ECE 602 A3

Meng Tianao 20810553

Ebore Thomas 20868189

Exercise 1

(a) Max function $f(x) = \max_{i=1 \cdots n} x_i$ on R^n

we assume that when $i=t$, we get the max value

$$f^*(y) = \sup_{x \in R^n} (y^T x - \max_{i=1 \cdots n} x_i)$$
$$= \sup_{x \in R^n} (y^T x - x_t)$$
$$= \sup_{x \in R^n} (\sum_{i=1}^{n} y_i x_i - x_t)$$

when $y_i < 0$, because $x \in R^n$, we can choose one of x_i very negative, while other components of x to be zero. Therefore, when $y_i < 0$, $y^T x - f(x)$ is unbounded.

when $y \geq 0$ and $\sum_{i=1}^n y_i > 1$, we can easily choose $x \in R^n$, which makes $y^T - f(x)$ ∞

similar for $y \geq 0$ and $\sum_{i=1}^n y_i < 1$ case.

Therefore, we should choose $y \geq 0$ and $\sum_{i=1}^n y_i = 1$, which makes the supremum is at most zero.

$\therefore f^*(y)$ is the indicator function of the set $S = \{y \geq 0 \mid \sum_{i=1}^n y_i = 1\}$

(b) Sum of largest elements: $f(x) = \sum_{i=1}^{r} x_{[i]}$ on R^n

This function returns r largest value of $x \in R^n$.

$$f^*(y) = \sup_{x \in R^n} \left(y^T x - \sum_{i=1}^{r} x_{[i]} \right)$$

$$= \sup_{x \in R^n} \left(\sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} \right)$$

when $y_i < 0$, because $x \in R^n$, we can choose one of x_i very negative, while other components of x to be zero. Therefore, when $y_i < 0$, $y^T x - f(x)$ is unbounded.

when $y \geq 0$ and $\sum_{i=1}^{n} y_i > r$, we can easily choose $x \in R^n$, which makes $y^T - f(x) = \infty$

similar for $y \geq 0$ and $\sum_{i=1}^{n} y_i < r$ case.

when $y \geq 0$, $\sum_{i=1}^{n} y_i = r$, but $\sum_{i=1}^{r} y_{[i]} \neq r$ (the sum of coefficient of largest value of x), which also can makes $y^T - f(x) = \infty$

Therefore, we should choose $y \geq 0$ and $\sum_{i=1}^{n} y_i = r$, and $\sum_{i=1}^{r} y_{[i]} = r$ which makes the supremum is at most zero.

$\therefore$ $f^*(y)$ is the indicator function of the set $S = \{ y \geq 0 \mid \sum_{i=1}^{n} y_i = r$ and $\sum_{i=1}^{r} y_{[i]} = r \}$

(1) p - norm: $f(x) = \|x\|_p$ on $\mathbb{R}^n$

$$f^*(y) = \begin{cases} 0 & \|y\|_q \le 1 \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1 \\ \infty, & \text{otherwise} \end{cases}$$

Exercise 2.

The Lagrangian is

$$L(x, z, v) = \sum_{k=1}^{n} x_k \log\left(\frac{x_k}{y_k}\right) - z(1^T x - 1) - v^T(Ax - b)$$

The optimal x of the Lagrangian has k-th component given by

$$x_k \frac{y_k}{x_k} + \log\left(\frac{x_k}{y_k}\right) - z - a_k^T v = 0.$$

$$\therefore\ x_k^* = y_k \times e^{z + a_k^T v - y_k} = \frac{1}{F} y_k e^{a_k^T v}$$

where F is a constant in terms of z that allows x to sum to one.

$$\sum_{k=1}^{n} x_k^* = \frac{1}{F} \sum_{k=1}^{n} y_k e^{a_k^T v} = 1 \Rightarrow F = \sum_{k=1}^{n} y_k e^{a_k^T v}$$

$$\therefore\ g(v) = L(x^*, z, v) = \sum_{k=1}^{n} \frac{1}{F} y_k e^{a_k^T v} \log \frac{e^{a_k^T v}}{F} - v^T\left(A \begin{bmatrix} \frac{1}{F} y_1 e^{a_1^T v} \\ \vdots \\ \frac{1}{F} y_n e^{a_n^T v} \end{bmatrix} - b\right)$$

$$= \frac{1}{F}\left(\sum_{k=1}^{n} y_k e^{a_k^T v} \times (a_k^T v - \log F)\right) - A^T v x^* + b^T v$$

$$= \frac{1}{F}\left(\sum_{k=1}^{n} y_k e^{a_k^T v} \times a_k^T v - \sum_{k=1}^{n} y_k e^{a_k^T v} \log F\right) - A^T v x^* + b^T v$$

$$= \frac{1}{F}\left(F \times A^T v - F \log F\right) - A^T v x^* + b^T v$$

$$= A^T v - \log F - A^T v x^* + b^T v$$

$$= b^T v - \log F = b^T v - \log \sum_{k=1}^{n} y_k e^{a_k^T v}$$

Exercise 3

we assume $ru = v$, $\|v\|_\infty \le r$, $r \in R$

$a_i(x_c + v) \le b_i$, $i = 1, 2, \dots, m$ holds for any v

$$\therefore \sup_{v: \|v\|_\infty \le r} \left(a_i^T (x_c + v) \right) = a_i^T x_c + \sup_{v: \|v\|_\infty \le r} a_i^T v \le b_i$$

Since $\sup\{a_i^T v \mid \|v\|_\infty \le r\} = \|a_i\|_1 r$,

As a result, we arrive an LP of the form.

$$\max_{x_c, r} \ r \quad \text{subject to} \ \begin{bmatrix} a_i \\ \|a_i\|_1 \end{bmatrix}^T \begin{bmatrix} x_c \\ r \end{bmatrix} \le b_i, \ i = 1, 2, \dots, m$$

$$\max \ [0, 1] \begin{bmatrix} x_c \\ r \end{bmatrix} \quad \text{subject to} \ \begin{bmatrix} a_i \\ \|a_i\|_1 \end{bmatrix}^T \begin{bmatrix} x_c \\ r \end{bmatrix} \le b_i, \ i = 1, \dots, m$$

Therefore, it can still be formulated as a linear program

Exercise 4. $\min \|Ax-b\|_2^2$ s.t. $Cx=h$

$$f_0(x) = \|Ax-b\|_2^2 = (Ax-b)^T(Ax-b) = (x^TA^T-b^T)(Ax-b)$$
$$= x^TA^TAx - x^TA^Tb - b^TAx + b^Tb$$

KKT conditions: ① primal feasibility

$$Cx^* = h \quad ①$$

② Dual feasibility (however there is no λ^* in this problem)

$$\lambda^* \geq 0$$

③ First-order optimality

$$\nabla f_0(x^*) = 2A^TAx^* - 2A^Tb$$

$$\therefore 2A^TAx^* - 2A^Tb + C^Tv^* = 0 \quad ②$$

④ Complementary slackness (however there is no λ^* in this problem)

$$\lambda^{*T}(Cx^* - h) = 0$$

Solving x^* and v^* results in

$$\begin{bmatrix} C & 0 \\ 2A^TA & C^T \end{bmatrix}\begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} h \\ 2A^Tb \end{bmatrix}$$

$\because C$ is wide, $x^* = C^T(CC^T)^{-1}h$

$\because C^T$ is thin, $(C^T)^{\dagger} = (CC^T)^{-1}C \Rightarrow v^* = (CC^T)^{-1}C\left(2A^Tb - 2A^TAC^T(CC^T)^{-1}h\right)$