Exercise 1 : | HW2 | | LEVY JAROD | (1)

1] Let's transform the problem into its standard form:

$$\underset{x}{\text{Minimize}} \; c^T x$$
subject to $b - Ax = 0$
$-x \leq 0$

→ Lagrangian with dual variable y for the equality constraint and λ for the inequality:
$\times \; L(x, y, \lambda) = c^T x + y^T(b - Ax) - \lambda^T x$

- Dual function: $g(y, \lambda) = \inf_x L(x, y, \lambda)$

→ Differenciation with respect to x gives: $\nabla_x L(x, y, \lambda) = c - A^T y - \lambda$

If x non positive, the problem is unbounded below.

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

- Dual Problem: $\underset{y}{\text{maximize}} \; b^T y$
subject to $c - A^T y - \lambda = 0$
$\lambda \geq 0$

The dual problem of (P) is (D)

2] Similarly, without all details for (D):

→ Standard form: $\underset{y}{\text{minimize}} \; - b^T y$
s. to $A^T y \leq c$

→ Lagrangian: $L(y, x) = -b^T y + x^T(A^T y - c)$
$= y^T(-b + Ax) - c^T x$

→ Differenciation $\nabla_y L(y, x) = -b + Ax$

→ $g(x) = \begin{cases} -c^T x & \text{if } Ax - b = 0 \\ -\infty & \text{otherwise} \end{cases}$ → x has to be non negative otherwise the problem is unbounded below.

→ Dual Problem: $\underset{x}{\max} \; c^T x$
s. to $Ax = b$
$x \geq 0$

The dual problem of (D) is (P)

3] Thanks to 1] and 2], we recognize that the left part (2)
will give the right part and vice versa if we look for the dual problem. Both objectives functions are linear so the problem can be rewritten as:

$$\min_x c^T x + \max_y b^T y$$
$$\text{s. to} \quad Ax = b$$
$$x \geq 0$$
$$A^T y \leq c$$

Re-doing the same computation, we will retrieve the same problem and so the problem is self dual.

4] 1] Thanks to LP strong duality property, the objective values of the primal and dual problem at these solutions will be equal:

$$c^T x^* = b^T y^*$$

- if (x^*, y^*) is the optimal solution for the self dual problem, x^* and y^* would also satisfy the constraints of (P) and (D) otherwise the self dual problem would not be feasible at these points.
- Self dual is the combination of (P) and (D) both LP and x^*, y^* satisfy the constraints and provide the optimal value for self dual problem. They also provide feasible solution for (P) and (D)

So (x^*, y^*) can be obtained by solving (P) and (D)

2] Strong Duality gives us $c^T x^* = b^T y^*$ then $c^T x^* - b^T y^* = 0$

So, $p^* = 0$ exactly

Exercise 3 :

1] (Sep.2) solves (Sep.1) because the loss function is represented implicitly by the constraints on z. The constraints ensure that if a data point x_i is misclassified, the corresponding variable z_i will be positive and if it's correctly classified z_i will be zero.

Multiplying by τ, we retrieve the form:

$$\min_{w,z} \frac{1}{N}\sum_{i=1}^{\tilde{}} z_i + \frac{\tau}{2}\|w\|_2^2$$

~~The sum~~ of z_i and the 2 constraints $\begin{cases} z_i \geq 1 - y_i(w^T x_i) \\ z \geq 0 \end{cases}$

Correspond to $L(w, x, z)$ in (Sep.1)

In summary, both (Sep.1) and (Sep.2) aim to minimize misclassification and regularization. However (Sep.1) does this through the explicit loss function while (Sep.2) does it implicitly through the constraints.

2] Lagrangian :

$$L(w, z, \lambda, \pi) = \frac{1}{N\tau} 1^T z + \frac{1}{2}\|w\|_2^2 + \sum_{i=1}^{\tilde{}} \lambda_i (1 - y_i w^T x_i - z_i) - \pi_i z_i$$

Differenciation:

- $\nabla_w L = w - \sum_{i=1}^{N} \lambda_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^{N} \lambda_i y_i x_i$
- $\nabla_z L = \frac{1}{N\tau} - \lambda_i - \pi_i = 0$ for all i $\Rightarrow \pi_i = \frac{1}{N\tau} - \lambda_i$ for all i

π has to be positive so we have:

$$0 \leq \lambda_i \leq \frac{1}{N\tau} \text{ for all } i$$

Rearranging by injecting the new expressions of π and w:

$$g(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2}\|\sum_{i=1}^{N} \lambda_i y_i x_i\|_2^2$$

The dual problem is : $\max_{\lambda} \text{imize } g(\lambda)$

subject to : $0 \leq \lambda \leq 1^T \frac{1}{N\tau}$

The variable w in the primal problem become a linear combination of datapoints in the dual, reducing the number of variables significantly especially if the number of support vectors ($\lambda_i > 0$) is small.

Exercise 4 :

Let's use the hint and find the dual of the problem :

$$\begin{aligned} &\underset{a}{\text{maximize}} \ a^T x \\ &\text{s.to} \ C^T a \le d \end{aligned}$$

- **Lagrangian** : $L(a, z) = a^T x + z^T (d - C^T a)$

$\nabla_a L(a, z) = x - Cz = 0 \Rightarrow C^T z = x$

- **Dual Problem** :

$$\begin{aligned} &\underset{z}{\text{Minimize}} \ d^T z \\ &\text{s.to} \ C^T z = x \\ &\quad z \geq 0 \end{aligned}$$

Now consider : $\underset{x}{\text{minimize}} \ c^T x$

s.t $\sup_{a \in P} a^T x \le b$

Given the definition of the supremum, x is feasible for the original problem if and only if $a^T x \le b$ for all $a \in P$. This is equivalent of saying that the optimal value of the dual problem (ie $d^T z$) is less than or equal to b. Hence : $d^T z \le b$

$C^T z = x$

$z \geq 0$

Given those results, the 2 following problems are indeed equivalent

$$\begin{aligned} &\underset{x}{\text{Minimize}} \ c^T x \\ &\text{s.to} \ \sup_{a \in P} a^T x \le b \end{aligned} \qquad \begin{aligned} &\underset{x}{\text{Minimize}} \ c^T x \\ &\text{s.to} \ d^T z \le b \\ &\quad C^T z = x \\ &\quad z \geq 0 \end{aligned}$$

- Robust linear programming problem can be transformed into another linear programming problem as described.

Exercise 2 :

1] $f^*(y) = \sup_{x \in \text{dom} f} (y^T x - f(x))$

$= \sup_{x \in \text{dom} f} (y^T x - \|x\|_1) = \sup_{x \in \text{dom} f} \left(\sum_{i=1}^{d} y_i x_i - \sum_{i=1}^{d} |x_i| \right)$

- $y_i x_i$ maximized when x_i and y_i have same sign.
- The term $-\|x\|_1$ implies that the magnitude should not exceed 1.
- If $y_i > 1$ or $y_i < -1$, $y_i x_i$ dominates and goes to infinity
- If $-1 \leq y_i \leq 1$ it maximizes $y_i x_i - |x_i|$

$\sup_{x \in \text{dom} f} \{ y_i x_i - |x_i| \} = |y_i| - 1$

The conjugate is : $f^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$

2] $\text{Minimize}_x \ \|Ax - b\|_2^2 + \|x\|_1$

- Let's introduce $z = Ax$, the problem becomes: $\text{Minimize}_{x,z} \ \|z - b\|_2^2 + \|x\|_1$

  Subject to $z = Ax$

- Lagrangian: $L(x,z,\mu) = \|z-b\|_2^2 + \|x\|_1 + \mu^T (Ax - z)$

μ is the dual variable introduced for the equality constraint.

$L(x,z,\mu) = \|z-b\|_2^2 - \mu^T z + \|x\|_1 + \mu^T A x$

- Differentiation.

  → with respect to z : $\nabla_z L = 2(z-b) - \mu$

  $\nabla_z L = 0 \iff z = \frac{\mu}{2} + b$

  → with respect to x :

We can introduce : $L_1(x,\mu) = \|x\|_1 + \mu^T A x = |x| + \mu^T A x$

For each i : $L_{1,i}(x_i,\mu) = (1 + \text{sgn}(x_i) \mu^T A) |x_i|$

So we have an unbounded problem if $\|\mu^T A\|_\infty > 1$

Putting all together, we replace z by the value found with differentiation and we add the constraint found with respect to x - (6)

$$g(\mu) = \inf_{x,z} L(x, z, \mu)$$
$$= \| \tfrac{\mu}{2} + b - b \|_2^2 - \mu (b + \tfrac{\mu}{2})$$

Dual Problem :

$$\underset{\mu}{\text{Maximize}} \quad -\frac{1}{4} \| \mu \|_2^2 - \mu^T b$$
$$\text{subject to} \quad \| \mu^T A \|_\infty \leq 1$$

Exercise 5:

1] $\underset{x}{\text{Minimize}} \quad c^T x$

Subject to $\quad Ax \leq b$

$x_i(1-x_i) = 0$ for all i

The Lagrangian is: λ and μ are the dual variables corresponding to the constraints

$$L(x,\lambda,v) = c^T x + \lambda^T(Ax-b) + v^T \sum_{i=1}^{N} x_i(1-x_i)$$

Let's introduce diag(v) to be able to write it in a "vectorized" way

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

For each: $\nabla_x L_i(x_i,\lambda,v_i) = 2v_i^T x_i + c + \lambda^T A - v_i = 0$

$$x_i = -\frac{(c_i + \lambda^T A_i - v_i)^2}{2v_i}$$

Replacing in the Lagrangian: $g(\lambda,v) = \begin{cases} -\lambda^T b - \frac{1}{4}\sum_{i=1}^{N} \frac{(c_i + \lambda^T A - v_i)^2}{v_i} \\ -\infty \quad \text{otherwise} \end{cases}$

Using the hint provided, we obtain the dual problem:

$\underset{\lambda}{\text{Maximize}} \quad -\lambda^T b + \sum_{i=1}^{N} \min\{0, c_i + \lambda^T A\}$

Subject to $\quad \lambda \geq 0$

2] Let's derive the dual of the LP relaxation

$$L(x,\lambda,\mu,v) = c^T x + \lambda^T(Ax-b) - \mu^T x + v^T(x - \underset{\text{vector}}{\mathbb{1}})$$

$$= (c + A^T\lambda - \mu + v)^T x - b^T\lambda - \mathbb{1}^T v$$

$$g(\lambda,\mu,v) = \begin{cases} -b^T\lambda - \mathbb{1}^T v & A^T\lambda - \mu + v + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

It is equivalent to the other problem. The 2 relaxation gives the same value:

<u>Dual problem</u>: $\underset{\lambda,v}{\text{Maximize}} \quad -b^T\lambda - \mathbb{1}^T v$

Subject to $\quad A^T\lambda - \mu + v + c = 0$

$\lambda \geq 0, \; \mu \geq 0, \; v \geq 0$

v is affine

$v = \mu - c - A^T\lambda$

Corresponds to

$\sum_{i=1}^{N} \min\{0, c_i + \lambda^T A\}$ *

* In problem 1.