ВЫПУКЛЫЙ АНАЛИЗ

ЛЕКЦИЯ 19

7. СУБГРАДИЕНТ И СУБДИФФЕРЕНЦИАЛ ФУНКЦИИ (ПРОДОЛЖЕНИЕ)

7. СУБГРАДИЕНТ И СУБДИФФЕРЕНЦИАЛ ФУНКЦИИ (ПРОДОЛЖЕНИЕ)

7.5. Критерий минимума для субдифференцируемых функций.

7.5. Критерий минимума для субдифференцируемых функций. Для произвольных множеств $U \subset R^n$ и функций $I : U \to R^1$, справедливо следующее утверждение.

Теорема 5. ***Пусть*** $I : U \to R^1, U \subset R^n$. ***Для того, чтобы точка*** $u_* \in U$ ***была точкой минимума функции*** I ***на множестве*** U ***необходимо и достаточно выполнение включения*** $0 \in \partial I(u_*)$.

Доказательство. Пусть u_* – точка минимума функции I на множестве U.

$$I(u_*) = \min_{u \in U} I(u) \Rightarrow \quad I(u) \ge I(u_*) \; \forall u \in U \Rightarrow$$

$$I(u) \ge I(u_*) + \langle 0, u - u_* \rangle, \; \forall u \in U \Rightarrow \; 0 \in \partial I(u_*).$$

Обратно

$$0 \in \partial I(u_*) \Rightarrow \quad I(u) \ge I(u_*) + \langle 0, u - u_* \rangle, \; \forall u \in U \Rightarrow$$

$$I(u) \ge I(u_*) \; \forall u \in U \Rightarrow \; I(u_*) = \min_{u \in U} I(u) \Rightarrow$$

Теорема доказана.

Для выпуклых функций, определенных на выпуклых множествах, справедлив более содержательный результат.

Теорема 6. ***Пусть*** $W \subset R^n$ ***- открытое выпуклое множество,*** $I : W \to R^1$ ***выпуклая функция и*** $U \subset W$ ***- выпуклое подмножество. Для того, чтобы функция*** I ***достигала своей нижней грани на множестве*** U ***в точке*** $u_* \in U$ ***необходимо и достаточно, чтобы существовал субградиент*** $c_* = c(u_*) \in \partial I(u_*)$ ***такой, что***

$$\langle c_*, u - u_* \rangle \geq 0, \forall u \in U.$$

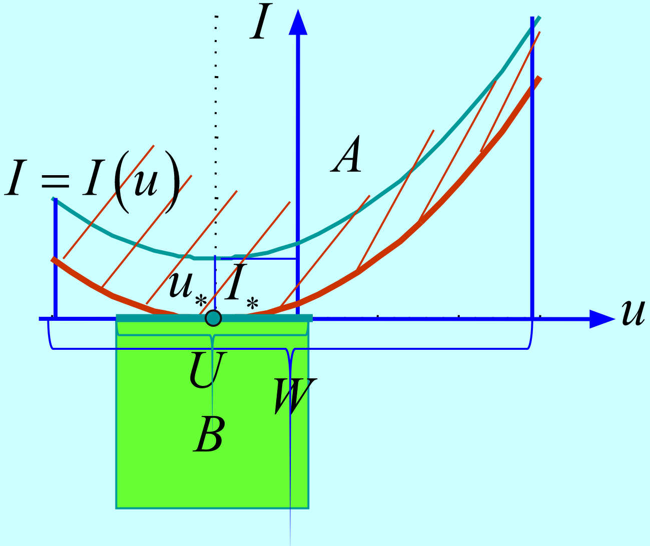
Необходимость. Пусть

$$u_* \in U_* = \left\{ u \in U \,\middle|\, I(u_*) = \inf_{u \in U} I(u) = I_* > -\infty \right\}.$$

В пространстве R^{n+1} введем множества

$$A = \left\{ a = \begin{pmatrix} u \\ a_0 \end{pmatrix} \in R^{n+1} \,\middle|\, u \in W,\, a_0 \geq I(u) - I_* \right\},$$

$$B = \left\{ b = \begin{pmatrix} v \\ b_0 \end{pmatrix} \in R^{n+1} \,\middle|\, v \in U,\, b_0 < 0 \right\} \Rightarrow \bar{B} = \left\{ b = \begin{pmatrix} v \\ b_0 \end{pmatrix} \in R^{n+1} \,\middle|\, v \in \bar{U},\, b_0 \leq 0 \right\}.$$



$$A = \left\{ a = \begin{pmatrix} u \\ a_0 \end{pmatrix} \in R^{n+1} \,\middle|\, u \in W,\ a_0 \ge I(u) - I_* \right\},$$

$$B = \left\{ b = \begin{pmatrix} v \\ b_0 \end{pmatrix} \in R^{n+1} \,\middle|\, v \in U,\ b_0 < 0 \right\},$$

Введенные множества A и B являются выпуклыми.

Факт выпуклости множества A доказывается аналогично выпуклости надграфика выпуклой функции **теорема 4.1.** Выпуклость множества B очевидна. Покажем, что $A \cap B = \varnothing.$

Действительно, пусть $\begin{pmatrix} u \\ a_0 \end{pmatrix} \in A.$ Тогда, либо $u \in U \Rightarrow I(u) - \overset{\min\limits_{u\in U} I(u)}{I_*} \ge 0 \Rightarrow$ $a_0 \ge 0,$ либо $u \in W \setminus U.$ В обоих случаях $a \notin B$ и $A \cap B = \varnothing.$

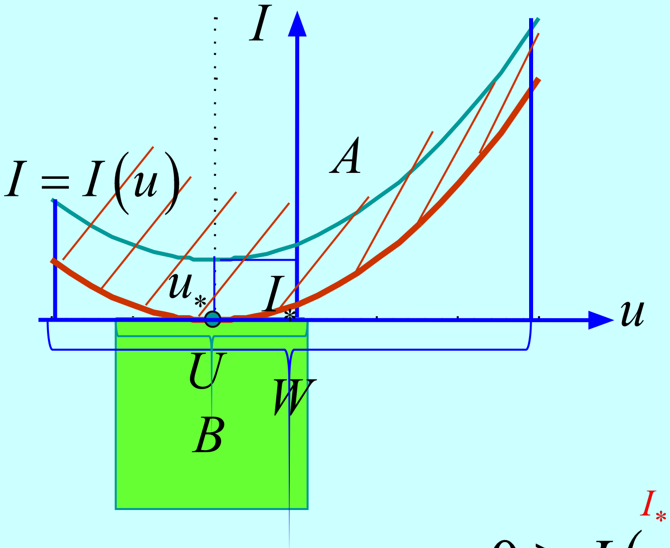
Тогда существует гиперплоскость с нормальным вектором $g = \begin{pmatrix} d \\ \mu \end{pmatrix} \ne 0,$ отделяющая множества $\bar{A} \supset A$ и $\bar{B},$ т.е

$$\langle g, a \rangle \le \gamma \le \langle g, b \rangle,\ a \in A,\ b \in \bar{B}.$$

$$\langle g, a\rangle \le \gamma \le \langle g, b\rangle,\ a\in A, b\in \overline{B}, \quad g=\begin{pmatrix} d\\ \mu\end{pmatrix}\ne 0$$

Отсюда выводим

$$\left\langle \begin{pmatrix} d\\ \mu\end{pmatrix}, \overset{a\in A}{\begin{pmatrix} u\\ a_0\end{pmatrix}} \right\rangle \le \gamma \le \left\langle \begin{pmatrix} d\\ \mu\end{pmatrix}, \overset{b\in \overline{B}}{\begin{pmatrix} v\\ b_0\end{pmatrix}} \right\rangle, \quad \begin{pmatrix} u\\ a_0\end{pmatrix}\in A, \quad \begin{pmatrix} v\\ b_0\end{pmatrix}\in \overline{B}. \quad (1)$$



$$A=\left\{ a=\begin{pmatrix} u\\ a_0\end{pmatrix}\in R^{n+1} \,\middle|\, u\in W,\ a_0\ge I(u)-I_* \right\},$$

$$\overline{B}=\left\{ b=\begin{pmatrix} v\\ b_0\end{pmatrix}\in R^{n+1} \,\middle|\, v\in \overline{U},\ b_0\le 0 \right\}.$$

Заметим, что для точки $\begin{pmatrix} u_*\\ 0\end{pmatrix}\in \overline{B}$ выполнено

$$0\ge \overset{I_*}{I(u_*)}-I_* \Rightarrow \begin{pmatrix} u_*\\ 0\end{pmatrix}\in A \Rightarrow \begin{pmatrix} u_*\\ 0\end{pmatrix}\in A\cap \overline{B}.$$

Тогда из (1) выводим

$$\left\langle \begin{pmatrix} d\\ \mu\end{pmatrix}, \begin{pmatrix} u_*\\ 0\end{pmatrix} \right\rangle \le \gamma \le \left\langle \begin{pmatrix} d\\ \mu\end{pmatrix}, \begin{pmatrix} u_*\\ 0\end{pmatrix} \right\rangle \Rightarrow \gamma=\langle d, u_*\rangle.$$

Неравенство (1) $\left\langle \begin{pmatrix} d \\ \mu \end{pmatrix}, \begin{pmatrix} u \\ a_0 \end{pmatrix} \right\rangle \le \overset{\langle d, u_* \rangle}{\gamma} \le \left\langle \begin{pmatrix} d \\ \mu \end{pmatrix}, \begin{pmatrix} v \\ b_0 \end{pmatrix} \right\rangle \quad (1)$ перепишем в виде

$$\left\langle \begin{pmatrix} d \\ \mu \end{pmatrix}, \begin{pmatrix} u \\ a_0 \end{pmatrix} \right\rangle \le \langle d, u_* \rangle \le \left\langle \begin{pmatrix} d \\ \mu \end{pmatrix}, \begin{pmatrix} v \\ b_0 \end{pmatrix} \right\rangle, \quad \Rightarrow$$

$$\langle d, u \rangle + \mu a_0 \le \langle d, u_* \rangle \le \langle d, v \rangle + \mu b_0, \quad \begin{pmatrix} u \\ a_0 \end{pmatrix} \in A, \begin{pmatrix} v \\ b_0 \end{pmatrix} \in \bar{B} \qquad (2)$$

Из правого неравенства в (2) при $b = \begin{pmatrix} u_* \\ -1 \end{pmatrix} \in \bar{B} = \left\{ b = \begin{pmatrix} v \\ b_0 \end{pmatrix} \in R^{n+1} \,\middle|\, \begin{matrix} v \in \bar{U}, \\ b_0 \le 0 \end{matrix} \right\}$

получим

$$\langle d, u_* \rangle \le \left\langle d, \overset{u_*}{v} \right\rangle + \mu \overset{-1}{b_0} = \langle d, u_* \rangle - \mu \Rightarrow \langle d, u_* \rangle \le \langle d, u_* \rangle - \mu \Rightarrow \quad \mu \le 0.$$

Покажем, что $\mu < 0.$ Пусть все же $\mu = 0.$ Тогда из левого неравенства в (2) выводим

$$\langle d, u \rangle \le \langle d, u_* \rangle, \ \forall u \in W. \qquad (3)$$

Полагаем в (3) $\langle d, u\rangle \le \langle d, u_*\rangle, \ \forall u \in W \quad (3)$ $u = u_* + \varepsilon\, d,$ где

$\varepsilon > 0$ настолько мало, что $u = u_* + \varepsilon\, d \in W.$ Из (3) выводим

$$\langle d, u_* + \varepsilon\, d\rangle \le \langle d, u_*\rangle \Rightarrow \overset{>0}{\varepsilon}\, d^2 \le 0 \Rightarrow \ d = 0.$$

Получили противоречие с тем, что $g = \begin{pmatrix} d \\ \mu \end{pmatrix} \ne 0.$ Таким образом , $\mu < 0.$

Разделим (2) $\langle d, u\rangle + \mu\, a_0 \le \langle d, u_*\rangle \le \langle d, v\rangle + \mu\, b_0 \quad (2)$ на $\mu < 0.$

В результате получим

$$\left\langle \frac{d}{\mu}, u \right\rangle + a_0 \ge \left\langle \frac{d}{\mu}, u_* \right\rangle \ge \left\langle \frac{d}{\mu}, v \right\rangle + b_0.$$

Полагаем $c_* = -\dfrac{d}{\mu},$ Тогда

$$-\langle c_*, u\rangle + a_0 \ge -\langle c_*, u_*\rangle \ge -\langle c_*, v\rangle + b_0,$$

$$\forall u \in W,\ a_0 \ge I(u) - I_*,\ \forall v \in \bar{U},\ b_0 \le 0 \Rightarrow$$

$$-\langle c_*, u\rangle + a_0 \ge -\langle c_*, u_*\rangle \ge -\langle c_*, v\rangle + b_0, \quad -\langle c_*, u-u_*\rangle + a_0 \ge 0 \ge -\langle c_*, v-u_*\rangle + b_0, \qquad (4)$$

$$\forall u \in W,\ a_0 \ge I(u) - I_*,\ \forall v \in \bar{U},\ b_0 \le 0.$$

При $a_0 = I(u) - I_*$ из левого неравенства в (4) $-\langle c_*, u-u_*\rangle + \overset{I(u)-I_*}{a_0} \ge 0$ (4) находим

$$-\langle c_*, u-u_*\rangle + I(u) - I_* \ge 0 \Rightarrow$$

$$I(u) - \overset{I(u_*)}{I_*} \ge \langle c_*, u-u_*\rangle,\ \forall u \in W \Rightarrow c_* \in \partial I(u_*)$$

При $b_0 = 0$ из правого неравенства в (4) находим

$$\langle c_*, v-u_*\rangle \ge \overset{0}{b_0} = 0 \Rightarrow \langle c_*, v-u_*\rangle \ge 0, \quad \forall v \in U \subset \bar{U}.$$

Необходимость $\exists c_* \in \partial I(u_*): \langle c_*, u-u_*\rangle \ge 0, \forall u \in U$ доказана.

Достаточность. Пусть для некоторой точки $u_* \in U$ и $c_* \in \partial I(u_*)$ выполнено

$$\langle c_*, u-u_*\rangle \ge 0, \forall u \in U. \qquad (5)$$

По определению субградиента $I(u) \ge I(\overset{u_*}{v}) + \langle \overset{c_*}{c}, u-v\rangle, c \in \partial I(v)$ тогда

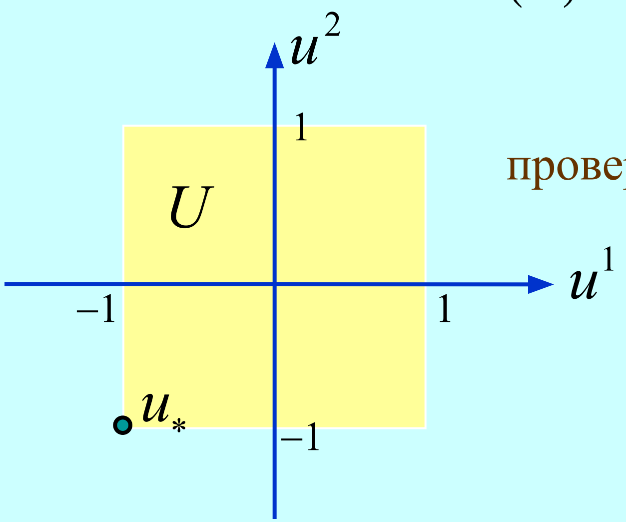
$$I(u) - I(u_*) \overset{(5)}{\ge} \langle c_*, u-u_*\rangle \ge 0, \forall u \in U \Rightarrow u_* \in U_*.$$

Достаточность доказана. Теорема доказана полностью.

Замечание. Как видно из доказательства теоремы множества $A, B \subset R^{n+1}$ вектор

$$g = \begin{pmatrix} d \\ \mu \end{pmatrix} \in R^{n+1}, \quad \text{субградиент} \quad c_* = -\frac{d}{\mu} \in \partial I(u_*) \quad \text{не зависят от выбора} \quad u_* \in U_*.$$

Упражнение. Для $I(u) = \max\left\{\begin{pmatrix} u^1 \\ u^2 \end{pmatrix}\right\}$, и $U = \left\{\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \middle| \, |u^1| \le 1, |u^2| \le 1\right\}$.

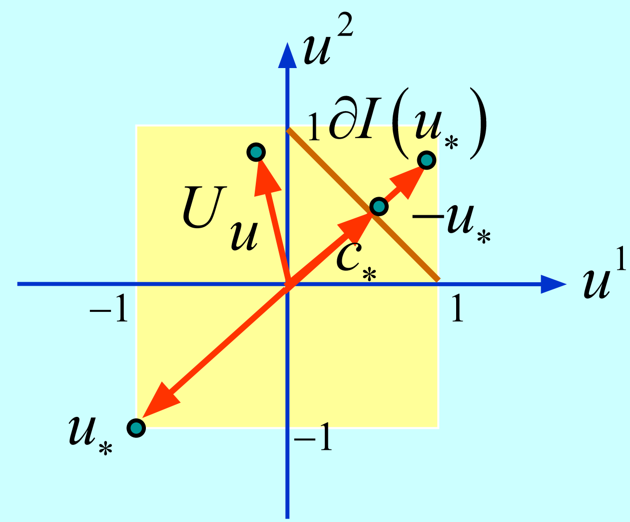


проверить выполнение **теоремы 6** $\exists c_* \in \partial I(u_*): \langle c_*, u - u_* \rangle \ge 0$

Решение.

Здесь $u_* = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Тогда $z(u_*) = \{1, 2\}$ и

$$\partial I(u_*) = \left\{ c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \middle| \, c_1 + c_2 = 1, c_1 \ge 0, c_2 \ge 0 \right\}.$$



В качестве $c_* \in \partial I(u_*)$ возьмем вектор $c_* = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Тогда

$$\left\langle c_*, u-u_* \right\rangle = \left\langle \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} u^1-(-1) \\ u^2-(-1) \end{pmatrix} \right\rangle = \frac{1}{2}\left(u^1+u^2\right)+1 \geq 0$$

для всех $\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \in U = \left\{ \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \middle| \left|u^1\right| \leq 1, \left|u^2\right| \leq 1 \right\}.$

Теорема 6. ***Пусть*** $W \subset R^n$ ***- открытое выпуклое множество,*** $I : W \to R^1$ ***выпуклая функция и*** $U \subset W$ ***- выпуклое подмножество. Для того, чтобы функция*** I ***достигала своей нижней грани на множестве*** U ***в точке*** $u_* \in U$ ***необходимо и достаточно, чтобы существовал субградиент*** $c_* = c\left(u_*\right) \in \partial I\left(u_*\right)$ ***такой, что***

$$\left\langle c_*, u-u_* \right\rangle \geq 0, \forall u \in U. \quad (5)$$

Следствие. ***Пусть в условиях доказанной теоремы*** $u_* \in \operatorname{int} U$. ***Тогда субградиент, обеспечивающий неравенство (5), может быть только нулевым вектором.***

Действительно, $u_* \in \operatorname{int} U \Rightarrow u = u_* + \overset{<0}{\varepsilon}\, c_* \in U$, при $|\varepsilon| \le \varepsilon_0$, где $\varepsilon_0 > 0$ достаточно мало. Из (5) $\langle c_*, u - u_* \rangle \ge 0, \forall u \in U \quad (5)$ для такого $u \in U$ выводим

$$\left\langle c_*, u_* + \overset{<0}{\varepsilon}\, c_* - u_* \right\rangle \ge 0 \Rightarrow \overset{<0}{\varepsilon}\, c_*^2 \ge 0 \Rightarrow c_* = 0.$$

Таким образом, проверка точек $u_* \in \operatorname{int} U$ на оптимальность сводится к проверке включения $0 \in \partial I(u_*)$.

Почему в упражнении $c_* = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \ne 0$?

Пример 4. Пусть $I(u) = |u|,\ u \in U = R^1$. Здесь $U_* = \{u_*\} = \{0\}$ и $u_* = 0 \in \operatorname{int} U$. Очевидно, что $c_* = 0 \in \partial I(0) = \{c \in R^1 \,|\, |c| \le 1\}$. Покажем, что других субградиентов $c_* \in \partial I(0) = \{c \in R^1 \,|\, |c| \le 1\}$, обеспечивающих неравенство (5),

$\langle c_*, u - u_* \rangle \ge 0, \forall u \in U \quad (5)$ $\left\langle c_*, u - \overset{0}{u_*} \right\rangle \ge 0, \Rightarrow c_* \cdot u \ge 0, \forall u \in R^1$, не существует.

$> \quad <$

$< \quad >$

Действительно, $c_* \cdot u < 0.$