этой теоремы основано на модифицированном методе Ньютона.

А) Не ограничивая общности, для краткости записи считаем, что $\hat{x} = 0$ и $F(\hat{x}) = 0$. Тогда надо доказать, что существуют отображение $\varphi\colon U \subset X \to X$ некоторой окрестности нуля U и число $K > 0$ такие, что

$$F\big(x + \varphi(x)\big) = 0, \quad \|\varphi(x)\| \leq K\|F(x)\| \quad \forall\, x \in U.$$

По лемме о правом обратном операторе для оператора $F'(\hat{x})\colon X \overset{\text{на}}{\to} Z$ существуют обратное отображение $M\colon Z \to X$ и константа (не ограничивая общности, можем считать ее большей 1) $C > 1$ такие, что

$$F'(\hat{x}) \circ M = I_Z, \quad \|Mz\| \leq C\|z\| \quad \forall\, z \in Z.$$

Отображение $F \in SD(\hat{x})$, поэтому для $\varepsilon = \dfrac{1}{2C}$ существует $\delta > 0$ такое, что

$$\|F(x') - F(x'') - F'(\hat{x})[x' - x'']\| \leq \frac{1}{2C}\|x' - x''\| \tag{1}$$

при $\|x'\| < \delta$, $\|x''\| < \delta$. Из строгой дифференцируемости в точке $\hat{x} = 0$ следует также непрерывность F в некоторой окрестности нуля. Выберем δ' столь малым, чтобы $\|x\| + C\|F(x)\| < \dfrac{\delta}{2}$ при $\|x\| < \delta'$.

Возьмем фиксированный элемент $x \in U := \overset{o}{B}(0, \delta')$. Для него определим последовательность элементов $\{\xi_n\}$ с помощью рекуррентного соотношения

$$\xi_{n+1} := \xi_n - M\big(F(\xi_n)\big),\ n \geq 0,\ \xi_0 = x. \tag{2}$$

В) Докажем по индукции, что $\|\xi_n\| < \delta$ для любого $n \geq 0$. Действительно, $\|\xi_0\| = \|x\| < \dfrac{\delta}{2}$. При $n = 1$ из (2) и леммы о правом обратном операторе получаем оценку

$$\|\xi_1 - x\| = \|MF(x)\| \leq C\|F(x)\|, \tag{3}$$

откуда по неравенству треугольника для норм $\|\xi_1\| \leq \|x\| + C\|F(x)\| < \dfrac{\delta}{2}$.

Пусть $\|\xi_i\| < \delta$ при $i = 0, 1, \ldots, k$ $(k \geq 1)$. Выведем отсюда, что $\|\xi_{k+1}\| < \delta$. Из соотношения (2) следует, что $\xi_{i+1} - \xi_i + MF(\xi_i) = 0$.

Применяя к обеим частям этого равенства оператор $F'(\hat{x})$, получим для $i = 0, 1, \ldots, k$

$$F'(\hat{x})[\xi_{i+1} - \xi_i] + F'(\hat{x})[MF(\xi_i)] = 0 \iff F'(\hat{x})[\xi_{i+1} - \xi_i] + F(\xi_i) = 0, \tag{4}$$

откуда

$$\|\xi_{i+1} - \xi_i\| \overset{(2)}{=} \|MF(\xi_i)\| \le C\|F(\xi_i)\| =$$

(вычтем из $F(\xi_i)$ выражение $F(\xi_{i-1}) + F'(\hat{x})(\xi_i - \xi_{i-1})$ равное нулю в силу соотношения (4))

$$= C\|F(\xi_i) - F(\xi_{i-1}) - F'(\hat{x})(\xi_i - \xi_{i-1})\| \le$$

$$\overset{(1)}{\le} \frac{1}{2}\|\xi_i - \xi_{i-1}\| \le \text{ (аналогично) } \frac{1}{2^2}\|\xi_{i-1} - \xi_{i-2}\| \le \ldots$$

$$\Rightarrow \|\xi_{i+1} - \xi_i\| \le \frac{1}{2^i}\|\xi_1 - x\| \overset{(3)}{\le} \frac{1}{2^i}C\|F(x)\| < \frac{1}{2^i} \cdot \frac{\delta}{2},\ i = 1, \ldots, k. \tag{5}$$

Отсюда в силу неравенства треугольника получаем

$$\|\xi_{k+1}\| = \|\xi_{k+1} - \xi_k + \xi_k - \xi_{k-1} + \cdots + \xi_2 - \xi_1 + \xi_1\| \le$$

$$\le \|\xi_{k+1} - \xi_k\| + \|\xi_k - \xi_{k-1}\| + \cdots + \|\xi_2 - \xi_1\| + \|\xi_1\| <$$

$$< \left(\frac{1}{2^k} + \frac{1}{2^{k-1}} + \cdots + \frac{1}{2}\right)\frac{\delta}{2} + \frac{\delta}{2} < \delta.$$

Таким образом, мы получили, что $\|\xi_{k+1}\| < \delta$, откуда по индукции следует, что $\|\xi_n\| < \delta$ для любого $n \ge 0$ и неравенство (5) выполняется для всех элементов последовательности ξ_n.

C) Из неравенства (5) следует, что

$$\|\xi_{n+m} - \xi_n\| = \|\xi_{n+m} - \xi_{n+m-1} + \xi_{n+m-1} - \xi_{n+m-2} + \cdots + \xi_{n+1} - \xi_n\| \le$$

$$\le \|\xi_{n+m} - \xi_{n+m-1}\| + \|\xi_{n+m-1} - \xi_{n+m-2}\| + \cdots + \|\xi_{n+1} - \xi_n\| \le$$

$$\le \left(\frac{1}{2^{n+m-1}} + \frac{1}{2^{n+m-2}} + \cdots + \frac{1}{2^n}\right)\frac{\delta}{2} < \frac{2}{2^n} \cdot \frac{\delta}{2} \to 0 \text{ при } n \to \infty,$$

т. е. $\{\xi_n\}_{n \ge 0}$ — фундаментальная последовательность и, значит, она сходится в силу банаховости X. Обозначим $\varphi(x) = \lim\limits_{n\to\infty} \xi_n - x$. Поскольку

$$\|\xi_n - x\| = \|\xi_n - \xi_{n-1} + \xi_{n-1} - \xi_{n-2} + \cdots + \xi_2 - \xi_1 + \xi_1 - x\| \le$$

$$\leq \|\xi_n - \xi_{n-1}\| + \|\xi_{n-1} - \xi_{n-2}\| + \dots + \|\xi_2 - \xi_1\| + \|\xi_1 - x\| \leq$$
$$\overset{(5)}{\leq} \|\xi_1 - x\| \left(\frac{1}{2^{n-1}} + \frac{1}{2^{n-2}} + \dots + 1 \right) \leq 2\|\xi_1 - x\| \overset{(3)}{\leq} 2C\|F(x)\|,$$

то

$$\|\varphi(x)\| = \lim_{n\to\infty} \|\xi_n - x\| \leq 2\|\xi_1 - x\| \overset{K:=2C}{=} K\|F(x)\|,$$
$$\|x + \varphi(x)\| \leq \|x\| + \|\varphi(x)\| \leq \|x\| + 2C\|F(x)\| < \delta.$$

Отображение F, как объяснялось выше, непрерывно в окрестности $\hat{x} = 0$, и, значит, непрерывно в точке $x + \varphi(x) = \lim\limits_{n\to\infty} \xi_n$ и поэтому

$$F\big(x+\varphi(x)\big) = F(\lim \xi_n) = \lim_{n\to\infty} F(\xi_n) \overset{(4)}{=} -\lim_{n\to\infty} F'(\hat{x})(\xi_{n+1}-\xi_n) = 0 = F(\hat{x}). \triangleright$$

Пусть X — линейное нормированное пространство, M — некоторое его подмножество. Вектор $h \in X$ называется *односторонним касательным (полукасательным) вектором к множеству* M *в точке* $\hat{x} \in X$, если существуют $\varepsilon > 0$ и отображение $r\colon (0, \varepsilon] \to X$, такие, что

a) $\hat{x} + th + r(t) \in M \ \forall\ t \in (0, \varepsilon]$;

b) $\|r(t)\| = o(t)$ при $t \to +0$.

Вектор h называется *касательным к множеству* M *в точке* $\hat{x} \in X$, если векторы h и $-h$ являются односторонними касательными векторами к M в $\hat{x}$. Иными словами, элемент $h \in X$ называется *касательным вектором к множеству* M *в точке* $\hat{x}$, если существуют $\varepsilon > 0$ и отображение $r\colon [-\varepsilon, \varepsilon] \to X$ такие, что

a) $\hat{x} + th + r(t) \in M \ \forall\ t \in [-\varepsilon, \varepsilon],\ t \neq 0$;

b) $\|r(t)\| = o(t)$ при $t \to 0$.

Множество всех касательных векторов к M в точке $\hat{x}$ обозначается $T_{\hat{x}}M$, множество односторонних касательных векторов обозначается $T^+_{\hat{x}}M$. Очевидно, что $T_{\hat{x}}M$ и $T^+_{\hat{x}}M$ — конусы. Если множество $T_{\hat{x}}M$ является подпространством в X, то оно называется *касательным пространством к множеству* M *в точке* $\hat{x}$.

Если точка $\hat{x} \in M$, то в обоих определениях t может равняться нулю.

Во многих случаях, в том числе и представляющих значительный интерес для теории экстремальных задач, множество касательных векторов может быть найдено при помощи такого следствия из теоремы Люстерника.

Теорема (о касательном пространстве). *Пусть* X, Z — *банаховы пространства,* $F\colon X \to Z$, $F \in SD(\hat{x})$, *оператор* $F'(\hat{x})$ — *эпиморфизм, множество* $M = \{x \in X \mid F(x) = F(\hat{x})\}$. *Тогда*

$$T_{\hat{x}}M = \operatorname{Ker} F'(\hat{x}).$$

Доказательство. A) Докажем, что $T_{\hat{x}}M \subset \operatorname{Ker} F'(\hat{x})$. Пусть $h \in T_{\hat{x}}M$, тогда существуют $\varepsilon > 0$ и отображение $r\colon [-\varepsilon, \varepsilon] \to X$, такие, что $\hat{x}+th+r(t) \in M$ $\forall\, t \in [-\varepsilon, \varepsilon]$, $\|r(t)\| = o(t)$ при $t \to 0$. При малых t

$$F(\hat{x}) = F\big(\hat{x} + th + r(t)\big) = F(\hat{x}) + tF'(\hat{x})[h] + o(t).$$

Отсюда $tF'(\hat{x})[h] + o(t) = 0$ и, значит, $F'(\hat{x})[h] = 0$, т.е. $h \in \operatorname{Ker} F'(\hat{x}) \Rightarrow T_{\hat{x}}M \subset \operatorname{Ker} F'(\hat{x})$.

B) Докажем, что $T_{\hat{x}}M \supset \operatorname{Ker} F'(\hat{x})$. Пусть $h \in \operatorname{Ker} F'(\hat{x})$. Положим $r(t) = \varphi(\hat{x} + th)$, где φ — отображение, построенное в теореме Люстерника. Тогда

$$F\big(\hat{x}+th+r(t)\big) \overset{\mathrm{def}}{=} F\big(\hat{x}+th+\varphi(\hat{x}+th)\big) = F(\hat{x}) \Longrightarrow \hat{x}+th+r(t) \in M,$$

$$F(\hat{x} + th) - F(\hat{x}) = tF'(\hat{x})[h] + o(t) \overset{F'(\hat{x})[h]=0}{=} o(t),$$

$$\|r(t)\| = \|\varphi(\hat{x}+th)\| \le K\|F(\hat{x}+th) - F(\hat{x})\| = K\|tF'(\hat{x})[h] + o(t)\| = K\|o(t)\| = o(t),$$

т. е. $h \in T_{\hat{x}}M \Rightarrow \operatorname{Ker} T_{\hat{x}}M \supset \operatorname{Ker} F'(\hat{x})$.

Таким образом, $T_{\hat{x}}M = \operatorname{Ker} F'(\hat{x})$. $\triangleright$