南京大学智能科学与技术学院期末试卷

学年学期 24-25（1） 开课单位 智科 课程性质 必修 课程号 90111203

课程名称 最优化方法导论 (A 卷) 任课老师 杨林 考试时长 2 小时 总分 100 分

题号	一	二	三	四	五	六	七	八	九	总分
满分	12	8	8	20	10	10	10	10	12	100

得分	

一、证明下面命题。

（1）证明一个集合是凸集当且仅当它与任意直线的交是凸的。

（2）证明一个集合是仿射的, 当且仅当它与任意直线的交是仿射的。

总共 12 分，每问 6 分。

证明:

（1）必要性:

假设 C 是一个凸集。根据凸集的定义，对于任意的 $x, y \in C$ 和任意的 $\lambda \in [0,1]$，都有:

$$\lambda x + (1-\lambda)y \in C.$$

现在考虑任意的直线 L。设 $x, y \in C \cap L$，这意味着 $x, y \in C$ 且 $x, y \in L$。我们需要证明 $C \cap L$ 也是凸的。

取 $\lambda \in [0,1]$，根据直线的定义，x 和 y 也在直线 L 上，因此可以表示为:

$$z = \lambda x + (1-\lambda)y.$$

由于 $x, y \in C$ 且 C 是凸的，我们有:

$$z \in C.$$

我们还需要证明 $z \in L$。因为 x 和 y 在直线 L 上，z 也是在 L 上。因此:

$$z \in C \cap L.$$

因此，$C \cap L$ 是凸的。

充分性：

假设 $C \cap L$ 对任意直线 L 都是凸的。

取任意 $x, y \in C$，我们考虑通过 x 和 y 的直线 L，即：

$$L = \{z = \lambda x + (1-\lambda)y : \lambda \in \mathbb{R}\}.$$

显然，$x, y \in L$，并且因为 $x, y \in C$，所以 $x, y \in C \cap L$。

根据假设 $C \cap L$ 是凸的，因此对于任意的 $\lambda \in [0,1]$，我们有：

$$z = \lambda x + (1-\lambda)y \in C \cap L.$$

这意味着 $z \in C$ 且 $z \in L$，所以：

$$z \in C.$$

因此，对于任意的 $\lambda \in [0,1]$，我们得出 $\lambda x + (1-\lambda)y \in C$，这表明 C 是凸的。

（2）必要性：

假设 A 是一个仿射集。根据仿射集的定义，对于任意的 $x, y \in A$ 和任意的 $\lambda \in \mathbb{R}$，有：

$$\lambda x + (1-\lambda)y \in A.$$

考虑任意的直线 L。设 $x, y \in A \cap L$，这意味着 $x, y \in A$ 且 $x, y \in L$。我们需要证明 $A \cap L$ 也是仿射的。

取 $\lambda \in \mathbb{R}$，我们可以表示直线上的任意点 z 为：

$$z = \lambda x + (1-\lambda)y.$$

由于 $x, y \in A$ 且 A 是仿射的，因此 $z \in A$。

我们还需要证明 $z \in L$。由于 x 和 y 在直线 L 上，z 也在 L 上。因此：

$$z \in A \cap L.$$

这表明 $A \cap L$ 是仿射的。

充分性：

假设对于任意直线 L，$A \cap L$ 是仿射的。

取任意 $x, y \in A$，考虑通过 x 和 y 的直线 L，即：

$$L = \{z = \lambda x + (1-\lambda)y : \lambda \in \mathbb{R}\}.$$

显然，$x, y \in L$，并且因为 $x, y \in A$，所以 $x, y \in A \cap L$。

根据假设 $A \cap L$ 是仿射的，因此对于任意的 $\lambda \in \mathbb{R}$，我们有：

$$z = \lambda x + (1-\lambda)y \in A \cap L.$$

这意味着 $z \in A$ 且 $z \in L$，所以：

$$z \in A.$$

因此，对于任意的 $\lambda \in \mathbb{R}$，我们得出 $\lambda x + (1-\lambda)y \in A$，这表明 A 是仿射的。

得分	

二. 证明集合 S 的凸包是所有包含 S 的凸集的交。

总共 8 分，根据步骤的完整程度酌情给分。

证明：

首先，我们定义集合 S 的凸包 $\mathrm{conv}(S)$ 为所有可以通过 S 中点的凸组合得到的点的集合。换句话说，$\mathrm{conv}(S)$ 包含所有形如

$$x = \sum_{i=1}^{n} \lambda_i s_i \quad (s_i \in S, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1)$$

的点 x。

设 C 是一个包含 S 的凸集。因为 C 包含 S 中的所有点，对于任意的 $s_1, s_2 \in S$，它们的凸组合 $\lambda s_1 + (1-\lambda)s_2$ 也属于 C（因 C 是凸的）。

因此，我们得出结论：

$$\mathrm{conv}(S) \subseteq C.$$

设所有包含 S 的凸集的集合为 $\mathcal{C}$，即

$$\mathcal{C} = \{C \mid S \subseteq C \text{ 且} C \text{ 是凸的}\}.$$

那么所有这些 C 的交为

$$\bigcap_{C \in \mathcal{C}} C.$$

我们已经证明了任意包含 S 的凸集 C 都包含 $\mathrm{conv}(S)$，因此

$$\mathrm{conv}(S) \subseteq \bigcap_{C \in \mathcal{C}} C.$$

接下来，我们证明交是凸包。考虑任意点 $x \in \bigcap_{C \in \mathcal{C}} C$，这意味着 x 属于所有包含 S 的凸集 C。特别地，x 也必须是所有包含 S 的凸组合的极限点，这意味着 x 必须可以表示为某些 S 中点的凸组合。因此，$x \in \mathrm{conv}(S)$。

结合以上两点，我们得出结论：

$$\mathrm{conv}(S) = \bigcap_{C \in \mathcal{C}} C.$$

得分	

三. 假设 $K \subset X$ 是一个锥体。证明 K 是凸的，当且仅当对所有 $x, y \in K$，都有 $x + y \in K$。

总共 8 分，根据步骤的完整程度酌情给分。

证明:

必要性：假设 K 是一个凸集。根据凸集的定义，对于任意 $x, y \in K$ 和任意 $\lambda \in [0, 1]$，都有：

$$\lambda x + (1 - \lambda)y \in K.$$

取 $\lambda = \frac{1}{2}$，则我们得到：

$$\frac{1}{2}x + \frac{1}{2}y \in K.$$

这可以重写为：

$$\frac{1}{2}(x + y) \in K.$$

由于 K 是锥体，包含所有非负倍数的线性组合，因此我们可以得到 $x + y \in K$。

充分性：假设对于所有 $x, y \in K$，都有 $x + y \in K$。我们需要证明 K 是一个凸集，即对于任意 $x, y \in K$ 和任意 $\lambda \in [0, 1]$，都有 $\lambda x + (1 - \lambda)y \in K$。

设 $x, y \in K$ 且 $\lambda \in [0, 1]$。我们可以通过以下步骤证明 $\lambda x + (1 - \lambda)y \in K$。

1. 由于 K 是锥体，若 $x \in K$ 且 $\lambda > 0$，则 $\lambda x \in K$。2. 设 $\mu = 1 - \lambda$，则 $\mu \in [0, 1]$（因为 $\lambda \in [0, 1]$）。3. 同样，由于 K 是锥体，若 $y \in K$ 且 $\mu > 0$，则 $\mu y \in K$。

现在，我们可以利用 $x + y \in K$ 的假设。我们考虑 $x + y \in K$ 并且也是 K 的元素：

$$x + y \in K \implies \lambda x + \mu y \in K.$$

因此，通过加法和缩放的性质，我们得到：

$$\lambda x + (1 - \lambda)y = \lambda x + \mu y \in K.$$

得分	

四. 验证以下函数是否为凸函数（给出分析过程）。

（1） $f(x,y)=x^2y^2+\frac{x}{y},\ x>0,y>0$

（2） $f(x,y)=\log\left(e^x+e^y\right)-\log x,\ x>0$

(3) $f(x,y)=\exp\left\{x^2+e^{-y}\right\},\ x>0,y>0$

(4) $f(x,y)=xy\log xy, x>0,y>0$

(5) $f(x,y)=-\log(cx+dy),\ c,d\in\mathbb{R}$

总一小题 4 分，结果 1 分，分析过程 3 分。

(1) $f(x,y)=x^2y^2+\frac{x}{y},\ x>0,\ y>0$

分析：计算 Hessian 矩阵：

$$H=\begin{bmatrix}2y^2 & 4xy-\frac{1}{y^2}\\ 4xy-\frac{1}{y^2} & 2x^2+\frac{2x}{y^3}\end{bmatrix}$$

取点 $(1,1)$，Hessian 行列式为 $\det(H)=(2)(2)-(4-1)^2=4-9=-5<0$，故矩阵不定。

答案：$\boxed{\text{非凸}}$

(2) $f(x,y)=\log\left(e^x+e^y\right)-\log x,\ x>0$

分析：分解为 $\log(e^x+e^y)$（凸）和 $-\log x$（凸）。Hessian 矩阵为：

$$H=\begin{bmatrix}\frac{e^xe^y}{(e^x+e^y)^2}+\frac{1}{x^2} & -\frac{e^xe^y}{(e^x+e^y)^2}\\ -\frac{e^xe^y}{(e^x+e^y)^2} & \frac{e^xe^y}{(e^x+e^y)^2}\end{bmatrix}$$

对任意向量 $v=(v_1,v_2)$，验证 $v^THv=\frac{1}{x^2}v_1^2+\frac{e^xe^y}{(e^x+e^y)^2}(v_1-v_2)^2\geq 0$。

答案：$\boxed{\text{凸}}$

(3) $f(x,y)=\exp\left(x^2+e^{-y}\right),\ x>0,\ y>0$

分析：令 $g(x,y)=x^2+e^{-y}$（凸），则 $f=e^g$。计算 Hessian：

$$H=e^g\begin{bmatrix}4x^2+2 & -2xe^{-y}\\ -2xe^{-y} & e^{-2y}+e^{-y}\end{bmatrix}$$

主对角线元素 $4x^2+2>0$，$e^{-2y}+e^{-y}>0$，且 $\det(H)>0$，故正定。

答案：$\boxed{\text{凸}}$

(4) $f(x,y)=xy\log(xy),\ x>0,\ y>0$

分析：计算 Hessian 矩阵：

$$H = \begin{bmatrix} \frac{y}{x} & \log(xy)+2 \\ \log(xy)+2 & \frac{x}{y} \end{bmatrix}$$

行列式 $\det(H) = \frac{y}{x} \cdot \frac{x}{y} - (\log(xy)+2)^2 = 1 - (\log(xy)+2)^2$。当 $xy = e^{-1}$ 时，$\det(H) = 1 - 1^2 = 0$；当 $xy = e^1$ 时，$\det(H) = 1 - 9 = -8 < 0$。

答案：$\boxed{\text{非凸}}$

(5) $f(x, y) = -\log(cx + dy),\ c, d \in \mathbb{R}$

分析：$\log(cx + dy)$ 在定义域 $cx + dy > 0$ 上为凹函数（Hessian 矩阵为负半定），其负数 $-\log(cx + dy)$ 为凸函数。

答案：$\boxed{\text{凸}}$

得分	

五. 证明函数 $f:\mathrm{R}^n\to\mathrm{R}$ 是凸函数，其中 $f(x)=\sum_{i=1}^{r}|x|_{[i]}$, 其定义域为 R^n, 其中 $|x|$ 表示向量, 其分量 $|x|_i=|x_i|$ (即 $|x|$ 的每个分量是向量 x 对应分量的绝对值)，$|x|_{[i]}$ 表示向量 $|x|$ 中第 i 大的分量。换言之, $|x|_{[1]},|x|_{[2]},\cdots,|x|_{[n]}$ 是向量 x 的分量的绝对值按照非增的顺序进行排列。总共 10 分，根据步骤的完整程度酌情给分。

证明:

设 $a_i=|x|_i$ 和 $b_i=|y|_i$，其中 a_i 和 b_i 分别是向量 x 和 y 的绝对值按非增顺序排列的分量。由于绝对值的性质，$|x_i|$ 和 $|y_i|$ 是非负的。

设 $z=\lambda x+(1-\lambda)y$，我们需要分析 $f(z)$ 的形式。首先，计算 $|z|$ 的分量:

$$|z|_i=|\lambda x+(1-\lambda)y|_i.$$

根据绝对值的性质，排序后的绝对值的和可以用以下不等式来估计:

$$|z|_i\le\lambda|x|_{[k]}+(1-\lambda)|y|_{[k]}$$

对于每一个 i。这个不等式基于三角不等式和绝对值的性质。

由于 $|z|_{[i]}$ 是 $|z|$ 的排序分量，我们有:

$$f(z)=\sum_{i=1}^{r}|z|_{[i]}\le\sum_{i=1}^{r}\left(\lambda|x|_{[i]}+(1-\lambda)|y|_{[i]}\right)=\lambda\sum_{i=1}^{r}|x|_{[i]}+(1-\lambda)\sum_{i=1}^{r}|y|_{[i]}=\lambda f(x)+(1-\lambda)f(y).$$

因此，对于任意 $x,y\in\mathbb{R}^n$ 和 $\lambda\in[0,1]$，我们有

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y),$$

从而证明了函数 f 是凸函数。

得分	

六. （1）验证点 $(2,4)$ 在如下问题中满足最优性条件：

$$\begin{aligned} \min \quad & (x_1-4)^2+(x_2-6)^2 \\ \text{s.t.} \quad & x_1^2 \le x_2, \\ & x_2 \le 4. \end{aligned}$$

（2）考虑同一点，其中第二个不等式约束变为 $x_2 \le 5$。

总共 10 分，第一小题 6 分，第二小题 4 分，结果错分别减 1 分，过程步骤视情况酌情给分。

(1) **证明**：

- **可行性验证**：$2^2=4\le 4$ 且 $4\le 4$，满足约束
- **拉格朗日函数**：

$$L=(x_1-4)^2+(x_2-6)^2+\lambda_1(x_1^2-x_2)+\lambda_2(x_2-4)$$

- **梯度条件**：

$$\begin{cases} \frac{\partial L}{\partial x_1}=2(x_1-4)+2\lambda_1 x_1=0 \\ \frac{\partial L}{\partial x_2}=2(x_2-6)-\lambda_1+\lambda_2=0 \end{cases}$$

代入 $(2,4)$ 得：

$$\begin{cases} -4+4\lambda_1=0 \Rightarrow \lambda_1=1 \\ -4-\lambda_1+\lambda_2=0 \Rightarrow \lambda_2=5 \end{cases}$$

- **非负性**：$\lambda_1=1\ge 0$，$\lambda_2=5\ge 0$
- **互补松弛**：$\lambda_1(x_1^2-x_2)=0$，$\lambda_2(x_2-4)=0$

结论：$\boxed{\text{满足最优性条件}}$

(2) **证明**：

- **可行性验证**：$2^2=4\le 4$ 且 $4\le 5$，仍可行
- **梯度条件**（此时 $\lambda_2=0$）：

$$\begin{cases} -4+4\lambda_1=0 \Rightarrow \lambda_1=1 \\ -4-\lambda_1=0 \Rightarrow \lambda_1=-4 \quad \text{（矛盾）} \end{cases}$$

- **非负性破坏**：λ_1 无法同时满足 $\lambda_1=1$ 和 $\lambda_1=-4$

结论: $\boxed{\text{不满足最优性条件}}$

得分	

七. 给定线性规划问题

$$\begin{array}{ll} \min & 5x_1 + 21x_3 \\ \text{s. t.} & x_1 - x_2 + 6x_3 \geqslant b_1, \\ & x_1 + x_2 + 2x_3 \geqslant 1, \\ & x_1, x_2, x_3 \geqslant 0, \end{array}$$

其中 b_1 是某一个正数，已知这个问题的一个最优解为 $(x_1, x_2, x_3) = \left(\frac{1}{2}, 0, \frac{1}{4}\right)$.

(1) 给出该问题的对偶问题。

(2) 求出该对偶问题的最优解。

总共 10 分，每小题 5 分，根据步骤的完整程度酌情给分。

证：

(1) 根据对偶规则，对偶问题为：

$$\begin{array}{ll} \max & b_1 y_1 + y_2 \\ \text{s.t.} & y_1 + y_2 \leqslant 5 \\ & -y_1 + y_2 \leqslant 0 \\ & 6y_1 + 2y_2 \leqslant 21 \\ & y_1, y_2 \geqslant 0 \end{array}$$

(2) 利用互补松弛条件：

原问题变量非零对应的对偶约束取等：

$$\begin{cases} y_1 + y_2 = 5 \quad (\because x_1 > 0) \\ 6y_1 + 2y_2 = 21 \quad (\because x_3 > 0) \end{cases}$$

解方程组：

$$\begin{aligned} &\text{由第一式得：} y_2 = 5 - y_1 \\ &\text{代入第二式：} 6y_1 + 2(5 - y_1) = 21 \\ &\Rightarrow 4y_1 = 11 \Rightarrow y_1 = \frac{11}{4} \\ &\Rightarrow y_2 = 5 - \frac{11}{4} = \frac{9}{4} \end{aligned}$$

验证约束：

$$-y_1 + y_2 = -\frac{11}{4} + \frac{9}{4} = -\frac{1}{2} \leqslant 0 \quad (\text{满足})$$

强对偶定理验证：原问题最优值：

$$5 \times \frac{1}{2} + 21 \times \frac{1}{4} = \frac{31}{4}$$

对偶问题目标值：

$$b_1 \times \frac{11}{4} + \frac{9}{4} = \frac{31}{4} \Rightarrow b_1 = 2$$

答案：对偶问题最优解为

$$\boxed{(y_1, y_2) = \left(\frac{11}{4}, \frac{9}{4}\right)}$$

得分	

八. 求解以下优化问题

$$\begin{array}{ll} \min_{x,y\in\mathbb{R}^2} & (x-1)^2+(y-2)^2 \\ \text{s.t.} & x \le y, \\ & x+2y \le 2 \end{array}$$

总共 10 分，根据步骤的完整程度酌情给分。

证明：目标函数表示点 (x,y) 到 $(1,2)$ 的欧氏距离平方。需在可行域内找到最近点。无约束最小值点 $(1,2)$ 不满足 $x+2y\le 2$（因 $1+4=5>2$），故最优解必在约束边界。

- 在约束 $x+2y=2$ 上使用拉格朗日乘数法：

  $$\begin{cases} \frac{\partial L}{\partial x} = 2(x-1)+\lambda = 0 \\ \frac{\partial L}{\partial y} = 2(y-2)+2\lambda = 0 \\ x+2y=2 \end{cases}$$

  解得：

  $$\lambda = \frac{6}{5}, \quad x = \frac{2}{5}, \quad y = \frac{4}{5}$$

  验证 $x \le y$：$\frac{2}{5} \le \frac{4}{5}$ 成立。

- 在约束 $x=y$ 上，极值点为 $(\frac{2}{3},\frac{2}{3})$，但目标值 $\frac{17}{9}\approx 1.888$ 大于 $\frac{9}{5}=1.8$

目标值比较：

$$f\left(\frac{2}{5},\frac{4}{5}\right) = \left(\frac{2}{5}-1\right)^2 + \left(\frac{4}{5}-2\right)^2 = \frac{9}{5}$$
$$f\left(\frac{2}{3},\frac{2}{3}\right) = \frac{17}{9}$$

答案：最优解为

$$\boxed{\left(\frac{2}{5},\frac{4}{5}\right)}$$

得分	

九. 点 $x_0 \in \mathrm{R}^n$ 在多面体 $\mathcal{P} = \{x \mid Ax \le b\}$ 上的投影（在 ℓ_∞-范数下），可以被定义为下面优化问题的解：

$$\begin{array}{ll} \text{minimize} & \|x - x_0\|_\infty \\ \text{subject to} & Ax \le b \end{array}$$

其中变量为 $x \in \mathrm{R}^n$。我们假设 $\mathcal{P}$ 是非空的。

(a) 将此问题表述为线性规划（LP）形式。

(b) 推导其对偶问题，并尽可能简化。

(c) 证明如果 $x_0 \notin \mathcal{P}$，则可以从对偶问题的最优解构造出一个超平面将 x_0 和 $\mathcal{P}$ 分离。总共 12 分，前两小题每题 2 分，最后一题 8 分。最后一题 4 个步骤，每个步骤 2 分。酌情给分。

解： (a) 将此问题表述为线性规划（LP）形式。

首先，定义 $\|x - x_0\|_\infty = \max_i |x_i - (x_0)_i|$。为了将此问题转换为线性规划形式，可以引入新的辅助变量 t，表示最大绝对误差，得到以下形式的线性规划：

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & -t \le x_i - (x_0)_i \le t, \quad i = 1, 2, \ldots, n \\ & Ax \le b \end{array}$$

其中，变量为 $x \in \mathbb{R}^n$ 和 $t \in \mathbb{R}$。

(b)

构造对偶问题时，我们首先写出拉格朗日对偶函数。拉格朗日乘子对应于不等式约束 $Ax \le b$ 和 $-t \le x_i - (x_0)_i \le t$。

拉格朗日函数为：

$$L(x, t, \lambda, \mu) = t + \sum_{i=1}^{n} \mu_i (x_i - (x_0)_i) + \sum_{i=1}^{n} \mu_i (-x_i + (x_0)_i) + \lambda^\top (Ax - b)$$

然后求对偶问题：

$$\text{maximize} \quad \text{minimize}_x t + \lambda^\top (Ax - b)$$

$$
\begin{aligned}
&\max\left(\lambda_3-\lambda_2\right)^{\top} x_0-\lambda_1^{\top} b \\
&\text{sit. } 1+1^{\top} \lambda_3=1^{\top} \lambda_2 \\
&A^{\top} \lambda_1 \leqslant \lambda_3 \\
&\lambda_1 \geqslant 0 \\
&\lambda_3 \geqslant 0
\end{aligned} \tag{1}
$$

（c）如果该对偶问题可行且存在最优解，则

$$
\lambda_1^{\top} A X_0-\lambda_1^{\top} b=\left(\lambda_1^*\right)^{\top}\left(A x_0-b\right)
$$

时 $x_0 \notin P$ 则

$$
A x_0>b. \qquad \lambda_1^* \geqslant 0
$$

则

$$
\lambda_1^{*T} A x_0-\lambda_1^* b \geqslant 0.
$$

而对于 $\forall x \in P$，有

$$
\begin{aligned}
&Ax-b \leqslant 0. \\
&\lambda_1^*(Ax-b) \leqslant 0 \leqslant \quad \lambda_1^{*T} A x_0-\lambda_1^* b.
\end{aligned}
$$

则超平面

$$
\left\{x \mid \lambda_1^{*T} A x-\lambda_1^{*T} b\right\} \text{ 分离} x_0 \text{ 和} P\text{。}
$$