6 Гладкая задача без ограничений

В этом пункте даются необходимые и достаточные условия экстремума функционалов в нормированных пространствах.

6.1 Постановка задачи

Пусть $f\colon X \to \mathbb{R}$ — отображение линейного нормированного пространства X во множество действительных чисел (в этом случае обычно говорим *функционал* на пространстве X), обладающее некоторой гладкостью, т. е. определенными свойствами дифференцируемости. *Гладкой задачей без ограничений* называется задача об отыскании экстремумов этого функционала:

$$f(x) \to \mathrm{extr}.$$

6.2 Необходимые условия I порядка

Теорема 1 (аналог теоремы Ферма в нормированных пространствах). *Пусть $\hat{x} \in locextr f$ — точка локального экстремума функционала f и функционал f дифференцируем по Фреше (имеет вариацию по Лагранжу) в точке $\hat{x}$. Тогда*

$$f'(\hat{x}) = 0 \qquad (\delta f(\hat{x}, h) = 0 \quad \forall\ h \in X).$$

Доказательство. Возьмем произвольный, но фиксированный элемент $h \in X$. Рассмотрим функцию $\varphi(\lambda) = f(\hat{x} + \lambda h)$. Поскольку $\hat{x} \in \mathrm{locextr} f$, то $0 \in \mathrm{locextr} \varphi$ — локальный экстремум функции φ. По теореме Ферма для функций одной переменной $\varphi'_\lambda(0) = 0$. По определению вариации по Лагранжу это эквивалентно тому, что $\delta f(\hat{x}, h) = 0$. В силу произвольности h $\delta f(\hat{x}, h) = 0\ \forall\ h \in X$.

Если функционал f дифференцируем по Фреше в точке $\hat{x}$, то в этой точке он имеет вариацию по Лагранжу и $f'(\hat{x})[h] = \delta f(\hat{x}, h)$. Поскольку из уже доказанного следует, что эта вариация $\delta f(\hat{x}, \cdot) = 0$, то и $f'(\hat{x}) = 0$ в силу определения дифференцируемости по Фреше. $\triangleright$

6.3 Необходимые и достаточные условия II порядка

Теорема 2. `Необходимые условия экстремума:` *если $\hat{x}$ — точка локального минимума функционала f $\big(\hat{x} \in locmin\, f\big)$ и функционал f*

дважды дифференцируем в точке $\hat{x}$ $\big(f \in D^2(\hat{x})\big)$, *то*

$$f'(\hat{x}) = 0,\ f''(\hat{x})[h,h] \geq 0 \quad \forall\ h \in X.$$

`Достаточные условия экстремума:` *если* $f'(\hat{x}) = 0$ *и*

$$f''(\hat{x})[h,h] \geq \alpha\|h\|^2 \quad \forall\ h \in X \tag{$*$}$$

при некотором $\alpha > 0$, *то* $\hat{x} \in \mathit{locmin}\, f$.

Доказательство. По формуле Тейлора

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

Необходимость. Поскольку $\hat{x} \in \operatorname{locmin} f$, то по аналогу теоремы Ферма $f'(\hat{x}) = 0$. Поэтому в силу формулы Тейлора

$$f(\hat{x}+\lambda h) - f(\hat{x}) = \frac{\lambda^2}{2}f''(\hat{x})[h,h] + r(\lambda h), \quad r(\lambda h) = o(|\lambda|^2)$$

при достаточно малых λ и фиксированном h. Поскольку $\hat{x} \in \operatorname{locmin} f$, то $f(\hat{x}+\lambda h) - f(\hat{x}) \geq 0$ при малых λ. Поэтому

$$\tfrac{\lambda^2}{2}f''(\hat{x})[h,h] + r(\lambda h) \geq 0.$$

Разделим обе части последнего неравенства на λ^2 и устремим λ к нулю. Поскольку $\frac{r(\lambda h)}{\lambda^2} \to 0$, то отсюда

$$f''(\hat{x})[h,h] \geq 0 \quad \forall\ h \in X.$$

Достаточность. Так как $f'(\hat{x}) = 0$, то по формуле Тейлора в силу заданного условия $f''(\hat{x})[h,h] \geq \alpha||h||^2$ имеем:

$$f(\hat{x}+h) - f(\hat{x}) = \frac{1}{2}f''(\hat{x})[h,h] + r(h) \geq \frac{\alpha}{2}||h||^2 + r(h) \geq 0$$

при достаточно малых h, так как $r(h) = o(\|h\|^2)$. Следовательно, $\hat{x} \in \operatorname{locmin} f$.

Для локального максимума неравенства имеют противоположный вид: $f''(\hat{x})[h,h] \leq 0$ и $f''(\hat{x})[h,h] \leq -\alpha\|h\|^2$ $(\alpha > 0)$ соответственно.

Условие $(*)$ называется условием *строгой положительности* второй производной Фреше функционала f.

Отметим, что в конечномерных пространствах условие положительной определенности симметричной матрицы A, т. е. условие

$$\langle Ah, h\rangle > 0 \quad \forall\ h \in \mathbb{R}^n,\ h \neq 0,$$

гарантирует ее строгую положительность (и, значит, является достаточным условием минимума в стационарной точке). В бесконечномерных пространствах это не так.

Пример 1. *В бесконечномерных пространствах условие положительной определенности отображения не гарантирует строгой положительности отображения.*

Рассмотрим отображение

$$f\colon l_2 \to \mathbb{R},\ x = (x_1, ..., x_n, \ldots),\ \|x\|_{l_2} = \left(\sum_{n=1}^{+\infty} x_n^2\right)^{\frac{1}{2}},\ f(x) = \sum_{n=1}^{\infty} \frac{x_n^2}{n}.$$

Первая производная Фреше $f'(x)[h] = 2\sum\limits_{n=1}^{\infty} \dfrac{x_n h_n}{n}$. Вторая производная Фреше в нуле $f''(0)[h,h] = 2\sum\limits_{n=1}^{\infty} \dfrac{h_n^2}{n} > 0\ \forall\ h \neq 0$ и, следовательно, является положительно определенной. Но вместе с тем вторая производная отображения f в нуле не является строго положительной. Действительно, неравенство $f''(0)[h,h] \geq \alpha\|h\|^2\ \forall\ h$ не может выполняться ни для какой константы $\alpha > 0$, поскольку на последовательности векторов $\{h^n\} = e_n$, $n = 1, 2, \ldots$ (e_n — n-й базисный вектор пространства l_2), $f''(0)[h_n, h_n] = \dfrac{2}{n} = \dfrac{2}{n}\|h_n\|^2 \ngeq \alpha\|h_n\|^2$.

Пример 2. *В бесконечномерных пространствах условие равенства нулю первой производной и положительной определенности второй производной не гарантируют локального экстремума отображения.*

Рассмотрим отображение

$$f\colon l_2 \to \mathbb{R}, \quad f(x) = \sum_{n=1}^{\infty}\left(\frac{x_n^2}{n^3} - x_n^4\right),$$

и задачу без ограничений $f(x) \to \min$. Точка $\hat{x} = 0$ является стационарной. Действительно,

$$f'(x)[h] = \sum_{n=1}^{\infty}\left(\frac{2x_n h_n}{n^3} - 4\hat{x}_n^3 h_n\right) \Longrightarrow f'(0) = 0.$$

Вторая производная Фреше $f''(x)[h,h] = \sum\limits_{n=1}^{\infty}\left(\dfrac{2h_n^2}{n^3} - 12x_n^2h_n^2\right)$. Значит, в нуле $f''(0)[h,h] = 2\sum\limits_{n=1}^{\infty}\dfrac{h_n^2}{n^3} > 0 \ \forall \ h \neq 0$ и, следовательно, является положительно определенным оператором, но вместе с тем, как и в предыдущем примере, не является строго положительным. В задаче на минимум точка $\hat{x} = 0 \notin \mathrm{locmin} f$, поскольку на последовательности векторов $\{x^n\} = \dfrac{e_n}{n}$, $n = 1, 2, \ldots$ (e_n — n-й базисный вектор пространства l_2), $f(x^n) = \dfrac{1}{n^5} - \dfrac{1}{n^4} < 0 = f(0)$ при $n > 1$, а сама последовательность $\left\{\dfrac{e_n}{n}\right\} \to \hat{x} = 0$ в пространстве l_2 при $n \to +\infty$.