Convex Optimization – Homework 2

Boubacar SOW

Exercice 1 (LP duality): For a given $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times d}$ consider the following linear optimization problems,

$$\begin{aligned} \min\ & c^T x \\ \text{s.t}\ & Ax = b \qquad (P) \\ & x \geq 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} \max_y\ & b^T y \qquad (D) \\ \text{s.t}\ & A^T y \leq c \end{aligned}$$

1. Compute the dual of problem (P) and simplify it if possible

$$L(x, \lambda, \nu) = f_0(x) - \lambda^T x + \nu(Ax - b)$$

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu) = \inf_{x \in D} c^T x - \lambda x + \nu(Ax - b)$$

$$g(\lambda, \nu) = \inf -b^T \nu + (c + A^T \nu - \lambda)^T x$$

This function is linear in x, we have thus

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

The dual problem is: maximize $-b^T \nu$
subject to $A^T \nu + c \succeq 0$

2. Compute the dual of problem (D)

$$L(x, \lambda, \nu) = f_0(x) + \lambda(A^T y - c) = -b^T y + \lambda(A^T y - c)$$

$$\begin{aligned} g(\lambda, \nu) &= \inf -b^T y + \lambda^T (A^T y - c) = \inf -b^T y + \lambda^T A y - \lambda^T c \\ &= \inf (\lambda^T A - b)^T y - \lambda^T c \\ g(\lambda, \nu) &= \inf \underbrace{(A\lambda - b)\, y - c^T \lambda}_{L(\lambda, \nu)} \end{aligned}$$

The function $L(\lambda, \nu)$ is linear in y, we have thus

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

The dual problem is: maximize $-c^T \lambda$
subject to $A\lambda - b = 0$, $\lambda \geq 0$

3. Prove that the following problem is self dual:

$$\begin{aligned} \min\ & c^T x - b^T y \\ \text{s.t}\ & Ax = b \qquad\qquad (\text{self-dual}) \\ & x \geq 0 \\ & A^T y \leq c \end{aligned}$$

The problem is equivalent to

$$\begin{aligned} \text{Min}\ & c^T x - b^T y \\ \text{st.}\ & Ax - b = 0 \\ & -x \leq 0 \\ & A^T y - c \leq 0 \end{aligned}$$

The lagrangian function is:

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

The domain of the feasible set is: $D = \{(x,y) \mid Ax = b,\ x \geq 0;\ A^Ty \leq c\}$

$$\Rightarrow g(\lambda_1,\lambda_2,\nu) = \inf_{(x,y)\in D} L(x,y,\lambda_1,\lambda_2,\nu)$$

$$= \inf\ (c-\lambda_1+A^T\nu)^Tx + (A\lambda_2 - b)^T y - (\lambda_2^T c + \nu^T b)$$

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

The dual problem is:

$$\begin{aligned} &\underset{\lambda_2,\nu}{\text{maximize}} && -c^T\lambda_2 - b^T\nu \\ &\text{subject to} && c-\lambda_1 + A^T\nu = 0 \\ &&& A\lambda_2 - b = 0 \\ &&& \lambda_1 \geq 0,\ \lambda_2 \geq 0 \end{aligned} \iff \begin{aligned} &\underset{\lambda_2,\nu}{\text{maximize}} && -c^T\lambda_2 - b^T\nu \\ &\text{subject to} && A^T\nu + c \geq 0 \\ &&& A\lambda_2 - b = 0 \\ &&& \lambda_2 \geq 0 \end{aligned}$$

$$\iff \begin{aligned} &\underset{\lambda_2,t}{\text{minimize}} && c^T\lambda_2 - b^Tt \\ &\text{subject to} && A^Tt - c \leq 0 \\ &&& A\lambda_2 - b = 0 \\ &&& \lambda_2 \geq 0 \end{aligned} \iff \boxed{\begin{aligned} &\underset{\lambda_2,t}{\text{minimize}} && c^T\lambda_2 - b^Tt \\ &\text{subject to} && A\lambda_2 - b = 0 \\ &&& A^Tt - c = 0 \\ &&& -\lambda_2 \leq 0 \end{aligned}}$$

4. Assume the above problem feasible and bounded and let $[x^*, y^*]$ the optimal solution. Using the strong duality property of linear programs. Show that:

* The vector $[x^*, y^*]$ can also be obtained by solving (P) and (D)

The feasible set of (self-Dual) is:

$\{(x,y) \mid Ax = b;\ x \geq 0;\ A^Ty \leq c\} = \{x \mid Ax = b;\ x\geq 0\} \cup \{y \mid A^Ty \leq c\}$

We can thus decompose the problem into two minimization problems

$$\Rightarrow \begin{aligned} &\underset{x,y}{\text{Min}} && c^Tx - b^Ty \\ &\text{subject to} && Ax = b \\ &&& x \geq 0 \\ &&& A^Ty \leq c \end{aligned} \iff \begin{aligned} &\underset{x}{\text{Min}} && c^Tx \\ &\text{subject to} && Ax = b \\ &&& x \geq 0 \end{aligned} + \begin{aligned} &\underset{y}{\text{Min}} && -b^Ty \\ &\text{subject to} && A^Ty \leq c \end{aligned}$$

$$\iff \begin{aligned} &\underset{x}{\text{Min}} && c^Tx \\ &\text{subject to} && Ax = b \\ &&& x \geq 0 \end{aligned} + \begin{aligned} &\underset{y}{\text{Max}} && b^Ty \\ &\text{subject to} && A^Ty \leq c \end{aligned}$$

$$\equiv \quad (P) \quad + \quad (D)$$

Since we know that $[x^*, y^*]$ is a feasible solution for the (self-Dual) problem, we can see here that x^* is an optimal solution for (P) $\underbrace{[(P) \text{ is convex}]}_{\text{linear objective and constraints}}$ and y^* is an optimal solution for (D) $\underbrace{[(D) \text{ is convex}]}_{\text{linear objective and constraints}}$

* The optimal of (self-Dual) is exactly 0

$$\begin{aligned} &\text{Min} && c^Tx - b^Ty \\ &\text{st} && Ax = b \\ &&& x \geq 0 \\ &&& A^Ty > c \end{aligned} = \begin{aligned} &\text{Min } c^Tx - \text{Min } b^Ty \\ &\text{st } Ax = b \quad A^Ty \leq c \\ &\quad x \geq 0 \end{aligned}$$

Strong duality $\quad c^Tx^* = y^{*T} b$

So the self dual optimal value is $c^Tx^* - b^Ty^* = 0$

$$\Rightarrow \text{Self-Dual} = 0$$

Exercise 2: Regularized least-square:

For a given $A \in \mathbb{R}^{n\times d}$ and $b \in \mathbb{R}^n$, consider the following optimization problem

$$\min_x \|Ax-b\|_2^2 + \|x\|_1$$

1. Compute the conjugate of $\|x\|_1$.

let's $f(x) = \|x\|_1$

$$f^*(y) = \sup \langle y, x\rangle - f(x) = \sup\ y^T x - \|x\|_1 = \sup_x\ y^T x - \sum_{i=1}^d |x_i|$$
$$= \sup_x \sum_{i=1}^d x_i y_i - \sum_{i=1}^d |x_i|$$

We have to explore all the possible values of y

- If $y_i \leq 1$, we have:

$$x_i y_i \leq |x_i y_i| \leq |x_i|$$
$$\Rightarrow \sum_{i=1}^d x_i y_i \leq \sum_{i=1}^d |x_i| \Rightarrow y^T x - \|x\|_1 \leq 0$$

$$\Rightarrow \sup_x\ y^T x - \|x\|_1 = 0 \Rightarrow f^*(y) = 0$$

- If $y_i > 1$:

let's have a case where $x_i = t > 0$ and $x_k = 0\ \forall k \neq i$

$$\Rightarrow y^T x - \|x\|_1 = y_i t - t = \underbrace{t(y_i - 1)}_{>0} \xrightarrow{t \to +\infty} +\infty$$

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

$$\Rightarrow \boxed{f^*(y) = \begin{cases} 0 & \text{if } |y_i| \leq 1 \ (\text{or } \|y\|_\infty \leq 1) \\ \infty & \text{otherwise} \end{cases}}$$

2. Compute the dual of regularized least-squares (LRS)

(RLS): $$\underset{x}{\text{minimize}}\ \|Ax-b\|_2^2 + \|x\|_1$$

This problem is equivalent to:

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & \|y\|_2^2 + \|x\|_1 \\ \text{subject to} & Ax - b = y \end{array} \iff \begin{array}{ll} \underset{x,y}{\text{minimize}} & \|y\|_2^2 + \|x\|_1 \\ \text{subject to} & Ax - b - y = 0 \end{array}$$

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

Domain $D = \{(x,y) \mid Ax - y - b = 0\}$

Dual lagrangian: $$g(v) = \inf_{x,y\in D} \|x\|_1 + v^T A x + \|y\|_2^2 - v^T b - v^T y$$

$$= \underbrace{\inf_{x\in D} \{\|x\|_1 + v^T A x\}}_{A} + \underbrace{\inf_{y\in D}\{\|y\|_2^2 - v^T y\}}_{B} - b^T v$$

$$A = \inf_x \{\|x\|_1 + v^T A x\}$$
$$= -\sup_x \{-\|x\|_1 - v^T A x\}$$
$$= -\sup_x \{(-v^T A)^T x - \|x\|_1\} = \begin{cases} 0 & \text{if } \|-A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

(from the previous question)

$$B = \inf_{y\in D}\{y^T y - v^T y\}$$

$$\nabla_y (y^T y - v^T y) = 0 \Rightarrow 2y - v = 0 \Rightarrow y^* = \frac{1}{2} v$$
$$\Rightarrow B = \frac{1}{4} v^T v - \frac{1}{2} v^T v = -\frac{1}{4} v^T v$$

$$\Rightarrow g(v) = \begin{cases} -\frac{1}{4} v^T v - b^T v & \text{if } \|-A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

concave function

$\Rightarrow$ The dual problem is: $$\begin{array}{ll} \text{maximize} & -\frac{1}{4} v^T v - v^T b \\ \text{subject to} & \|-A^T v\|_\infty \leq 1 \end{array}$$

Exercise 3: Data separation :

$y_i \in \{-1,1\}$, w^T : normal vector. We would like $w^T x_i \le -1 \Rightarrow y_i = -1$ and $w^T x_i \ge 1 \Rightarrow y_i = 1$.

loss function $\mathcal{L}(w, x_i, y_i) = \max\{0, 1 - y_i(w^T x_i)\}$; Objective : $\min\limits_{w} \frac{1}{n}\sum\limits_{i=1}^{n} \mathcal{L}(w, x_i, y_i) + \overset{\text{regularization parameter}}{\frac{\tau}{2}} \|w\|_2^2$ (sep. 1)

1. Consider the following quadratic optimization problem :

$$\begin{aligned} \text{minimize} \quad & \frac{1}{n\tau}\mathbf{1}^T z + \frac{1}{2}\|w\|_2^2 \\ \text{s.t} \quad & z_i \ge 1 - y_i(w^T x_i) \quad \forall i = 1,..n \quad (\lambda_i) \\ & z \ge 0 \quad (\pi) \end{aligned} \qquad \text{(Sep.2)}$$

Explain why (Sep.2) solves (Sep.1)

$$\text{(sep1)} : \min_{w} \frac{1}{n}\sum_{i=1} \max\{0, 1 - y_i(w^T x_i)\} + \frac{\tau}{2}\|w\|_2^2$$

In (sep1), the loss function includes a max function that makes the problem hard to solve because the max operation is not differentiable, so we introduce the auxiliary variable z_i and we add two constraints $z_i \ge 1 - y_i(w^T x_i)$ and $z \ge 0$

$$\Rightarrow \begin{cases} \text{If } 1 - y_i(w^T x_i) > 0 \text{ then } z_i \ge 0 \\ \text{If } 1 - y_i(w^T x_i) \le 0 \text{ then } z_i \ge 1 - y_i(w^T x_i) \ge 0 \end{cases} \Rightarrow z_i = \max\{0 ; 1 - y_i(w^T x_i)\}$$

$$\text{(sep 1)} \Leftrightarrow \begin{aligned} \min_{w,z} \quad & \frac{1}{n}\sum_{i=1}^{n} z_i + \frac{\tau}{2}\|w\|_2^2 \\ \text{s.t.} \quad & z_i \ge 1 - y_i(w^T x_i) \\ & z_i \ge 0 \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \min \quad & \frac{1}{n\tau}\mathbf{1}^T z + \frac{1}{2}\|w\|^2 \\ \text{s.t} \quad & z_i \ge 1 - y_i(w^T x_i) \\ & z_i \ge 0 \end{aligned} \quad \text{(sep 2)}$$

We have showed that by solving (sep2) we solve (sep1)

2. Compute the dual of (sep.2) and try to reduce the number of variables

$$\begin{aligned} \min \quad & \frac{1}{n\tau}\mathbf{1}^T z + \frac{1}{2}\|w\|^2 \\ \text{s.t} \quad & z_i \ge 1 - y_i(w^T x_i) \\ & z_i \ge 0 \end{aligned} \qquad \text{(sep 2)} \Leftrightarrow \begin{aligned} \text{minimize} \quad & \mathbf{1}^T z + \frac{1}{2}\|w\|_2^2 \\ \text{subject to} \quad & 1 - y_i(w^T x_i) - z_i \le 0 \quad \forall i = 1,\dots,n \\ & -z \le 0 \end{aligned}$$

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

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

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

Dual Lagrangian $g(\lambda, \pi) = \inf\limits_{z,w \in \mathcal{D}} \left(\frac{1}{n\tau} - \pi - \lambda\right)^T z + \frac{1}{2}\|w\|_2^2 + \sum_{i=1}^{n} \lambda_i(1 - y_i(x_i^T w))$

$$\mathcal{D} = \{(z, w) \mid -z \le 0 ; 1 - y_i(x_i^T w) - z_i \le 0 \quad \forall i = 1,\dots,n\}$$

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

$$= \mathbf{1}^T\lambda + \inf_{w} \underbrace{\frac{1}{2}w^T w + \sum_{i=1}^{n} -\lambda_i y_i(x_i^T w)}_{A} + \underbrace{\inf\left(\frac{1}{n\tau} - \pi - \lambda\right)^T z}_{B}$$

$$A(w) = \frac{1}{2}w^T w - \sum_{i=1}^{n} \lambda_i y_i(x_i^T w) \Rightarrow \nabla_w A = w - \sum_{i=1}^{n} \lambda_i y_i x_i$$

$$\nabla_w A = 0 \Rightarrow w = \sum_{i=1}^{n} \lambda_i y_i x_i$$

$B(z) = (\frac{1}{n\gamma} - \pi - \lambda)^T z$ [linear function]

$$\inf B(z) = \begin{cases} 0 & \text{if } \frac{1}{n\gamma} - \pi - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

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

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

Dual problem:

$$\begin{aligned} \underset{\lambda,\pi}{\text{maximize}} \quad & -\frac{1}{2}\|\sum_{i=1}^n \lambda_i y_i x_i\|_2^2 + 1^T\lambda \\ \text{subject to} \quad & \frac{1}{n\gamma} - \pi - \lambda = 0 \\ & \lambda \geq 0 \\ & \pi \geq 0 \end{aligned} \iff \begin{aligned} \text{maximize} \quad & -\frac{1}{2}\|\sum_{i=1}^n \lambda_i y_i x_i\|_2^2 + 1^T\lambda \\ \text{s.t} \quad & \frac{1}{n\gamma} - \lambda \geq 0 \\ & \lambda \geq 0 \end{aligned}$$

$$\iff \begin{aligned} \text{maximize} \quad & -\frac{1}{2}\|\sum_{i=1}^n \lambda_i y_i x_i\|_2^2 + 1^T\lambda \\ \text{s.t} \quad & \frac{1}{n\gamma} \geq \lambda \\ & \lambda \geq 0 \end{aligned}$$

Exercise 4: Robust linear programming:

Consider the following robust LP: $\begin{cases} \min_x c^T x \\ \text{s.t} \sup_{a \in \mathcal{P}} a^T x \leq b \end{cases}$, with $x \in \mathbb{R}^n$, $\mathcal{P} = \{a \mid C^T a \preceq d\}$

Show that this problem is equivalent to:

$$\boxed{\begin{aligned} & \min c^T x \\ \text{s.t} \quad & d^T z \leq b \\ & C^T z = x \\ & z \geq 0 \end{aligned}}$$