第二章（上）

2.1

根据范数函数的性质（三角不等式和齐次性）：

$$f(\Theta x+(1-\Theta)y)\leqslant f(\Theta x)+f((1-\Theta)y)=\Theta f(x)+(1-\Theta)f(y)$$

满足定义一，得证。
零范数不是凸函数，考虑在一维情况下：

$$\begin{gathered}
x_1=(1)\quad x_2=(0)\\
f(\frac{1}{2}*1+\frac{1}{2}*0)=1\\
\frac{1}{2}f(1)+\frac{1}{2}f(0)=\frac{1}{2}
\end{gathered}$$

显然不满足凸函数定义。

2.2

$$\begin{gathered}
\|\Theta x+(1-\Theta)y-x_c\|\\
=\|\Theta x+(1-\Theta)y-\Theta x_c-(1-\Theta)x_c\|\\
=\|\Theta(x-x_c)+(1-\Theta)(y-x_c)\|\\
\leqslant\|\Theta(x-x_c\|+\|(1-\Theta)(y-x_c)\|\\
=\Theta\,\|(x-x_c\|+(1-\Theta)\,\|(y-x_c)\|\\
\leqslant\Theta r+(1-\Theta)r\\
=r
\end{gathered}$$

满足凸集的定义，得证。

2.3

定义一推定义二：

$$\begin{gathered}
\square\, g(\Theta t_1+(1-\Theta)t_2)\\
=f(x+[\Theta t_1+(1-\Theta)t_2]v)\\
=f(x+\Theta t_1v+(1-\Theta)t_2v)\\
=f(\Theta x+(1-\Theta)x+\Theta t_1v+(1-\Theta)t_2v)\\
=f(\Theta(x+t_1v)+(1-\Theta)(x+t_2v))\\
\leqslant\Theta f((x+t_1v))+(1-\Theta)f(x+t_2v)\\
=\Theta g(t_1)+(1-\Theta)g(t_2)
\end{gathered}$$

定义二推定义一：

$$\begin{gathered}
\text{令}\,y=x+tv,\\
\text{可证明}\,f(\Theta y_1+(1-\Theta)y_2)\leqslant\Theta f(y_1)+(1-\Theta)f(y_2)
\end{gathered}$$

2.4

(1)

$$\triangledown f(x)=(a^Tx)\frac{\partial(b^Tx)}{\partial x}+\frac{\partial(a^Tx)}{\partial x}(b^Tx)=a^Txb+ab^Tx$$

(2)

$$\frac{\partial^2 f}{\partial x \partial x^T} = \frac{\partial(a^T x b + a b^T x)}{\partial x^T} = \frac{\partial(a^T x b)}{\partial x^T} + \frac{\partial(b^T x a)}{\partial x^T} = a^T b + b^T a$$

2.5

如果X随机变量是满足离散概率分布，则有：

$$f(\Theta_1 x_1 + \Theta_2 x_2 + \ldots + \Theta_n x_n) \leqslant \Theta_1 f(x_1) + \Theta_2 f(x_2) + \ldots + \Theta_n f(x_n), \quad \sum_{i=1}^{n} \Theta_i = 1$$
$$E(x) = \Theta_1 x_1 + \Theta_2 x_2 + \ldots + \Theta_n x_n$$
$$E(f(x)) = \Theta_1 f(x_1) + \Theta_2 f(x_2) + \ldots + \Theta_n f(x_n)$$
$$\text{因此，} \quad f(E(x)) \leqslant E(f(x))$$

如果x随机变量是满足连续概率分布，则可以通过中值定理将E（x）化为离散概率分布的形式，后面过程如上。

第二章（下）

2.6

(1)

$$a^T(\Theta x + (1-\Theta) y) = \Theta a^T x + (1-\Theta) a^T y \geqslant \Theta b + (1-\Theta) b = b$$

满足凸集定义，得证。

(2)构造拉格朗日乘数

$$L(x, \lambda) = x^T x - \lambda(a^T x - b)$$

由KTT定理可得：

$$2x - \lambda a = 0$$
$$\lambda(a^T x - b) = 0$$
$$\lambda \geqslant 0$$
$$\text{若} \lambda = 0\text{，则约束条件不起作用，显然不符合题意；因此} (a^T x - b) = 0$$
$$\text{即} a^T x^* = b$$

(3)

由(2)得：

$$x = \frac{\lambda a}{2}$$
$$\text{带入} (2) \text{式，得：}$$
$$\lambda = \frac{2b}{a^T a}$$
$$\text{则} x = \frac{ab}{a^T a}$$

2.7

首先构造广义拉格朗日乘子函数：

$$L(x_1, x_2, x_3, \lambda, \theta) = x_1^2 - 4x_1 x_2 + x_2^2 + \lambda(x_1 + x_2 + x_3 - 1) + \theta(x_1 - x_2 + x_3) \quad \theta \leq 0. \tag{1}$$

对x_1 求偏导数，有

$$2x_1 - 4x_2 + \lambda + \theta = 0 \qquad (2)$$

对x_2求偏导数，有：

$$-4x_1 + 2x_2 + \lambda - \theta = 0 \qquad (3)$$

对x_3求偏导数有：

$$\lambda + \theta = 0 \qquad (4)$$
$$\lambda = -\theta$$

由(1)和(2)可以解得：

$$x_2 = \frac{\lambda}{2} + \frac{\theta}{6}$$
$$x_1 = \frac{\lambda}{2} - \frac{\theta}{6} \qquad (5)$$

将(4)和（5）带入到（1）得：

$$L(\theta) = \frac{1}{3}\theta^2 + \theta$$

所以，此时的对偶问题为：

$$\max \frac{1}{3}\theta^2 + \theta$$
$$s.t. \quad \theta \le 0$$

2.8

答案见课堂讲解s.t.~~~