7 Гладкая задача с равенствами

7.1 Постановка задачи

Пусть X, Y — линейные нормированные пространства. Функционал $f\colon X \to \mathbb{R}$ и отображение $F\colon X \to Y$ обладают определенной гладкостью. *Гладкой экстремальной задачей с ограничениями типа равенств* в нормированных пространствах называется следующая задача:

$$f(x) \to \mathrm{extr}; \quad F(x) = 0. \tag{P}$$

Отметим, что отображение типа равенства в бесконечномерных пространствах может содержать в себе как конечное, так и бесконечное число равенств.

7.2 Необходимые условия I порядка

Теорема (принцип Лагранжа)**.** *Пусть $\hat{x} \in locextr\, P$ — точка локального экстремума в задаче (P), X, Y — банаховы пространства* (условие банаховости)*, функционал f и отображение $F \in SD(\hat{x})$ — строго дифференцируемы в точке $\hat{x}$* (условие гладкости)*, $Im\, F'(\hat{x})$ — замкнутое подпространство в Y* (ослабленное условие регулярности)*. Тогда существуют множители Лагранжа $\lambda_0 \in \mathbb{R}$ и функционал $y^* \in Y^*$ не равные одновременно нулю, $(\lambda_0, y^*) \neq 0$, и такие, что для функции Лагранжа задачи (P)*

$$\mathscr{L}(x) = \lambda_0 f(x) + \langle y^*, F(x) \rangle$$

выполняется условие стационарности:

$$\mathscr{L}'(\hat{x}) = 0 \quad \Big(\Leftrightarrow \lambda_0 \langle f'(\hat{x}), h \rangle + \langle y^*, F'(\hat{x})[h] \rangle = 0 \;\; \forall\, h \in X \Leftrightarrow$$

$$\langle \lambda_0 f'(\hat{x}) + (F'(\hat{x}))^* y^*, h \rangle = 0 \;\; \forall\, h \in X \Leftrightarrow \lambda_0 f'(\hat{x}) + (F'(\hat{x}))^* y^* = 0 \Big).$$

Доказательство. Не ограничивая общности, считаем, что $f(\hat{x}) = 0$. Определим отображение $\mathcal{F}(x) := (f(x), F(x))$. Тогда отображение $\mathcal{F}$ строго дифференцируемо в точке $\hat{x}$, производная Фреше $\mathcal{F}'(\hat{x}) = \big(f'(\hat{x}),\, F'(\hat{x})\big)$, оба отображения $\mathcal{F}$ и $\mathcal{F}'$ отображают пространство X в пространство $\mathbb{R} \times Y$.

Поскольку подпространство $\operatorname{Im} F'(\hat{x})$ замкнуто по условию, образ $f'(\operatorname{Ker} F'(\hat{x}))$ тоже замкнут (в $\mathbb{R}$ имеются всего два подпространства: 0 и $\mathbb{R}$, и оба они замкнутые), то к отображению $\mathcal{F}'(\hat{x})$ можно применить лемму о замкнутости образа. По этой лемме образ $\operatorname{Im} \mathcal{F}'(\hat{x})$ замкнут в $\mathbb{R} \times Y$.

Перепишем условие стационарности в виде

$$\langle(\lambda_0, y^*), (\langle f'(\hat{x}), h\rangle, F'(\hat{x})[h])\rangle = 0 \iff$$

$$\langle(\lambda_0, y^*), \mathcal{F}'(\hat{x})[h]\rangle = 0 \quad \forall\, h \in X \iff \langle(\lambda_0, y^*), \operatorname{Im} \mathcal{F}'(\hat{x})\rangle = 0$$

$$\iff (\lambda_0, y^*) \in (\operatorname{Im} \mathcal{F}'(\hat{x}))^{\perp} \iff (\operatorname{Im} \mathcal{F}'(\hat{x}))^{\perp} \neq 0.$$

Таким образом, условие стационарности означает, что аннулятор $(\operatorname{Im} \mathcal{F}'(\hat{x}))^{\perp}$ непуст. По лемме о нетривиальности аннулятора замкнутое подпространство $\operatorname{Im} \mathcal{F}'(\hat{x})$ является собственным (иначе оно состояло бы).

Предположим, что условие стационарности не выполняется. Тогда замкнутое подпространство $\operatorname{Im} \mathcal{F}'(\hat{x})$ не является собственным, то есть совпадает со всем пространством.

Итак, пусть $\operatorname{Im} \mathcal{F}'(\hat{x}) = \mathbb{R} \times Y$. По теореме об обратном отображении существуют отображение $\mathcal{F}^{-1} \colon W \subset \mathbb{R} \times Y \to X$ некоторой окрестности W точки $(\hat{\alpha}, \hat{y})$ ($(\hat{\alpha}, \hat{y}) = (0, 0)$) и константа $K > 0$ такие, что $\mathcal{F}^{-1}(\hat{\alpha}, \hat{y}) = \hat{x}$,

$$\mathcal{F}(\mathcal{F}^{-1}(\alpha, y)) = (\alpha, y), \quad \|\mathcal{F}^{-1}(\alpha, y) - \mathcal{F}^{-1}(\hat{\alpha}, \hat{y})\| \leq K\|(\alpha, y) - (\hat{\alpha}, \hat{y})\|$$

$$\iff \|\mathcal{F}^{-1}(\alpha, y) - \hat{x}\| \leq K\|(\alpha, y)\| \quad \forall\, (\alpha, y) \in W.$$

Положим $x(\varepsilon) = \mathcal{F}^{-1}(\varepsilon, 0)$ для достаточно малого по модулю ε. Тогда

$$\mathcal{F}(x(\varepsilon)) = (\varepsilon, 0) \iff f(x(\varepsilon)) - f(\hat{x}) = \varepsilon,\ F(x(\varepsilon)) = 0,$$

$$\|x(\varepsilon) - \hat{x}\| = \|\mathcal{F}^{-1}(\varepsilon, 0) - \hat{x}\| \leq K\|(\varepsilon, 0)\| = K|\varepsilon|.$$

Из этих соотношений следует, что вектор $\hat{x}$ не доставляет в задаче экстремума, ибо вблизи его существуют допустимые векторы $x(\varepsilon)$, на которых функционал f принимает значения как большие так и меньшие чем $f(\hat{x})$. Получили противоречие с тем, что $\hat{x} \in \mathrm{locextr} P$. Таким образом, невырожденный случай невозможен, и тем самым теорема доказана. $\triangleright$

Замечание. *Если в условиях теоремы выполнено условие регулярности отображения F в точке $\hat{x}$, т. е. $Im\,F'(\hat{x}) = Y$, то множитель $\lambda_0 \neq 0$, и, следовательно, можем считать его равным единице: $\lambda_0 = 1$.*

$\lhd$ Действительно, если $\lambda_0 = 0$, то $y^* \neq 0$ в силу того, что множители Лагранжа одновременно в ноль не обращаются. Поэтому условие стационарности приобретает вид: $\langle y^*, F'(\hat{x})[h] \rangle = 0 \ \forall\ h \in X \Leftrightarrow \langle y^*, Y \rangle = 0 \Leftrightarrow y^* = 0$. Получили противоречие. $\rhd$

7.3 Необходимые условия II порядка

Теорема. *Пусть $\hat{x} \in loc\min P$ — точка локального минимума в задаче (P), X, Y — банаховы пространства* (условие банаховости), *функционал f и отображение F имеют в этой точке вторые производные Фреше $(f, F \in D^2(\hat{x}))$* (условие гладкости), *$Im\,F'(\hat{x}) = Y$* (условие регулярности)*. Тогда существует множитель Лагранжа — функционал $y^* \in Y^*$ такой, что для функции Лагранжа с множителем Лагранжа $\lambda_0 = 1$ задачи (P)*

$$\mathscr{L}(x) = f(x) + \langle y^*, F(x) \rangle$$

выполняются условия стационарности:

$$\mathscr{L}'(\hat{x}) = 0 \quad \Big(\Leftrightarrow f'(\hat{x}) + (F'(\hat{x}))^* y^* = 0 \Big) \tag{1}$$

и неотрицательности:

$$\mathscr{L}''(\hat{x})[h, h] \geq 0 \quad \forall\ h \in \operatorname{Ker} F'(\hat{x}). \tag{2}$$

Доказательство. Напомним, что по Замечанию к Следствию 2 п. 5.3.3 существование второй производной Фреше в точке гарантирует строгую дифференцируемость отображения в этой точке.

Поэтому условие стационарности (1) с множителем Лагранжа $\lambda_0 = 1$ вытекает из правила множителей Лагранжа для гладкой задачи с равенствами и замечания к нему (см. предыдущий пункт).

Докажем условие неотрицательности. Возьмем $h \in \operatorname{Ker} F'(\hat{x})$. Тогда по теореме о касательном пространстве $\operatorname{Ker} F'(\hat{x}) = T_{\hat{x}}M$, где $M = \{x \in X \mid |F(x) = F(\hat{x}) = 0\}$. Следовательно, $h \in T_{\hat{x}}M$

и, значит, существуют $\varepsilon > 0$ и отображение $r\colon [-\varepsilon; \varepsilon] \to X$ такие, что $F(\hat{x} + th + r(t)) = 0$ для любого $t \in [-\varepsilon; \varepsilon]$ и норма $\|r(t)\| = o(t)$ при $t \to 0$. Таким образом, $\hat{x} + th + r(t)$ — допустимый элемент в задаче (P) при $t \in [-\varepsilon; \varepsilon]$ и, так как $\hat{x} \in \operatorname{locmin} P$, то $f(\hat{x}) \leq f(\hat{x} + th + r(t))$. Поэтому по определению функции $\mathscr{L}$ и по формуле Тейлора

$$f(\hat{x}) \leq f(\hat{x} + th + r(t)) = f(\hat{x} + th + r(t)) +$$

(добавим и вычтем $\langle y^*, F(\hat{x} + th + r(t))\rangle$)

$$+\langle y^*, F(\hat{x}+th+r(t))\rangle - \langle y^*, F(\hat{x}+th+r(t))\rangle = \mathscr{L}(\hat{x}+th+r(t)) - \langle y^*, F(\hat{x}+th+r(t))\rangle =$$

$$= \mathscr{L}(\hat{x}) + \mathscr{L}'(\hat{x})[th+r(t)] + \frac{1}{2}\mathscr{L}''(\hat{x})[th+r(t), th+r(t)] + o(\|th+r(t)\|^2) =$$

$$= f(\hat{x}) + \frac{t^2}{2}\mathscr{L}''(\hat{x})[h, h] + o(t^2).$$

Отсюда $\dfrac{t^2}{2}\mathscr{L}''(\hat{x})[h, h] + o(t^2) \geq 0$ при малых t. Разделим обе части последнего неравенства на t^2 и устремим t к нулю. Получим $\mathscr{L}''(\hat{x})[h, h] \geq 0$. $\triangleright$

7.4 Достаточные условия II порядка

Теорема. *Пусть выполняются условия теоремы о необходимых условиях II порядка* (банаховость, гладкость, регулярность, стационарность для функции Лагранжа $\mathscr{L}(x) = f(x) + \langle y^*, F(x)\rangle$ с множителем $\lambda_0 = 1$) *и для некоторого $\alpha > 0$ выполняется условие строгой положительности:*

$$\mathscr{L}''(\hat{x})[h, h] \geq \alpha\|h\|^2 \quad \forall\, h \in Ker F'(\hat{x}).$$

Тогда $\hat{x} \in locmin\, P$ — точка локального минимума в задаче (P).

Доказательство. По лемме о правом обратном отображении к отображению $F'(\hat{x})\colon X \overset{\text{на}}{\to} Y$ существуют отображение $M\colon Y \to X$ и константа $C > 0$ такие, что,

$$F'(\hat{x}) \circ M = I_Y, \quad \|My\| \leq C\|y\| \quad \forall\, y \in Y.$$

Возьмем $\hat{x}+h$ — допустимый элемент в задаче ($F(\hat{x}+h)=0$). Положим $h_2 := M(F'(\hat{x})[h])$ и обозначим $h_1 := h - h_2$. Тогда $F'(\hat{x})[h_1] = F'(\hat{x})[h-h_2] = F'(\hat{x})[h] - F'(\hat{x})M(F'(\hat{x})[h]) = F'(\hat{x})[h] - F'(\hat{x})[h] = 0$. Значит, $h_1 \in \operatorname{Ker} F'(\hat{x})$. По формуле Тейлора

$$0 = F(\hat{x}+h) = F(\hat{x}) + F'(\hat{x})[h] + \frac{1}{2}F''(\hat{x})[h,h] + o(\|h\|^2) \Rightarrow$$

$$F'(\hat{x})[h] = -\frac{1}{2}F''(\hat{x})[h,h] - o(\|h\|^2).$$

Отсюда существует $\delta > 0$ такое, что

$$\|F'(\hat{x})[h]\| = \frac{1}{2}\|F''(\hat{x})[h,h]\| + o(\|h\|^2) \le C_1\|h\|^2 \quad \forall\ \|h\| < \delta$$

с некоторой константой $C_1 > 0$. Поэтому $\|h_2\| = \|M(F'(\hat{x})[h])\| \le C\|F'(\hat{x})[h]\| \le CC_1\|h\|^2 \le CC_1\delta\|h\| = \varepsilon\|h\|$ при $\varepsilon = CC_1\delta$ и по неравенству треугольника для норм $\|h\| - \|h_2\| \le \|h_1\| = \|h-h_2\| \le \|h\| + \|h_2\| \Leftrightarrow (1-\varepsilon)\|h\| \le \|h_1\| \le (1+\varepsilon)\|h\|$.

Вновь по формуле Тейлора

$$f(\hat{x}+h) = \mathscr{L}(\hat{x}+h) = \mathscr{L}(\hat{x}) + \mathscr{L}'(\hat{x})[h] + \frac{1}{2}\mathscr{L}''(\hat{x})[h,h] + o(\|h\|^2) =$$

$$\overset{\mathscr{L}'(\hat{x})=0}{=} f(\hat{x}) + \frac{1}{2}\mathscr{L}''(\hat{x})[h,h] + o(\|h\|^2).$$

Отсюда, обозначая $B := \|\mathscr{L}''(\hat{x})\|$, имеем

$$f(\hat{x}+h) - f(\hat{x}) = \frac{1}{2}\mathscr{L}''[h_1+h_2, h_1+h_2] + o(\|h\|^2) =$$

$$= \frac{1}{2}\Big(\mathscr{L}''[h_1,h_1] + 2\mathscr{L}''[h_1,h_2] + \mathscr{L}''[h_2,h_2]\Big) + o(\|h\|^2) \ge$$

$$\ge \frac{1}{2}\Big(\alpha\|h_1\|^2 - 2B\|h_1\|\cdot\|h_2\| - B\|h_2\|^2\Big) + o(\|h\|^2) \ge$$

$$\ge \frac{1}{2}\|h\|^2\Big(\alpha(1-\varepsilon)^2 - 2B(1+\varepsilon)\varepsilon - B\varepsilon^2\Big) + o(\|h\|^2) \ge 0$$

при достаточно малых $\varepsilon > 0$ (при $\varepsilon = 0$ множитель в круглых скобках равен $\alpha > 0$). Из последнего соотношения следует, что $\hat{x} \in \mathrm{locmin} P$. $\triangleright$