Homework - 2 -

Student: Imane SI SALAH

Course: Convex optimization

Exercice 1: (LP duality)

for given $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times d}$ we have the two linear optimization problems:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \qquad (P)$$

and

$$\begin{aligned} \text{maximize} \quad & b^T y \\ \text{subject to} \quad & A^T y \leq c \end{aligned} \qquad (D)$$

① Dual of (P)

→ the Lagrangian of (P) is $L(x, \lambda, \mu) = c^T x - \lambda^T x + \mu^T (Ax - b)$

and the dual function is

$$g(\lambda, \mu) = \inf_x L(x, \lambda, \mu)$$

$$= \inf_x (c - \lambda + A^T \mu) x + \mu^T b$$

we are trying to minimize a linear function of x

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

therefore the dual problem is:

$$\begin{aligned} \max \quad & -\mu^T b \\ \text{st} \quad & c - \lambda + A^T \mu = 0 \\ & \lambda \geq 0 \end{aligned}$$

the equality constraint can be rearranged and we drop λ

$$c + A^T\mu = \lambda \geq 0$$

so the simplified dual problem of (P) is

$$\boxed{\begin{array}{ll} \max & -\mu^T b \\ \text{st} & A^T\mu + c \geq 0 \end{array}}$$

② the dual problem of (D)

→ the Lagrangian is given by

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

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

so the dual function is given by

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

We are trying to minimiz a linear function of y

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

therefore the dual problem of (D) is

$$\boxed{\begin{array}{ll} \max & -\lambda^T c \\ \text{st} & A\lambda = b \\ & \lambda \geq 0 \end{array}}$$

We clearly notice that the dual of problem (D) is (P) itself.

③ Prove that the problem is self dual

$$\min_{x,y} \; c^T x - b^T y$$

$$\text{st} \quad Ax = b, \quad x \geq 0, \quad A^T y \leq c \qquad \text{(Self-Dual)}$$

→ the Lagrangian of this problem is given by

$$L(x, y, \lambda_1, \lambda_2, \mu) = c^T x - b^T y - \lambda_1^T x + \lambda_2^T (A^T x - c) + \mu^T (b - Ax)$$

$$= (c - \lambda_1 - A^T \mu)^T x + (A \lambda_2 - b)^T y - c^T \lambda_2 + b^T \mu$$

(this is linear of x and y)

the lagrange dual function is:

$$g(\lambda_1, \lambda_2, \mu) = \inf_{x,y} L(x, y, \lambda_1, \lambda_2, \mu)$$

$$= \begin{cases} -c^T \lambda_2 + b^T \mu & \text{if } \begin{cases} c - \lambda_1 - A^T \mu = 0 \\ A \lambda_2 - b = 0 \end{cases} \\ -\infty & \text{elsewhere} \end{cases}$$

therefore the dual problem of (Self-Dual) is

$$\max \; -c^T \lambda_2 + b^T \mu$$

$$\text{st} \quad c - \lambda_1 - A^T \mu = 0$$

$$A \lambda_2 - b = 0$$

$$\lambda_1 \geq 0$$

$$\lambda_2 \geq 0$$

this is equivalent to:

$$\min \; c^T \lambda_2 - b^T \mu$$

$$\text{st} \quad A \lambda_2 = b$$

$$c - A^T \mu \geq 0$$

$$\lambda_2 \geq 0$$

$$\xrightarrow{\text{set } \lambda_2 = x,\ \mu = y}$$

$$\boxed{\begin{aligned} &\min \; c^T x - b^T y \\ &\text{st} \quad Ax = b \\ &\quad A^T y \leq c \\ &\quad x \geq 0 \end{aligned}}$$

Which is the same as (Self-Dual)

④ Assuming that the (Self-Dual) problem is feasible and bounded and $[x^*, y^*]$ is the optimal solution
we show that the optimal solution can be obtained by solving (P) and (D)

we have the constraints of the (Self-Dual) problem are completely independent, so it can be devided into two inep problems as follows:

$$\begin{array}{lll} \min\ c^T x - b^T y & & \min\ c^T x & & \max\ b^T y \\ \text{st}\quad Ax = b & \Rightarrow & \text{s.t.}\quad Ax = b & + & \text{st}\quad A^T y \leq c \\ \qquad x \geq 0 & & \qquad x \geq 0 & & \\ \qquad A^T y \leq c & & & & \\ \text{(Self-Dual)} & & \text{(P)} & & \text{(D)} \end{array}$$

So the optimal solution of the (Self Dual) problem is (x^*, y^*) where x^* is the optimal solution of (P) and y^* is the optimal solution of (D)

- Show that the optimal value of (Self-Dual) is 0

Now from question 2 we saw that (P) is the dual of (D), this means that $p^* = d^*$ where p^* and d^* are the optimal values of (P) and (D) Respectivly

so $c^T x^* = b^T y^* \longrightarrow \boxed{c^T x^* - b^T y^* = 0}$

Exercice 2-1 (Regularized least Square)

For given $A \in \mathbb{R}^{n \times d}$ and $b \in \mathbb{R}^n$, we have the following op :

$$\min_x \|Ax - b\|_2^2 + \|x\|_1 \qquad (RLS)$$

1 - the conjugate of $\|x\|_1$

in general, the conjugate of a function $f : \mathbb{R}^n \to \mathbb{R}$

is $f^*(y) = \max_x y^T x - f(x)$

and we have the dual of a norm is $\|y\|_\infty = \sup_{\|x\| < 1} x^T y$

$= \|y\|_\infty$

so take $f(x) = \|x\|_1$

then $f^*(y) = \max_x y^T x - \|x\|_1$

Now we want to evaluate $f^*(y)$ we get 2 case

case 1 : $\|y\|_\infty \leq 1$ then $y^T x \leq \|x\| \, \|y\|_\infty \leq \|x\|$

so we get equality if $x = 0 \Rightarrow f^*(y) = 0$

case 2 : $\|y\|_\infty > 1$ and by definition we have

$$\|y\|_* = \max_{\|x\| < 1} x^T y > 1$$

so $\exists x$ with $\|x\| \leq 1$ and $x^T y > 1$

so $f^*(y) \geq y^T (tx) - \|tx\| = t \left(y^T x - \|x\| \right)$

as $t \to \infty$ $f^*(y) \to +\infty$

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

② Compute the dual of (RLS)

we rewrite the (RLS) problem as follows by performing a change of variable :

$$\min \|y\|_2^2 + \|x\|_1$$
$$\text{st} \quad y = Ax - b$$

→ the lagrangian function is :

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

so the dual function is :

$$g(\mu) = \inf_y \left(\|y\|_2^2 + \mu^T y \right) + \inf_x \left(\|x\|_1 - \mu^T A x \right) + \mu^T b$$

• Now we minimize $h(y, \mu) = \|y\|_2^2 + \mu^T y$

$$\nabla_y h(y, \mu) = 2y + \mu = 0 \Rightarrow \boxed{y = -\frac{\mu}{2}}$$

so $\inf_y h(y, \mu) = -\frac{\|\mu\|_2^2}{4}$

• then we minimize $k(x, \mu) = \|x\|_1 - \mu^T A x$

which is eq to maximizing $-k(x, \mu) = \mu^T A x - \|x\|_1$

$$= (A^T \mu)^T x - \|x\|_1$$

so $\max_x -k(x, \mu) = \max_x (A^T \mu)^T x - \|x\|_1$

$$= \begin{cases} 0 & \|A^T \mu\|_\infty \leq 1 \\ +\infty & \|A^T \mu\|_\infty > 1 \end{cases}$$ from qst 1

therefore the dual problem is

$$\boxed{\begin{aligned} &\max \ \mu^T b - \frac{1}{4} \|\mu\|_2^2 \\ &\text{s.t.} \quad \|A^T \mu\|_\infty \leq 1 \end{aligned}}$$

Exercice 3 : (Data Separation)

data $x_i \in \mathbb{R}^d$, $y_i \in \{-1, 1\}$ are labels

we are searching for a hyperplane defined by its normal w which separates the data points according to their label

st $w^T x_i \leq -1 \Rightarrow y_i = -1$ and $w^T x_i \geq 1 \Rightarrow y_i = 1$

- we introduce the loss function

$$\mathcal{L}(w, x_i, y_i) = \max\{0, 1 - y_i(w^T x_i)\}$$

→ to improve the performance we optimize this problem

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

1. We explain why the following problem solves (Sep 1)

$$\min \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2$$

$$\text{st} \quad z_i \geq 1 - y_i(w^T x_i) \quad \forall x_i = 1, \dots, n$$

→ starting from (Sep1) we perform a change variable $z = L(w, x_i, y_i)$ and scaling of $\frac{1}{\tau}$

this results in a problem that has explicit constraints on z : $\begin{cases} z \geq 1 - y_i(w^T x_i) \\ z \geq 0 \end{cases}$

So the lower bounds introduced by the implicit constraints are the same as those introduced by the original loss function, therefore the solution of (Sep 2) is the solution of (Sep 1) up to a scaling factor.

② the dual of (sep 2)

We first compute the Lagrangian:

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

Now the dual of the Lagrangian is

$$g(\lambda, \pi) = \inf_z \left[\frac{1}{nC} 1^T z - \sum_{i=1}^{n} \lambda_i z_i - \pi^T z\right] + \inf_w \left[\frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i)\right] + \sum_{i=1}^{n} \lambda_i$$

take $h(w, \lambda, \pi) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i)$

$$\nabla_w h(w, \lambda, \pi) = w - \sum_{i=1}^{n} \lambda_i y_i x_i = 0$$

we get $w = \sum_{i=1}^{n} \lambda_i y_i x_i$

so $\inf_w h(w, \lambda, \pi) = -\frac{1}{2}\|\sum_{i=1}^{n} \lambda_i y_i x_i\|_2^2$

also $h(z, \lambda, \pi) = \frac{1}{nC} 1^T z - \sum_{i=1}^{n} \lambda_i z_i - \pi^T z$

$$\inf_z h(z, \lambda, \pi) = \begin{cases} 0 & \text{if } \frac{1}{nC} 1 - \sum_{i=1}^{n} \lambda_i - \pi \\ -\infty & \text{elsewhere.} \end{cases}$$

so $$g(\lambda, \pi) = \begin{cases} \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\|\sum_{i=1}^{n} \lambda_i y_i x_i\|_2^2 & \text{if } \frac{1}{nC} 1 - \sum_{i=1}^{n} \lambda_i - \pi = 0 \\ -\infty & \text{elsewhere} \end{cases}$$

So the dual problem is:

$$\boxed{\begin{array}{ll} \max & \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\|\sum_{i=1}^{n} \lambda_i y_i x_i\|_2^2 \\ \text{st} & \frac{1}{nC} 1 - \sum_{i=1}^{n} \lambda_i \geq 0 \\ & \lambda_i \geq 0 \quad i = 1, \dots, n \end{array}}$$