Homework 2

Megi Derrishi

Exercise 1 (LP Duality)

1) The Lagrangian of problem P is:

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

which is an affine function of x. It follows that the dual function is:

$$g(\lambda, v) = \begin{cases} -b^T v & \text{if } c + A^T v - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem is to maximize $g(\lambda, v)$ subject to $\lambda \geq 0$ and after making the implicit constraints explicit we obtain:

$$\begin{array}{ll} \max\limits_{v} & -b^T v \\ \text{s.t} & A^T v + c \geq 0 \end{array} \qquad \cdots\cdots\cdots (1)$$

2) The problem (D) is equivalent to: $\begin{array}{ll} \min\limits_{y} & -b^T y \\ \text{s.t} & A^T y \leq c \end{array}$

The Lagrangian is hence:

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

which is affine in y; It follows the dual function is:

$$g(\lambda) = \begin{cases} -c^T \lambda & \text{if } A\lambda - b = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Then the dual problem is to maximize $g(\lambda)$ subject to $\lambda \geq 0$:

$$\begin{array}{ll} \max\limits_{\lambda} & -c^T \lambda \\ \text{s.t.} & A\lambda = b \\ & \lambda \geq 0 \end{array} \qquad \cdots\cdots\cdots (2)$$

3) The Lagrangian is:

$$\mathcal{L}(x,y,\lambda_1,\lambda_2,v) = c^T x - b^T y + v^T(b - Ax) - \lambda_1^T x + \lambda_2^T (A^T y - c)$$
$$= b^T v - c^T \lambda_2 + (c - A^T v - \lambda_1)^T x + (A\lambda_2 - b)^T y$$

which is affine in x and y, the dual function is hence:

$$g(\lambda_1, \lambda_2, v) = \begin{cases} b^T v - c^T \lambda_2 & \text{if } c - A^T v - \lambda_1 = 0 \text{ and } A\lambda_2 - b = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Hence, the dual problem after making the implicit constraints explicit we obtain:

$$\begin{array}{ll} \max\limits_{\lambda_1,\lambda_2,v} & -c^T\lambda_2 + b^T v \\ \text{s.t.} & c - A^T v - \lambda_1 = 0 \\ & A\lambda_2 = b \\ & \lambda_2 \geq 0 \\ & \lambda_1 \geq 0 \end{array} \iff \begin{array}{ll} \max\limits_{\lambda_2,v} & -c^T\lambda_2 + b^T v \\ \text{s.t.} & c \geq A^T v \\ & A\lambda_2 = b \\ & \lambda_2 \geq 0 \end{array} \iff \begin{array}{ll} \min\limits_{\lambda_2,v} & c^T\lambda_2 - b^T v \\ \text{s.t} & c \geq A^T v \\ & A\lambda_2 = b \\ & \lambda_2 \geq 0 \end{array} \qquad \cdots\cdots\cdots (3)$$

We observe that by doing a change of variables λ_2 to x and v to y (and the max into a min) we have the (Self-Dual) problem.

4) (a) Since the constraints of the (Self-Dual) problem are independent and the problem is bounded and feasible then we observe that (Self-Dual) is the problem (P) minus the problem (D):

$$\min_{x,y} c^T x - b^T y \quad \text{s.t.} \quad Ax = b,\ x \geq 0,\ A^T y \leq c \;=\; \min_x c^T x \ \text{s.t.}\ Ax=b,\ x \geq 0 \;-\; \min_y -b^T y \ \text{s.t.}\ A^T y \leq c \;=\; \min_x c^T x \ \text{s.t.}\ Ax=b,\ x \geq 0 \;-\; \max_y b^T y \ \text{s.t.}\ A^T y \leq c$$

So by solving (P) and obtain the optimal point x^* and resp. solving for (D) we obtain y^* we hence also have solved the problem (Self-Dual) i.e $[x^*, y^*]$.

(b) Note that the dual of P is D and the dual of D is P up to a change of variables. (P) is a convex problem and feasible, bounded by hypothesis, hence strong duality holds. As a result (let p^* (resp. d^*) be the optimal value of P (resp. D)):

$$\min_{x,y} c^T x - b^T y \quad Ax = b,\ x \geq 0,\ A^T y \leq c \;=\; p^* - d^* = p^* - p^* = 0$$

Exercise 2 (RLS)

1) Compute $\|\cdot\|_1^*$. Let $f = \|\cdot\|_1$ by definition of the conjugate we have:

$$f^*(y) = \sup_x y^T x - f(x) = \sup_x y^T x - \|x\|_1 = \sup_x \sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \quad \text{......}(4)$$

We look at 3 cases.

<u>case 1</u>: Suppose $\exists y_i > 1$. We choose $x_i = t > 0$ and $x_j = 0$ $(j \neq i)$. Then:

$$y^T x - \|x\|_1 = y_i t - t = t(y_i - 1)$$

$$\lim_{t \to \infty} t(y_i - 1) \longrightarrow +\infty \quad \text{hence} \quad f^*(y) = +\infty$$

<u>case 2</u>: Suppose $\exists y_i < -1$. We choose $x_i = t < 0$ and $x_j = 0$ $(j \neq i)$. Then:

$$y^T x - \|x\|_1 = y_i t + t = t(y_i + 1)$$

$$\lim_{t \to -\infty} t(y_i + 1) \to +\infty \quad \text{hence} \quad f^*(y) = +\infty$$

<u>case 3</u>: Suppose $\forall y_i \in [\![-1, 1]\!] \Leftrightarrow \|y\|_\infty \leq 1$. We have then:

$$\begin{aligned} y^T x - f(x) &\leq \sum_{i=1}^d |y_i x_i| - \sum_{i=1}^d |x_i| \\ &\leq \sum_{i=1}^d \underbrace{(|y_i| - 1)}_{\leq 0} |x_i| \\ &\leq 0 \end{aligned}$$

If we choose $x = 0$ then $f^*(y) = 0$.

<u>Conclusion</u>

$$\|\cdot\|_1^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise.} \end{cases}$$

2) $\min_x \|Ax - b\|_2^2 + \|x\|_1$ is equivalent to

$$\min_x \|y\|_2^2 + \|x\|_1 \quad \text{s.t} \quad y = Ax - b$$

The Lagrangian is:

$$\begin{aligned} \mathcal{L}(x, y, v) &= \|y\|_2^2 + \|x\|_1 + v^T(y - Ax + b) \\ &= b^T v + \|y\|_2^2 + v^T y + \|x\|_1 - (A^T v)^T x \end{aligned}$$

It follows the dual function is:

$$g(v) = \inf_{x,y} \mathcal{L}(x, y, v) = b^T v + \inf_y(\|y\|_2^2 + v^T y) + \inf_x(\|x\|_1 - (A^T v)^T x)$$

a) Since $f'(y) = \|y\|_2^2 + v^T y$ is convex and diffrentiable then we have:

$$\nabla f'(y) = 2y + v = 0 \Leftrightarrow y = -\tfrac{1}{2} v \quad \text{so the minimum value of } f'(y) \text{ is } \|-\tfrac{1}{2} v\|_2^2 - \tfrac{1}{2}\|v\|_2^2 = -\tfrac{1}{4}\|v\|_2^2$$

b)

$$\inf(\|x\|_1 - (A^T v)^T x) = \sup_x\left((A^T v)^T x - \|x\|_1\right) = \|A^T v\|_1^* \quad (\text{by definition } (4))$$

$$\therefore \; g(v) = b^T v - \tfrac{1}{4}\|v\|_2^2 + \|A^T v\|_1^* = \begin{cases} b^T v - \frac{1}{4}\|v\|_2^2 & \text{if } \|A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem is then:

$$\max_v \; b^T v - \frac{1}{4}\|v\|_2^2 \quad \text{s.t} \quad \|A^T v\|_\infty \leq 1$$

Exercise 3 (Data Separation)

1)
$$\min_{\omega,z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|\omega\|_2^2$$
$$\text{s.t.} \quad \forall i \in [\![1,n]\!], z_i \geq 1 - y_i(\omega^T x_i) \quad (\lambda_i)$$
$$z \geq 0 \quad (\pi)$$

We have that the (Sep 2.) problem is:

$$\min_{z} \frac{1}{n\tau} \underbrace{\mathbb{1}^T z}_{\text{— linear in } z} + \min_{w} \frac{1}{2}\|\omega\|_2^2$$
s.t
$$\forall i \in [\![1,n]\!], z_i \geq \underbrace{1 - y_i(\omega^T x_i)}_{\text{lower bound}}$$
$$z \geq 0$$

$$\Leftrightarrow \quad \frac{1}{n\tau} \min_{w} \sum_{i=1}^{n} \max(0, 1 - y_i(\omega^T x_i)) + \min_{w} \frac{1}{2}\|\omega\|_2^2$$

$$\Leftrightarrow \quad \frac{1}{n\tau} \min_{w} \sum_{i=1}^{n} \mathcal{L}(\omega, x_i, y_i) + \frac{1}{2}\|\omega\|_2^2$$

$$\overset{\tau>0}{\Leftrightarrow} \quad \frac{1}{n} \min_{w} \sum_{i=1}^{n} \mathcal{L}(\omega, x_i, y_i) + \frac{\tau}{2}\|\omega\|_2^2 \qquad (\text{Sep.1})$$

2) The Lagrangian of (Sep 2) is:

$$\mathcal{L}(z,\omega,\lambda_i,\pi) = \frac{1}{n\tau}\mathbb{1}^T z + \frac{1}{2}\|\omega\|_2^2 - \pi^T z + \sum_{i=1}^{n} \lambda_i \left(1 - y_i(\omega^T x_i) - z_i\right)$$
$$= \frac{1}{2}\|\omega\|_2^2 + \left(\frac{1}{n\tau}\mathbb{1} - \pi\right)^T z + \mathbb{1}^T\lambda - \lambda^T z - \sum_{i=1}^{n} \lambda_i y_i(\omega^T x_i)$$
$$= \frac{1}{2}\|\omega\|_2^2 + \left(\frac{1}{n\tau}\mathbb{1} - \pi - \lambda\right)^T z + \mathbb{1}^T\lambda - \sum_{i=1}^{n} \lambda_i y_i(\omega^T x_i)$$

The dual function:

$$g(\lambda,\pi) = \inf_{z,w}\left(\frac{1}{2}\|\omega\|_2^2 + \left(\frac{1}{n\tau}\mathbb{1} - \pi - \lambda\right)^T z + \mathbb{1}^T\lambda - \sum_{i=1}^{n} \lambda_i y_i(\omega^T x_i)\right)$$

a) $\inf_{w} g'(\omega) = \inf_{w}\left(\frac{1}{2}\|\omega\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i(\omega^T x_i)\right)$. $\quad \nabla g'(\omega) = w - \sum_{i=1}^{n} \lambda_i y_i x_i$

So the minimum value is $+w = \sum_{i=1}^{n} \lambda_i y_i x_i$ ....(5)

Note $g'(\omega)$ is convex + differentiable.

b) $$\inf_{z} g''(z) = \inf_{z}\left(\frac{1}{n\tau}\mathbb{1} - \pi - \lambda\right)^T z := \begin{cases} 0 & \text{if } \frac{1}{n\tau}\mathbb{1} - \pi - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

We report (5) into $g'(\omega)$. Hence
$$\frac{1}{2}\left\|\sum_{i}^{n} \lambda_i y_i x_i\right\|_2^2 - \sum_{i}^{n} \lambda_i y_i\left(\sum \lambda_j y_j x_j^T x_i\right)$$
$$= \frac{1}{2}\left(\sum_i \lambda_i y_i x_i\right)^T\left(\sum_j \lambda_j y_j x_j\right) - \sum_i\sum_j \lambda_i y_i \lambda_j y_j x_j^T x_i$$
$$= \frac{1}{2}\sum_{ij}^{n} \lambda_i y_i x_i^T \lambda_j y_j x_j - \sum_{ij} \lambda_i y_i \lambda_j y_j x_j^T x_i$$
$$= -\frac{1}{2}\sum_{ij} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$g(\lambda,\pi) = \begin{cases} \mathbf{1}^T\lambda - \frac{1}{2}\sum_{ij} \lambda_i \lambda_j y_i y_j x_i^T x_j & \text{if } \frac{1}{n\gamma}\mathbf{1} - \pi - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

It follows that the dual problem is to max $g(\lambda,\pi)$ subject to $\lambda \geq 0$ and $\pi \geq 0$:

$$\begin{aligned} \max_{\lambda,\pi} \quad & \mathbf{1}^T\lambda - \frac{1}{2}\sum_{ij} \lambda_i \lambda_j y_i y_j x_i^T x_j \\ \text{s.t.} \quad & \frac{1}{n\gamma} \geq \lambda \geq 0 \end{aligned}$$

QED.