

ДЗ 6 по Методам Оптимизации.

Выпуклые функции. Сильно выпуклые функции.

Соколов Игорь, группа 573

16 декабря 2017 г.

1

Выпуклы ли следующие функции:

1. $f(x) = e^x - 1, x \in \mathbb{R};$
2. $f(x_1, x_2) = x_1 x_2, x \in \mathbb{R}_{++}^2;$
3. $f(x_1, x_2) = 1/(x_1 x_2), x \in \mathbb{R}_{++}^2?$

Решение:

1.1

$f(x) = e^x - 1$, тогда $f''(x) = e^x > 0$ на $\mathbb{R} \Rightarrow$ по дифференциальному критерию 2-го порядка f выпукла на \mathbb{R} .

1.2

$f(x_1, x_2) = x_1 x_2$.

$$f''(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Эта форма не является положительно определенной.

$\Rightarrow f(x_1, x_2)$ не выпукла.

1.3

$f(x_1, x_2) = 1/(x_1 x_2)$.

$$f''(x) = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}.$$

$$\Delta_1 = \frac{2}{x_1^3 x_2} > 0 \text{ и } \Delta_2 = \det(f''(x)) = \frac{3}{x_1^4 x_2^4} > 0 \text{ на } \mathbb{R}_{++}^2.$$

Значит $f''(x)$ положительно определена \Rightarrow по критерию 2-го порядка f выпукла на R_{++}^2

2

Докажите, что множество $S = \left\{ x \in \mathbb{R}_{++}^n \mid \prod_{i=1}^n x_i \geq 1 \right\}$ выпукло.

Доказательство.

$$S = \left\{ x \in \mathbb{R}_{++}^n \mid \prod_{i=1}^n x_i \geq 1 \right\} = \left\{ x \in \mathbb{R}_{++}^n \mid x_n \geq \frac{1}{\prod_{i=1}^{n-1} x_i} \right\}$$

$$\Rightarrow S\text{- надграфик функции } f(x_1, \dots, x_{n-1}) = \frac{1}{\prod_{i=1}^{n-1} x_i}$$

$$\frac{\partial f}{\partial x_k} = - \frac{1}{x_k^2 \prod_{i=1, i \neq k}^{n-1} x_i}$$

$$\frac{\partial^2 f}{\partial x_k \partial x_k} = \frac{2}{x_k^3 \prod_{i=1, i \neq k}^{n-1} x_i} = \frac{2}{x_k^2 \prod_{i=1}^{n-1} x_i} = \frac{2f}{x_k^2}$$

$$\frac{\partial^2 f}{\partial x_k \partial x_p} = \frac{1}{x_k^2 x_p^2 \prod_{i=1, i \neq k, p}^{n-1} x_i} = \frac{2}{x_k x_p \prod_{i=1}^{n-1} x_i} = \frac{2f}{x_k x_p}$$

$$f''(x) = 2f \begin{pmatrix} \frac{1}{x_1^2} & \frac{1}{x_1 x_2} & \frac{1}{x_1 x_3} & \frac{1}{x_1 x_4} & \cdots & \frac{1}{x_1 x_n} \\ \frac{1}{x_2 x_1} & \frac{1}{x_2^2} & \frac{1}{x_2 x_3} & \frac{1}{x_2 x_4} & \cdots & \frac{1}{x_2 x_n} \\ \frac{1}{x_3 x_1} & \frac{1}{x_3 x_2} & \frac{1}{x_3^2} & \frac{1}{x_3 x_4} & \cdots & \frac{1}{x_3 x_n} \\ \frac{1}{x_4 x_1} & \frac{1}{x_4 x_2} & \frac{1}{x_4 x_3} & \frac{1}{x_4^2} & \cdots & \frac{1}{x_4 x_n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n x_1} & \frac{1}{x_n x_2} & \frac{1}{x_n x_3} & \frac{1}{x_n x_4} & \cdots & \frac{1}{x_n^2} \end{pmatrix}.$$

$$\Delta_1 = \begin{vmatrix} \frac{1}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_2 x_1} & \frac{1}{x_2^2} \end{vmatrix} = 0$$

$$\Delta_2 = \begin{vmatrix} \frac{1}{x_1^2} & \frac{1}{x_1 x_2} & \frac{1}{x_1 x_3} \\ \frac{1}{x_2 x_1} & \frac{1}{x_2^2} & \frac{1}{x_2 x_3} \\ \frac{1}{x_3 x_1} & \frac{1}{x_3 x_2} & \frac{1}{x_3^2} \end{vmatrix} = \frac{1}{x_1^2} \begin{vmatrix} \frac{1}{x_2^2} & \frac{1}{x_2 x_3} \\ \frac{1}{x_3 x_2} & \frac{1}{x_3^2} \end{vmatrix} - \frac{1}{x_1 x_2} \begin{vmatrix} \frac{1}{x_2 x_1} & \frac{1}{x_2 x_3} \\ \frac{1}{x_3 x_1} & \frac{1}{x_3^2} \end{vmatrix} + \frac{1}{x_1 x_3} \begin{vmatrix} \frac{1}{x_2 x_1} & \frac{1}{x_2^2} \\ \frac{1}{x_3 x_1} & \frac{1}{x_3 x_2} \end{vmatrix} =$$

$$= \frac{1}{x_1^2} \cdot 0 - \frac{1}{x_1 x_2} \cdot 0 + \frac{1}{x_1 x_3} \cdot 0 = 0 \quad (1)$$

И так далее можно показать, что $\forall i \rightarrow \Delta_i = 0$

То есть: Определитель матрицы $n \times n$ можно подсчитать разложением по строке. После одного разложения надо подсчитать сумму n определителей матриц размера $n - 1 \times n - 1$

Применим эту операцию рекурсивно для подсчета определителей каждой из матриц (рекурсивно опускаемся вниз матрицы).

После проделывания этой процедуры $n - 2$ раза ($n - 2$ разложения по строке), необхо-

димо будет подсчитать определители вида $\begin{vmatrix} \frac{1}{x_{n-1}x_j} & \frac{1}{x_{n-1}x_{j+k}} \\ \frac{1}{x_n x_j} & \frac{1}{x_n x_{j+k}} \end{vmatrix} = 0 \quad \forall j \in [1, n]$

Затем поднимаясь вверх по рекурсии получим увидим, что $\forall i \rightarrow \Delta_i = 0$

$\Rightarrow f''(x)$ положительно полуопределенная матрица

\Rightarrow по критерию 2-го порядка $f(x)$ выпуклая функция.

\Rightarrow Надграфик $f(x)$ - выпуклое множество.

$\Rightarrow S-$ выпуклое множество.

□

3

Докажите, что функция

1. $f(X) = \text{tr}(X^{-1})$, $X \in S_{++}^n$ выпукла;
2. $f(X) = (\det X)^{1/n}$, $X \in S_{++}^n$ вогнута.

3.1

Доказательство. Пусть $g(t) = \text{tr}(Z + tV)$, где $Z \succ 0$ и $V \in S_{++}^n$

$$\begin{aligned}
g(t) &= \text{tr}((Z + tV)^{-1}) = \\
&= \text{tr}((Z^{-1}(\mathbb{E} + tZ^{-1/2}VZ^{-1/2})^{-1}) = \\
&= \text{tr}((Z^{-1}Q(\mathbb{E} + t\Lambda)^{-1}Q^T) = \\
&= \text{tr}(Q^T Z^{-1}Q(\mathbb{E} + t\Lambda)^{-1}) = \\
&= \sum_{i=1}^n \frac{(Q^T Z^{-1}Q)_{ii}}{1 + t\lambda_i} \quad (1)
\end{aligned}$$

В серии равенств выше мы воспользовались следующим:

- $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^{-1}$, где Q - матрица перехода к базису из собственных векторов.
- $\text{tr}(\mathbb{E} + t\Lambda)^{-1} = \sum_{i=1}^n \frac{1}{1 + t\lambda_i}$
- $\text{tr}(Q^T Z^{-1}Q(\mathbb{E} + t\Lambda)^{-1}) = \text{tr}(Q^T Z^{-1}Q[(\mathbb{E} + t\Lambda)^{-1}]^T) = \sum_i \frac{(Q^T Z^{-1}Q)_{ii}}{1 + t\lambda_i}$, так как $\text{tr}(XY^T) = \sum_{i,j} X_{ij}Y_{ij}$ и $(\mathbb{E} + t\Lambda)^{-1}$ диагональна

В последнем равенстве выражения (1) видим неотрицательную комбинацию выпуклых функций $\phi(t) = \frac{1}{1 + \lambda_i t}$ (их выпуклость легко показать, через критерий 2-го порядка)
 $\Rightarrow f(X)$ — выпуклая функция. \square

3.2

Доказательство. Пусть $g(t) = f(Z + tV)$, где $Z \succ 0$ и $V \in \mathbf{S}_{++}^n$

$$(\det(Z + tV))^{1/n} = (\det Z)^{1/n} (\det(I + tZ^{-1/2}VZ^{-1/2}))^{1/n} = (\det Z)^{1/n} \left(\prod_{i=1}^n (1 + t\lambda_i) \right)^{1/n}.$$

Где λ_i - собственные значения матрицы $Z^{-1/2}VZ^{-1/2}$ и $\det(Z) > 0$

Легко показать по критерию 2-го порядка, что $\left(\prod_{i=1}^n (1 + t\lambda_i) \right)^{1/n}$ вогнутая функция.

Действительно:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{f}{n^2} A \quad \text{где} \quad A_{ij} = \begin{cases} (1 - n)x_i^{-2} & \text{if } i = j \\ x_i^{-1}x_j^{-1} & \text{if } i \neq j \end{cases}$$

Пусть $y_i = 1/x_i$. Нам надо доказать, что $v^T A v \leq 0 \forall v$

Действительно: $v^T A v = (\sum_{i=1}^n y_i v_i)^2 - n \sum_{i=1}^n y_i^2 v_i^2 \leq 0$ по формуле Коши-Буняковского примененного к $1 \times (y_i v_i)$

$$\Rightarrow \text{показали вогнутость функции} \left(\prod_{i=1}^n (1 + t\lambda_i) \right)^{1/n}$$

$\Rightarrow f(X)$ — вогнутая. \square

4

Расстоянием Кульбака - Лейблера между $p, q \in \mathbb{R}_{++}^n$ называется:

$$D(p, q) = \sum_{i=1}^n (p_i \log(p_i/q_i) - p_i + q_i)$$

Докажите, что $D(p, q) \geq 0 \forall p, q \in \mathbb{R}_{++}^n$ и $D(p, q) = 0 \Leftrightarrow p = q$ Подсказка:

$$D(p, q) = f(p) - f(q) - \nabla f(q)^T(p - q), \quad f(p) = \sum_{i=1}^n p_i \log p_i$$

Доказательство.

$$f(p) = \sum_{i=1}^n p_i \log p_i \quad (1)$$

$$\frac{\partial f}{\partial x_k} = 1 + \log p_k$$

$$\frac{\partial^2 f}{\partial x_k^2} = \frac{1}{p_k} > 0 \quad \forall p \in \mathbb{R}_{++}^n$$

$\Rightarrow f(p)$ - строго выпуклая функция.

$$\Rightarrow f(p) > f(q) + \nabla f(q)^T(p - q) \quad \forall p, q \in \mathbb{R}_{++}^n : p \neq q \quad \text{критерий 1-го порядка} \quad (2)$$

$$\sum_{i=1}^n p_i \log p_i > \sum_{i=1}^n q_i \log q_i + \sum_{i=1}^n (1 + \log q_i)(p_i - q_i) = \sum_{i=1}^n (p_i \log q_i + p_i - q_i)$$

$$\sum_{i=1}^n p_i \log p_i - \sum_{i=1}^n (p_i \log q_i - p_i + q_i) > 0$$

$$\sum_{i=1}^n (p_i \log(p_i/q_i) - p_i + q_i) > 0$$

$$\Rightarrow D(p, q) > 0, p \neq q$$

Так как $D(p, q) = f(p) - f(q) - \nabla f(q)^T(p - q) \Leftrightarrow D(p, q) = 0$ при $p = q$

□

5

Пусть x - действительная переменная, принимающая конечный набор значений $a_1 < a_2 < \dots < a_n$ с вероятностями $\mathbb{P}(x = a_i) = p_i$. Оцените выпуклость и вогнутость следующих функций от p на множестве $\left\{ p \mid \sum_{i=1}^n p_i = 1, p_i \geq 0 \right\}$

1. $\mathbb{E}x$

2. $\mathbb{P}\{x \geq \alpha\}$
3. $\mathbb{P}\{\alpha \leq x \leq \beta\}$
4. $\sum_{i=1}^n p_i \log p_i$
5. $\mathbb{V}x = \mathbb{E}(x - \mathbb{E}x)^2$
6. $\mathbf{quartile}(x) = \inf \{\beta \mid \mathbb{P}\{x \leq \beta\} \geq 0.25\}$

Решение:

1. $\mathbb{E}x = p_1 a_1 + \dots + p_n a_n$ — неотрицательная комбинация константных функций a_i
 $\Rightarrow \mathbb{E}x$ — выпукла и вогнута.
2. Пусть $k = \min\{i \mid a_i \geq \alpha\}$. Тогда $\mathbb{P}(x \geq \alpha) = \sum_{i=k}^n p_i$ — линейная функция от переменной $\mathbf{p} \Rightarrow$ выпукла и вогнута (по критерию 2-го порядка, так как $f''(p) = 0$).
3. Пусть $k = \min\{i \mid a_i \geq \alpha\}$ и $m = \max\{i \mid a_i \leq \beta\}$.
Тогда $\mathbb{P}\{\alpha \leq x \leq \beta\} = \sum_{i=k}^m p_i$ — линейная функция от переменной $\mathbf{p} \Rightarrow$ выпукла и вогнута.
4. В предыдущей задаче (1) было показано, что $\sum_{i=1}^n p_i \log p_i$ — строго выпуклая.
 \Rightarrow не является вогнутой.
5. $\mathbb{V}x = \mathbb{E}(x - \mathbb{E}x)^2 = \mathbb{E}$
- 6.

$$\begin{aligned}
\mathbf{quartile}(x) &= \inf \{\beta \mid \mathbb{P}\{x \leq \beta\} \geq 0.25\} \\
&= \inf \{\beta \mid 1 - \mathbb{P}\{\beta < x\} \geq 0.25\} \\
&= \inf \{\beta \mid \mathbb{P}\{\beta < x\} \leq 0.75\}
\end{aligned}$$

Заметим, что функция $f(x) = \mathbf{quartile}(x)$ не является непрерывной (принимает значения из дискретного множества $\{a_1, \dots, a_n\}$).

По теореме 3.1.5 из Жадана (стр 100): если $f(x)$ выпуклая функция на выпуклом множестве $X \subseteq \mathbb{R}^n$, то $f(x)$ непрерывна $\forall x \in ri(X)$.

Значит $f(x) = \mathbf{quartile}(x)$ не является выпуклой. Заметим, что $-f(x)$ тоже не является выпуклой, значит $f(x)$ не является вогнутой.