Homework:

1. Convex set.

(1). a polyhedron $\{x\in R: Ax\le b\}$ for some $A\in R^{m\times n}$, $b\in R^m$ is convex.

证明如下:

Hyperplanes: 设 $x_1, x_2 \in n$. $a^T x_1 = b$, $a^T x_2 = b$.

$\because a^T(\theta x_1 + (1-\theta)x_2) = \theta a^T x_1 + (1-\theta) a^T x_2 = b$

$\therefore \theta x_1 + (1-\theta) x_2 \in n$.

Halfspaces: 设 $x_1, x_2 \in n$. $a^T x_1 \le b$. $a^T x_2 \le b$

$\therefore a^T(\theta x_1 + (1-\theta)x_2) = \theta a^T x_1 + (1-\theta) a^T x_2 \le b$

$\therefore \theta x_1 + (1-\theta) x_2 \in n$.

$\because$ 凸集的交集也为凸集. 故. 证得.

(2). $f: R^n \to R$. $\{(x,t) | f(x)\le t, x\in R^n, t\in R\}$ is convex.

证明如下: $\because f(x)$是凸函数.

$\therefore f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$

设 x, y 有 $f(x)\le t$, $f(y)\le t$

$\theta f(x) + (1-\theta) f(y) \le t$.

$\therefore f(\theta x + (1-\theta) y) \le t$.

$\therefore \{(x,t) | f(x)\le t, x\in R^n, t\in R\}$是凸集 证得.

(3). Show that $\{x\in R^n_+ | \prod_{i=1}^n x_i \ge 1\}$ is convex.

证明: 设有 $x_1, x_2, \cdots, x_n \in R^n$, 使 $\prod_{i=1}^n x_i \ge 1$

$y_1, y_2, \cdots, y_n \in R_+^n$. 使 $\prod_{i=1}^n y_i \ge 1$.

$\theta \prod_{i=1}^n x_i + (1-\theta) \prod_{i=1}^n y_i \ge \prod_{i=1}^n x_i^\theta \prod_{i=1}^n y_i^{1-\theta} \ge 1$.

故得证

(4) show that set $\{x | \|x-a\|_2 \le \theta \|x-b\|_2\}$, where $a\ne b$ and $0\le\theta\le 1$. is convex.

证明如下: $(x-a)^T(x-a) \le \theta (x-b)^T (x-b)$.

$x^T x - a^T x - a x + a^T a \le \theta [x^T x - b^T x - b x + b^T b]$

$(1-\theta) x^T x - 2(a^T - \theta b^T) x + a^T a - \theta b^T b \le 0$

$\begin{cases} \theta=1 \text{时. 为仿射集} \\ 0\le\theta\le 1 \text{时. 也为凸集.} \end{cases}$

2. Convex function.

(1). Prove that entropy function, defined as $f(x) = -\sum_{i=1}^{n} x_i \log(x_i)$ with $\operatorname{dom} f = \{x \in R_{++}^n : \sum_{i=1}^{n} x_i = 1\}$, is strictly concave.

证明如下: $y = \sum_{i=1}^{n} x_i \log(x_i) = x^T \log x$

$\nabla y = \log x + 1.$ $\quad \nabla^2 y = \operatorname{diag}\{\frac{1}{x_1}, \frac{1}{x_2}, \cdots \frac{1}{x_n}\} > 0$

$\therefore f = -y$ is strictly concave.

(2). Show that $f(x_1, x_2) = \frac{1}{x_1 x_2}$ on R_{++}^2 is convex.

证明如下: $f = \frac{1}{x_1 x_2}.$ $\quad \nabla f = \left[-\frac{1}{x_1^2 x_2}, -\frac{1}{x_1 x_2^2}\right].$

$$\nabla^2 f = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix}$$

$\because x_1, x_2 > 0 \quad \therefore |\nabla^2 f| = \frac{3}{x_1^4 x_2^4} > 0. \quad \therefore f(x)$ is strictly convex

(3). Show that $f(x) = tr(X^{-1})$ is convex on $\operatorname{dom} f = S_{++}^n$.

证明如下: if $g(t) = tr\{X^{-1}(I + tQ\Sigma Q^T)^{-1}\} = tr\{X^{-1}Q(I+t\Sigma)^{-1}Q^T\}$

$= tr[Q^T X^{-1} Q (I + t\Sigma)^{-1}] = \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \frac{1}{1+t\lambda_i}$

$\therefore g''(t) = \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \frac{2\lambda_i^2}{(1+t\lambda_i)^3}.$

$(1+t\lambda_i) > 0.$ $\quad Q^T X^{-1} Q$ is PD. $\Rightarrow g''(t) > 0.$ is always true

$\therefore g(t)$ is convex in $t.$

$\therefore f(x) = tr(X^{-1})$ is convex.

3. Dual problem.

(1). $\min c^T x$ s.t. $f(x) \le 0$ $\xRightarrow{\text{dual}}$ $\max_{\lambda} \min_{x} \{\lambda f(x) + c^T x\}$ s.t. $\lambda \ge 0$

12). min $c^T x$.

s.t. $Gx \le h$.

$Ax = b$

$\xRightarrow{\text{dual}}$ max $hy_1 + by_2$

s.t. $Gy_1 + Ay_2 = c^T$.

$y_1 \le 0$.

4. KKT condition:

min $x_1^2 + x_2^2$.

s.t. $(x_1-1)^2 + (x_2-1)^2 \le 1$

$(x_1-1)^2 + (x_2+1)^2 \le 1$.

$$L(x_1, x_2, \lambda_1, \lambda_2) = x_1^2 + x_2^2 + \lambda_1\left[(x_1-1)^2 + (x_2-1)^2\right] + \lambda_2\left[(x_1-1)^2 + (x_2+1)^2 - 1\right]$$

$$g_1(x) = (x_1-1)^2 + (x_2-1)^2 - 1 \le 0$$

$$g_2(x) = (x_1-1)^2 + (x_2+1)^2 - 1 \le 0$$

KKT condition: $g_1(x) \le 0$, $g_2(x) \le 0$, $\lambda_1, \lambda_2 \ge 0$, $\lambda_1 g_1(x) = 0$, $\lambda_2 g_2(x) = 0$

$$\nabla f_0(x) + \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) = 0$$

$\because f_0(x) = x_1^2 + x_2^2$.

$\therefore \nabla f_0(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$ $\quad \nabla g_1(x) = \begin{bmatrix} 2(x_1-1) \\ 2(x_2-1) \end{bmatrix}$

$\nabla g_2(x) = \begin{bmatrix} 2(x_1-1) \\ 2(x_2+1) \end{bmatrix}$

$$\therefore \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2(x_1-1) \\ 2(x_2-1) \end{bmatrix} + \lambda_2 \begin{bmatrix} 2(x_1-1) \\ 2(x_2+1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(2). $\min \|Ax-b\|_2^2$

s.t. $Gx=h$.

$\frac{\partial L}{\partial x}\Big|_{x^*} = 2A^T(Ax^*-b)+G^Tv=0$

$Gx^*-h=0$. 解得.

$v^* = 2(GG^T)^{-1}GA^T(b-Ax^*)$.