Выпуклые функции и глобальные минимумы

Выпуклые множества

Определение 5 *Множество D называется* ***выпуклым****, если для всех $s, v \in D$ и любого $\alpha \in [0, 1]$ выполнено $\alpha s + (1-\alpha)v \in D$.*

Геометрически выпуклость множества означает, что отрезок, соединяющий любые две точки из множества D, целиком лежит в множестве D.
Отрезок, соединяющий точки s и v, это множество точек вида

$$\{y \in \mathbf{R}^n : \; y = \alpha s + (1-\alpha)v, \, 0 \le \alpha \le 1\}$$

Пустое и одноточечное множества считаются выпуклыми по определению.

Пример 7 *Множество точек вида*

$$\Gamma = \{ x \in \mathbf{R}^n | \, (c, x) = \gamma \} , \quad c \in \mathbf{R}^n, \quad c \neq 0, \quad \gamma \in \mathbf{R},$$

называемое гиперплоскостью в $\mathbf{R}^n$*, является выпуклым.*

Пример 8 *Замкнутые полупространства, порожденные гиперплоскостью* Γ*,*

$$\Gamma^+ = \{ x \in \mathbf{R}^n | \, (c, x) \geq \gamma \} , \quad \Gamma^- = \{ x \in \mathbf{R}^n | \, (c, x) \leq \gamma \} ,$$

являются выпуклыми множествами в $\mathbf{R}^n$*.*

Определение 6 *Множества вида*

$$X = \left\{ x \in \mathbf{R}^n \middle| \left(c^i, x\right) \le \gamma_i, \quad i = 1, \ldots, m \right\},$$

где $c^i = (c^i_1, \ldots, c^i_n)$, $\gamma_i \in \mathbf{R}$ $(i = 1, \ldots, m)$ *называются полиэдральными или полиэдрами.*

То есть полиэдр — это пересечение конечного числа полупространств. Из следующей теоремы будет следовать, что полиэдр — выпуклое множество.

Теорема 9
Пересечение (конечное или бесконечное) выпуклых множеств — выпуклое множество.

Доказательство. Пусть множества W_a выпуклы в $\mathbf{R}^n$ для всех $a \in A$. Обозначим

$$W = \bigcap_{a \in A} W_a.$$

Пусть w_1, $w_2 \in W$, тогда w_1, $w_2 \in W_a$ для всех $a \in A$. В силу выпуклости каждого W_a имеем

$$\alpha w_1 + (1 - \alpha)\, w_2 \in W_a, \quad \forall a \in A, \quad \forall \alpha \in [0, 1].$$

Отсюда следует, что $\alpha w_1 + (1 - \alpha)\, w_2 \in W$, то есть множество W выпукло. □

Объединение выпуклых множеств не обязательно выпукло.

Выпуклые функции

Пусть $f : D \to \mathbf{R}$, где D – выпуклое множество.

Определение 7 *Функция $f(x)$ называется* ***выпуклой*** *на множестве D, если $\forall\, s, v \in D$ и $\forall\, \alpha \in [0,1]$ выполнено* ***неравенство Йенсена****:*

$$f(\alpha s + (1-\alpha)v) \le \alpha f(s) + (1-\alpha) f(v). \qquad (12)$$

Если (12) выполнено строго для всех $s, v \in D$, $s \neq v$, и для всех $0 < \alpha < 1$, то функция f называется строго выпуклой на множестве D.
Если неравенство (12) выполнено в противоположную сторону, т.е. $\ge$, то функция $f(x)$ называется вогнутой.
Если $f(x)$ — вогнутая функция, то $-f(x)$ — выпуклая функция.

***Пример* 9** *Функция $f\ : \mathbf{R}^n \to \mathbf{R}$, определенная формулой $f\,(x) = (c, x)$, $c \in \mathbf{R}^n$, является выпуклой.*

Дадим эквивалентное определение выпуклой функции через ее надграфик.

Определение 8 *Рассмотрим функцию* $f : D \to \mathbf{R}$, *где* $D \subset \mathbf{R}^n$.

Надграфиком функции f *назовем множество*

$$epi\, f = \left\{(x, c) \in \mathbf{R}^{n+1} \ : x \in D,\, c \in \mathbf{R},\, c \geq f\left(x\right)\right\}.$$

Теорема 10 *Пусть* D *— выпуклое множество,* $f \, : D \to \mathbf{R}$.

Функция f *выпукла тогда и только тогда,*
когда $epi\, f$ *— выпуклое множество в* $\mathbf{R}^{n+1}$.

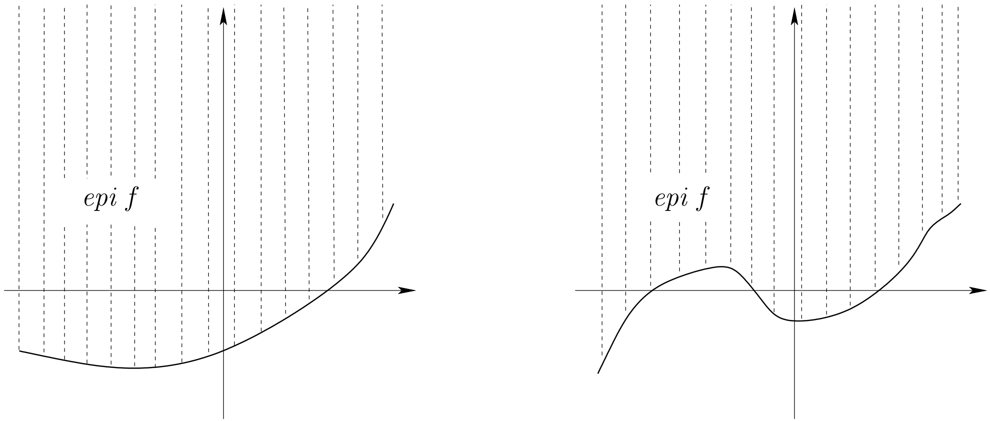

Рис. 1: слева надграфик выпуклой функции, справа надграфик функции, которая не является выпуклой

Критерий выпуклости дифференцируемой функции

Теорема 11

Пусть $f\ : D \to \mathbf{R}$ *— дифференцируемая функция,*
$D \subset \mathbf{R}^n$ *— выпуклое множество.*
Функция $f\left(x\right)$ *является выпуклой тогда и только тогда, когда*

$$f\left(x\right) - f\left(y\right) \geq \left(f'\left(y\right), x - y\right) \,\forall\, x,\, y \in D.$$

Критерий выпуклости

дважды дифференцируемой функции

Теорема 12

Пусть $D \subset \mathbf{R}^n$ *— выпуклое множество,*
$f\ : D \to \mathbf{R}$ *дважды дифференцируемая функция.*
Функция f *выпукла* $\Longleftrightarrow$ $f''\left(x\right)$ *неотрицательно определена* $\forall x \in D$.

Доказательство.

$\Rightarrow$ Пусть f — выпуклая функция.

$y, x \in D$, обозначим $h = x - y$, тогда

$$y + \alpha h = y + \alpha\,(x - y) = \alpha x + (1 - \alpha)\,y \in D \quad (0 \le \alpha \le 1)$$

В силу критерия выпуклости для дифференцируемой функции (Т. 11)

$$\forall\, y,\, x \in D \quad f\,(y + \alpha h) - f\,(y) - (f'\,(y)\,, \alpha h) \ge 0.$$

С другой стороны, применяя формулу Тейлора, имеем при $\alpha \to +0$:

$$f\,(y + \alpha h) - f\,(y) - (f'\,(y)\,, \alpha h) = \frac{1}{2}\,(f''\,(y)\,\alpha h, \alpha h) + o\left(\alpha^2\right).$$

Отсюда

$$\frac{1}{2}\,(f''\,(y)\,\alpha h, \alpha h) + o\left(\alpha^2\right) \ge 0.$$

Разделим обе части неравенства на α^2 и рассмотрим предел при $\alpha \to +0$, получим $(f''\,(y)\,h, h) \ge 0$.

$\Leftarrow$ Пусть $f''(s)$ — неотрицательно определенная матрица $\forall\, s \in D$.

$y,\, x \in D$, $h = x - y$. Запишем для функции f формулу Тейлора с остаточным членом в форме Лагранжа

$$f(x) - f(y) = \left(f'(y), h\right) + \frac{1}{2}\left(f''(y + \alpha h)\, h, h\right),$$

где α – некоторое число, $0 < \alpha < 1$.

Точка $y + \alpha h \in D$ в силу выпуклости множества D.

Так как $f''(y + \alpha h)$ неотрицательно определенная матрица, то

$$f(x) - f(y) - \left(f'(y), x - y\right) \geq 0,$$

Применяем критерий выпуклости дифференцируемой функции

$\Rightarrow f(x)$ — выпуклая функция. $\square$

***Пример* 10** *Исследовать на выпуклость функцию*

$$f\left(x,y,z\right)=x^2+xy+y^2+2z.$$

Решение.

$$f''\left(x,y,z\right)=\begin{pmatrix}2&1&0\\1&2&0\\0&0&0\end{pmatrix}$$

Критерий Сильвестра $\Rightarrow$ угловые миноры:

$$\triangle_1=2>0,\quad \triangle_{12}=2\cdot 2-1\cdot 1=3>0,\quad \triangle_{123}=0.$$

Матрица $f''\left(x,y,z\right)$ может быть неотрицательно определенной, поэтому вычислим главные миноры:

$$\delta_1=0,\quad \delta_2=0,\quad \delta_3=3,\quad \delta_{12}=0,\quad \delta_{13}=2,\quad \delta_{23}=2,\quad \delta_0=0.$$

Так как все главные миноры неотрицательны, то $f''\left(x,y,z\right)$ неотрицательно определена $\Rightarrow$ функция $f\left(x,y,z\right)$ выпуклая.

Выпуклая задача оптимизации

$$f_0(x) \to \min, \quad x \in D \subset \mathbf{R}^n \quad (\mathbf{z_c})$$

Задача (z_c) называется выпуклой задачей, если D – выпуклое множество, f_0 – выпуклая на D функция.

Теорема 13 *Пусть x^* — локальный минимум в выпуклой задаче (z_c). Тогда x^* — абсолютный минимум в (z_c).*

Доказательство.
$x^* \in locmin \Rightarrow$ найдется $U_\varepsilon(x^*)$: $f_0(x) \geq f_0(x^*)$ для всех $x \in D \bigcap U_\varepsilon(x^*)$. Возьмем любую точку $\tilde{x} \in D$. Соединим точки x^* и $\tilde{x}$ отрезком:
$\alpha\tilde{x} + (1-\alpha)x^* \in D$ для всех $0 \leq \alpha \leq 1$. Пусть $\alpha > 0$ и $\alpha \to 0$.
Тогда точки $x_\alpha = \alpha\tilde{x} + (1-\alpha)x^*$ попадут в $U_\varepsilon(x^*)$.
Для этих точек $f_0(x^*) \leq f_0(x_\alpha)$. Используем выпуклость $f_0 \quad \Rightarrow$

$$f_0(x^*) \;\leq\; f_0(\alpha\tilde{x} + (1-\alpha)x^*) \leq \alpha f_0(\tilde{x}) + (1-\alpha) f_0(x^*)$$

$$\Rightarrow \quad \alpha f_0(x^*) \leq \alpha f_0(\tilde{x}) \quad \Rightarrow \quad f_0(x^*) \leq f_0(\tilde{x})$$

$$\Rightarrow \quad x^* \in \text{absmin}$$

$\square$

Задача1. *Пусть* $f_0(x)$ *— выпуклая на* $\mathbf{R}^n$ *функция, дифференцируемая в точке* x^*. *Доказать, что если* $f_0'(x^*) = 0$, *то* x^* *является точкой глобального минимума* $f_0(x)$ *на* $\mathbf{R}^n$, *а следовательно, и на любом множестве* $D \subset \mathbf{R}^n$, *содержащем точку* x^*.