南京大学智能科学与技术学院期末试卷

学年学期 23-24（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	十	总分
满分	8	12	8	8	8	14	12	9	8	13	100

得分	

一、(1) 假设 $\|\cdot\|$ 是一个定义在 $\mathbb{R}^m$ 上的范数，定义 $\|x\|_A=\|Ax\|$，$A\in\mathbb{R}^{m\times n}$。证明如果 $\operatorname{rank}(A)=n$, 则 $\|\cdot\|_A$ 也是一个范数。(4 分)

(2) 定义函数 $f:\mathbb{C}^n\to\mathbb{R}, f(x)=\sum_{i=1}^n(|\operatorname{Re}x_i|+|\operatorname{Im}x_i|)$，$f$ 是范数吗？试证明。(4 分)

证明 (1): $\|x\|_A=\|Ax\|\geqslant 0$ (非负性)

因为 $\operatorname{rank}(A)=n$，所以 $\|x\|_A=\|Ax\|=0$ 当且仅当 $x=0$

$\|tx\|_A=\|tAx\|=|t|\cdot\|Ax\|=|t|\|x\|_A, t\in R$（齐次性）

$$\|x+y\|_A=\|Ax+Ay\|\leq\|Ax\|+\|Ay\|=\|x\|_A+\|y\|_A \quad (x\in R^n, y\in R^n)$$

满足三角不等式。

综上，$\|\cdot\|_A$ 是一个范数

（2）f 是范数; 证明: $f(x)=\sum_{i=1}^n(|\operatorname{Re}x_i|+|\operatorname{Im}x_i|)\geqslant 0$ （非负性）

$$f(tx)=\sum_{i=1}^n(|\operatorname{Re}tx_i|+|\operatorname{Im}tx_i|)=|t|\sum_{i=1}^n(|\operatorname{Re}x_i|+|\operatorname{Im}x_i|)=|t|f(x),\forall t\in R$$

满足齐次性。

$$f(x+y)=\sum_{i=1}^n(|\operatorname{Re}(x_i+y_i)|+|\operatorname{Im}(x_i+y_i)|)\leqslant\sum_{i=1}^n(|\operatorname{Re}x_i|+|\operatorname{Im}x_i|+|\operatorname{Re}y_i|+|I_my_i|)$$

$$=f(x)+f(y),\forall x,y\in C^n$$

满足三角不等式。综上，f 是范数。

得分	

二. (1) 将闭凸集 $\left\{x \in \mathbb{R}_+^2 \mid x_1 x_2 \geqslant 1\right\}$ 表示为半空间的交集。(4 分)

(2) 设 $K \subseteq \mathbb{R}^n, K \neq \emptyset$。证明 K 是一个闭凸集当且仅当它是包含 K 的所有闭半空间的交集。(8 分)

解: (1) 对任何满足 $mn=1$ 的 $(m,n),(m,n>0)$ 有 $x_2 - n \geqslant -\frac{1}{m^2}(x_1 - m)$。

即 $-m^2 x_2 - x_1 \leqslant -m^2 n - m$

则原集合可表示为

$$\bigcap_{\substack{m\cap\in R^2\\ mn=1}} \left\{(x_1,x_2) \mid \left(-1,-m^2\right)^\top x \leqslant \left(-m^2 n - m\right)\right\}$$

即半空间交集。

证: (2) 一方面, 若 K 是包含 K 的所有闭半空间的交集。由交集是保凸运算, 故 K 凸。又 K 是半空间交集, 知 K 闭, 故 K 闭凸集。

另一方面, 若 K 是一个闭凸集, 去证

$$K = \bigcap_{\substack{K \subseteq A\\ A\text{ 是闭半空间}}} A$$

任取 $x \in K$, 由 $K \subseteq A$ 知 $x \in \bigcap_{\substack{K \subseteq A\\ A\text{ 是闭半空间}}} A$。

任取 $x \in \bigcap_{\substack{K \subseteq A\\ A\text{ 是闭半空间}}} A$, 若 $x \notin K$ 可知存在一个半空间 A_0 包含 K, 但 $x \notin A_0$ 与 $x \in \cap A$ 即对任意 $A_0\,(K \subseteq A_0, A_0$ 是闭半空间) $x \in A_0$ 矛盾。

则 $K = \bigcap_{\substack{K \subseteq A\\ A\text{ 是闭半空间}}} A$。

综合两方面得证。

得分	

三. 假设存在一个超平面将两个凸集合 C_1 和 C_2 严格分离（其中 C_2 为一个锥）。证明存在一个过原点的超平面，其构成的一个半空间包含了 C_2，但与 C_1 没有交集。(8 分)

证明: $\exists a \neq 0, \quad a^\top x = b$

$$\forall x \in C_1, \quad a^\top x < b$$

$$\forall x \in C_2, a^\top x > b$$

$\because C_2$ 为凸锥, 过原点

$$\therefore b < 0$$

$\therefore a^\top x = b$ 不过原点

将 $a^\top x = b$ 平移，使其经过原点，得到新的超平面 $a^T x = 0$

显然，$\forall x \in C_1$ $a^\top x < 0$，$\forall x \in C_2, a^T x \geqslant 0$，$\therefore a^\top x \geqslant 0$

这个半空间包含了 C_2 但与 C_1 没有交集

$\therefore$ 存在一个过原点的超平面，其一个半空间包含 C_2 与 C_1 无交集

得分	

四. 一个集合 C 被叫做 "mid-point convex"，如果对于所有的 $x, y \in C$，我们有 $\frac{1}{2}(x+y) \in C$。证明一个集合 C 如果是 "mid-point convex" 并且是个闭集合，那么 C 是凸集合。(8 分)

证明: 若 $\forall x, y \in C$, $\frac{1}{2}(x+y) \in C$ 且 C 是闭的，证 C 为凸集。

$\because \forall x, y \in C$, $\frac{1}{2}(x+y) \in C$ 即 $x + \frac{1}{2}(y-x) \in C$

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

下面进行一步步调整: 此时 $x_0 = x, y_0 = y, \quad \theta_s = 0, \theta_t = 1, \theta_m = \frac{1}{2}$

(1) $\theta_m = \frac{\theta_t + \theta_s}{2}$

若 $\theta > \theta_m$ 则 $x_{n+1} = \frac{x_n + y_n}{2}, \quad \theta'_m = \frac{3\theta_t}{4} + \frac{1}{4}\theta_s, \quad \theta_s = \theta_m$

若 $\theta \leq \theta_m$ 则 $y_{n+1} = \frac{x_n + y_n}{2}, \quad \theta'_m = \frac{3}{4}\theta_s + \frac{1}{4}\theta_t, \theta_t = \theta_m$

(2) 若 $\theta'_m \neq \theta$，重复 (1)，直到 $\theta'_m = \theta$

(3) 依次类推，无限逼近 θ (夹逼定理) $\theta - \left(\frac{1}{4}\right)^n \leq \theta \leq \theta + \left(\frac{1}{4}\right)^n$

$\therefore$ 总能找到 θ，使 $x + \theta(y-x) \in C$

得分	

五. 设 C 和 D 为 $\mathbb{R}^n$ 的不相交的子集。考虑集合 $(a,b)\in\mathbb{R}^{n+1}$，它满足对任意 $x\in C$ 有 $a^Tx\le b$；对任意 $x\in D$ 有 $a^Tx\ge b$。证明这个集合是一个凸锥（并且如果没有分离 C 和 D 的超平面，那么它是单点集 $\{0\}$）。(8 分)

证明: 对 $\forall x\in C$ 有 $a^\top x\le b$ ，对 $\forall x\in D$ 有 $a^\top x\ge b$ 设该集合为 H

1. 若存在分离 C 和 D 的超平面

(1) 有两个恰好与 C 和 D 相切的超平面，$a_1^\top x=b_1, a_2^\top x=b_2$

则对 $\forall x\in C$，$\forall\theta_1,\theta_2\ge 0,\quad \theta_1a_1^\top x+\theta_2a_2^\top x\le\theta_1b_1+\theta_2b_2$

对 $\forall x\in D,\theta_1a_1^\top x+\theta_2a_2^\top x\ge\theta_1b_1+\theta_2b_2$

$\therefore\theta_1(a_1,b_1)+\theta_2(a_2,b_2)\in H$

$\therefore H$ 是凸锥。

(2) 只有一个分离 C 和 D 的超平面, $a^\top x=b$

则 $\forall\theta\ge 0,\forall x\in C,\theta a^\top x\le\theta b,\quad\forall x\in D,\quad\theta a^\top x=\theta b$

$\therefore$ 也是凸锥。

2. 无分离心与 C 和 D 的超平面

显然单点集 $\{0\}$ 是唯一解。

$\therefore$ 得证。

得分	

六. 设函数 f: $\mathbb{R} \to \mathbb{R}$ 是凸函数。试证明:

(a)(单调性质)利用凸函数的定义证明，对于三变量 $x_1 < x_2 < x_3$, 函数 f 满足如下公式:

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

(4 分)

(b) 利用上述结果证明当 x 增长到 $+\infty$ 时，$f(x)$ 有如下四种情况:（1）$f(x)$ 单调递减直到 $-\infty$,（2）$f(x)$ 单调递减至某个常数,（3）$f(x)$ 随着 x 增长到达某个值并保持不变,（4）$f(x)$ 单调递增至 $+\infty$。(10 分)

(a) 证明:

由一阶条件

$$f\left(x_3\right) \geq f\left(x_2\right)+\nabla f\left(x_2\right)\left(x_3-x_2\right)$$

$$f\left(x_1\right) \geq f\left(x_2\right)+\nabla f\left(x_2\right)\left(x_1-x_2\right).$$

$$\frac{f\left(x_2\right)-f\left(x_1\right)}{x_2-x_1} \leq \nabla f\left(x_2\right).$$

$$\frac{f\left(x_3\right)-f\left(x_2\right)}{x_3-x_2} \geq \nabla f\left(x_2\right) \geq \frac{f\left(x_2\right)-f\left(x_1\right)}{x_2-x_1}$$

所以得证

(b) 证明:

(1) 若 $\frac{f(x_2)-f(x_1)}{x_2-x_1} = \frac{f(x_3)-f(x_1)}{x_3-x_2} = k\,(\forall x_1 < x_2 < x_3)$

则 $f(x)$ 为线性函数 ; 若斜率 $k < 0$, 则 $\lim_{x\to+\infty} f(x) = -\infty$

(2) 若 $\frac{f(x_2)-f(x_1)}{x_2-x_1} = k < 0$ 且 $x \to +\infty$ 时 $k = 0$

则 $f(x)$ 几乎不变

又 $f(x)$ 在 R 上单调递减，所以 $f(x)$ 单调递减至某个常数。

(3) 当 $k > 0$ 且 $x \to +\infty$ 时 $k = 0$

即为第 (3) 种情况。

(4) $\frac{f(x_2)-f(x_1)}{x_2-x_1} = k$ 为常数且 $k > 0$，则 $\lim_{x\to+\infty} f(x) = +\infty$

得分	

七. 判断凸函数。

（1）验证 $\sqrt{e^x + e^{-y}}$ 是否为凸函数；（6 分）

（2）验证几何平均数是凹函数:

$$f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}, x \in (0, \infty)^n.$$

提示：使用柯西-施瓦茨不等式 $\left(a^T a\right)\left(b^T b\right) \geq \left(a^T b\right)^2$。（6 分）

证：（1）Hessian 矩阵为

$$\frac{1}{2}\left(e^x + e^y\right)^{-\frac{3}{2}} \begin{pmatrix} e^x\left(\frac{e^x}{2} + e^{-y}\right) & \frac{1}{2} e^{x-y} \\ \frac{1}{2} e^{x-y} & e^{-y}\left(e^x + \frac{e^{-y}}{2}\right) \end{pmatrix}$$

注意 $e^x\left(\frac{e^x}{2} + e^{-y}\right) > 0$

$$e^{x-y}\left(\frac{e^x}{2} + e^{-y}\right)\left(e^x + \frac{e^{-y}}{2}\right) - \frac{1}{4} e^{2x-2y} = \frac{1}{4} e^{2x-2y}\left(2e^{x+y} + 5 + 2e^{-x-y}\right) > 0.$$

则为正定矩阵, 是凸函数。

(2)

$$\begin{aligned} \frac{\partial f(x)}{\partial x_i} &= \frac{1}{n}\left(x_1 \cdots x_n\right)^{\frac{1}{n}-1} \cdot \frac{x_1 x_2 \cdots x_n}{x_i} = \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{n x_i} \\ \frac{\partial^2 f(x)}{\partial^2 x_i} &= \frac{1-n}{n^2} \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{x_i^2} \\ \frac{\partial^2 f(x)}{\partial x_i \partial x_j} &= \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{n^2 x_i x_j} \end{aligned}$$

则

$$\begin{aligned} v^{\top} f(x) v &= \sum_{i=1}^{n} \frac{1-n}{n^2} \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{x_i^2} v_i^2 + 2 \sum_{1 \leqslant i<j \leqslant n} \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{n^2 x_i x_j} v_i v_j \\ &= \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{n^2}\left(\sum_{1 \leqslant i<j \leqslant n}\left(\frac{v_i^2}{x_i^2} + 2 \frac{v_i v_j}{x_i x_j}\right) - \sum n \frac{v_i^2}{x_i^2}\right) \\ &= \frac{\left(x_1 \cdots x_n\right)^{\frac{1}{n}}}{n^2}\left(\left(\sum_{i=1}^{n} \frac{v_i}{x_i}\right)^2 - \sum_{i=1}^{n} n \frac{v_i^2}{x_i^2}\right) (\text{ 完全平方}) \\ &\leqslant 0 (\text{ 柯西}) \end{aligned}$$

故凹函数。

得分	

八. 使用 KKT 条件找到集合

$$M=\left\{x\in\mathbb{R}^2: x_1+x_2\geq 4, 2x_1+x_2\geq 5\right\}$$

中最接近点 $(0,0)$ 的点。

可以存在多个点（解）吗?（9 分）

解：我们使用目标中的欧几里得距离来构建一个非线性规划问题：

$$\begin{aligned} &\min x_1^2+x_2^2 \\ &\text{s.t.}\ -x_1-x_2+4\leq 0, \\ &\qquad -2x_1-x_2+5\leq 0. \end{aligned}$$

写出拉格朗日函数

$$L\left(x_1,x_2,u_1,u_2\right)=x_1^2+x_2^2+u_1\left(-x_1-x_2+4\right)+u_2\left(-2x_1-x_2+5\right), u_1,u_2\geq 0.$$

推导 KKT 条件 i) 可行性，

ii)

$$u_1\left(-x_1-x_2+4\right)=0, u_1\geq 0$$

$$u_2\left(-2x_1-x_2+5\right)=0, u_2\geq 0$$

iii) $\frac{\partial L}{\partial x_1}=2x_1-u_1-2u_2=0,$

$$\frac{\partial L}{\partial x_2}=2x_2-u_1-u_2=0$$

现在，我们将通过分析最优条件来尝试找到 KKT 点，按照互补条件进行：

1. 设 $u_1=0,u_2=0$ ： 从 iii) 我们得到 $x_1=0,x_2=0$，这不是一个可行点。

2. 设 $x_1+x_2=4,u_2=0$ ： 结合 iii) 我们解

$$\begin{aligned} &2x_1-u_1=0, \\ &2x_2-u_1=0, \\ &x_1+x_2=4, \end{aligned}$$

我们得到 $x_1=2,x_2=2,u_1=4>0$，即我们有 KKT 点 $(2,2,4,0)$。

3. 设置 $u_1=0,2x_1+x_2=5$：求解

$$\begin{aligned} &2x_1-2u_2=0, \\ &2x_2-u_2=0, \\ &2x_1+x_2=5 \end{aligned}$$

我们得到 $x_1 = 2, x_2 = 1, u_2 = 2$，这不是一个可行点。

4. 设置 $x_1 + x_2 = 4, 2x_1 + x_2 = 5$：我们得到 $x_1 = 1, x_2 = 3$ 并通过解算来计算拉格朗日乘数

$$\begin{aligned} u_1 + 2u_2 &= 2, \\ u_1 + u_2 &= 6 \end{aligned}$$

我们得到 $u_1 = 10, u_2 = -4 < 0$，即拉格朗日乘数不是非负的，$(1, 3, 10, -4)$ 不是 KKT 点。

由于集合 M 是凸的，对应于投影 $(2, 2)$ 的最近点必须是唯一的。

得分	

九. 计算最小化分段线性函数（Piecewise-linear minimization）

$$\text{minimize} f(x)=\max_{i=1,\dots,m}\left(a_i^T x+b_i\right)$$

的线性规划（LP formulation）与对偶线性规划（dual LP）。(8 分)

解: 该问题可转化为线性规划形式

$\min t$

s.t. $a_i^\top x+b_i \leqslant t \quad i=1,\cdots,m.$

$$\begin{aligned}
L(t,x,\lambda) &= t+\sum_{i=1}^m \lambda_i\left(a_i^\top x+b_i-t\right)\\
&= t-\sum_{i=1}^m \lambda_i t+\left(\sum_{i=1}^m \lambda_i a_i^\top\right)x+\sum_{i=1}^m b_i^\top \lambda_i\\
&= \left(1-\sum_{i=1}^m \lambda_i\right)t+\left(\sum_{i=1}^m \lambda_i a_i^\top\right)x+\sum_{i=1}^m b_i^\top \lambda_i
\end{aligned}$$

则

$$\begin{aligned}
& g(\lambda)=\inf_{t,x}\left(1-\sum_{i=1}^m \lambda_i\right)t+\left(\sum_{i=1}^m \lambda_i a_i^\top\right)x+\sum_{i=1}^m b_i^\top \lambda_i\\
& =\sum_{i=1}^m b_i^\top \lambda_i+\inf_t(1-\sum_{i=1}^m x_i)t+\inf_x\left(\sum_{i=1}^m \lambda_i a_i^\top\right)x\\
& \text{若}\sum_{i=1}^m \lambda_i a_i^\top \neq 0, \text{则} g(\lambda)=-\infty\\
& \text{若} 1-\sum_{i=1}^m \lambda_i \neq 0, \text{则} g(\lambda)=-\infty
\end{aligned}$$

则其对偶线性规划为

$\max \sum_{i=1}^n b_i{}^\top \lambda$

s.t. $\sum_{i=1}^m \lambda_i a_i=0$

$\sum_{i=1}^m \lambda_i=1 \quad i=1,\cdots,m.$

$\lambda_i \geq 0$

得分	

十. 已知下面是一个强对偶性成立的非凸优化问题。分析最优解（充分或必要）条件，以及如何寻找其最优解。可以存在多个最优解吗？

$$\begin{array}{ll} \text{minimize}_x & -2(x_1-2)^2 - x_2^2 \\ \text{subject to} & x_1^2 + x_2^2 \leq 25 \\ & x_1 \geq 0 \end{array}$$

（13 分）

解：首先注意到约束集是凸的，并且 $(1,1)$ 是一个 Slater 点，确保了各处的资格条件。拉格朗日函数表示为

$$\mathcal{L}(x_1, x_2, \mu_1, \mu_2) = -2(x_1-2)^2 - x_2^2 + \mu_1\left(x_1^2 + x_2^2 - 25\right) - \mu_2 x_1$$

KKT 条件表示为

$$\begin{cases} -4(x_1-2) + 2\mu_1 x_1 - \mu_2 = 0 \\ -2x_2 + 2\mu_1 x_2 = 0 \\ x_1^2 + x_2^2 \leq 25 \\ x_1 \geq 0 \\ \mu_1, \mu_2 \geq 0 \\ \mu_1 = 0 \quad \text{或} \quad x_1^2 + x_2^2 = 25 \\ \mu_2 = 0 \quad \text{或} \quad x_1 = 0 \end{cases}$$

如果 $\mu_1 = \mu_2 = 0$，我们有 $x_1 = 2$ 和 $x_2 = 0$，满足原始约束。因此 $\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ 是一个满足 KKT 条件的原始-对偶点，相关值为 0。如果 $\mu_1 = 0$ 且 $\mu_2 > 0$，我们有 $x_1 = x_2 = 0$ 且 $\mu_2 = 8 > 0$，这是一个原始-对偶点，值为 -8。如果 $\mu_2 = 0$ 且 $\mu_1 > 0$，我们有

$$\begin{cases} -4(x_1-2) + 2\mu_1 x_1 = 0 \\ -2x_2 + 2\mu_1 x_2 = 0 \\ x_1 \geq 0 \\ \mu_1 > 0 \\ x_1^2 + x_2^2 = 25 \end{cases}$$

因此，$x_2 = 0$ 或 $\mu_1 = 1$。在第一种情况下，我们得到 $x_1 = 5, x_2 = 0$，因此 $\mu_1 = 6/5 > 0$ 且 $\mu_2 = 0$，这是一个 KKT 点，值为 -18。在第二种情况下，我们得到 $x_1 = 4$ 和 $x_2 = \pm 3$，$\mu_1 = 1$ 和 $\mu_2 = 0$，这是两个 KKT 点，值为 -17。最后，如果 $\mu_2 > 0$ 且 $\mu_1 > 0$，我们有 $x_1 = 0$ 和 $x_2 = \pm 5$，$\mu_1 = 1$ 和 $\mu_2 = 8$，这是两个 KKT 点，值为 -33，因此是全局最小值。