最优化导论第二次作业题答案

1. 一个集合 C 被称为 **中点凸的**，如果当任意两个点 $a, b \in C$ 时，它们的平均值或中点 $(a+b)/2$ 也在 C 中。显然，一个凸集必然是中点凸的。可以证明，在一些温和条件下，中点凸性蕴含凸性。作为一个简单情形，请证明：如果 C 是闭集并且是中点凸的，那么 C 是凸的。

证：若 $C \subset \mathbb{R}^n$ 是闭集并且**中点凸**（即 $\forall x, y \in C,\ (x+y)/2 \in C$），下证：

$$\forall\, x, y \in C,\ \forall\, \theta \in [0,1], \quad \theta x + (1-\theta) y \in C.$$

首先证明形如 $m/2^k$（其中 $k \in \mathbb{N},\ m = 0, 1, \ldots, 2^k$）的数满足如下条件：
$\forall\, x, y \in C, \forall\, k \in \mathbb{N}, \forall\, m \in \{0, \ldots, 2^k\}$, 有

$$\frac{m}{2^k} x + \left(1 - \frac{m}{2^k}\right) y \in C.$$

使用数学归纳法证明：

- $k = 1$ 时，$m \in \{0, 1, 2\}$：
  - $m = 0$ 得 $y \in C$;
  - $m = 2$ 得 $x \in C$;
  - $m = 1$ 得 $\frac{x+y}{2} \in C$（由中点凸）。
- 设结论对某个 k 成立。对 $k+1$：
  - 若 m 为偶数，则

    $$\frac{m}{2^{k+1}} = \frac{m/2}{2^k},$$

    由归纳假设可得结论。
  - 若 m 为奇数，记 $m_\pm = \frac{m \pm 1}{2}$。注意

    $$\frac{m}{2^{k+1}} x + \left(1 - \frac{m}{2^{k+1}}\right) y = \frac{1}{2}\left[\frac{m_+}{2^k} x + \left(1 - \frac{m_+}{2^k}\right) y\right] + \frac{1}{2}\left[\frac{m_-}{2^k} x + \left(1 - \frac{m_-}{2^k}\right) y\right].$$

    方括号内两点各自由归纳假设在 C 中；两者的**中点**仍在 C（中点凸），故结论成立。

由数学归纳法，结论成立。

给定任意 $x, y \in C$ 与 $\theta \in [0,1]$。对每个 $k \in \mathbb{N}$，取

$$\theta_k := \frac{m_k}{2^k}, \qquad m_k = \mathrm{round}(\theta \cdot 2^k),$$

round是四舍五入函数，即最接近 θ 的二进制分数（等价地，满足 $|\theta - \theta_k| \le 2^{-(k+1)}$）。
由第一步，点

$$z_k := \theta_k x + (1 - \theta_k) y \in C \quad (\forall k).$$

又

$$\|z_k - (\theta x + (1-\theta) y)\| = |\theta_k - \theta|\, \|x - y\| \le 2^{-(k+1)} \|x - y\| \xrightarrow[k \to \infty]{} 0,$$

故 $z_k \to \theta x + (1-\theta) y$。

由于 C 是**闭集**，极限点也在 C。于是 $\theta x+(1-\theta)y\in C$。
这对任意 $x,y\in C$、任意 $\theta\in[0,1]$ 成立，故 C 为凸集。

2. 集合 $C\subseteq\mathbf{R}^n$ 的支撑函数定义为

$$S_C(y)=\sup\{y^Tx\mid x\in C\}.$$

（我们允许 $S_C(y)$ 取值为 $+\infty$。）假设 C 和 D 是 $\mathbf{R}^n$ 中的闭凸集。证明：当且仅当它们的支撑函数相等时，$C=D$。

证：“$\Rightarrow$”（显然）

若 $C=D$，则对一切 y，$\sup_{x\in C}y^\top x=\sup_{x\in D}y^\top x$，故 $s_C=s_D$。

“$\Leftarrow$”（关键）

设 $s_C(y)=s_D(y)$ 对所有 y 成立。我们证明 $C\subseteq D$ 与 $D\subseteq C$。

(a) 证 $D\subseteq C$

反证法。若存在 $x_0\in D\setminus C$。由于 C 是**闭且凸**，由强分离定理可得存在 $a\in\mathbb{R}^n\setminus\{0\}$ 与常数 b 使得

$$a^\top x\le b\quad(\forall x\in C),\qquad a^\top x_0>b.$$

于是

$$s_C(a)=\sup_{x\in C}a^\top x\le b<a^\top x_0\le\sup_{x\in D}a^\top x=s_D(a).$$

从而 $s_C(a)<s_D(a)$，与 $s_C=s_D$ 矛盾。故 $D\subseteq C$。

(b) 证 $C\subseteq D$

互换 C 与 D 的角色，重复上面的论证，可得 $C\subseteq D$。

由 (a)(b)，得 $C=D$。

注：**强分离定理:**

设 $C\subset\mathbb{R}^n$ 是**非空闭凸集**，$x_0\notin C$。
则存在一个非零向量 $a\in\mathbb{R}^n$ 和一个实数 b，使得

$$a^\top x_0>b\quad\text{且}\quad a^\top x\le b\quad\forall x\in C.$$

3. 考虑集合

$$C=\{(x_1,x_2)\in\mathbb{R}^2\mid |x_1|+\sqrt{|x_2|}\le 1\}.$$

判断集合 C 是否为凸集

解：C 的凸性判断：**非凸**

把约束改写一下：由 $|x_1|+\sqrt{|x_2|}\le 1$ 得

$$\sqrt{|x_2|}\le 1-|x_1| \quad \Longleftrightarrow \quad |x_2|\le (1-|x_1|)^2,\ \ |x_1|\le 1.$$

记 $g(x_1):=(1-|x_1|)^2$（在 $[-1,1]$ 上是**凸**函数）。集合可写为

$$C=\{(x_1,x_2):\ |x_1|\le 1,\ -g(x_1)\le x_2\le g(x_1)\}.$$

若要使 $\{(x_1,x_2):x_2\le g(x_1)\}$ 为凸集，需要 g **凹**（因为它是一个上图（hypograph））。但这里 g 是**凸**的，因此直观上“上边界向里拱”的带状区域通常是**非凸**的。

给出具体反例：
取

$$a=(0,1),\quad b=(1,0),$$

有 $|0|+\sqrt{|1|}=1$、$|1|+\sqrt{0}=1$，所以 $a,b\in C$。但它们的中点

$$m=\tfrac{a+b}{2}=\left(\tfrac{1}{2},\tfrac{1}{2}\right)$$

满足

$$|m_1|+\sqrt{|m_2|}=\tfrac{1}{2}+\sqrt{\tfrac{1}{2}}\approx 0.5+0.7071=1.2071>1,$$

故 $m\notin C$。因此 C **不是凸集。**

4. 设 $\mathbb{R}^2$ 中的凸集 $S_1=\{(x_1,x_2)\mid x_1^2+x_2^2\le 1\}$（单位闭圆盘）， $S_2=\{(x_1,x_2)\mid x_1\ge 1\}$（右半闭半空间）。 求 $S=S_1\cap S_2$，并证明 S 是凸集。

证：若 $(x_1,x_2)\in S$，则需同时满足：

1. $x_1^2+x_2^2\le 1$;
2. $x_1\ge 1$。

由 $x_1\ge 1$ 得 $x_1^2\ge 1$。因此

$$1\le x_1^2\le 1-x_2^2 \quad\Longrightarrow\quad x_2^2\le 0 \quad\Longrightarrow\quad x_2=0.$$

代回得 $x_1^2\le 1$，结合 $x_1\ge 1$，可知 $x_1=1$。
因此，$S=\{(1,0)\}$.
取任意 $x,y\in S$，必有 $x=y=(1,0)$。对任意 $\lambda\in[0,1]$，

$$\lambda x+(1-\lambda)y=(1,0)\in S.$$

故 S 满足凸性的定义。

5. 设 a 和 b 是 $\mathbf{R}^n$ 中的两个不同点。证明所有距离 a 比距离 b 更近的点（欧几里得范数意义下），即

$$\{x\mid \|x-a\|_2\le\|x-b\|_2\},$$

构成一个半空间。将其显式地写成形如 $c^Tx\le d$ 的不等式。

证：从条件出发：

$$\|x-a\|_2 \le \|x-b\|_2 \quad\Longleftrightarrow\quad \|x-a\|_2^2 \le \|x-b\|_2^2.$$

展开两边平方：

$$(x-a)^\top (x-a) \le (x-b)^\top (x-b).$$

即

$$x^\top x - 2a^\top x + a^\top a \;\;\le\;\; x^\top x - 2b^\top x + b^\top b.$$

整理为：

$$2(b-a)^\top x \;\;\le\;\; b^\top b - a^\top a.$$

记

$$c = 2(b-a), \qquad d = b^\top b - a^\top a,$$

则集合可写成

$$\{x \mid c^\top x \le d\}.$$

因此所有距离 a 比距离 b 更近的点构成一个半空间。

6. 设 $C \subseteq \mathbf{R}^n$ 为以下二次不等式的解集：

$$C = \{x \in \mathbf{R}^n \mid x^T A x + b^T x + c \le 0\},$$

其中 $A \in \mathbf{S}^n$，$b \in \mathbf{R}^n$，$c \in \mathbf{R}$。

证明：若 $A \succeq 0$，则 C 是凸集。

证：定义二次函数：

$$q(x) = x^\top A x + b^\top x + c.$$

取任意 $x_1, x_2 \in C$ 与 $\theta \in [0,1]$，记
$z = \theta x_1 + (1-\theta)x_2$。利用对称性可得恒等式

$$z^\top A z = \theta x_1^\top A x_1 + (1-\theta) x_2^\top A x_2 - \theta(1-\theta)(x_1-x_2)^\top A (x_1-x_2).$$

同时 $b^\top z = \theta b^\top x_1 + (1-\theta) b^\top x_2$。于是

$$q(z) = \theta q(x_1) + (1-\theta) q(x_2) - \theta(1-\theta)(x_1-x_2)^\top A(x_1-x_2).$$

当 $A \succeq 0$ 时，$(x_1-x_2)^\top A(x_1-x_2) \ge 0$，故上式右端

$$q(z) \ \leq \ \theta q(x_1) + (1-\theta)q(x_2) \ \leq \ 0,$$

因为 $x_1, x_2 \in C \Rightarrow q(x_1) \leq 0,\, q(x_2) \leq 0$。
从而 $z \in C$，即 C 凸。

$\square$

7. 证明：如果 S_1 和 S_2 是 $\mathbf{R}^{m \times n}$ 中的凸集，那么它们的部分和

$$S = \{(x, y_1 + y_2) \mid x \in \mathbf{R}^m,\ y_1, y_2 \in \mathbf{R}^n,\ (x, y_1) \in S_1,\ (x, y_2) \in S_2\}$$

也是凸集。

证：取任意两点 $z^{(1)}, z^{(2)} \in S$。
则存在 $x^{(1)}, x^{(2)} \in \mathbb{R}^m$ 与 $y_1^{(1)}, y_2^{(1)}, y_1^{(2)}, y_2^{(2)} \in \mathbb{R}^n$ 使

$$z^{(1)} = (x^{(1)},\, y_1^{(1)} + y_2^{(1)}), \quad z^{(2)} = (x^{(2)},\, y_1^{(2)} + y_2^{(2)}),$$

并且

$$(x^{(1)}, y_1^{(1)}),\ (x^{(2)}, y_1^{(2)}) \in S_1, \qquad (x^{(1)}, y_2^{(1)}),\ (x^{(2)}, y_2^{(2)}) \in S_2.$$

对任意 $\theta \in [0,1]$，考虑它们的凸组合：

$$\theta z^{(1)} + (1-\theta) z^{(2)} = \Big(\theta x^{(1)} + (1-\theta) x^{(2)},\ \theta y_1^{(1)} + (1-\theta) y_1^{(2)} \ + \ \theta y_2^{(1)} + (1-\theta) y_2^{(2)}\Big).$$

由于 S_1 凸且 $(x^{(1)}, y_1^{(1)}), (x^{(2)}, y_1^{(2)}) \in S_1$，得

$$\Big(\theta x^{(1)} + (1-\theta) x^{(2)},\ \theta y_1^{(1)} + (1-\theta) y_1^{(2)}\Big) \in S_1.$$

同理，因 S_2 凸且 $(x^{(1)}, y_2^{(1)}), (x^{(2)}, y_2^{(2)}) \in S_2$，得

$$\Big(\theta x^{(1)} + (1-\theta) x^{(2)},\ \theta y_2^{(1)} + (1-\theta) y_2^{(2)}\Big) \in S_2.$$

于是根据 S 的定义，

$$\Big(\theta x^{(1)} + (1-\theta) x^{(2)},\ \big[\theta y_1^{(1)} + (1-\theta) y_1^{(2)}\big] + \big[\theta y_2^{(1)} + (1-\theta) y_2^{(2)}\big]\Big) \in S.$$

也即

$$\theta z^{(1)} + (1-\theta) z^{(2)} \in S.$$

因此，对任意两点及任意 $\theta \in [0,1]$ 均成立，S 为凸集。$\square$

8. 设 $C, D \subset \mathbb{R}^n$ 为两个不相交的凸集。考虑集合

$$S = \{(a,b) \in \mathbb{R}^{n+1} \mid a^\top x \le b,\ \forall x \in C;\ a^\top y \ge b,\ \forall y \in D\}.$$

证明：集合 S 是一个凸集。

证：取任意两点 $(a_1, b_1), (a_2, b_2) \in S$，以及任意 $\theta_1, \theta_2 \ge 0$。
记

$$(a,b) = (\theta_1 a_1 + \theta_2 a_2,\ \theta_1 b_1 + \theta_2 b_2).$$

任取 $x \in C$，因为 $(a_1, b_1), (a_2, b_2) \in S$，有

$$a_1^\top x \le b_1, \quad a_2^\top x \le b_2.$$

于是

$$(\theta_1 a_1 + \theta_2 a_2)^\top x = \theta_1 a_1^\top x + \theta_2 a_2^\top x \le \theta_1 b_1 + \theta_2 b_2 = b.$$

故对所有 $x \in C$，都有 $a^\top x \le b$。

任取 $y \in D$，同理有

$$a_1^\top y \ge b_1, \quad a_2^\top y \ge b_2,$$

于是

$$(\theta_1 a_1 + \theta_2 a_2)^\top y = \theta_1 a_1^\top y + \theta_2 a_2^\top y \ge \theta_1 b_1 + \theta_2 b_2 = b.$$

故对所有 $y \in D$，都有 $a^\top y \ge b$。

因此 S 是凸集。□

9. 设 K^* 为凸锥 K 的对偶锥。证明以下性质：

(a) K^* 确实是一个凸锥。

(b) 若 $K_1 \subseteq K_2$，则 $K_2^* \subseteq K_1^*$。

(c) K^* 是闭集。

(d) K^* 的内部由下式给出： $\operatorname{int} K^* = \{y \mid y^T x > 0 \ \ \forall x \in K \setminus \{0\}\}.$

(e) 如果 K 有非空内部，则 K^* 是尖的（pointed）。

(f) K^{**} 是 K 的闭包。（因此，如果 K 是闭集，则 $K^{**} = K$。）

证：（a）K^* 是凸锥：
由定义：

$$K^* = \{y \mid y^\top x \ge 0,\ \forall x \in K\}.$$

若 $y_1, y_2 \in K^*$，$\theta_1, \theta_2 \ge 0$，则

$$(\theta_1 y_1 + \theta_2 y_2)^\top x = \theta_1 y_1^\top x + \theta_2 y_2^\top x \ge 0, \quad \forall x \in K.$$

故 $\theta_1 y_1 + \theta_2 y_2 \in K^*$。因此 K^* 为凸锥。

(b) 若 $K_1 \subset K_2$，取 $y \in K_2^*$，则 $\forall x \in K_2,\ y^\top x \geq 0$。

由于 $K_1 \subset K_2$，$\forall x \in K_1$，亦有 $y^\top x \geq 0$。

故 $y \in K_1^*$。因此 $K_2^* \subset K_1^*$。

(c) 设 $\{y^k\} \subset K^*$，且 $y^k \to y$。

对任意 $x \in K$，有 $(y^k)^\top x \geq 0$，取极限得 $y^\top x \geq 0$。

因此 $y \in K^*$。所以 K^* 闭

(d) 若 $y \in \operatorname{int} K^*$，则存在 $\epsilon > 0$ 使得 $B(y, \epsilon) \subset K^*$。

若存在 $x \in K \setminus \{0\}$ 使 $y^\top x = 0$，则取 $u = -x/\|x\|$，有 $(y + \delta u)^\top x < 0$（当 $\delta > 0$ 足够小时），矛盾。

因此必须对所有 $x \in K$，有 $y^\top x > 0$。

反之，若 y 满足 $\forall x \in K, y^\top x > 0$，则由连续性可知在某邻域内仍保持非负，从而 $y \in \operatorname{int} K^*$。
故

$$\operatorname{int} K^* = \{y \in \mathbb{R}^n \mid y^\top x > 0,\ \forall x \in K \setminus \{0\}\}.$$

(e) 若 K 有非空内部，反证法：若 K^* 不是尖锥，则存在 $y \neq 0$ 使 $y, -y \in K^*$。
则 $\forall x \in K$，有

$$y^\top x \geq 0, \quad (-y)^\top x \geq 0 \implies y^\top x = 0.$$

特别地，对 $x \in \operatorname{int} K$ 也成立。由内点性质，存在方向使 $y^\top x > 0$，矛盾。
故 K^* 是尖锥。

(f) 由定义，

$$K^{**} = \{x \in \mathbb{R}^n \mid y^\top x \geq 0,\ \forall y \in K^*\}.$$

K^*是所有包含 K 的闭半空间的法向量的交集。一个凸集总可以表示为所有包含它的闭半空间的交集，于是

$$K^{**} = \overline{K}.$$

特别地，当 K 是闭凸锥时，$K^{**} = K$。

10. 设 D 为 $\mathbb{R}^n$ 中的非空闭集，并且是凸集，且 $y \notin D$。证明存在唯一的点 $\bar{x} \in D$，使得
$\|y - \bar{x}\| = \inf_{x \in D} \|y - x\|$。

证：令 $r := \inf_{x \in D} \|y - x\| \in [0, +\infty)$。

取极小化序列 $\{x^k\} \subset D$ 使 $\|y - x^k\| \downarrow r$ （调递减收敛到r）。
对任意 k, m，记中点 $z^{k,m} = (x^k + x^m)/2$。由 D 的凸性，$z^{k,m} \in D$。

利用平行四边形恒等式：

$$\|y - z^{k,m}\|^2 = \tfrac{1}{2}\|y - x^k\|^2 + \tfrac{1}{2}\|y - x^m\|^2 - \tfrac{1}{4}\|x^k - x^m\|^2.$$

因 $z^{k,m} \in D$，有 $\|y - z^{k,m}\|^2 \geq r^2$。于是

$$\tfrac{1}{4}\|x^k - x^m\|^2 \leq \tfrac{1}{2}\left(\|y - x^k\|^2 + \|y - x^m\|^2\right) - r^2 \xrightarrow[k,m\to\infty]{} 0,$$

从而 $\{x^k\}$ 是 Cauchy 列。

$\mathbb{R}^n$ 完备 ⇒ $\{x^k\}$ 收敛到某点 $\bar{x} \in \mathbb{R}^n$。
又因 D 闭且 $x^k \in D$，故 $\bar{x} \in D$。因此

$$\|y - \bar{x}\| = \lim_k \|y - x^k\| = r,$$

存在性成立。

下证唯一性：设 $x_1, x_2 \in D$ 都达到最小值：$\|y - x_1\| = \|y - x_2\| = r$。

对中点 $z = (x_1 + x_2)/2 \in D$，同样有

$$\|y - z\|^2 = \tfrac{1}{2}\|y - x_1\|^2 + \tfrac{1}{2}\|y - x_2\|^2 - \tfrac{1}{4}\|x_1 - x_2\|^2 = r^2 - \tfrac{1}{4}\|x_1 - x_2\|^2.$$

由于 $z \in D$，应有 $\|y - z\| \geq r$，从而 $\|x_1 - x_2\| = 0$，即 $x_1 = x_2$。唯一性成立。