Ахмаджонов Мумтозбек Б05-927б

Домашнее задание №3

Задача 1.

Являются ли следующие функции выпуклыми или вогнутыми:

- $f(x) = x_1x_2$ на $\mathbb{R}^2$;
- $f(x) = \dfrac{1}{x_1x_2}$ на $\mathbb{R}^2_+$;
- $f(x) = x_1^\alpha x_2^{1-\alpha}$, где $0 \leqslant \alpha \leqslant 1, x \in \mathbb{R}^2_+$?

Решение:

1) $f(x) = x_1x_2$ на $\mathbb{R}^2$

Воспользуемся утверждением, что фукнция выпукла <=> ее гессиан неотрицательно определен.

$$\nabla^2 f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Пусть $a = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, тогда:

$$a^T \nabla^2 f a = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2 < 0$$

Значит, данная функция не выпукла. Аналогичным образом, можем показать, что $-f(x)$ - также не выпукла. Тогда, данная функция ни выпукла, ни вогнута.

2) $f(x) = \dfrac{1}{x_1x_2}$ на $\mathbb{R}^2_+$

$$\nabla^2 f = \begin{bmatrix} \dfrac{2}{x_1^3x_2} & \dfrac{1}{x_1^2x_2^2} \\ \dfrac{1}{x_1^2x_2^2} & \dfrac{2}{x_1x_2^3} \end{bmatrix}$$

Рассмотрим $\forall \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$:

$$\begin{bmatrix} a & b \end{bmatrix} \nabla^2 f \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \dfrac{2}{x_1^3x_2} & \dfrac{1}{x_1^2x_2^2} \\ \dfrac{1}{x_1^2x_2^2} & \dfrac{2}{x_1x_2^3} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \dfrac{2a}{x_1^3x_2} + \dfrac{b}{x_1^2x_2^2} & \dfrac{a}{x_1^2x_2^2} + \dfrac{2b}{x_1x_2^3} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} =$$

$$= \frac{2a^2}{x_1^3x_2} + \frac{ab}{x_1^2x_2^2} + \frac{ab}{x_1^2x_2^2} + \frac{2b^2}{x_1x_2^3} = \frac{a^2}{x_1^3x_2} + \frac{b^2}{x_1x_2^3} + \frac{1}{x_1x_2}\left(\frac{a^2}{x_1^2} + \frac{2ab}{x_1x_2} + \frac{b^2}{x_2^2}\right) =$$

$$= \frac{a^2}{x_1^3x_2} + \frac{b^2}{x_1x_2^3} + \frac{1}{x_1x_2}\left(\frac{a}{x_1} + \frac{b}{x_2}\right)^2 \geqslant 0$$

Получили, что функция выпукла, т.к. гессиан положительно полуопределен.

3) $f(x) = x_1^{\alpha} x_2^{1-\alpha}$, где $0 \leqslant \alpha \leqslant 1, x \in \mathbb{R}_+^2$

$$\nabla^2 f = \begin{bmatrix} -\alpha(1-\alpha)\dfrac{x_2^{1-\alpha}}{x_1^{2-\alpha}} & \alpha(1-\alpha)\dfrac{1}{x_1^{1-\alpha}x_2^{\alpha}} \\ \alpha(1-\alpha)\dfrac{1}{x_1^{1-\alpha}x_2^{\alpha}} & -\alpha(1-\alpha)\dfrac{x_1^{\alpha}}{x_2^{\alpha+1}} \end{bmatrix} = -\frac{\alpha(1-\alpha)}{x_1^{-\alpha}x_2^{\alpha-1}}\begin{bmatrix} \dfrac{1}{x_1^2} & -\dfrac{1}{x_1x_2} \\ -\dfrac{1}{x_1x_2} & \dfrac{1}{x_2^2} \end{bmatrix}$$

Рассмотрим $\forall \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$:

$$\begin{bmatrix} a & b \end{bmatrix} \nabla^2 f \begin{bmatrix} a \\ b \end{bmatrix} = -\frac{\alpha(1-\alpha)}{x_1^{-\alpha}x_2^{\alpha-1}} \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \dfrac{1}{x_1^2} & -\dfrac{1}{x_1x_2} \\ -\dfrac{1}{x_1x_2} & \dfrac{1}{x_2^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} =$$

$$= -\frac{\alpha(1-\alpha)}{x_1^{-\alpha}x_2^{\alpha-1}} \begin{bmatrix} \dfrac{a}{x_1^2} - \dfrac{b}{x_1x_2} & -\dfrac{a}{x_1x_2} + \dfrac{b}{x_2^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} =$$

$$= -\frac{\alpha(1-\alpha)}{x_1^{-\alpha}x_2^{\alpha-1}} \left(\frac{a^2}{x_1^2} - \frac{ab}{x_1x_2} - \frac{ab}{x_1x_2} + \frac{b^2}{x_2^2} \right) = -\frac{\alpha(1-\alpha)}{x_1^{-\alpha}x_2^{\alpha-1}} \left(\frac{a}{x_1} + \frac{b}{x_2} \right)^2 \leqslant 0$$

Значит, $-f(x)$ выпукла $\Longrightarrow f(x)$ вогнута.

Ответ:

1. не является ни выпуклой, ни вогнутой
2. выпукла
3. вогнута

Задача 2.

Покажите, что:

- $f(X) = tr\left(X^{-1}\right)$ выпукла на S_{+}^{n}
- $f(X) = \det\left(X^{-1}\right)^{\frac{1}{n}}$ выпукла на S_{+}^{n}

Решение:

1) $f(X) = tr\left(X^{-1}\right)$

Рассмотрим $X = Z + tV$ и функцию $g(t)$ на интервале, где $Z \succ 0$. Пусть P - матрица перехода матрицы $Z + tV$ к диагональной матрице D. Тогда:

$$D = P^{-1}XP \Longrightarrow X = PDP^{-1} \Longrightarrow X^{-1} = PD^{-1}P^{-1} \Longrightarrow (Z+tV)^{-1} = PD^{-1}P^{-1}$$

$$g(t) = tr\left((Z+tV)^{-1}\right) = tr\left(PD^{-1}P^{-1}\right) = tr\left(D^{-1}P^{-1}P\right) = tr\left(D^{-1}\right) = \sum_{i=1}^{n}\frac{1}{\lambda_i},$$

где λ_i - собственные числа матрицы $Z + tV$, тогда $\lambda_i = z_i + tv_i$, где z_i -с.з. матрицы Z, а v_i - с.з. матрицы V.

$$g(t) = \sum_{i=1}^{n}\frac{1}{\lambda_i} = \sum_{i=1}^{n}\frac{1}{z_i + tv_i}$$

$$\nabla g = g'(t) = -\sum_{i=1}^{n}\frac{v_i}{(z_i + tv_i)^2}$$

$$\nabla^2 g = g''(t) = \sum_{i=1}^{n}\frac{2v_i^2}{(z_i + tv_i)^3}$$

Так как, $Z + tV \succ 0 \Longrightarrow \forall i\ \lambda_i > 0 \Longrightarrow z_i + tv_i > 0 \Longrightarrow \nabla^2 g > 0 \Longrightarrow g(t)$ - выпукла. А значит и данная функция выпукла.

2) $f(X) = \det\left(X^{-1}\right)^{\frac{1}{n}}$

Снова рассмотрим $X = Z + tV$ и функцию $g(t)$ на интервале, где $Z \succ 0$. Пусть P - матрица перехода матрицы $Z + tV$ к диагональной матрице D. Тогда:

$$D = P^{-1}XP \Longrightarrow X = PDP^{-1} \Longrightarrow X^{-1} = PD^{-1}P^{-1} \Longrightarrow (Z+tV)^{-1} = PD^{-1}P^{-1}$$

$$g(t) = \det\left((Z+tV)^{-1}\right)^{\frac{1}{n}} = \det\left(PD^{-1}P^{-1}\right)^{\frac{1}{n}} = \det(D)^{-\frac{1}{n}} = \left(\prod_{i=1}^{n}\frac{1}{\lambda_i}\right)^{\frac{1}{n}},$$

где λ_i - собственные числа матрицы $Z + tV$, тогда $\lambda_i = z_i + tv_i$, где z_i -с.з. матрицы Z, а v_i - с.з. матрицы V.

$$g(t) = \left(\prod_{i=1}^{n}\frac{1}{\lambda_i}\right)^{\frac{1}{n}} = \left(\prod_{i=1}^{n}\frac{1}{z_i + tv_i}\right)^{\frac{1}{n}}$$

$$\nabla g = g'(t) = -\frac{1}{n}\left(\prod_{i=1}^{n}\frac{1}{z_i + tv_i}\right)^{\frac{1-n}{n}}\left(\prod_{i=1}^{n}\frac{1}{z_i + tv_i}\sum_{j=1}^{n}\frac{v_j}{z_j + tv_j}\right) =$$

$$= -\frac{1}{n}\left(\prod_{i=1}^{n}\frac{1}{z_i+tv_i}\right)^{\frac{1}{n}}\sum_{j=1}^{n}\frac{v_j}{z_j+tv_j}$$

$$\nabla^2 g = g''(t) = -\frac{1}{n}\sum_{j=1}^{n}\frac{v_j}{z_j+tv_j}\frac{d}{dt}\left(\prod_{i=1}^{n}\frac{1}{z_i+tv_i}\right)^{\frac{1}{n}} - \frac{1}{n}\left(\prod_{i=1}^{n}\frac{1}{z_i+tv_i}\right)^{\frac{1}{n}}\frac{d}{dt}\left(\sum_{j=1}^{n}\frac{v_j}{z_j+tv_j}\right) =$$

$$= \frac{1}{n^2}\left(\prod_{i=1}^{n}\frac{1}{z_i+tv_i}\right)^{\frac{1}{n}}\left(\sum_{j=1}^{n}\frac{v_j}{z_j+tv_j}\right)^2 + \frac{1}{n}\left(\prod_{i=1}^{n}\frac{1}{z_i+tv_i}\right)^{\frac{1}{n}}\left(\sum_{j=1}^{n}\frac{v_j^2}{(z_j+tv_j)^2}\right)$$

Так как, $Z+tV \succ 0 \Longrightarrow \forall i\ \lambda_i > 0 \Longrightarrow z_i+tv_i > 0 \Longrightarrow \nabla^2 g > 0 \Longrightarrow g(t)$ - выпукла. А значит и данная функция выпукла. $\square$

Задача 3.

$a_1, a_2, \ldots, a_n \in \mathbb{R} : P(x = a_i) = p_i$. Являются ли следующие функции от p на множестве $\left\{ p \mid \sum_{i=1}^{n} p_i = 1, p_i \geqslant 0 \right\}$ выпуклыми или вогнутыми:

- Ex;
- $P\{x \geqslant \alpha\}$;
- $P\{\alpha \leqslant x \leqslant \beta\}$;
- $\sum_{i=1}^{n} p_i \ln p_i$;
- $Dx = E(x - Ex)^2$;
- $\mathbf{quartile}(x) = \inf\{\beta \mid P\{x \leqslant \beta\} \geqslant 0.25\}$?

Решение:

1) $f(p) = Ex = \sum_{i=1}^{n} a_i p_i$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(p)}{\partial p_1^2} & \ldots & \frac{\partial^2 f(p)}{\partial p_1 \partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(p)}{\partial p_n \partial p_1} & \ldots & \frac{\partial^2 f(x)}{\partial p_n^2} \end{bmatrix} = 0$$

Значит, функция и вогнута, и выпукла.

2) и 3)

тут функции по сути являются частичными суммами основной суммы Ex. Т.е. без ограничения общности, пусть $\{a_n\}$ - отсортированная по возрастанию посоедовательность, тогда

$$P\{x \geqslant \alpha\} = \begin{cases} \sum_{i=m}^{n} p_i & a_1 \leqslant \alpha \leqslant a_m \\ 1 & \alpha < a_1 \\ 0 & a > a_n \end{cases}$$

$$P\{\alpha \leqslant x \leqslant \beta\} = \begin{cases} \ldots & \ldots \\ \sum_{i=k}^{m} a_i p_i & a_1 \leqslant \alpha \leqslant a_k \leqslant \beta \leqslant a_m \\ \ldots & \ldots \end{cases}$$

Таким, образом, легко можно убедиться, что аналогично предыдущему пункту, получим нулевой гессиан.

4) $f(p) = \sum_{i=1}^{n} p_i \ln p_i$

$$\frac{\partial f}{\partial p_k} = \ln p_k + 1$$

$$\frac{\partial^2 f}{\partial p_k^2} = \frac{1}{p_k}$$
$$\frac{\partial^2 f}{\partial p_k \partial p_l} = 0$$
Значит, собственные значения гессиана все > 0, значит гессиан сам положительно полуопределен.

5) $f(p) = Dx = Ex^2 - (Ex)^2 = \sum_{i=1}^{n} a_i^2 p_i - \left(\sum_{i=1}^{n} a_i p_i \right)^2$

$$\frac{\partial f}{\partial p_k} = a_k^2 - 2a_k Ex$$
$$\frac{\partial^2 f}{\partial p_k^2} = -2a_k^2$$
$$\frac{\partial^2 f}{\partial p_k \partial p_l} = -2a_k a_l$$
Получаем такой гессиан:
$$\nabla^2 f = -2 \begin{bmatrix} a_1^2 & \dots & a_1 a_n \\ \vdots & \ddots & \vdots \\ a_n a_1 & \dots & a_n^2 \end{bmatrix} = -2aa^T,$$
где $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, тогда
$$\forall y \in \mathbb{R}^n : y^T \nabla^2 f y = -2y^T a a^T y = -2\langle y, a \rangle^2 \leqslant 0$$
Значит, данная функция вогнута.

6) $f(p) = \textbf{quartile}(x) = \inf\{\beta \mid P\{x \leqslant \beta\} \geqslant 0.25\}$
Можно доказать,что любая выпуклая (а в следствии и вогнутая) функция непрерывна.
Достаточно легко убедиться, что $\inf\{\beta \mid P\{x \leqslant \beta\} \geqslant 0.25\} \in \{a_1, a_2, \dots, a_n\}$. Получается, данная функция не является непрерывной, т.к. принимает значения из конечного множества, значит она не является ни выпуклой, ни вогнутой.

Ответ:

- и вогнута, и выпукла
- и вогнута, и выпукла
- и вогнута, и выпукла
- выпукла
- вогнута
- ни выпукла, ни вогнута

Задача 4.

Пусть f - выпуклая функция. Покажите выпуклость функции $\psi(x) = \inf_y\{f(y) : Ay = x\}$.

Решение:

Введем новую функцию:

$$h(x, y) = \begin{cases} f(y) & Ay = x \\ +\infty & \text{иначе} \end{cases} \tag{1}$$

Тогда, $\psi(x) = \inf_y h(x, y)$, также, очевидно, что $h(x, y)$ - выпукла по $\begin{bmatrix} x \\ y \end{bmatrix}$. Осталось, доказать, что минимум по y функции (1) выпуклый по x.

Применим определение выпуклости к $\psi(x)$:

$$\psi(\lambda x_1 + (1-\lambda)x_2) = \inf_y h(\lambda x_1 + (1-\lambda)x_2, y) \leqslant$$
$$\leqslant h(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \leqslant \lambda h(x_1, y_1) + (1-\lambda)h(x_2, y_2) \leqslant$$
$$\leqslant \lambda\psi(x_1) + (1-\lambda)\psi(x_2) + \epsilon\ \forall\epsilon > 0 \Longrightarrow \psi(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda\psi(x_1) + (1-\lambda)\psi(x_2)$$

$\square$