ДЗ 6 по Методам Оптимизации. Выпуклые функции. Сильно выпуклые функции.

Соколов Игорь, группа 573

12 ноября 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 $R \Rightarrow$ по дифференциальному критерию 2-го порядка f выпукла на 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

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

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

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

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

2

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

Доказательство.

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

$\Rightarrow S-$ надграфик функции $f(x_1, \ldots, x_{n-1}) = \dfrac{1}{\prod\limits_{i=1}^{n-1} x_i}$

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

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

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

$$f''(x) = 2f \begin{pmatrix} \frac{1}{x_1^2} & \frac{1}{x_1x_2} & \frac{1}{x_1x_3} & \frac{1}{x_1x_4} & \cdots & \frac{1}{x_1x_n} \\ \frac{1}{x_2x_1} & \frac{1}{x_2^2} & \frac{1}{x_2x_3} & \frac{1}{x_2x_4} & \cdots & \frac{1}{x_2x_n} \\ \frac{1}{x_3x_1} & \frac{1}{x_3x_2} & \frac{1}{x_3^2} & \frac{1}{x_3x_4} & \cdots & \frac{1}{x_3x_n} \\ \frac{1}{x_4x_1} & \frac{1}{x_4x_2} & \frac{1}{x_4x_3} & \frac{1}{x_4^2} & \cdots & \frac{1}{x_4x_n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_nx_1} & \frac{1}{x_nx_2} & \frac{1}{x_nx_3} & \frac{1}{x_nx_4} & \cdots & \frac{1}{x_n^2} \end{pmatrix}.$$

$$\Delta_1 = \begin{vmatrix} \frac{1}{x_1^2} & \frac{1}{x_1x_2} \\ \frac{1}{x_2x_1} & \frac{1}{x_2^2} \end{vmatrix} = 0$$

$$\Delta_2 = \begin{vmatrix} \frac{1}{x_1^2} & \frac{1}{x_1x_2} & \frac{1}{x_1x_3} \\ \frac{1}{x_2x_1} & \frac{1}{x_2^2} & \frac{1}{x_2x_3} \\ \frac{1}{x_3x_1} & \frac{1}{x_3x_2} & \frac{1}{x_3^2} \end{vmatrix} = \frac{1}{x_1^2}\begin{vmatrix} \frac{1}{x_2^2} & \frac{1}{x_2x_3} \\ \frac{1}{x_3x_2} & \frac{1}{x_3^2} \end{vmatrix} - \frac{1}{x_1x_2}\begin{vmatrix} \frac{1}{x_2x_1} & \frac{1}{x_2x_3} \\ \frac{1}{x_3x_1} & \frac{1}{x_3^2} \end{vmatrix} + \frac{1}{x_1x_3}\begin{vmatrix} \frac{1}{x_2x_1} & \frac{1}{x_2^2} \\ \frac{1}{x_3x_1} & \frac{1}{x_3x_2} \end{vmatrix} =$$

$$= \frac{1}{x_1^2}\cdot 0 - \frac{1}{x_1x_2}\cdot 0 + \frac{1}{x_1x_3}\cdot 0 = 0 \quad (1)$$

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

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

Затем поднимаясь вверх по рекурсии получим увидем, что $\forall i \to \Delta_i = 0$

$\Rightarrow f''(x)$ положительно полуопределенная матрица
$\Rightarrow$ по критерию 2-го порядка $f(x)$ выпуклая функция.
$\Rightarrow$ Надграфик $f(x)$ - выпуклое множество.
$\Rightarrow S-$ выпуклое множество.

$\square$

3

Докажите, что функция
1. $f(X) = \mathbf{tr}(X^{-1}), X \in S^n_{++}$ выпукла;
2. $f(X) = (\det X)^{1/n}, X \in S^n_{++}$ вогнута.

3.1

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

$$g(t) = \mathbf{tr}(\left(Z+tV\right)^{-1}\right) =$$
$$= \mathbf{tr}\left((Z^{-1}(\mathbb{E}+tZ^{-1/2}VZ^{-1/2})^{-1}\right) =$$
$$= \mathbf{tr}\left((Z^{-1}Q(\mathbb{E}+t\Lambda)^{-1}Q^T\right) =$$
$$= \mathbf{tr}\left(Q^TZ^{-1}Q(\mathbb{E}+t\Lambda)^{-1}\right) =$$
$$= \sum_{i=1}^{n} \frac{\left(Q^TZ^{-1}Q\right)_{ii}}{1+t\lambda_i} \quad (1)$$

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.
- $\mathbf{tr}(\mathbb{E}+t\Lambda)^{-1} = \sum\limits_{i=1}^{n} \dfrac{1}{1+t\lambda_i}$
- $\mathbf{tr}(Q^TZ^{-1}Q(\mathbb{E}+t\Lambda)^{-1}) = \mathbf{tr}(Q^TZ^{-1}Q\left[(\mathbb{E}+t\Lambda)^{-1}\right]^T) = \sum\limits_{i} \dfrac{\left(Q^TZ^{-1}Q\right)_{ii}}{1+t\lambda_i}$, так как $\mathbf{tr}(XY^T) = \sum\limits_{i,j} X_{ij}Y_{ij}$ и $(\mathbb{E}+t\Lambda)^{-1}$ диагональна

В последнем равенстве выражения (1) видим неотрицательную комбинацию выпуклых функций $\phi(t) = \dfrac{1}{1+\lambda_i t}$ (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 $V \in \mathbf{S}^n_{++}$

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

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

Легко показать по критерию 2-го порядка, что $\left(\prod\limits_{i=1}^{n}(1+t\lambda_i)\right)^{1/n}$ 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 $v^TAv \le 0 \forall v$

Действительно: $v^TAv = \left(\sum_{i=1}^n y_iv_i\right)^2 - n\sum_{i=1}^n y_i^2v_i^2 \le 0$ по формуле Коши-Буняковского примененного к $1 \times (y_iv_i)$

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

$\Rightarrow f(X)-$ вогнутая.

$\square$

4

Расстоянием Кульбака - Лейблера между 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.

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

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

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

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

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

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

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

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

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

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

1. $\mathbb{E}x$

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

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 $\mathbb{P}(x \geq \alpha) = \sum\limits_{i=k}^{n} p_i$. - 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 $\mathbb{P}\{\alpha \leq x \leq \beta\} = \sum\limits_{i=k}^{m} p_i$ - линейная функция от переменной $\mathbf{p} \Rightarrow$ выпукла и вогнута.

4. В предыдущей задаче (1) было показано, что $\sum\limits_{i=1}^{n} p_i \log p_i$ - строго выпуклая.

   $\Rightarrow$ не является вогнутой.

5. $\mathbb{V}x = \mathbb{E}(x - \mathbb{E}x)^2 = \mathbb{E}$

6.
$$\begin{aligned} \textbf{quartile}(x) &= \inf\{\beta \mid \mathbb{P}\{x \leq \beta\} \geq 0.25\} \\ &= \inf\{\beta \mid 1 - \mathbb{P}\{\beta < x\} \geq 0.25\} \\ &= \inf\{\beta \mid \mathbb{P}\{\beta < x\} \leq 0.75\} \end{aligned}$$

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