南京大学智能科学与技术学院
最优化方法导论期中试卷

自觉遵守考试规则，诚信考试，绝不作弊

装订线内不要答题

院 (系) ____________ 班级 ____________ 姓名 ____________ 学号 ____________

题号	一	二	三	四	五	六	七	八	满分
得分									

1. 设 D 是 $\mathbb{R}^n$ 中的一个非空凸集, f 是定义在 D 上的凸函数, α 是一个实数, 则水平集 $D_\alpha=\{\boldsymbol{x}\mid \boldsymbol{x}\in D, f(\boldsymbol{x})\leqslant\alpha\}$ 是凸集. (10 分)

证明　对于任意的 $\boldsymbol{x}_1,\boldsymbol{x}_2\in D_\alpha$, 根据 D_α 定义, 有

$$f\left(\boldsymbol{x}_1\right)\leqslant\alpha, f\left(\boldsymbol{x}_2\right)\leqslant\alpha.$$

由于 D 为凸集, 因此对每个数 $\lambda\in[0,1]$, 必有

$$\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in D.$$

又由于 $f(\boldsymbol{x})$ 是 D 上的凸函数, 则有

$$\begin{aligned} f\left(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\right) &\leqslant \lambda f\left(\boldsymbol{x}_1\right)+(1-\lambda)f\left(\boldsymbol{x}_2\right)\\ &\leqslant \lambda\alpha+(1-\lambda)\alpha=\alpha \end{aligned}$$

因此 $\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in D_\alpha$, 故 D_α 为凸集。

2. 用定义验证下列各集合是凸集: (1) $S=\{(x_1,x_2)\mid x_1+2x_2\geqslant 1, x_1-x_2\geqslant 1\}$; (2) $S=\{(x_1,x_2)\,|x_2\geqslant|\,x_1\,|\}$; (3) $S=\{(x_1,x_2)\mid x_1^2+x_2^2\leqslant 10\}$. (10 分)

证明　(1) 对集合 S 中任意两点 $x^{(1)}=\begin{bmatrix}x_1^{(1)}\\x_2^{(1)}\end{bmatrix}, x^{(2)}=\begin{bmatrix}x_1^{(2)}\\x_2^{(2)}\end{bmatrix}$ 及每个数 $\lambda\in[0,1]$, 有

$$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=\begin{bmatrix}\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\\ \lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\end{bmatrix}.$$

由题设, 有

$$\begin{aligned}&\left[\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\right]+2\left[\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\right]\\=&\lambda\left(x_1^{(1)}+2x_2^{(1)}\right)+(1-\lambda)\left(x_1^{(2)}+2x_2^{(2)}\right)\geqslant\lambda+(1-\lambda)=1,\\&\left[\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\right]-\left[\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\right]\\=&\lambda\left(x_1^{(1)}-x_2^{(1)}\right)+(1-\lambda)\left(x_1^{(2)}-x_2^{(2)}\right)\geqslant\lambda+(1-\lambda)=1,\end{aligned}$$

因此, $\lambda x^{(1)}+(1-\lambda)x^{(2)}\in S$, 故 S 是凸集. (2) 对集合 S 中任意两点 $\boldsymbol{x}^{(1)}=\begin{bmatrix}x_1^{(1)}\\x_2^{(1)}\end{bmatrix}$ 和 $\boldsymbol{x}^{(2)}=\begin{bmatrix}x_1^{(2)}\\x_2^{(2)}\end{bmatrix}$ 及每个数 $\lambda\in[0,1]$, 有

$$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=\begin{bmatrix}\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\\\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\end{bmatrix}.$$

由题设, 有

$$\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\geqslant\lambda\left|x_1^{(1)}\right|+(1-\lambda)\left|x_1^{(2)}\right|\geqslant\left|\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\right|,$$

因此 $\lambda x^{(1)}+(1-\lambda)x^{(2)}\in S$, 故 S 是凸集.

(3) 对集合 S 中任意两点 $x^{(1)}=\begin{bmatrix}x_1^{(1)}\\x_2^{(1)}\end{bmatrix}$ 和$x^{(2)}=\begin{bmatrix}x_1^{(2)}\\x_2^{(2)}\end{bmatrix}$ 及每个数$\lambda\in[0,1]$, 有

$$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=\begin{bmatrix}\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\\\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\end{bmatrix}.$$

由题设, 有

$$\begin{aligned}&\left[\lambda x_1^{(1)}+(1-\lambda)x_1^{(2)}\right]^2+\left[\lambda x_2^{(1)}+(1-\lambda)x_2^{(2)}\right]^2\\=&\lambda^2x_1^{(1)2}+2\lambda(1-\lambda)x_1^{(1)}x_1^{(2)}+(1-\lambda)^2x_1^{(2)2}+\lambda^2x_2^{(1)2}+2\lambda(1-\lambda)x_2^{(1)}x_2^{(2)}\\&+(1-\lambda)^2x_2^{(2)2}=\lambda^2\left[x_1^{(1)2}+x_2^{(1)2}\right]+(1-\lambda)^2\left[x_1^{(2)2}+x_2^{(2)2}\right]+\lambda(1-\lambda)\left[2x_1^{(1)}x_1^{(2)}\right.\\&\left.+2x_2^{(1)}x_2^{(2)}\right]\leqslant10\lambda^2+10(1-\lambda)^2+\lambda(1-\lambda)\left[x_1^{(1)2}+x_1^{(2)2}+x_2^{(1)2}+x_2^{(2)2}\right]\\\leqslant&10\lambda^2+10(1-\lambda)^2+20\lambda(1-\lambda)=10,\end{aligned}$$

因此 $\lambda x^{(1)}+(1-\lambda)x^{(2)}\in S$, 故 S 是凸集.

3.$f(x)$ 为凸函数的充要条件是对任意的 $\boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$, 一元函数 $\varphi(\alpha)=f(\boldsymbol{x}+\alpha\boldsymbol{y})$ 是关于 α 的凸函数. (10 分)

证明 必要性. 设 $\lambda_1 \geqslant 0, \lambda_2 \geqslant 0$, 且 $\lambda_1+\lambda_2=1$, 由 $\varphi(\alpha)$ 的定义和 $f(x)$ 的凸性, 有

$$\begin{aligned}
\varphi\left(\lambda_1\alpha_1+\lambda_2\alpha_2\right) &= f\left(\boldsymbol{x}+\left(\lambda_1\alpha_1+\lambda_2\alpha_2\right)\boldsymbol{y}\right) \\
&= f\left(\lambda_1\boldsymbol{x}+\lambda_2\boldsymbol{x}+\left(\lambda_1\alpha_1+\lambda_2\alpha_2\right)\boldsymbol{y}\right) \\
&= f\left(\lambda_1\left(\boldsymbol{x}+\alpha_1\boldsymbol{y}\right)+\lambda_2\left(\boldsymbol{x}+\alpha_2\boldsymbol{y}\right)\right) \\
&\leqslant \lambda_1 f\left(\boldsymbol{x}+\alpha_1\boldsymbol{y}\right)+\lambda_2 f\left(\boldsymbol{x}+\alpha_2\boldsymbol{y}\right) \\
&= \lambda_1\varphi\left(\alpha_1\right)+\lambda_2\varphi\left(\alpha_2\right).
\end{aligned}$$

由定义知 $\varphi(\alpha)$ 是凸函数.

充分性. 任取 $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$, 记 $\boldsymbol{x}=\boldsymbol{x}+0(\boldsymbol{y}-\boldsymbol{x}), \boldsymbol{y}=\boldsymbol{x}+(\boldsymbol{y}-\boldsymbol{x}), \varphi(\alpha)=f(\boldsymbol{x}+\alpha(\boldsymbol{y}-\boldsymbol{x}))$, 则

$$\begin{aligned}
f\left(\lambda_1\boldsymbol{x}+\lambda_2\boldsymbol{y}\right) &= f\left(\lambda_1(\boldsymbol{x}+0(\boldsymbol{y}-\boldsymbol{x}))+\lambda_2(\boldsymbol{x}+(\boldsymbol{y}-\boldsymbol{x}))\right) \\
&= f\left(\boldsymbol{x}+\left(\lambda_1 0+\lambda_2 1\right)(\boldsymbol{y}-\boldsymbol{x})\right) \\
&= \varphi\left(\lambda_1 0+\lambda_2 1\right) \\
&\leqslant \lambda_1\varphi(0)+\lambda_2\varphi(1) \\
&= \lambda_1 f(\boldsymbol{x})+\lambda_2 f(\boldsymbol{y}),
\end{aligned}$$

故知 $f(\boldsymbol{x})$ 是凸函数.

4. 设 f 是定义在 $\mathbb{R}^n$ 上的凸函数, $x^{(1)}, x^{(2)}, \cdots, x^{(k)}$ 是 $\mathbb{R}^n$ 中的点, $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是非负数, 且满足 $\lambda_1+\lambda_2+\cdots+\lambda_k=1$, 证明:

$$f\left(\lambda_1 x^{(1)}+\lambda_2 x^{(2)}+\cdots+\lambda_k x^{(k)}\right) \leqslant \lambda_1 f\left(x^{(1)}\right)+\lambda_2 f\left(x^{(2)}\right)+\cdots+\lambda_k f\left(x^{(k)}\right).$$

(10 分)

证明 用数学归纳法. 当 $k=2$ 时, 根据凸函数的定义, 必有

$$f\left(\lambda_1 x^{(1)}+\lambda_2 x^{(2)}\right) \leqslant \lambda_1 f\left(x^{(1)}\right)+\lambda_2 f\left(x^{(2)}\right).$$

设 $k=m$ 时不等式成立. 当 $k=m+1$ 时, 有

$$\begin{aligned}
& f\left(\lambda_1 x^{(1)}+\lambda_2 x^{(2)}+\cdots+\lambda_m x^{(m)}+\lambda_{m+1} x^{(m+1)}\right) \\
=& f\left(\sum_{i=1}^m \lambda_i\left(\frac{\lambda_1}{\sum_{i=1}^m \lambda_i} x^{(1)}+\frac{\lambda_2}{\sum_{i=1}^m \lambda_i} x^{(2)}+\cdots+\frac{\lambda_m}{\sum_{i=1}^m \lambda_i} x^{(m)}\right)+\lambda_{m+1} x^{(m+1)}\right).
\end{aligned}$$

记

$$\hat{\boldsymbol{x}}=\frac{\lambda_1}{\sum_{i=1}^m \lambda_i}\boldsymbol{x}^{(1)}+\frac{\lambda_2}{\sum_{i=1}^m \lambda_i}\boldsymbol{x}^{(2)}+\cdots+\frac{\lambda_m}{\sum_{i=1}^m \lambda_i}\boldsymbol{x}^{(m)}.$$

自觉遵守考试规则，诚信考试，绝不作弊

装订线内不要答题

由于 $f(\boldsymbol{x})$ 是凸函数, $\sum_{i=1}^{m}\lambda_i+\lambda_{m+1}=1, \lambda_i\geqslant 0$, 根据凸函数定义, 有

$$f\left(\left(\sum_{i=1}^{m}\lambda_i\right)\hat{\boldsymbol{x}}+\lambda_{m+1}\boldsymbol{x}^{(m+1)}\right)\leqslant\left(\sum_{i=1}^{m}\lambda_i\right)f(\hat{\boldsymbol{x}})+\lambda_{m+1}f\left(\boldsymbol{x}^{(m+1)}\right).$$

根据归纳法假设, 有

$$f(\hat{x})\leqslant\frac{\lambda_1}{\sum_{i=1}^{m}\lambda_i}f\left(x^{(1)}\right)+\frac{\lambda_2}{\sum_{i=1}^{m}\lambda_i}f\left(x^{(2)}\right)+\cdots+\frac{\lambda_m}{\sum_{i=1}^{m}\lambda_i}f\left(\boldsymbol{x}^{(m)}\right).$$

代人上式, 则有

$$f\left(\lambda_1x^{(1)}+\lambda_2x^{(2)}+\cdots+\lambda_{m+1}x^{(m+1)}\right)\leqslant\lambda_1f\left(x^{(1)}\right)+\lambda_2f\left(x^{(2)}\right)+\cdots+\lambda_{m+1}f\left(x^{(m+1)}\right),$$

即 $k=m+1$ 时, 不等式也成立. 从而得证.

5. 试判定下述非线性规划是否为凸规划:

(1) $\begin{cases}\min f(X)=x_1^2+x_2^2+8\\ x_1^2-x_2\geq 0\\ -x_1-x_2^2+2=0\\ x_1,x_2\geq 0\end{cases}$ (2) $\begin{cases}\min f(X)=2x_1^2+x_2^2+x_3^2-x_1x_2\\ x_1^2+x_2^2\leq 4\\ 5x_1^2+x_3=10\\ x_1,x_2,x_3\geq 0\end{cases}$

(3) $\begin{cases}\max \quad f(X)=x_1+x_2\\ \text{s.t.}\ \begin{cases}x_1^2+x_2^2\leqslant 1\\ x_1,x_2\geqslant 0\end{cases}\end{cases}$ (15 分)

解 (1) $\begin{cases}\min f(X)=x_1^2+x_2^2+8\\ g_1(X)=x_1^2-x_2\geq 0\\ g_2(X)=-x_1-x_2^2+2=0\\ x_1,x_2\geq 0\end{cases}$ $f(X),g_1(X),g_2(X)$ 的海赛矩阵的行列式:

$$|H|=\begin{vmatrix}\frac{\partial^2f(X)}{\partial x_1^2} & \frac{\partial^2f(X)}{\partial x_1\partial x_2}\\ \frac{\partial^2f(X)}{\partial x_2\partial x_1} & \frac{\partial^2f(X)}{\partial x_2{}^2}\end{vmatrix}=\begin{vmatrix}2&0\\0&2\end{vmatrix}>0$$

$$|g_1|=\begin{vmatrix}\frac{\partial^2g_1(X)}{\partial x_1{}^2} & \frac{\partial^2g_1(X)}{\partial x_1\partial x_2}\\ \frac{\partial^2g_1(X)}{\partial x_2\partial x_1} & \frac{\partial^2g_1(X)}{\partial x_2{}^2}\end{vmatrix}=\begin{vmatrix}2&0\\0&0\end{vmatrix}\geq 0$$

$$|g_2|=\begin{vmatrix}\frac{\partial^2g_2(X)}{\partial x_1^2} & \frac{\partial^2g_2(X)}{\partial x_1\partial x_2}\\ \frac{\partial^2g_2(X)}{\partial x_2\partial x_1} & \frac{\partial^2g_2(X)}{\partial x_2{}^2}\end{vmatrix}=\begin{vmatrix}0&0\\0&-2\end{vmatrix}\leq 0$$

知 $f(X)$ 为严格凸函数, $g_1(X)$, 为凸函数, $g_2(X)$ 为凹函数, 所以不是一个凸规划问题.

自觉遵守考试规则，诚信考试，绝不作弊

装订线内不要答题

(2)

$$\begin{cases} \min f(X) = 2x_1^2 + x_2^2 + x_3^2 - x_1x_2 \\ g_1(X) = x_1^2 + x_2^2 \le 4 \Leftrightarrow g_1(X) = -\left(x_1^2 + x_2^2\right) + 4 \ge 0 \\ g_2(X) = 5x_1^2 + x_3 = 10 \\ x_1, x_2, x_3 \ge 0 \end{cases}$$

同上有 $f(X), g_1(X), g_2(X)$ 的海赛矩阵的行列式

$$|H| = \begin{vmatrix} 4 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} > 0$$

$|g_1| = \begin{vmatrix} -2 & & \\ & -2 & \\ & & 0 \end{vmatrix}$, 是凹函数, $|g_2| = \begin{vmatrix} 10 & & \\ & 0 & \\ & & 0 \end{vmatrix}$ 是凸函数, 不是凸规划问题。

(3)

$$\min(-f(X)) = -x_1 - x_2$$

$$\text{s.t.} \begin{cases} g_1(X) = 1 - x_1^2 - x_2^2 \geqslant 0 \\ g_2(X) = x_1 \geqslant 0 \\ g_3(X) = x_2 \geqslant 0 \end{cases}$$

$$H_{-f}(X) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \geqslant 0, H_{-g_1}(X) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0,$$

$$H_{-g_2}(X) = H_{-g_3}(X) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \geqslant 0$$

说明 $-f(X)$ 是凸函数, $g_1(X), g_2(X), g_3(X)$ 是凹函数. 因此, 本模型是一个凸规划.

6. 试证明在一个集合上最小化一个线性函数等价于在其凸包上最小化该函数, 即:

$$\inf_{x \in conv(X)} c'x = \inf_{x \in X} c'x,$$

其中 $X \subset R^n$ 和 $c \in R^n$。此外，当且仅当等式右侧的下确界可以达到时，等式左侧的下确界才能达到。(15 分)

证明　由于 $X \subset conv(X)$ 可得

$$\inf_{x \in conv(X)} c'x \le \inf_{x \in X} c'x.$$

当然，任意 $\bar{x} \in conv(X)$ 均可写为 $\bar{x} = \sum_{i=1}^{m} \alpha_i x_i$，其中 $x_1, \ldots, x_m \in X$ 且参数 $\alpha_1, \ldots, \alpha_m \geq 0$ 满足 $\sum_{i=1}^{m} \alpha_i = 1$。因此，由 $c'x_i \geq \inf_{x\in X} c'x$ 可得

$$c'\bar{x} = \sum_{i=1}^{m} \alpha_i c' x_i \geq (\sum_{i=1}^{m} \alpha_i) \inf_{x\in X} c'x = \inf_{x\in X} c'x, \quad \forall \bar{x} \in conv(X).$$

关于 $\bar{x} \in conv(X)$ 取左侧的下确界为

$$\inf_{\bar{x}\in conv(X)} c'\bar{x} \geq \inf_{x\in X} c'x.$$

结合上两式可得

$$\inf_{x\in conv(X)} c'x = \inf_{x\in X} c'x.$$

由于 $X \subset conv(X)$ 且 $\inf_{x\in conv(X)} c'x = \inf_{x\in X} c'x$，集合 X 上每个可以到达 $c'x$ 的下确界的点都可以在集合 $conv(X)$ 中找到。相反，假设函数 $c'x$ 在集合 $conv(X)$ 的下确界可以在 $\bar{x} = conv(X)$ 上找到。对于由变量 $x_1, \ldots, x_m$ 和满足 $\sum_{i=1}^{m} \alpha_i = 1$ 的标量 $\alpha_1, \ldots, \alpha_m \geq 0$ 构成的变量 $\bar{x} = \sum_{i=1}^{m} \alpha_i x_i$，有

$$\inf_{x\in X} c'x = (\sum_{i=1}^{m} \alpha_i) \inf_{x\in X} c'x \leq \sum_{i=1}^{m} \alpha_i c' x_i = c'\bar{x} = \inf_{x\in conv(X)} c'x = \inf_{x\in X} c'x.$$

由于上式的左边部分与右边部分相等，由此可见，对于所有的 i 和 $\alpha_i > 0$，当且仅当 $c'x_i = \inf_{x\in conv(X)} c'x$ 时，等式成立。因此，集合 X 的 $c'x$ 的下确界可以到达。

7. 令函数 $f: R^n \to R$ 满足

$$f(x) = \frac{1}{p} \sum_{i=1}^{n} |x_i|^p$$

其中 $1 < p$。请证明上述函数的共轭函数为

$$f^{\star}(y) = \frac{1}{q} \sum_{i=1}^{n} |y_i|^q,$$

其中 q 的定义如下

$$\frac{1}{p} + \frac{1}{q} = 1.$$

(15 分)

证明　首先考虑 $n = 1$ 的情况。定义 x 和 y 为标量。通过将 $xy - (\frac{1}{p})|x|^p$ 的导数设为零，并且我们可以得到当 $sgn(x)|x|^{p-1} = y$ 成立时，关于 x 的上确界可以达到，这意味着 $xy = |x|^p$ 和 $|x|^{p-1} = |y|$。通过将公式代入共轭函数可得

$$f^{\star}(y) = |x|^p - \frac{1}{p}|x|^p = (1 - \frac{1}{p})|x|^p = (1 - \frac{1}{p})|x|^p = \frac{1}{q}|y|^{\frac{p}{p-1}} = \frac{1}{q}|y|^q. \qquad (\star)$$

我们现在注意到对于任意函数 $f: R^n \to (-\infty, \infty]$ 均存在以下形式

$$f(x) = f_1(x_1) + \cdots + f_n(x_n),$$

其中 $x = (x_1, \ldots, x_n)$ 且 $f_i : R \to (-\infty, \infty], i = 1, \ldots, n$，其共轭函数为

$$f^\star(y) = f_1^\star(y_1) + \cdots + f_n^\star(y_n),$$

其中 $f_i^\star : R \to (-\infty, \infty]$ 是 f_i 的共轭函数，$i = 1, \ldots, n$。通过结合上述事实和 $(\star)$ 式可得，$f^\star(y)$ 为 $f(x)$ 的共轭函数。

8. 令 C 为一个在 R^n 上的非空闭合凸锥，令 $x \in R^n$。试证明当且仅当 $\hat{x} \in C, (x - \hat{x})'\hat{x} = 0, x - \hat{x} \in C^\star$ 时，$\hat{x}$ 是 x 在 C 上的投影。(15 分)

证明　令 $\hat{x}$ 是 x 在 C 上的投影（由于 C 是凸的且闭合的，所以必然且唯一存在 $\hat{x}$ 满足条件）。通过投影定理可得

$$(x - \hat{x})'(y - \hat{x}) \le 0, \qquad \forall y \in C.$$

由于 C 是一个锥体，可以推断出 $(\frac{1}{2})\hat{x} \in C$ 和 $2\hat{x} \in C$，并且通过取 $y = (\frac{1}{2})\hat{x}$ 和 $y = 2\hat{x}$，可以推断出

$$(x - \hat{x})'\hat{x} = 0.$$

通过结合上述两式可得

$$(x - \hat{x})'y \le 0, \qquad \forall y \in C,$$

这说明 $x - \hat{x} \in C^\star$。
相反的，若 $\hat{x} \in C, (x - \hat{x})'\hat{x} = 0$，且 $x - \hat{x} \in C^\star$，则可以推断出

$$(x - \hat{x})'(y - \hat{x}) \le 0, \qquad \forall y \in C,$$

并且根据投影定理可得，$\hat{x}$ 是 x 在 C 上的投影。

自觉遵守考试规则，诚信考试，绝不作弊

装订线内不要答题