ВЫПУКЛЫЙ АНАЛИЗ

ЛЕКЦИЯ 13

3. ВЫПУКЛЫЕ ФУНКЦИИ (ПРОДОЛЖЕНИЕ)

3. ВЫПУКЛЫЕ ФУНКЦИИ

3.5. Критерии выпуклости гладких функций

3.5. Критерии выпуклости гладких функций. Проверка произвольной функции на выпуклость непосредственно по определению выпуклости обычно сопряжено со значительными трудностями. Для гладких функций существуют более конструктивные критерии выпуклости, упрощающие эту проверку.

Теорема 8. (Первый критерий выпуклости дифференцируемой функции). ***Пусть*** $U \subset R^n$ ***выпуклое множество,*** $I \in C^1(U)$, ***тогда для выпуклости функции*** I ***на множестве*** U ***необходимо и достаточно, чтобы***

$$I(u) \geq I(v) + \left\langle I'(v), u - v \right\rangle, \forall u, v \in U. \qquad (1)$$

Доказательство. Необходимость. Для всех $u, v \in U$ и $\alpha \in [0,1]$ имеем

$$I\left(\alpha u + (1-\alpha)v\right) \overset{\text{выпуклость } I}{\leq} \alpha I(u) + (1-\alpha) I(v) \Rightarrow$$

$$I\left(v + \alpha(u - v)\right) \leq I(v) + \alpha\left(I(u) - I(v)\right) \Rightarrow$$

$$I\left(v + \alpha(u - v)\right) - I(v) \leq \alpha\left(I(u) - I(v)\right) \Rightarrow$$

$=\alpha\left\langle I'\left(v+\theta\alpha(u-v)\right),u-v\right\rangle$

$$I\left(v+\alpha(u-v)\right)-I(v)\le\alpha\left(I(u)-I(v)\right)\Rightarrow$$

0

$$\alpha\left\langle I'\left(v+\theta\alpha(u-v)\right),u-v\right\rangle\le\alpha\left(I(u)-I(v)\right),\ \theta\in[0,1].\qquad(2)$$

Разделим неравенство (2) на $\alpha>0$ и устремим α к нулю

$$\left\langle I'(v),u-v\right\rangle\le I(u)-I(v)$$

В результате получим (1) $I(u)\ge I(v)+\left\langle I'(v),u-v\right\rangle \quad (1).$ Необходимость доказана.

Достаточность. Пусть $u,v\in U$ и $\alpha\in[0,1].$ Положим

$$w_\alpha=\alpha u+(1-\alpha)v\in U.$$

Из (1) $I(u)\ge I(v)+\left\langle I'(v),u-v\right\rangle,\ \forall u,v\in U \quad (1)$ находим

для $u,w_\alpha\in U:\quad I(u)-I(w_\alpha)\ge\left\langle I'(w_\alpha),u-w_\alpha\right\rangle,$

для $v,w_\alpha\in U:\quad I(v)-I(w_\alpha)\ge\left\langle I'(w_\alpha),v-w_\alpha\right\rangle.$

Первое из этих неравенств умножим на $\alpha,$ второе на $(1-\alpha)$ и сложим их почленно.

Имеем

$$I(u)-I(w_\alpha)\geq\langle I'(w_\alpha),u-w_\alpha\rangle\Big|\,\alpha$$

$$I(v)-I(w_\alpha)\geq\langle I'(w_\alpha),v-w_\alpha\rangle\Big|\,(1-\alpha)$$

$$\alpha\, I(u)+(1-\alpha)\,I(v)-I(w_\alpha)\geq$$

$$\geq\langle I'(w_\alpha),\alpha(u-w_\alpha)+(1-\alpha)(v-w_\alpha)\rangle=$$

$$=\left\langle I'(w_\alpha),\overbrace{\alpha u+(1-\alpha)v}^{=w_\alpha}-w_\alpha\right\rangle=\langle I'(w_\alpha),w_\alpha-w_\alpha\rangle=0.$$

Таким образом,

$$\alpha\, I(u)+(1-\alpha)\,I(v)-I(w_\alpha)\geq 0\Rightarrow$$

$$I\left(\overbrace{w_\alpha}^{\alpha u+(1-\alpha)v}\right)=I(\alpha u+(1-\alpha)v)\leq\alpha\, I(u)+(1-\alpha)I(v).$$

и теорема доказана.

Теорема 9. (Второй критерий выпуклости дифференцируемой функции). *Пусть* $U \subset R^n$ *выпуклое множество,* $I \in C^1(U)$, *тогда для выпуклости функции* I *на множестве* U *необходимо и достаточно, чтобы*

$$\langle I'(u) - I'(v), u - v \rangle \geq 0, \forall u, v \in U. \qquad (3)$$

Доказательство. Необходимость. В силу выпуклости функции I на множестве U по первому критерию выпуклости $I(u) \geq I(v) + \langle I'(v), u - v \rangle, \forall u, v \in U \quad (1)$ (**теорема 8**) следует, что для любой пары точек $u, v \in U$ справедливы неравенства

$$I(u) \geq I(v) + \langle I'(v), u - v \rangle, \; I(v) \geq I(u) + \langle I'(u), v - u \rangle.$$

Сложим эти неравенства почленно

$$I(u) + I(v) \geq \quad I(v) + I(u) + \langle I'(v), u - v \rangle + \langle I'(u), v - u \rangle \Rightarrow$$

$$-\langle I'(v), u - v \rangle - \langle I'(u), v - u \rangle \geq 0 \Rightarrow$$

$$-\langle I'(v), u - v \rangle + \langle I'(u), u - v \rangle \geq 0 \Rightarrow \langle I'(u) - I'(v), u - v \rangle \geq 0$$

Необходимость доказана.

Достаточность. Для всех $u, v \in U, \alpha \in [0,1]$ следует доказать справедливость неравенства

$$I\left(\alpha u+(1-\alpha)v\right) \le \alpha I(u)+(1-\alpha)I(v) \Rightarrow$$

$$\alpha I(u)+(1-\alpha)I(v)-I\left(\alpha u+(1-\alpha)v\right) \ge 0 \Rightarrow$$

$$\alpha I(u)+(1-\alpha)I(v)-I\left(\alpha u+(1-\alpha)v\right)+\alpha I\left(\alpha u+(1-\alpha)v\right)-$$

$$-\alpha I\left(\alpha u+(1-\alpha)v\right) = \alpha I(u)-\alpha I\left(\alpha u+(1-\alpha)v\right)+$$

$$+(1-\alpha)I(v)+\alpha I\left(\alpha u+(1-\alpha)v\right)-I\left(\alpha u+(1-\alpha)v\right)=$$

$$=\alpha\left[I(u)-I\left(\alpha u+(1-\alpha)v\right)\right]+(1-\alpha)I(v)+$$

$$+I\left(\alpha u+(1-\alpha)v\right)(\alpha-1) = \alpha\underbrace{\left[I(u)-I\left(\alpha u+(1-\alpha)v\right)\right]}_{I_1}+$$

$$+(1-\alpha)\underbrace{\left[I(v)-I\left(\alpha u+(1-\alpha)v\right)\right]}_{I_2}$$

Таким образом следует доказать. что

$$\alpha I_1+(1-\alpha)I_2 \ge 0.$$

Для всех $w, \Delta w \in R^n$ имеет место формула конечных приращений

$$I(w+\Delta w) - I(w) = \int_0^1 \langle I'(w+t\Delta w),\ \Delta w\rangle dt \qquad (4)$$

последовательно вычисляем выражения

$$I_1 = I(u) - I(\alpha u + (1-\alpha)\ v),\ I_2 = I(v) - I(\alpha u + (1-\alpha)v)$$

Для выражения I_1 полагаем

$$w_1 = \alpha u + (1-\alpha)v,\ \ w_1 + \Delta w_1 = u \Rightarrow \Delta w_1 = u - \overset{\alpha u+(1-\alpha)v}{w_1} = u -$$

$$-\alpha u - (1-\alpha)v = u(1-\alpha) - (1-\alpha)v = (1-\alpha)(u-v). \quad (5)$$

Тогда в силу (4) и (5) выводим

$$I_1 = I\left(\overset{w_1+\Delta w_1}{u}\right) - I\left(\alpha u + \overset{w_1}{(1-\alpha)v}\right) = I(w_1+\Delta w_1) - I(w_1) \overset{(4)}{=}$$

$$\overset{(4)}{=} \int_0^1 \left\langle I'\left(\overset{\alpha u+(1-\alpha)v}{w_1} + t\overset{(1-\alpha)(u-v)}{\Delta w_1}\right),\ \overset{(1-\alpha)(u-v)}{\Delta w_1}\right\rangle dt =$$

$$= (1-\alpha)\int_0^1 \langle I'(\alpha u + (1-\alpha)v + t(1-\alpha)(u-v)), (u-v)\rangle dt. \quad (6)$$

Для выражения I_2 полагаем $w_2 = \alpha u + (1-\alpha)v,\ \ w_2 + \Delta w_2 = v \Rightarrow$

$$\Delta w_2 = v - \overbrace{w_2}^{\alpha u-(1-\alpha)v} = v - \alpha u - (1-\alpha)v = \alpha(v-u). \quad (7)$$

Тогда в силу (4) $I(w+\Delta w) - I(w) = \int_0^1 \langle I'(w + t\Delta w),\ \Delta w\rangle dt \quad (4)$ находим

$$I_2 = I\left(\overbrace{v}^{w_2+\Delta w_2}\right) - I\left(\overbrace{\alpha u + (1-\alpha)v}^{w_2}\right) = I(w_2 + \Delta w_2) - I(w_2) \overset{(4)}{=}$$

$$\overset{(4)}{=} \alpha \int_0^1 \left\langle I'\left(\overbrace{w_2}^{\alpha u+(1-\alpha)v} + t\,\overbrace{\Delta w_2}^{\alpha(v-u)}\right), v - u \right\rangle dt =$$

$$= \alpha \int_0^1 \langle I'(\alpha u + (1-\alpha)v + t\alpha(v-u)), v - u\rangle dt.$$

Таким образом,

$$\Delta I(u,v,\alpha) = \alpha I_1 + (1-\alpha) I_2 =$$

$$= \alpha \overbrace{I_1}^{(1-\alpha)\int_0^1 \langle I'(\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)),(u-v)\rangle dt} + (1-\alpha) \overbrace{I_2}^{\alpha\int_0^1 \langle I'(\alpha u+(1-\alpha)v+t\alpha(v-u)),v-u\rangle dt} =$$

$$=\alpha(1-\alpha)\int_0^1\left\langle I'\big(\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)\big),u-v\right\rangle dt+$$

$$+(1-\alpha)\alpha\int_0^1\left\langle I'\big(\alpha u+(1-\alpha)v+t\alpha(v-u)\big),v-u\right\rangle dt=$$

$$=\alpha(1-\alpha)\cdot\left[\int_0^1\left\langle I'\left(\overbrace{\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)}^{z_1}\right)-\right.\right.$$

$$\left.\left.-I'\left(\overbrace{\alpha u+(1-\alpha)v+t\alpha(v-u)}^{z_2}\right),u-v\right\rangle dt\right]. \qquad (8)$$

Обозначим

$$z_1=\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)=$$

$$=\alpha u+(1-\alpha)v+t(1-\alpha)u-t(1-\alpha)v=$$

$$=\big[\overbrace{\alpha+t(1-\alpha)}^{\beta_{11}}\big]u+\big[\overbrace{(1-\alpha)-t(1-\alpha)}^{\beta_{12}}\big]v=\beta_{11}u+\beta_{12}v.$$

Заметим, что

$$\beta_{11} = \alpha + t\left(1-\alpha\right) \geq 0, \;\; \beta_{12} = \left(1-\alpha\right) - t\left(1-\alpha\right) \geq 0.$$

При этом

$$\beta_{11} + \beta_{12} = \alpha + t\left(1-\alpha\right) + \left(1-\alpha\right) - t\left(1-\alpha\right) =$$

$$= \alpha + 1 - \alpha + t\left(1-\alpha\right) - t\left(1-\alpha\right) = 1$$

Из выпуклости множества U отсюда следует, что $z_1 = \beta_{11} u + \beta_{12} v \in U.$

Аналогично, обозначим

$$z_2 = \alpha u + \left(1-\alpha\right)v + t\alpha\left(v-u\right) =$$

$$= \alpha u + \left(1-\alpha\right)v + t\alpha v - t\alpha u = \overbrace{\alpha\left(1-t\right)}^{\beta_{21}}u + \overbrace{\left(1-\alpha\left(1-t\right)\right)}^{\beta_{22}}v =$$

$$= \beta_{21} u + \beta_{22} v.$$

Заметим, что

$$\beta_{21} = \alpha\left(1-t\right) \ge 0, \qquad \beta_{22} = 1-\alpha\left(1-t\right) \ge 0$$

При этом

$$\beta_{21} + \beta_{22} = \alpha\left(1-t\right) + 1 - \alpha\left(1-t\right) = 1.$$

Из выпуклости множества U отсюда следует, что $z_2 = \beta_{21}u + \beta_{22}v \in U.$

Кроме того

$$\overset{\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)}{z_1} - \overset{\alpha u+(1-\alpha)v+t\alpha(v-u)}{z_2} =$$

$$= \alpha u + \left(1-\alpha\right)v + t\left(1-\alpha\right)\left(u-v\right) - \alpha u - \left(1-\alpha\right)v - t\alpha\left(v-u\right) =$$

$$= t\left(1-\alpha\right)\left(u-v\right) - t\alpha\left(v-u\right) =$$

$$= t\left(u-v\right)\left(1-\alpha+\alpha\right) = t\left(u-v\right) \Rightarrow$$

$$u - v = \frac{1}{t}\left(z_1 - z_2\right). \qquad (9)$$

Из (8)

$$\Delta I(u,v,\alpha)=\alpha(1-\alpha)\times\int_0^1\left\langle I'\left(\overbrace{\alpha u+(1-\alpha)v+t(1-\alpha)(u-v)}^{z_1}\right)-\right.$$

$$\left.-I'\left(\overbrace{\alpha u+(1-\alpha)v+t\alpha(v-u)}^{z_2}\right),\ \overbrace{u-v}^{\frac{1}{t}(z_1-z_2)}\right\rangle dt \qquad (8)$$

$$z_1=\alpha u+(1-\alpha)v+t(1-\alpha)(u-v),$$

в силу (9) $$z_2=\alpha u+(1-\alpha)v+t\alpha(v-u)$$ имеем

$$u-v=\frac{1}{t}(z_1-z_2) \qquad (9)$$

$$\Delta I(u,v,\alpha)=\alpha(1-\alpha)\int_0^1\overbrace{\left\langle I'(z_1)-I'(z_2),\ z_1-z_2\right\rangle}^{\geq 0\ \text{по условию теоремы}}\ \overbrace{\frac{1}{t}}^{\geq 0}\,dt.$$

По доказанному $z_1, z_2\in U.$ Тогда для этих точек выполнено условие (3)

$$\left\langle I'(u)-I'(v),\ u-v\right\rangle\geq 0,\ \forall\, u,v\in U. \qquad (3)$$

Следовательно $\Delta I(u,v,\alpha)\geq 0.$ Теорема доказана.

Теорема 10. (Критерий выпуклости дважды дифференцируемой функции). *Пусть* $U \subset R^n$ *выпуклое множество,* $I \in C^2(U)$, *тогда для выпуклости функции* I *на множестве* U *достаточно, чтобы*

$$\left\langle I''(u)\xi, \xi \right\rangle \geq 0, \quad \forall u \in U, \xi \in R^n, \qquad (10)$$

если $\operatorname{int} U \neq \varnothing$, *то это условие является необходимым.*

Доказательство. Необходимость. Пусть $\operatorname{int} U \neq \varnothing$. Возьмем $u \in \operatorname{int} U$ и $\xi \in R^n$. Тогда найдется число $\varepsilon_0 > 0$, что $u + \varepsilon\xi \in U$, если $|\varepsilon| \leq \varepsilon_0$.

Из второго критерия выпуклости $\left\langle I'(\overset{u+\varepsilon\xi}{u}) - I'(\overset{u}{v}), \overset{\varepsilon\xi}{u - v} \right\rangle \geq 0, \forall u, v \in U \quad (3)$

выводим

$$\left\langle I'(u+\varepsilon\xi) - I'(u), \varepsilon\xi \right\rangle = \left\langle I''(u+\theta\varepsilon\xi)\varepsilon\xi, \varepsilon\xi \right\rangle =$$

$$= \varepsilon^2 \left\langle I''(u+\theta\varepsilon\xi)\xi, \xi \right\rangle \overset{(3)}{\geq} 0, \quad \theta \in [0,1], \varepsilon \in [-\varepsilon_0, \varepsilon_0]. \qquad (11)$$

Сокращая в неравенстве (11) $\varepsilon^2 \left\langle I''\left(u+\theta\varepsilon\xi\right)\xi,\xi\right\rangle \geq 0 \quad (11)$ на $\varepsilon^2 > 0$

и устремляя ε к нулю, получим (10). $\left\langle I''(u)\xi,\xi\right\rangle \geq 0, \quad (10)$

Если $u \in U \setminus \operatorname{int} U$, то так как выпуклое множество не может содержать изолированных точек, найдется последовательность

$$\{u_k\} \to u,\ u_k \in \operatorname{int} U,\ k = 1, 2, \ldots \ .$$

По доказанному выше будет выполняться

$$\left\langle I''(u_k)\xi,\xi\right\rangle \geq 0,\ \ \xi \in R^n.$$

Переходя здесь к пределу при $k \to \infty$, получим условие (9) и для точек $u \in U \setminus \operatorname{int} U$. Необходимость доказана.

Достаточность. В силу второго критерия выпуклости достаточно доказать неравенство

$$\left\langle I'(u) - I'(v), u - v\right\rangle \geq 0, \quad u, v \in U.$$

Имеем

$$\left\langle I'(u) - I'(v), u - v\right\rangle = \left\langle \overbrace{I'(v + (u - v)) - I'(v)}^{I''(v+\theta(u-v))(u-v),\ \theta\in[0,1]}, u - v \right\rangle =$$

$$= \left\langle I''\left(\overbrace{v+\theta(u-v)}^{w}\right)\overbrace{(u-v)}^{\xi},\ \overbrace{u-v}^{\xi} \right\rangle = \left\langle I''(w)\xi,\xi \right\rangle,$$

Где обозначено

$$w = v + \overset{\theta\in[0,1]}{\theta}\,(u-v) = \theta\,\overset{\in U}{u} + (1-\theta)\,\overset{\in U}{v} \Rightarrow w \in U,$$

$$\xi = u - v \in R^n.$$

Тогда из условия (10) $\left\langle I''(u)\xi,\xi\right\rangle \geq 0, \quad \forall u \in U, \xi \in R^n \quad (10)$ получим искомое

$$\left\langle I'(u) - I'(v), u - v\right\rangle = \left\langle I''(w)\xi,\xi\right\rangle \overset{(10)}{\geq} 0.$$

Таким образом, для функции I выполнен второй критерий выпуклости дифференцируемых функций и, следовательно, она выпукла. Теорема доказана.

Пример 10. Определить значения параметров, для которых функция $I : R^3 \to R^1$, определенная равенством

$$I(x,y,z) = x^2 + 2axy + by^2 + cz^2$$

будет выпуклой на всем пространстве R^3.

Решение.

Функция I дважды непрерывно дифференцируема на всем пространстве R^3.

По **теореме 10** для ее выпуклости требуется положительность матрицы

$$I''(x,y,z)=\begin{pmatrix} 2 & 2a & 0 \\ 2a & 2b & 0 \\ 0 & 0 & 2c \end{pmatrix}.$$

Вычислим ее главные миноры и выясним, при каких значениях параметров они будут не отрицательны.

$$2>0,\ \begin{vmatrix} 2 & 2a \\ 2a & 2b \end{vmatrix}=4b-4a^2\geq 0,\ \begin{vmatrix} 2 & 2a & 0 \\ 2a & 2b & 0 \\ 0 & 0 & 2c \end{vmatrix}=\left(4b-4a^2\right)2c\geq 0.$$

Таким образом, искомая область изменения параметров определяется неравенствами

$a^2\leq b,\ c\geq 0.$ Например, значения параметров $a=2,\ b=5,\ c=1$ этим неравенствам удовлетворяют.

Упражнение. Построить область изменения значений параметров a и b, для которых функция

$$I(x,y)=\frac{1}{2}\left[(1-b)x^2+4axy+(1+b)y^2\right]$$

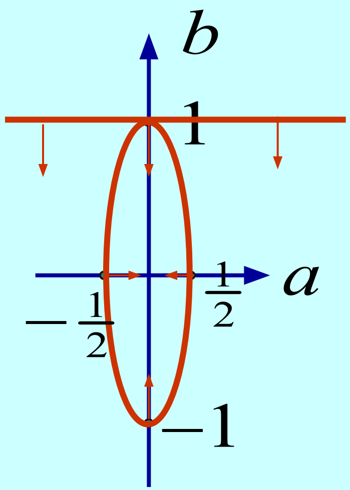
будет выпуклой на всем пространстве R^2.

Решение.

$$I'(x,y)=\begin{pmatrix}(1-b)x+2ay\\(1+b)y+2ax\end{pmatrix}\Rightarrow I''(x,y)=\begin{pmatrix}1-b & 2a\\2a & 1+b\end{pmatrix}$$

Условия выпуклости функции I:

$$\begin{cases}1-b\ge 0,\\1-b^2-4a^2\ge 0\end{cases}\Rightarrow\begin{cases}b\le 1,\\b^2+4a^2\le 1\end{cases}\Rightarrow$$



$$\Rightarrow\begin{cases}b\le 1,\\\dfrac{a^2}{\left(\frac{1}{2}\right)^2}+\dfrac{b^2}{1^2}\le 1\end{cases}$$