HomeWork 2
Convex Optimization
2023

Gabriel Singer

Exercice 1

1. First, lets write (P) in standard form wich is

$$(P_s) : \begin{cases} \min_{x\in\mathbb{R}^n} c^T x \\ -x \leq 0 \\ Ax = 0 \end{cases} .$$

The associated lagragian is

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

Then

$$\inf_{x\in\mathbb{R}^n} \mathcal{L}(x, \lambda, \nu) = \inf_{x\in\mathbb{R}^n} (c - \lambda + A^T\nu)^T x = \begin{cases} 0 & \text{if} \quad c - \lambda + A^T\nu = 0 \\ -\infty & \text{elif} \end{cases}$$

Which is :

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

The dual problem is then

$$\begin{cases} \max g(\lambda, \nu) \\ \lambda \geq 0 \end{cases} \longleftrightarrow \begin{cases} \max 0 \\ c + A^T\nu \geq 0 \end{cases}$$

2. Noticing that for any function f we have the following relation : $\max f = -\min -f$ the first line of (D) becomes $\min -b^T y$.

The standard form is :

$$\begin{cases} \min_{y\in\mathbb{R}^n} -b^T y \\ A^T y - c \leq 0 \end{cases} \longleftrightarrow \begin{cases} \min_{x\in\mathbb{R}^n} -b^T x \\ A^T x - c \leq 0 \end{cases}$$

Since there are no equality constraints we omit the dependence on η.

The associated Lagragian is :

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

The function $\mathcal{L}(., \lambda)$ is affine, it follows that if $-b + A\lambda = 0$ then $g(\lambda) = -\lambda^T c$ and $-\infty$ elsewhere, so $dom(g) = \{\lambda \in \mathbb{R}^n \quad A\lambda - b = 0\}$.

The dual of (D) is by definition

$$\begin{cases} \max_{\lambda\in\mathbb{R}^n} g(\lambda) \\ \lambda \geq 0 \end{cases} \longleftrightarrow \begin{cases} \max_{\lambda\in\mathbb{R}^n} -\lambda^T c \\ \lambda \geq 0 \end{cases}$$

3. The lagrangian associated is :

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

Wich is affine, so $g(\lambda_1,\lambda_2,\nu) = \nu^T b - \lambda_2^T c$ if $c - A^T\nu - \lambda_1 = 0$ and $A\lambda_2 - b = 0$ and $-\infty$ else. The dual of SD is then,

$$\begin{cases} \max \nu^T b - \lambda_2^T c \\ (\lambda_1,\lambda_2) \geq 0 \quad c - A^T\nu - \lambda_1 = 0 \quad A\lambda_2 - b = 0 \end{cases}$$

which is equivalent to

$$\begin{cases} \min -\nu^T b + \lambda_2^T c \\ (\lambda_1,\lambda_2) \geq 0 \quad c - A^T\nu - \lambda_1 = 0 \quad A\lambda_2 - b = 0 \end{cases}$$

By the symmetrie of the scalar product we have : $\lambda_2^T c - \nu^T b = c^T\lambda_2 - b^T\nu$ And since $\lambda_1 = c - A^T\nu \geq 0$ we finally get to :

$$\begin{cases} \min -\nu^T b + \lambda_2^T c \\ \lambda_2 \geq 0 \quad A^T\nu \leq c \quad A\lambda_2 = b \end{cases} = (SD).$$

4. Assume that $x^\star$ solves (P) and $y^\star$ solves (D). We would like to show that $[x^\star, y^\star]$ solves (Self-Dual).

   First by hypothesis we have $Ax^\star = b; \quad x^\star \geq 0$ and $A^T y^\star \leq c$ which means that $[x^\star, y^\star]$ verifies the constraints of (Self-Dual).

   Then we have also by construction that

$$\forall x \quad c^T x \geq c^T x^\star \tag{1}$$

   and

$$\forall y \quad b^T y \leq b^T y^\star \tag{2}$$

   This means that for any x, y we have

$$\underbrace{c^T(x^\star - x)}_{\underbrace{\leq}_{(1)} 0} - \underbrace{b^T(y^\star - y)}_{\underbrace{\geq 0}_{(2)}} \leq 0 \Longleftrightarrow c^T x^\star - b^T y^\star \leq c^T x - b^T y$$

   Which is exactly that $[x^\star, y^\star]$ solves (Self-Dual).

5. Let's show that for any (x, y) solution of (Self-Dual) we have :

$$c^T x - b^T y \geq 0$$

   and that we can construct one $(\tilde{x}, \tilde{y})$ veryfing the constraints such that

$$c^T\tilde{x} - b^T\tilde{y} = 0.$$

Exercice 2

1. Let $||y||_\star := \sup_{||x||\leq 1} y^T x$. We have to show that if $f = ||.||_1$ then

$$f^\star(y) = \begin{cases} 0 & ||y||_\star \leq 1 \\ +\infty & ||y||_\star > 1 \end{cases} \tag{3}$$

Assume that $||y||_\star \leq 1$ so for any $x \in \mathbb{R}^n$ we have that

$$y^T x \leq ||x||\,||y||_\star \leq ||x|| \iff y^T x - ||x|| \leq 0$$

with equality if $||x|| = 0$ which means that

$$\sup_{x \in \mathbb{R}^n} y^T x - ||x|| = 0 \quad \text{if} \quad ||y||_\star \leq 1.$$

Let y be fixed such that $||y||_\star > 1$ holds. Since

$$K = \{x \in \mathbb{R}^n, \quad ||x|| \leq 1\}$$

is compact and $x \mapsto y^T x$ is continuous by the Hein's theorem there exists $x_0(y) \in K$ such that

$$||y||_\star = y^T x_0(y) > 1 \iff \exists \eta > 0 \quad y^T x_0(y) = 1 + \eta. \tag{4}$$

In the following, we omit the dependence on y. Then define

$$\forall n \geq 0 \quad x_n = n x_0.$$

It follows that,

$$y^T x_n - ||x_n|| = n\left(y^T x_0 - ||x_0||\right) \underbrace{\geq}_{(4)} n \underbrace{\eta}_{>0} \to \infty$$

This shows (3).

2. We introduce a new variable $y = Ax - b$ it follow (RLS) is :

$$min_{x,y}\, ||y||^2 + ||x|| \quad s.t \quad y - Ax - b = 0 \tag{5}$$

The associated lagrangian function is

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

Let first minimize according to y it follows from $\nabla_y \mathcal{L} = 2y + \nu$ that

$$\inf_y \mathcal{L}(x, y, \nu) = \inf_y ||y||^2 + ||x|| + \nu^T(y - Ax - b) = \mathcal{L}(x, \frac{-\nu}{2}, \nu)$$

Which is ;

$$\inf_y \mathcal{L}(x, y, \nu) = -\frac{\nu^T \nu}{4} + ||x|| - \nu^T Ax - \nu^T b.$$

Now we minimize according to x; $\inf_x ||x|| - \nu^T Ax = -\sup_x (A^T\nu)^T x - ||x||$ by the previous question keep the condition $\sup_{|x| \leq 1} \nu^T Ax \leq 1$ to have, in this condition that the dual problem is $g(\nu) = -\frac{\nu^T\nu}{4} - \nu^T b$ subject to $\sup_{|x|\leq 1} \nu^T Ax \leq 1$

Exercice 3

- We writte that $(sep1)$ is equivalent to

$$\min_{w,z} \underbrace{\frac{1}{n\tau} \sum_{1 \leq i \leq n} z_i}_{(a)} + \frac{1}{2}\, ||w||_2^2 \,\text{s.t} \quad \underbrace{\forall 1 \leq i \leq n \quad z_i = L(w, x_i, y_i)}_{(b)}$$

  (a) is equal to $\frac{1}{n\tau} 1^T z$ and since

$$z_i = L(w, x_i, y_i) = \max 0, 1 - y_i w^T x_i \geq 0 \quad \text{and} \quad z_i \geq 1 - y_i w^T x_i,$$

  (b) is written as the constraint in $(sep2)$ which is exactly $(sep2.)$

- The lagrangian is given by :

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

Minimizing first according to w and setting the gradient equal to 0 and noticing that for z the function is linear, after computation we find :

$$g(w,z,\lambda,\pi) = -2^{-1}\sum_{1\leq i\leq n} \lambda_i^2 y_i^2 \left|\left|x\right|\right|_2^2 + 1^T z \quad \text{if} \quad \frac{1}{n\tau}1 - \lambda - \pi = 0, \quad -\infty \quad \text{else.}$$

The dual problem is then

$$\max_{\lambda,\pi} -2^{-1}\sum_{1\leq i\leq n} \lambda_i^2 y_i^2 \left|\left|x\right|\right|_2^2 + 1^T z \quad \text{s.t} \quad \frac{1}{n\tau}1 - \lambda - \pi = 0 \quad \frac{1}{n\tau} - \lambda - \pi = 0 \quad \lambda \geq 0.$$

Exercice 5

1. Applying the definition of the Lagrangian function we find :

$$L(x,\mu,\nu) = c^T x + \mu^T(Ax - b) - \nu^T x + x^T \mathrm{D}(\nu)x = x^T \mathrm{D}(\nu)x + (c + A^T\mu - \nu)^T x - b^T\mu.$$

Where $D(\nu) = Diag(\nu_i)_{1\leq i\leq n}$ To find the dual function we minimize over x : Since $L(.,\mu,\nu)$ is a convex function it suffices to solve (when $\nu \geq 0$) :

$$\nabla_x L = 0 \implies 2D(\nu)x + (c + A^T\nu - \eta) = 0 \implies x = -\frac{1}{2}D((\nu_i^{-1})_i)(c + A^T\nu - \eta)$$

Plugging this into L gives us ;

$$g(\mu,\nu) = \begin{cases} -b^T\mu - \frac{1}{4}\sum_{i=1}^n (c_i + C_i^T\mu - \nu_i)^2(\nu_i)^{-1} & \text{if } \nu \geq 0, \\ -\infty & \text{otherwise,} \end{cases}$$

where $C_i := Col_i(A)$ The resulting dual problem is

$$\text{maximize } \ -b^T\mu - \frac{1}{4}\sum_{i=1}^n (c_i + C_i^T\mu - \nu_i)^2(\nu_i)^{-1}$$

$$\text{subject to } \nu \geq 0.$$

Then, we apply the suggestion given in the exercice to eliminate η :

$$\text{maximize } \ -b^T\mu + \sum_{i=1}^n \min\{0, c_i + C_i^T\mu\} \text{subject to } \mu \geq 0.$$

2. We writte the condition

$$0 \leq x_i \leq 1 \longleftrightarrow -x_i \leq 0 \quad x_i - 1 \leq 0$$

to be in the standard form, this gives us the lagrangian function :

$$L(x,\alpha,\beta,\omega) = c^T x + \alpha^T(Ax - b) - \beta^T x + \omega^T(x - 1) = (c + A^T\alpha - \beta + \omega)^T x - \beta^T\alpha - \omega^T.1$$

$$g(\alpha,\beta,\omega) = \begin{cases} -b^T\alpha - \omega^T 1 & \text{if } A^T\alpha - \beta + \omega + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual is therefore :

$$\begin{cases} \max_{\alpha,\beta,\omega\in\mathbb{R}^n} -b^T\alpha - \omega^T 1 \\ A^T\alpha - \beta + \omega + c = 0 \\ \alpha \geq 0, \beta \geq 0, \omega \geq 0 \end{cases}$$

Since $max(-f) = -min(f)$ the dual problem is equivalent to the lagrangian relaxation problem so they have the same lower bound.