ВЫПУКЛЫЙ АНАЛИЗ

ЛЕКЦИЯ 16

5. МИНИМУМ ВЫПУКЛОЙ ФУНКЦИИ

5. МИНИМУМ ВЫПУКЛОЙ ФУНКЦИИ

5.1. Локальный и глобальный минимум выпуклой функции.

5.2. Минимум дифференцируемой выпуклой функции.

5.3. Минимум выпуклой функции, дифференцируемой по всем возможным направлениям.

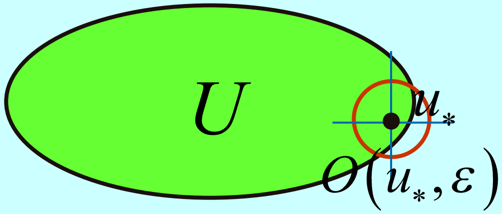
5.1. Локальный и глобальный минимум выпуклой функции. Выпуклые функции представляют собой удобные объекты исследования для анализа их значений на минимум. Это объясняется тем обстоятельством, что выпуклые функции не могут иметь локальных минимумов.

Теорема 1. ***Пусть функция*** $I: U \to R^1$, ***где*** $U \subset R^n$ ***выпуклое множество, выпукла. Тогда всякая точка локального минимума функции*** I ***на множестве*** U ***одновременно является точкой ее глобального минимума на этом множестве, причем множество***

$$U_* = \left\{ u_* \in U \,\middle|\, I(u_*) = \min_{u \in U} I(u) = I_* \right\}$$

выпукло. ***В случае, когда функция*** I ***строго выпукла на*** U, ***множество*** U_* ***содержит не более одной точки.***

Доказательство. Пусть u_* точка локального минимума функции I на множестве U. Тогда существует окрестность $O(u_*, \varepsilon)$ точки u_*, что



$$I(u_*) \le I(u),\ \forall u \in O(u_*,\varepsilon) \cap U.$$

Для любого $u \in U$ и достаточно малого $\alpha \in [0,1]$ будет выполнено $\alpha\|u-u_*\| < \varepsilon \Rightarrow \overbrace{u_* + \alpha(u-u_*)}^{\alpha\|u-u_*\|<\varepsilon} \in O(u_*,\varepsilon)$. С другой стороны, в силу выпуклости множества U имеем $u_* + \alpha(u-u_*) = \alpha \overset{\in U}{u} + (1-\alpha)\overset{\in U}{u_*} \in U$.

Таким образом, $\alpha u + (1-\alpha)u_* \in O(u_*,\varepsilon) \cap U$. В силу того, что u_* – точка локального минимума и из выпуклости функции I следует

$$\overset{\in U_*}{I(u_*)} \le I\left(\overbrace{u_* + \alpha(u-u_*)}^{\in O(u_*,\varepsilon)\cap U}\right) = I(\alpha u + (1-\alpha)u_*) \le \alpha I(u) + (1-\alpha) I(u_*)$$

или

$$I(u_*) \le \alpha I(u) + I(u_*) - \alpha I(u_*) \Rightarrow I(u_*) \le I(u),\ \forall u \in U \Rightarrow u_* \in U_*.$$

Таким образом, всякий локальный минимум одновременно является глобальным.

Пусть теперь $u_*, v_* \in U_*$ и $\alpha \in [0,1]$. Тогда

$$I_* \le I(\alpha u_* + (1-\alpha)v_*) \le \alpha \overset{=I_*}{I(u_*)} + (1-\alpha)\overset{=I_*}{I(v_*)} = \alpha I_* + (1-\alpha) I_* = I_*. \quad (1)$$

$$I\left(\alpha u_* + (1-\alpha)v_*\right) = I_* \Rightarrow \left(\alpha u_* + (1-\alpha)v_*\right) \in U_*,$$

и множество U_* выпукло. Если $u_* \neq v_*$, то для строго выпуклых функций неравенство (1)

$$I_* \leq I\left(\alpha u_* + (1-\alpha)v_*\right) \leq \alpha I(u_*) + (1-\alpha)I(v_*) = \alpha I_* + (1-\alpha)I_* = I_*. \qquad (1)$$

не может превратиться в равенство $I\left(\alpha u_* + (1-\alpha)v_*\right) = I_*$

при $\alpha \in (0,1)$. Следовательно, строго выпуклая функция не может достигать минимума на выпуклом множестве более чем в одной точке. Теорема доказана.

5.2. Минимум дифференцируемой выпуклой функции. Выведем условия, которым должна удовлетворять точка минимума, выпуклой дифференцируемой функции.

Теорема 2. ***Пусть функция*** $I : U \to R^1$, ***где*** $U \subset R^n$ ***– выпуклое множество, и*** $I \in C^1(U)$. ***Тогда в любой точке*** $u_* \in U_*$ ***выполняется неравенство***

$$\left\langle I'(u_*), u - u_* \right\rangle \geq 0, \forall u \in U, \qquad (1)$$

а в случае $u_* \in \operatorname{int} U$, равенство $I'(u_*) = 0$. ***Кроме того, если функция*** I

выпукла на множестве U, ***то условие (1)*** $\left\langle I'(u_*), u - u_* \right\rangle \geq 0, \forall u \in U \quad (1)$ ***является достаточным для того, чтобы*** $u_* \in U_*$.

Доказательство. Необходимость. Пусть $u_* \in U_*$. Для всех $u \in U, \alpha \in [0,1]$ имеем

$$0 \leq I\big(\alpha u + (1-\alpha) u_*\big) - I(u_*) = I\big(u_* + \alpha(u - u_*)\big) - I(u_*) \overset{I \in C^1(U)}{=}$$

$$\overset{I \in C^1(U)}{=} \left\langle I'(u_*), \alpha (u - u_*) \right\rangle + o(\alpha) = \alpha \left[\left\langle I'(u_*), u - u_* \right\rangle + \frac{o(\alpha)}{\alpha} \right] \Rightarrow$$

$$\left\langle I'(u_*), u - u_* \right\rangle + \frac{o(\alpha)}{\alpha} \geq 0.$$

$\mathbf{O}$

В последнем неравенстве устремим число $\alpha \in [0,1]$ к нулю. В пределе получим неравенство (1). $\left\langle I'(u_*), u - u_* \right\rangle \geq 0, \forall u \in U \quad (1)$

Если $u_* \in \operatorname{int} U$, то для любого вектора $e \in R^n, \|e\| = 1$ найдется число $\varepsilon_0 > 0$, что $u = u_* + \varepsilon e \in U, \forall \varepsilon \in (0, \varepsilon_0]$. В неравенстве (1) полагаем $u = u_* + \varepsilon e$. Тогда

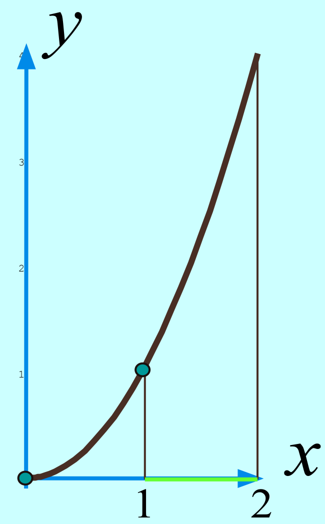
$$\left\langle I'(u_*), \overset{u_* + \varepsilon e}{u} - u_* \right\rangle = \left\langle I'(u_*), u_* + \varepsilon e - u_* \right\rangle = \overset{>0}{\varepsilon} \cdot \left\langle I'(u_*), e \right\rangle \geq 0 \Rightarrow \left\langle I'(u_*), e \right\rangle \geq 0$$

В силу произвольности $e \in R^n, \|e\| = 1$ отсюда выводим, что $I'(u_*) = 0$.

Необходимость доказана.

Заметим, что при доказательстве необходимости выпуклость функции I не использовалась. Для граничной точки u_* равенство $I'(u_*) = 0$ может выполняться, а может и не выполняться. Например, пусть

$$u \in R^1, \quad I(u) = u^2 \Rightarrow I'(u) = 2u$$



1) $U = [0, 1] \Rightarrow u_* = 0$

- равенство $I'(u_*) = 0$ выполняется;

2) $U = [1, 2] \Rightarrow u_* = 1$

- равенство $I'(u_*) = 0$ не выполняется.

Достаточность. Пусть функция I выпукла на множестве U и для некоторой точки $u_* \in U$ выполнено условие (1). $\langle I'(u_*), u - u_* \rangle \geq 0, \forall u \in U \quad (1)$ Тогда по первому критерию выпуклости гладких функций $I(u) \geq I(\overset{u_*}{v}) + \langle I'(\overset{u_*}{v}), u - \overset{u_*}{v} \rangle, \forall u, \overset{u_*}{v} \in U$

$$I(u) \geq I(u_*) + \overset{\geq 0}{\langle I'(u_*), u - u_* \rangle} \Rightarrow I(u) \geq I(u_*), \forall u \in U.$$

что и доказывает достаточность. Теорема доказана.

Из доказанной теоремы следует, что для выпуклых функций равенство $I'(u_*) = 0$ влечет за собой $u_* \in U_*$.

5.3. Минимум выпуклой функции, дифференцируемой по всем возможным направлениям. Требование существования производных по направлениям существенно менее жесткое, чем требование дифференцируемости. В связи с этим представляет интерес получить условия оптимальности в терминах производных по направлению.

Это тем более естественно, поскольку всякая выпуклая функция в любой внутренней точке области определения имеет производные по всем направлениям.

Теорема 3. ***Пусть*** $I : U \to R^1$, ***где*** $U \subset R^n$ ***выпуклое множество,*** $U_* \subset U$ ***множество точек минимума функции*** I ***на множестве*** U ***и в точке*** $u_* \in U_*$ ***функция*** I ***имеет производные по всем возможным направлениям. Тогда необходимо выполняется условие***

$$\frac{dI}{de}(u_*) \geq 0 \qquad (1)$$

для всех возможных направлений $e \left(\|e\| = 1 \right)$ ***в точке*** u_*. ***Кроме того, если функция*** I ***выпукла на*** U, ***то это условие достаточно для*** $u_* \in U_*$.

Доказательство. Необходимость. Пусть $u_* \in U_*$ и $e\left(\|e\|=1\right)$ - возможное направление в точке u_*. Тогда

$$I\big(\overset{\in U}{u_* + te}\big) \geq I(u_*) \quad \Rightarrow \quad \frac{I(u_*+te)-I(u_*)}{t} \geq 0, \;\; 0 < t \leq t_0, \quad (2)$$

где t_0 достаточно мало. Переходя в (2) к пределу при $t \to 0$, получим (1).

$$\frac{dI}{de}(u_*) \geq 0 \quad (1)$$

Необходимость доказана.

Достаточность. Пусть $I : U \to R^1$ - выпуклая функция и в некоторой точке $u_* \in U$ выполнено условие (1). Возьмем любую точку $u \in U, u \neq u_*$ и положим

$$e = \frac{u-u_*}{\|u-u_*\|}.$$

Направление e возможное в точке $u_* \in U$. Действительно, при малых $t > 0$, для которых выполнено $\alpha = \dfrac{t}{\|u-u_*\|} \leq 1,$ имеем

$$u_* + t\overset{\frac{u-u_*}{\|u-u_*\|}}{\overbrace{e}} = u_* + t \cdot \frac{u-u_*}{\|u-u_*\|} = u_* + \overbrace{\frac{t}{\|u-u_*\|}}^{=\alpha\leq 1}(u-u_*) = \overbrace{u_* + \alpha(u-u_*)}^{\alpha u+(1-\alpha)u_*} \in U,$$

Полагаем $f(t) = I\left(u_* + te\right), t \in \left[0, t_0\right], \quad t_0 = \left\| u - u_* \right\|.$

Функция $f : \left[0, t_0\right] \to R^1$ является выпуклой. Действительно, для всех $\alpha \in \left[0,1\right], t_1, t_2 \in \left[0, t_0\right]$ имеем

$$f\left(\overbrace{\alpha t_1 + \left(1-\alpha\right) t_2}^{t}\right) = I\left(u_* + \overbrace{\left(\alpha t_1 + \left(1-\alpha\right) t_2\right)}^{t} e\right) =$$

$$= I\left(\alpha u_* + \left(1-\alpha\right) u_* + \left(\alpha t_1 + \left(1-\alpha\right) t_2\right) e\right) =$$

$$= I\left(\alpha\left(u_* + t_1 e\right) + \left(1-\alpha\right)\left(u_* + t_2 e\right)\right) \overset{I \text{ выпукла}}{\le}$$

$$\le \alpha \overbrace{I\left(u_* + t_1 e\right)}^{f(t_1)} + \left(1-\alpha\right) \overbrace{I\left(u_* + t_2 e\right)}^{f(t_2)} = \alpha f\left(t_1\right) + \left(1-\alpha\right) f\left(t_2\right).$$

Выпуклость функции f доказана. Тогда по первому критерию выпуклости

$\overset{f(t)}{I(u)} \ge \overset{f(0)}{I(v)} + \left\langle \overset{f'(0)}{I'(v)}, \overset{t-0}{u - v} \right\rangle$ выводим

$$f(t) \ge f(0) + f'(0) \cdot \left(t - 0\right) \ge 0 \Rightarrow f(t) - f(0) \ge \overbrace{f'(0)}^{=\frac{dI}{de}(u_*) \ge 0 \, (1)} \cdot t \ge 0 \Rightarrow$$

$$f(t) \ge f(0), \forall t \in \left[0, t_0\right].$$

$f(t) \ge f(0), \forall t \in [0, t_0].$ В частности, $f(t_0) \ge f(0).$ Заметим, что

$$f(t_0) = I\left(u_* + \overset{\|u-u_*\|}{t_0} \cdot \overset{\frac{u-u_*}{\|u-u_*\|}}{e} \right) = I\left(u_* + \frac{\|u-u_*\|}{\|u-u_*\|}(u-u_*) \right) = I(u),$$

$$f(0) = I(u_*).$$

Тогда $\overset{I(u)}{f(t_0)} \ge \overset{I(u_*)}{f(0)} \Rightarrow I(u) \ge I(u_*)\ \forall u \in U,$ что и требовалось доказать.

В частности, если в точке u_* существует градиент $I'(u_*),$ то для направления

$e = \dfrac{u-u_*}{\|u-u_*\|}, \quad u \in U, u \ne u_*$ имеем

$$\overset{\langle I'(u_*), e\rangle}{\frac{dI}{de}(u_*)} = \left\langle I'(u_*), \overset{\frac{u-u_*}{\|u-u_*\|}}{e} \right\rangle = \left\langle I'(u_*), \frac{u-u_*}{\|u-u_*\|} \right\rangle = \frac{1}{\|u-u_*\|}\langle I'(u_*), u-u_* \rangle \Rightarrow$$

$$\frac{dI}{de}(u_*) = \overbrace{\frac{1}{\|u-u_*\|}}^{\geq 0}\left\langle I'(u_*), u-u_* \right\rangle$$

Тогда

$$\overbrace{\frac{dI}{de}(u_*) \geq 0}^{(1)} \Leftrightarrow \left\langle I'(u_*), u-u_* \right\rangle \geq 0.$$

Получили формулировку **теоремы 2.** $\left\langle I'(u_*), u-u_* \right\rangle \geq 0$ Таким образом, **теорема 3** является обобщением **теоремы 2** на существенно более широкий класс функций (функций дифференцируемых по всем возможным направлениям).

Упражнение. Точка $u_* = 0$ является точкой минимума функции $I : R^n \to R^1$, определенной формулой $I(u) = \|u\|.$

Доказать непосредственно, что эта точка и только она удовлетворяет неравенству

$$\frac{dI}{de}(u_*) \geq 0$$

для всех возможных направлений e $\left(\|e\|=1\right).$

Решение.

В точке $u = 0$ функция $I(u) = \|u\|$ дифференцируемая только по направлению, поэтому ее производную по направлению e вычисляем непосредственно по определению этой производной

$$\left.\frac{dI}{de}(u)\right|_{u=0} = \lim_{t\to+0}\frac{I(0+te)-I(0)}{t} = \lim_{t\to+0}\frac{\|te\|}{t} = \|e\| = 1.$$

Таким образом, $\left.\dfrac{dI}{de}(u)\right|_{u=0} = 1 > 0, \quad \forall \|e\| = 1.$

Для любой точки $u \in R^n$ любое направление e будет возможным. Вычислим производную по направлению e от функции $I(u) = \|u\|$ в произвольной точке $u \neq 0.$ В этих точках функция является дифференцируемой, поэтому

$$I'(u) = \frac{\partial\sqrt{\langle u,u\rangle}}{\partial u} = \frac{u}{\sqrt{\langle u,u\rangle}} = \frac{u}{\|u\|} \Rightarrow \frac{dI}{de}(u) = \langle I'(u), e\rangle = \frac{\langle u,\ e\rangle}{\|u\|}.$$

Покажем, что для любого $u \in R^n, u \neq 0$ невозможно

$$\frac{dI}{de}(u) \geq 0 \; \forall e, \quad \|e\| = 1.$$

Действительно, пусть $u \in R^n, u \neq 0.$ В силу $u \neq 0$ найдется $e,$ что $\langle e, u \rangle \neq 0.$

Направление $e_1 = -e$ также допустимо. Тогда одна из производных

$$\frac{dI}{de}(u) = \frac{\langle u,\ e \rangle}{\|u\|}, \quad \frac{dI}{de_1}(u) = \frac{\left\langle u,\ \overset{-e}{e_1} \right\rangle}{\|u\|} = -\frac{\langle u,\ e \rangle}{\|u\|}.$$

должна быть строго отрицательной.