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 (B 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										

得分	

一. 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:

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

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。(8 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 $\alpha_1, \ldots, \alpha_m \ge 0$ 满足 $\sum_{i=1}^m \alpha_i = 1$。因此，由 $c'x_i \ge \inf_{x\in X} c'x$ 可得

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

关于 $\bar{x} \in conv(X)$ 取左侧的下确界为

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

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 $c'x_i = \inf\limits_{x\in conv(X)} c'x$ 时，等式成立。因此，集合 X 的 $c'x$ 的下确界可以到达。

得分	

二. 给出两个不相交的闭凸集不能被严格分离的例子。(8 分)

例如，二维空间中集合 $\{(x,y)|y \geq 1/x\}$ 和 $\{(x,y)|y \leq 0\}$，相应的例子合理即可。

得分	

三. 令 $C = \{x \in X : \|x\| \leq 1\}$ 为空间中的单位球。证明 x 到 C 的投影可以表示为

$$\pi_C(x) = \frac{x}{\max\{1, \|x\|\}}.$$

(10 分)

证明：首先证明上述结论只需说明对于所有 C 中的点满足 $< x - \pi_C(x), z - x > \leq 0$。即

$$\left(x - \frac{x}{\max\{1, \|x\|\}}\right)(z - x) \leq 0$$

上式左边等于 $\left(\frac{x\max\{1,\|x\|\}-x}{\max\{1,\|x\|\}}\right)(z-x)$。因此只需证明 $ax(z-x)\leq 0$ 对所有的 $z\in C$ 都成立，其中 $a\geq 0$。根据范数的性质 $xz\leq\|x\|\|z\|\leq\|x\|^2=x^2$。得证。

得分	

四. 设函数 f:$\mathbb{R}^n\to\mathbb{R}$ 是凸函数，其定义域为 $\mathbb{R}^n$。函数在 $\mathbb{R}^n$ 上有上界。证明该函数是常数。(10 分)

证明：对于任意点 x。如果 $f'(x)\neq 0$，不妨设 $f'(x)>0$。

根据凸函数的性质，我们有 $f'(z)>f'(x)$ 对所有的 $z>x$ 都成立。

假设 $f'(x)=\delta>0$。那么我们有

$$f(z)-f(x)\geq\delta(z-x)$$

因此当 $z\to\infty$，$f(z)$ 趋向于正无穷。此时，$f(x)$ 不存在上界。因此 $f'(x)<0$ 不成立。同样，$f'(x)>0$ 也不成立。因此，只能有 $f'(x)=0$。得证。

得分	

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

（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：

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

（8 分）

证明:

由一阶条件

$$f(x_3) \ge f(x_2) + \nabla f(x_2)(x_3 - x_2)$$

$$f(x_1) \ge f(x_2) + \nabla f(x_2)(x_1 - x_2).$$

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

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

所以得证

得分	

六. 验证以下函数是否为凸函数。

（1）$f(x,y) = xy$ （8 分）

（2）$f(x) = \sum_{i=1}^{r} x_{[i]}$，$x_{[i]}$ 表示 x 中第 i 大的分量。（8 分）

（3）$\sqrt{e^x + e^{-y}}$；（8 分）

（1）容易得到其 Hessian 矩阵为 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$，因此该函数不是凸函数。

（2）$f(x)$ 可以重写为

$$f(x) = \max\{x_{i_1} + \cdots + x_{i_r} | 1 \le i_1 < i_2 \cdots \le n\}$$

显然，$f(x)$ 是凸函数。

（3）Hessian 矩阵为

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

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。

得分	

七. 定义最大值函数为 $f(x)=\max_{i=1,\ldots,n} x_i$，定义在 $\mathbb{R}^n$ 上。推导其共轭函数。(10 分)

解：我们下面将证明其共轭函数为

$$f^*(y)=\begin{cases} 0 & y\succeq 0, 1^T y=1 \\ \infty & 其他 \end{cases}$$

我们先验证共轭函数的定义域。假设 y 包含一个负元素，比如说 $y_k<0$。如果我们选择一个向量 x 使得 $x_k=-t$，$x_i=0$，$i\neq k$。当 t 趋于正无穷时，我们有

$$x^T y-\max_i x_i=-ty_k\rightarrow\infty,$$

因此 y 不在共轭函数的定义域中。下面，我们假设 $y\succeq 0$ 但是 $1^T y>1$。我们选择 $x=t1$ 并且令 t 趋于正无穷。因此我们有下面公式不存在一个上界。

$$x^T y-\max_i x_i=t1^T y-t$$

同样，对于 $y\succeq 0$ 以及 $1^T y<1$ 的情况，我们可以选择 $x=-t1$ 并且令 t 趋于正无穷。

剩下的情况则是当 $y\succeq 0$ 且 $1^T y=1$。在这种情况下，我们有

$$x^T y\leq\max_i x_i$$

因此我们有 $x^T y - \max_i x_i \leq 0$ 对所有的 x 都成立。因此，我们证明了上述共轭函数的成立。

得分	

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

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

的线性规划（LP formulation）与对偶线性规划（dual LP）。(10 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	

九. 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?

$$\begin{array}{ll} \text{maximize}_x & -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) \\ \text{subject to} & x_1^2 + x_2^2 + x_3^2 = 1 \end{array}$$

（12 分）

解：其 KKT 条件可以写为：

$$x_1^2 + x_2^2 + x_3^2 = 1$$
$$(-3\nu)x_1 + 1 = 0$$
$$(1 + \nu)x_2 + 1 = 0$$
$$(2 + \nu)x_3 + 1 = 0$$

通过观察上述的 KKT 条件容易得到 $\nu \neq 2$ $\nu \neq -1$，以及 $\nu \neq 3$。我们因此可以消去 x，得到下面关于 ν 的等式。

$$\frac{1}{(-3+\nu)^2} + \frac{1}{(1+\nu)^2} + \frac{1}{(2+\nu)^2} = 1 \tag{1}$$

上述公式包含四个解。分别是 -3.15，0.22，1.89，4.04。对应的 x 的解为

$$\begin{aligned} x &= (0.16, 0.47, -0.87), \\ x &= (0.36, -0.82, 0.45), \\ x &= (0.90, -0.35, 0.26), \\ x &= (-0.97, -0.20, 0.17) \end{aligned} \tag{2}$$

将上述解代入 $f_0(x)$ 可得四个值，分别为 1.17，0.67，-0.56，-4.70。比较大小即可得到其最优解。