Convex Optimisation
Home work 2

Exercise 1

1) (P) :

$$\min_{x \in \mathbb{R}^n} c^T x \quad Ax = b,\ x \geq 0$$

(P) <=>

$$\min c^T x \quad \text{s.t.}\quad Ax - b = 0,\ -x \leq 0$$

Therefore the Lagrangian associated is

$$\mathcal{L}_p(x, \lambda, \nu) = c^T x - \lambda^T x + \nu^T (Ax - b)$$
$$= (c - \lambda + A^T \nu)^T x - \nu^T b$$

And the dual function is:

$$g_p(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} \left((c - \lambda + A^T \nu)^T x \right) - \nu^T b$$

$$= \begin{cases} -\infty & \text{if } c - \lambda + A^T \nu \neq 0 \\ -\nu^T b & \text{otherwise.} \end{cases}$$

Therefore the dual problem of (P) is

$$\max_{\lambda, \nu,\ \lambda \geq 0} \left(g_p(\lambda, \nu) \right) \iff \max \left(-\nu^T b \right) \quad \text{s.t. } c - \lambda + A^T \nu = 0,\ \lambda \geq 0$$

$$\iff \max_{\lambda, \nu} \left(-b^T \nu \right) \quad c + A^T \nu = \lambda,\ \lambda \geq 0$$

$$\iff \max_{\nu} \left(-b^T \nu \right) \quad c + A^T \nu \geq 0 \overset{(y = -\nu)}{\iff} \max_{y} \left(b^T y \right) \quad A^T y \leq c \iff (D)$$

2) (D) $\iff$ $\max_y (b^T y)$ s.t. $A^T y \leq c$

$\iff$ $\min (-b^T y)$ s.t. $A^T y - c \leq 0$

$$\mathcal{L}_d(y, \lambda) = -b^T y + \lambda^T (A^T y - c)$$

$$= (-b + A\lambda)^T y - \lambda^T c$$

$$\Rightarrow g_d(\lambda) = \inf_y \mathcal{L}_d(y, \lambda)$$

$$= \inf_y (-b + A\lambda)^T y - \lambda^T c$$

$$= \begin{cases} -\infty & \text{if } A\lambda \neq b \\ -\lambda^T c & \text{otherwise} \end{cases}$$

The dual problem of D is therefore

$$\max_\lambda (-\lambda^T c) \quad \lambda \geq 0, \quad A\lambda - b = 0 \quad \overset{\lambda = x}{\iff} \quad \min (c^T x) \quad x \geq 0, \quad Ax = b$$

$\iff$ (P).

3) Let the following problem:

(P$_1$): $\min_{x,y} c^T x - b^T y$ s.t. $x \geq 0$, $Ax = b$, $A^T y \leq c$ $\iff$ $\min_{x,y} c^T x - b y$ s.t. $-x \leq 0$, $A^T y - c \leq 0$, $Ax - b = 0$

$$\Rightarrow \mathcal{L}_{P_1}(x, y, \lambda_1, \lambda_2, \nu) = (c - \lambda_1 + A^T \nu)^T x + (A\lambda_2 - b)^T y - \lambda_2^T c - b^T \nu$$

$$\Rightarrow g_{P_1}(\lambda_1, \lambda_2, \nu) = \begin{cases} -\infty & \text{if } c + A^T \nu \neq \lambda_1 \text{ or } A\lambda_2 \neq b \\ -\lambda_2^T c - b^T \nu & \text{otherwise} \end{cases}$$

The dual problem of P_1 is:

$$\max_{\lambda_1,\lambda_2,\nu} \left(-\lambda_2^T c - b^T \nu\right)$$
$$\text{st} \quad \lambda_1 \geq 0 \,,\quad \lambda_2 \geq 0$$
$$c + A^T\nu = \lambda_1$$
$$A\lambda_2 = b$$

$$\iff \max_{\lambda_2,\nu} \left(-\lambda_2^T c - b^T \nu\right)$$
$$\text{s.t} \quad \lambda_2 \geq 0 \,,\quad c + A^T\nu \geq 0$$
$$A\lambda_2 = b$$

$$\overset{\substack{x=\lambda_2\\ y=-\nu}}{\iff} \min_{x,y} \left(c^T x - b^T y\right)$$
$$\text{st} \quad x \geq 0 \quad c \geq A^T y$$
$$Ax = b$$

$$\iff (P_1).$$

4) $(P_1) \iff \min_{x\in K_x,\, y\in K_y} \left(c^T x - b^T y\right) \quad , K_x = \{x \mid x \geq 0,\, Ax = b\}$
$K_y = \{y \mid Ay \leq c\}$

$$\iff \min_{x\in K_x}(c^T x) - \max_{y\in K_y}(b^T y)$$

$$\iff (P) - (D).$$

~~Then~~ if p_1^* is the solution of (P_1), p^*, d^* solutions of $(P), (D)$.

Then $\underline{p_1^* = p^* - d^*}$.

As (P) is a linear problem the strong duality holds.
and as (D) is the dual of (P), $p^* = d^* \Rightarrow \underline{p_1^* = 0}$

Exercise 2

1) $f : x \mapsto \|x\|_1 \Rightarrow f^*(y) = \sup_x \left(y^T x - \|x\|_1 \right)$

$$= \sup_x \left(\sum_i y_i x_i - |x_i| \right)$$

if $\exists i$ st $|y_i| > 1$, then with $x_t = y_i t e_i$, $t > 0$

we have: $\sum_j x_j y_j - |x_j| = x_i y_i - |x_i|$

$$= y_i^2 t - |y_i| t$$

$$= \underbrace{\left(|y_i|^2 - |y_i| \right)}_{>0 \text{ as } |y_i| > 1} t$$

Therefore $y^T x_t - \|x_t\|_1 \xrightarrow[t \to +\infty]{} +\infty$

$f^*(y) = +\infty$.

if $\forall i$, $|y_i| \leq 1$. then either $y^T x < 0$ and $y^T x - \|x\|_1 < 0$ too

or $y^T x \geq 0 \Rightarrow y^T x = |y^T x| \leq \sum \underbrace{|y_i|}_{\leq 1} |x_i|$

$$\leq \sum |x_i|$$

$$\leq \|x\|_1$$

$\Rightarrow y^T x \leq \|x\|_1$

$y^T x - \|x\|_1 \leq 0$.

as for $x = 0$ $y^T x - \|x\|_1 = 0$

$f^*(y) = 0$.

Finally: $f^*(y) = \begin{cases} \cancel{+\infty}\ 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

Convex Optimization

Homework 2

Exercise 2.

2) (P) : $\min_{x} \|Ax - b\|_2^2 + \|x\|_1$

$$(P) \iff \min_{x,y} \|y\|_2^2 + \|x\|_1 \quad y = Ax - b \qquad \iff \min_{x,y} \|y\|_2^2 + \|x\|_1 \quad y - Ax + b = 0$$

The Langrangian is:

$$\mathcal{L}(x, y, \nu) = \|y\|_2^2 + \|x\|_1 + \nu^T(y - Ax + b)$$
$$= \nu^T y + y^T y + \|x\|_1 - \nu^T A x + \nu^T b.$$

$$g(\nu) = \inf_{x,y} (\mathcal{L}(x, y, \nu))$$
$$= \inf_{y} (\nu^T y + y^T y) + \inf_{x} (\|x\|_1 - \nu^T A x) + \nu^T b$$
$$= \inf_{y} (\nu^T y + y^T y) - f^*(\nu^T A) + \nu^T b$$

Let $h_\nu(y) = \nu^T y + y^T y$

$\nabla h_\nu (y) = \nu + 2y$. $\nabla h_\nu = 0 \iff y = -\frac{1}{2}\nu$

As h_ν is convex ($\|y\|_2^2$ is convex, $\nu^T y$ is convex, so is the sum)

$$\inf_{y} (h_\nu(y)) = h_\nu\left(-\frac{1}{2}\nu\right)$$
$$= -\frac{1}{4}\nu^T \nu$$

$$\Rightarrow g(\nu) = -\frac{1}{4}\nu^T\nu + \nu^T b - f^*(\nu^T A)$$

The dual is $\max_{\nu} \left(g(\nu) \right)$

$$\iff \max_{\nu} \left(-\frac{1}{4} \|\nu\|_2^2 + \nu^T b \right)$$
$$\text{s.t. } \|A^T \nu\|_\infty \leq 1$$

Exercise 3

1) S_1 : $\min_{w} \frac{1}{n} \sum_i \ell(w, x_i, y_i) + \frac{\Sigma}{2} \|w\|_2^2$

$$S_1 \iff \min_{w} \frac{1}{2n} \sum_i \ell(w, x_i, y_i) + \frac{1}{2} \|w\|_2^2$$

$$\iff \min_{z,w} \frac{1}{2n} \sum_i z_i + \frac{1}{2} \|w\|_2^2$$
$$z_i = \ell(w, x_i, y_i) = \max(0, 1 - y_i(w^T x_i))$$

$$\iff \min_{z,w} \frac{1}{2n} 1_n^T z + \frac{1}{2} \|w\|_2^2 \qquad \boxed{(S_1')}$$
$$z_i = \max(0, 1 - y_i(w^T x_i))$$
$$z_i \geq 0$$
$$z_i \geq 1 - y_i(w^T x_i)$$

~~One can observe than~~ $\cancel{S_1'}$ ~~is more constraint than~~ $\cancel{S_2}$

~~(it has one more constraint:~~ $\cancel{z_i = \max(0, 1 - y_i(w^T x_i))}$ ~~)~~

~~Therefore if~~ $\cancel{s_1^*}$ ~~is solution of~~ $\cancel{S_1'}$~~, and~~ $\cancel{s_2^*}$ ~~solution of~~ $\cancel{S_2}$

$\cancel{s_2^* \leq s_1^*}$

Now assume that s_2^* (z^*, w^*) which minimize S_2 is such that $\exists i$

$$z_i^* \neq \max(0, 1 - y_i(w^{*T} x_i))$$

as s_2^* is solution of S_2 it minimizes $\frac{1}{2n} \sum_i z_i + \frac{1}{2} \|w\|_2^2$

with $z_j^* \geq 0$ and $z_j^* \geq 1 - y_j(w^{*T} x_j)$

Therefore $\forall j \quad z^*_j \geq \max(0, 1-y_j(w^{*T}x_j))$

$$\Rightarrow z^*_i > \max(0, 1-y_i(w^{*T}x_i))$$

But then with z' defined as :

$$z'_j = \begin{cases} z^*_j & \text{if } j \neq i \\ \max(0, 1-y_i(w^{*T}x_i)) & \text{if } j = i \end{cases}$$

Then clearly (w^*, z') verifies $z'_j \geq 0$, $z'_j \geq 1-y_j(w^T x_j)$

and we have $\frac{1}{n\tau} 1_n^T z' + \frac{1}{2}\|w^*\|_2^2 < \frac{1}{n\tau} 1_n^T z^* + \frac{1}{2}\|w^*\|_2^2$

which is impossible by definition of (z^*, w^*)

Therefore z^*, w^* verifies : $z^*_j = \max(0, 1-y_j(w^T x_j))$

$\Rightarrow$ $\underline{S_2 \Leftrightarrow S_1' \Leftrightarrow S_1.}$

2) $\mathcal{L}(w, z, \lambda, \pi) = \frac{1}{\tau n}\sum_i z_i + \frac{1}{2}\sum w_i^2 - \pi^T z + \sum_i \lambda_i(1-y_i(w^T x_i) - z_i)$

$$= \sum_i \left(\frac{1}{\tau n} - \pi_i - \lambda_i\right) z_i + \sum_i \lambda_i + \|w\|_2^2 + \sum_i \lambda_i(-y_i(w^T x_i))$$

$$= \left(\frac{1}{\tau n} 1_n - \pi - \lambda\right)^T z + \sum \lambda_i + \frac{1}{2}\|w\|_2^2 - \sum_i \lambda_i y_i x_i^T w$$

$\Rightarrow g(\lambda, \pi) = \inf_{w,z} \mathcal{L}(w, z, \lambda, \pi)$

$$= \inf_z \left(\left(\frac{1}{\tau n} - \pi - \lambda\right)^T z\right) + \frac{1}{2}\inf_w \underbrace{\left(\|w\|_2^2 - 2(\lambda \circ y)^T X w\right)}_{\text{same form as in 2.2} \Rightarrow w = X^T(\lambda \circ y)} + \sum \lambda_i$$

$$= \begin{cases} -\infty & \text{if } \frac{1}{\tau n} - \pi - \lambda \neq 0 \\ \sum \lambda_i - \underbrace{(\lambda \circ y)^T X X^T (\lambda \circ y)}_{\sum_j \left(\sum_i \lambda_i y_i x_{ij}\right)^2} & \text{otherwise} \end{cases}$$

$$g(\lambda,\pi) = \inf_{w,z} \left(\mathcal{L}(w,z,\lambda,\pi) \right)$$

$$= \inf_z \left(\left(\frac{1}{\varepsilon n} - \pi - \lambda\right)^T z \right) + \frac{1}{2} \underbrace{\inf_w \left(\|w\|_2^2 - 2\left(\Sigma \lambda_i y_i x_i\right)^T w \right)}_{\substack{\text{same as in } 2.2 \\ \text{minimal for } w = \Sigma \lambda_i y_i x_i}} + \Sigma \lambda_i$$

$$= \begin{cases} -\infty & \text{if } \pi + \lambda \neq \frac{1}{\varepsilon n} 1_n \\ 1_n^T \lambda - \frac{1}{2} \left(\Sigma \lambda_i y_i x_i\right)^T \left(\Sigma \lambda_i y_i x_i\right) \end{cases}$$

The dual problem is:

$$\max_{\pi,\lambda} \left(1_n^T \lambda - \frac{1}{2} \|\Sigma \lambda_i y_i x_i\|_2^2 \right)$$
$$\pi + \lambda = \frac{1}{\varepsilon n} 1_n$$
$$\pi \geq 0$$
$$\lambda \geq 0$$

$$\Longleftrightarrow \quad \max_{\lambda} \left(1_n^T \lambda - \frac{1}{2} \|\Sigma \lambda_i y_i x_i\|_2^2 \right) \quad \lambda \geq 0 \;,\; \lambda \leq \frac{1}{\varepsilon n} 1_n$$