最优化方法导论第二次小测

1. 求下列各个锥的对偶锥 K^* 。

(a) $K = \{0\}$，在$\mathbb{R}^2$空间

(b) $K = \mathbb{R}^2$

(c) $K = \{(x_1, x_2) \mid |x_1| \le x_2\}$

(d) $K = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$

解： (a) $K = \{0\}$

$$\begin{aligned} K^* &= \{y \mid y^T x \ge 0, \ \forall x \in K\} \\ &= \{y \mid y^T 0 \ge 0\} \\ &= \mathbb{R}^2. \end{aligned}$$

因此：$K^* = \mathbb{R}^2$.

(b) $K = \mathbb{R}^2$

我们需要找出所有 $y \in \mathbb{R}^2$，使得

$$y^T x \ge 0, \quad \forall x \in \mathbb{R}^2.$$

但若 $y \neq 0$，取 $x = -y$，则

$$y^T x = -\|y\|_2^2 < 0,$$

不满足条件。因此唯一可能的选择是 $y = 0$。于是：$K^* = \{0\}$.

(c) $K = \{(x_1, x_2) \mid |x_1| \le x_2\}$

下证该锥是自对偶的，即$K^* = K$.。我们分别证明 $K^* \subseteq K$ 与 $K \subseteq K^*$。

（1）证明 $K^* \subseteq K$:

取任意 $y = (y_1, y_2) \in K^*$。按定义，对所有 $x \in K$ 有

$$y_1 x_1 + y_2 x_2 \ge 0.$$

令 x 取两条边界射线上的点（皆属于 K）：

- 取 $x = (t, t)$（即 $x_1 = x_2 = t > 0$），则

$$y^\top x = t(y_1 + y_2) \ge 0 \ \Rightarrow \ y_1 + y_2 \ge 0.$$

- 取 $x = (-t, t)$（即 $x_1 = -t, \ x_2 = t > 0$），则

$$y^\top x = t(-y_1 + y_2) \ge 0 \ \Rightarrow \ y_2 - y_1 \ge 0.$$

两式合并得

$$y_2 \ge |y_1|.$$

故 $y \in K$，从而 $K^* \subseteq K$。

（2）证明 $K \subseteq K^*$：

取任意 $y = (y_1, y_2) \in K$，即 $y_2 \ge |y_1|$。对任意 $x = (x_1, x_2) \in K$（即 $x_2 \ge |x_1|$），有

$$\begin{aligned} y^\top x = y_1x_1 + y_2x_2 \ &\ge\ -|y_1||x_1| + y_2x_2 \\ &\ge\ -|y_1|\,x_2 + y_2x_2 \quad (\text{因 } x_2 \ge |x_1| \ge 0) \\ &=\ (y_2 - |y_1|)\,x_2\ \ge\ 0 \quad (\text{因 } y_2 \ge |y_1|,\ x_2 \ge 0). \end{aligned}$$

于是对所有 $x \in K$ 都有 $y^\top x \ge 0$，即 $y \in K^*$。故 $K \subseteq K^*$。

由上两步得 $K^* = K$。证毕。

(d) $K = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$

这里 K 是一条过原点的直线。 其正交方向为 $(1,1)$，即所有与该直线垂直的向量形成对偶锥。 因此：

$$K^* = \{(x_1, x_2) \mid x_1 - x_2 = 0\}。$$

2. 设函数 $f : \mathbb{R}^n \to \mathbb{R}$，定义其上镜图（epigraph）为：

$$\operatorname{epi} f = \{(x,t) \in \mathbb{R}^{n+1} \mid x \in \operatorname{dom} f,\ f(x) \le t\}.$$

证明：函数 f 为凸函数，当且仅当其上镜图 $\operatorname{epi} f$ 为凸集。

证：

(⇒) 若 f 为凸函数。
任取 $(x_1, t_1), (x_2, t_2) \in \operatorname{epi} f$，即：

$$t_1 \ge f(x_1), \quad t_2 \ge f(x_2).$$

对任意 $\theta \in [0,1]$，考虑点：

$$(x_\theta, t_\theta) = \big(\theta x_1 + (1-\theta)x_2,\ \theta t_1 + (1-\theta)t_2\big)。$$

由 f 的凸性：

$$f(x_\theta) \le \theta f(x_1) + (1-\theta) f(x_2) \le \theta t_1 + (1-\theta)t_2 = t_\theta。$$

因此 $(x_\theta, t_\theta) \in \operatorname{epi} f$，说明 $\operatorname{epi} f$ 对凸组合封闭，故为凸集。

(⇐) 若 $\operatorname{epi} f$ 为凸集。
任取 $x_1, x_2 \in \operatorname{dom} f$，令

$$t_1 = f(x_1), \quad t_2 = f(x_2)。$$

则 $(x_1, t_1), (x_2, t_2) \in \operatorname{epi} f$。
由凸性假设，

$$(x_\theta, t_\theta) = \big(\theta x_1 + (1-\theta)x_2,\ \theta t_1 + (1-\theta)t_2\big) \in \operatorname{epi} f。$$

即有：

$$f(\theta x_1 + (1-\theta)x_2) \le t_\theta = \theta f(x_1) + (1-\theta) f(x_2),$$

这正是凸函数的定义。故 f 为凸函数。

3. 判断集合是否为凸集，并说明理由：

$$S = \{(x_1, x_2) \mid x_2 \ge |x_1|\}$$

解：

根据定义，若对任意 $x^{(1)}, x^{(2)} \in S$ 和任意 $\theta \in [0,1]$， 其凸组合 $\theta x^{(1)} + (1-\theta)x^{(2)} \in S$，则 S 为凸集。

设

$$x^{(1)} = (x_1^{(1)}, x_2^{(1)}), \quad x^{(2)} = (x_1^{(2)}, x_2^{(2)}),$$

且满足

$$x_2^{(1)} \ge |x_1^{(1)}|, \quad x_2^{(2)} \ge |x_1^{(2)}|.$$

对任意 $\theta \in [0,1]$，考虑

$$x = \theta x^{(1)} + (1-\theta)x^{(2)} = (\theta x_1^{(1)} + (1-\theta)x_1^{(2)},\ \theta x_2^{(1)} + (1-\theta)x_2^{(2)}).$$

由绝对值的凸性（即 $|\theta a + (1-\theta)b| \le \theta|a| + (1-\theta)|b|$）可得：

$$|\,\theta x_1^{(1)} + (1-\theta)x_1^{(2)}\,| \le \theta|x_1^{(1)}| + (1-\theta)|x_1^{(2)}|.$$

又因为

$$x_2^{(1)} \ge |x_1^{(1)}|, \quad x_2^{(2)} \ge |x_1^{(2)}|,$$

所以

$$\theta x_2^{(1)} + (1-\theta)x_2^{(2)} \ge \theta|x_1^{(1)}| + (1-\theta)|x_1^{(2)}| \ge |\,\theta x_1^{(1)} + (1-\theta)x_1^{(2)}\,|.$$

因此：

$$x_2 = \theta x_2^{(1)} + (1-\theta)x_2^{(2)} \ge |x_1|,$$

即 $x \in S$。因此该集合是凸集。

4. 设 f 是一个凸函数，定义函数 g 为

$$g(x) = \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha}.$$

(a) 证明 g 是齐次的，即对所有 $t \ge 0$，都有

$$g(tx) = t g(x)。$$

(b) 证明 g 是 f 的最大齐次下界函数： 若 h 是齐次的且满足 $h(x) \le f(x)$ 对所有 x 都成立，则有 $h(x) \le g(x)$ 对所有 x 成立。

(c) 证明 g 是凸函数。

解：(a)
当 $t > 0$ 时，

$$g(tx) = \inf_{\alpha>0} \frac{f(\alpha t x)}{\alpha} = t \inf_{\alpha>0} \frac{f(\alpha t x)}{t\alpha} = tg(x)。$$

当 $t = 0$ 时，$g(tx) = g(0) = 0$。
因此，g 是齐次的。

(b) 若 h 是一个齐次的下界函数，则

$$h(x) = \frac{h(\alpha x)}{\alpha} \le \frac{f(\alpha x)}{\alpha}, \quad \forall \alpha > 0。$$

对 α 取下确界，得到 $h(x) \le g(x)$。
因此，g 是 f 的最大齐次下界函数。

(c) 取任意 $x_1, x_2 \in \mathbb{R}^n$ 与 $\theta \in [0,1]$，设 $x_\theta = \theta x_1 + (1-\theta)x_2$。

由于

$$g(x) = \inf_{\alpha>0} \frac{f(\alpha x)}{\alpha} = \inf_{t>0} t f\left(\frac{x}{t}\right),$$

定义 $h(x,t) = tf\left(\frac{x}{t}\right)$，下面先证明，若 f 是凸函数，则 h 也是凸函数。

(1) 证明 h 的凸性

任取 $(u_1, s_1), (u_2, s_2) \in \operatorname{dom} h$ 与 $\theta \in [0,1]$，令

$$u_\theta = \theta u_1 + (1-\theta)u_2, \qquad s_\theta = \theta s_1 + (1-\theta)s_2 > 0.$$

记

$$\lambda_1 := \frac{\theta s_1}{s_\theta}, \qquad \lambda_2 := \frac{(1-\theta)s_2}{s_\theta},$$

则 $\lambda_1, \lambda_2 \ge 0$ 且 $\lambda_1 + \lambda_2 = 1$。于是

$$\frac{u_\theta}{s_\theta} = \frac{\theta u_1 + (1-\theta)u_2}{\theta s_1 + (1-\theta)s_2} = \lambda_1 \frac{u_1}{s_1} + \lambda_2 \frac{u_2}{s_2}.$$

由 f 的凸性，

$$f\left(\frac{u_\theta}{s_\theta}\right) = f\left(\lambda_1 \frac{u_1}{s_1} + \lambda_2 \frac{u_2}{s_2}\right) \le \lambda_1 f\left(\frac{u_1}{s_1}\right) + \lambda_2 f\left(\frac{u_2}{s_2}\right).$$

两边同乘以 $s_\theta > 0$，得

$$h(u_\theta, s_\theta) = s_\theta f\left(\frac{u_\theta}{s_\theta}\right) \le \theta\, s_1 f\left(\frac{u_1}{s_1}\right) + (1-\theta)\, s_2 f\left(\frac{u_2}{s_2}\right) = \theta\, h(u_1, s_1) + (1-\theta)\, h(u_2, s_2).$$

因此 h 为凸函数。

(2) 证明 g 的凸性

由 $g(x)=\inf_{t>0} h(x,t)$，取 $t_1,t_2>0$ 使得

$$h(x_1,t_1)\le g(x_1)+\varepsilon,\qquad h(x_2,t_2)\le g(x_2)+\varepsilon,\quad \varepsilon>0.$$

令 $t_\theta=\theta t_1+(1-\theta)t_2>0$。由 h 的凸性，

$$h(x_\theta,t_\theta)\le \theta h(x_1,t_1)+(1-\theta)h(x_2,t_2)\le \theta g(x_1)+(1-\theta)g(x_2)+\varepsilon.$$

又因 $g(x_\theta)=\inf_{t>0} h(x_\theta,t)\le h(x_\theta,t_\theta)$，故

$$g(x_\theta)\le \theta g(x_1)+(1-\theta)g(x_2)+\varepsilon.$$

令 $\varepsilon\to 0$，即得

$$g(x_\theta)\le \theta g(x_1)+(1-\theta)g(x_2)\,.$$

因此 g 为凸函数。