第五次作业

1. 判断以下集合是否为凸集。

(a) 到某定点 x_0 比到给定集合 $\mathcal{A}$ 更近的点构成的集合

$$\mathcal{S}_1 = \left\{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \le \|x - y\|_2, \forall y \in \mathcal{A}\right\}$$

(b) 到集合 $\mathcal{A}_1$ 比到集合 $\mathcal{A}_2$ 更近的点构成的集合

$$\mathcal{S}_2 = \{x \mid \operatorname{dist}(x, \mathcal{A}_1) \le \operatorname{dist}(x, \mathcal{A}_2)\}$$

其中 $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathbb{R}^n$，点到集合距离定义为

$$\operatorname{dist}(x, \mathcal{A}) = \min\{\|x - z\|_2 \mid z \in \mathcal{A}\}$$

(c) 已知集合 $\mathcal{A}_1, \mathcal{A}_2$ 是 $\mathbb{R}^{n+m}$ 中的凸集。定义新的集合

$$\mathcal{S}_3 = \left\{(x, y_1 + y_2) \,\middle|\, \begin{array}{l} x \in \mathbb{R}^m, y_1 \in \mathbb{R}^n, y_2 \in \mathbb{R}^n \\ (x, y_1) \in \mathcal{A}_1, (x, y_2) \in \mathcal{A}_2 \end{array}\right\}$$

解

(a) 对于任意给定的 y

$$\begin{aligned} & \|x - x_0\|_2 \le \|x - y\|_2 \\ \Leftrightarrow\ & (x - x_0)^\top (x - x_0) \le (x - y)^\top (x - y) \\ \Leftrightarrow\ & x^\top x - 2x_0^\top x + x_0^\top x_0 \le x^\top x - 2y^\top x + y^\top y \\ \Leftrightarrow\ & 2(y - x_0)^\top x \le y^\top y - x_0^\top x_0 \end{aligned}$$

$\mathcal{S}_1$ 可以表示成 (无穷多) 半空间的交集

$$\bigcap_{y \in \mathcal{A}} \{x \mid \|x - x_0\|_2 \le \|x - y\|_2\}$$

因此 $\mathcal{S}_1$ 是凸集

(b) $\mathcal{S}_2$ 不一定是凸集。取 $\mathcal{A}_1 = \{-1, 1\}$，$\mathcal{A}_2 = \{0\}$，则

$$\mathcal{S}_2 = \{x \in \mathbb{R} \mid x \le -1/2 \text{ or } x \ge 1/2\}$$

(c) 考察 $\mathcal{S}_3$ 中的两个点 $(\bar{x},\bar{y}_1+\bar{y}_2)$ 和 $(\tilde{x},\tilde{y}_1+\tilde{y}_2)$，即

$$(\bar{x},\bar{y}_1)\in\mathcal{A}_1,\ (\bar{x},\bar{y}_2)\in\mathcal{A}_2,\ (\tilde{x},\tilde{y}_1)\in\mathcal{A}_1,\ (\tilde{x},\tilde{y}_2)\in\mathcal{A}_2$$

对任意 $0\leq\theta\leq 1$：

$$\begin{aligned}&\theta(\bar{x},\bar{y}_1+\bar{y}_2)+(1-\theta)(\tilde{x},\tilde{y}_1+\tilde{y}_2)\\=&\theta\bar{x}+(1-\theta)\tilde{x},(\theta\bar{y}_1+(1-\theta)\tilde{y}_1)+(\theta\bar{y}_2+(1-\theta)\tilde{y}_2)\end{aligned}$$

由于集合 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 是凸集

$$\begin{aligned}\theta\bar{x}+(1-\theta)\tilde{x},\theta\bar{y}_1+(1-\theta)\tilde{y}_1\in\mathcal{A}_1\\ \theta\bar{x}+(1-\theta)\tilde{x},\theta\bar{y}_2+(1-\theta)\tilde{y}_2\in\mathcal{A}_2\end{aligned}$$

故 $\theta(\bar{x},\bar{y}_1+\bar{y}_2)+(1-\theta)(\tilde{x},\tilde{y}_1+\tilde{y}_2)\in\mathcal{S}_3$，所以 $\mathcal{S}_3$ 是凸集。

$\square$

2. 判断下列函数的凸性。

(a) $f(x)=e^x-1$，其中 $x\in\mathbb{R}$

(b) $f(x_1,x_2)=x_1x_2$，其中 $(x_1,x_2)\in\mathbb{R}^2_{++}$

(c) $f(x_1,x_2)=1/x_1x_2$，其中 $(x_1,x_2)\in\mathbb{R}^2_{++}$

(d) $f(x_1,x_2)=x_1/x_2$，其中 $(x_1,x_2)\in\mathbb{R}^2_{++}$

(e) $f(x_1,x_2)=x_1^2/x_2$，其中 $(x_1,x_2)\in\mathbb{R}\times\mathbb{R}_{++}$

(f) $h=f\cdot g$，其中 f 和 g 是非负单增凸函数。

解

(a) 凸，二阶导数大于 0

(b) 非凸，海森矩阵为

$$H=\begin{bmatrix}0&1\\1&0\end{bmatrix},\ \det(H)=-1<0$$

(c) 凸，海森矩阵为

$$H=\begin{bmatrix}\dfrac{2}{x_1^3x_2}&\dfrac{1}{x_1^2x_2^2}\\\dfrac{1}{x_1^2x_2^2}&\dfrac{2}{x_1x_2^3}\end{bmatrix}=\frac{1}{x_1x_2}\begin{bmatrix}\dfrac{2}{x_1^2}&\dfrac{1}{x_1x_2}\\\dfrac{1}{x_1x_2}&\dfrac{2}{x_2^2}\end{bmatrix}\succeq 0,\forall(x_1,x_2)\in\mathbb{R}^2_{++}$$

(d) 非凸，海森矩阵为

$$H=\frac{1}{x_2^2}\begin{bmatrix}0 & -1\\ -1 & 2x_1/x_2\end{bmatrix},\quad \det(H)=-\frac{1}{x_2^4}<0$$

(e) 凸，海森矩阵为

$$H=\begin{bmatrix}\dfrac{2}{x_2} & -\dfrac{2x_1}{x_2^2}\\ -\dfrac{2x_1}{x_2^2} & \dfrac{2x_1^2}{x_2^3}\end{bmatrix}=\frac{2}{x_2^3}\begin{bmatrix}x_2\\ -x_1\end{bmatrix}\begin{bmatrix}x_2\\ -x_1\end{bmatrix}^\top \succeq 0,\ \forall (x_1,x_2)\in\mathbb{R}\times\mathbb{R}_{++}$$

(f) 对任意 $0\le\theta\le 1$

$$\begin{aligned}
& f(\theta x+(1-\theta)y)\cdot g(\theta x+(1-\theta)y)\\
\le\ & (\theta f(x)+(1-\theta)f(y))\cdot(\theta g(x)+(1-\theta)g(y))\\
=\ & \theta f(x)g(x)+(1-\theta)f(y)g(y)\\
& +\theta(1-\theta)(f(y)-f(x))(g(x)-g(y))
\end{aligned}$$

由于 f 和 g 单调增，上式第 3 项非正，所以

$$\begin{aligned}
& f(\theta x+(1-\theta)y)\cdot g(\theta x+(1-\theta)y)\\
\le\ & \theta f(x)g(x)+(1-\theta)f(y)g(y)
\end{aligned}$$

因此 h 是凸函数。

□

3. 求解如下优化问题

$$\begin{aligned}
\min\quad & 2^x\cdot 4^y\\
\text{s.t.}\quad & \mathrm{e}^{-3x+y}\le 1\\
& |2x-y|\le 4
\end{aligned}$$

解 目标函数 $2^x\cdot 4^y=2^{x+2y}$，将目标函数替换为 $x+2y$ 不改变最优解；$\mathrm{e}^{-3x+y}\le 1$ 等价于 $-3x+y\le 0$; 绝对值约束可以写成线性不等式约束; 原问题和以下线性规划最优解相同

$$\begin{aligned}
\min\quad & x+2y\\
\text{s.t.}\quad & -3x+y\le 0\\
& -4\le 2x-y\le 4
\end{aligned}$$

最优解为 $(-4,-12)$ □

4. 考虑以下优化问题

$$\begin{aligned} \min \quad & (x_1-2)^2+(x_2-1)^2 \\ \text{s.t.} \quad & -x_1+x_2 \geqslant 0 \\ & -x_1-x_2+2 \geqslant 0 \end{aligned}$$

(a) 求解满足 KKT 条件的 (x_1^*, x_2^*) 及拉格朗日乘子。

(b) 写出该问题的拉格朗日对偶问题并求解。

解

(a) KKT 条件为

$$\begin{aligned} 2(x_1-2)+\lambda_1+\lambda_2 &= 0 \\ 2(x_2-1)-\lambda_1+\lambda_2 &= 0 \\ -x_1+x_2 &\geq 0 \\ -x_1-x_2+2 &\geq 0 \\ \lambda_1 \geq 0,\ \lambda_2 &\geq 0 \\ \lambda_1(x_1-x_2) &= 0 \\ \lambda_2(x_1+x_2-2) &= 0 \end{aligned}$$

若 $\lambda_1=\lambda_2=0$，则有 $x_1=2, x_2=1$，两个约束都不满足。

若 $\lambda_1=0, \lambda_2>0$，则第一个约束不起作用，KKT 条件变为

$$\begin{aligned} x_1+x_2 &= 2 \\ 2(x_1-2)+\lambda_2 &= 0 \\ 2(x_2-1)+\lambda_2 &= 0 \end{aligned}$$

解得 $x_1=3/2, x_2=1/2$，不满足第一个约束条件。

若 $\lambda_1>0, \lambda_2=0$，则第二个约束不起作用，KKT 条件变为

$$\begin{aligned} x_1 &= x_2 \\ 2(x_1-2)-\lambda_2 &= 0 \\ 2(x_2-1)+\lambda_2 &= 0 \end{aligned}$$

解得 $x_1=x_2=3/2$，不满足第二个约束条件。

若 $\lambda_1 > 0, \lambda_2 > 0$，两个约束都起作用，因此

$$x_1 - x_2 = 0$$
$$x_1 + x_2 = 2$$

解得 $x_1 = x_2 = 1$，代入 KKT 条件得 $\lambda_1 = \lambda_2 = 1$。

因此 $(x_1, x_2, \lambda_1, \lambda_2) = (1, 1, 1, 1)$ 满足 KKT 条件。

(b) 拉格朗日函数为

$$\begin{aligned} & L(x, \lambda) \\ =& (x_1 - 2)^2 + (x_2 - 1)^2 + \lambda_1(x_1 - x_2) + \lambda_2(x_1 + x_2 - 2) \\ =& x_1^2 + x_2^2 + (\lambda_1 + \lambda_2 - 4)x_1 + (\lambda_2 - \lambda_1 - 2)x_2 + 5 - 2\lambda_2 \end{aligned}$$

对偶函数为

$$\begin{aligned} D(\lambda) &= \min_x L(x, \lambda) \\ &= -\frac{(\lambda_1 + \lambda_2 - 4)^2 + (\lambda_1 - \lambda_2 + 2)^2}{4} + 5 - 2\lambda_2 \\ &= -\frac{\lambda_1^2 + \lambda_2^2}{2} + \lambda_1 + \lambda_2 \end{aligned}$$

对偶问题的最优解为 $\lambda_1 = \lambda_2 = 1$，最优值为 1。

□

5. 考虑凸二次优化问题

$$\min \left\{ \frac{x^\top Q x}{2} + q^\top x \;\middle|\; Ax \geq b \right\} \tag{QP}$$

(a) 若 Q 是正定矩阵，写出 (QP) 的拉格朗日对偶问题。

(b) 若 Q 是半正定矩阵，其逆矩阵不存在。试通过将 (QP) 表示为二阶锥优化，进一步通过锥优化的对偶形式推出 (QP) 的对偶问题 (本问选作)。

解

(a) 问题 (QP) 的拉格朗日函数为

$$\begin{aligned} L(x, \lambda) &= \frac{1}{2} x^\top Q x + q^\top x - \lambda^\top (Ax - b) \\ &= \frac{1}{2} x^\top Q x + (q^\top - \lambda^\top A)x + \lambda^\top b \end{aligned}$$

$L(x,\lambda)$ 是关于 x 的严格凸二次函数，当

$$x = Q^{-1}(A^\top\lambda - q)$$

时，拉格朗日函数取极小值, 故对偶函数

$$D(\lambda) = \lambda^\top b - \frac{(A^\top\lambda - q)^\top Q^{-1}(A^\top\lambda - q)}{2}$$

引入附加变量 $\mu = Q^{-1}(A^\top\lambda - q)$，则对偶问题为

$$\begin{aligned} \max \quad & \lambda^\top b - \frac{\mu^\top Q\mu}{2} \\ \text{s.t.} \quad & A^\top\lambda - Q\mu = q \\ & \lambda \geq 0 \end{aligned}$$

(b) 若 Q 是半正定矩阵，其逆矩阵不存在，上述推导不成立。将 (QP) 写为如下二阶锥优化问题

$$\begin{aligned} \min \quad & \frac{t}{2} + q^\top x \\ \text{s.t.} \quad & Ax \geq b : \lambda \\ & \begin{bmatrix} 2Dx \\ t \\ t \end{bmatrix} \geq_{\mathbb{L}^3} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} : \begin{matrix} \nu \\ \sigma \\ \rho \end{matrix} \end{aligned}$$

其中 D 是 Q 的平方根矩阵，满足 $Q = D^\top D$。对偶问题为

$$\begin{aligned} \max \quad & \sigma - \rho + b^\top\lambda \\ \text{s.t.} \quad & \rho^2 \geq \sigma^2 + \nu^\top\nu \\ & A^\top\lambda + 2D\nu = q \\ & \sigma + \rho = 1/2 \\ & \lambda \geq 0 \end{aligned}$$

由于 $\rho^2 - \sigma^2 = (\rho+\sigma)(\rho-\sigma) = (\rho-\sigma)/2$，对偶问题简化为

$$\begin{aligned} \max \quad & \sigma - \rho + b^\top\lambda \\ \text{s.t.} \quad & \sigma - \rho \leq -2\nu^\top\nu \\ & A^\top\lambda + 2D\nu = q \\ & \lambda \geq 0 \end{aligned}$$

消去目标函数中的 $\sigma-\rho$ 可得

$$\begin{aligned}\max\quad & b^\top\lambda-2\nu^\top\nu\\ \text{s.t.}\quad & A^\top\lambda+2D\nu=q\\ & \lambda\geq 0\end{aligned}$$

令 $\nu=-D\mu/2$，做变量替换可得最终的对偶形式

$$\begin{aligned}\max\quad & b^\top\lambda-\frac{\mu^\top Q\mu}{2}\\ \text{s.t.}\quad & A^\top\lambda-Q\mu=q\\ & \lambda\geq 0\end{aligned}$$

由于推导过程没有假设 Q 矩阵可逆，上述结论在 $Q\succ 0$ 和 $Q\succeq 0$ 两种情况下都适用。

$\square$