DM 2 — Optimisation Convexe

Titouan Le Breton

Exercice 1

1. **Dual de (P).**

   On introduit $\mu \in \mathbb{R}^n$ pour $Ax = b$ et $\lambda \in \mathbb{R}^d_+$ pour $x \geq 0$. Le Lagrangien est

   $$L(x, \lambda, \mu) = c^\top x - \lambda^\top x + \mu^\top (Ax - b) = (c - \lambda + A^\top \mu)^\top x - \mu^\top b.$$

   La fonction duale vaut

   $$g(\lambda, \mu) = \begin{cases} -\mu^\top b & \text{si } c - \lambda + A^\top \mu = 0,\ \lambda \geq 0, \\ -\infty & \text{sinon.} \end{cases}$$

   En éliminant $\lambda = c + A^\top \mu$ ($\lambda \geq 0$) et en posant $y = -\mu$, on obtient

   $$\boxed{(\mathrm{D})\ \max_y\ b^\top y \quad \text{s.c.} \quad A^\top y \leq c.}$$

2. **Dual de (D).**

   On a :

   $$\max_y\ b^\top y \quad \text{s.c.} \quad A^\top y \leq c = -(\min_y\ -b^\top y \quad \text{s.c.} \quad A^\top y \leq c)$$

   On pénalise $A^\top y - c \leq 0$ par $\lambda \in \mathbb{R}^d_+$. Le Lagrangien

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

   donne

   $$g(\lambda) = \inf_y L(y, \lambda) = \begin{cases} -\lambda^\top c & \text{si } A\lambda = b, \\ -\infty & \text{sinon.} \end{cases}$$

   Donc le dual de (D) est

   $$-\max_{\lambda \geq 0}\ -\lambda^\top c \quad \text{s.c.} \quad A\lambda = b,$$

   Soit :

   $$\boxed{\min_{\lambda \geq 0}\ \lambda^\top c \quad \text{s.c.} \quad A\lambda = b,}$$

   qui n'est autre que (P) (en identifiant λ à x).

3. **Auto-dualité du problème**

   $$(\mathrm{SD}) \quad \min_{x,y}\ c^\top x - b^\top y \text{ s.c. } x \geq 0, \quad A^\top y \leq c, \quad Ax = b.$$

   On peut séparer

   $$(\mathrm{SD}) = \underbrace{\min_x \{ c^\top x \mid x \geq 0,\ Ax = b \}}_{\text{valeur } p^\star} - \underbrace{\max_y \{ b^\top y \mid A^\top y \leq c \}}_{\text{valeur } d^\star}.$$

Le Lagrangien de (SD) est

$$L(x, y; \lambda, \tilde{\lambda}, \mu) = c^\top x - b^\top y \ - \ \lambda^\top x \ + \ \tilde{\lambda}^\top (A^\top y - c) \ + \ \mu^\top (Ax - b),$$

où $\lambda \geq 0$ est associé à $x \geq 0$, $\tilde{\lambda} \geq 0$ à $A^\top y \leq c$, et μ à $Ax = b$. En regroupant :

$$L = (c - \lambda + A^\top \mu)^\top x \ + \ (-b + A\tilde{\lambda})^\top y \ - \ \tilde{\lambda}^\top c \ - \ \mu^\top b.$$

L'infimum en (x, y) est fini ssi

$$c - \lambda + A^\top \mu = 0, \qquad -b + A\tilde{\lambda} = 0,$$

avec $\lambda, \tilde{\lambda} \geq 0$. La fonction duale vaut alors

$$g(\lambda, \tilde{\lambda}, \mu) = -\tilde{\lambda}^\top c \ - \ \mu^\top b.$$

Le **dual** de (SD) s'écrit donc

$$\max_{\lambda, \tilde{\lambda} \geq 0, \ \mu} \ -\mu^\top b - \tilde{\lambda}^\top c \quad \text{s.c.} \quad \lambda = c + A^\top \mu, \quad A\tilde{\lambda} = b.$$

En éliminant λ et en posant $x := \tilde{\lambda}$ et $y := -\mu$, on obtient

$$\max_{x \geq 0, \ y} \ b^\top y - c^\top x \quad \text{s.c.} \quad A^\top y \leq c, \quad Ax = b,$$

Puis en passant à l'opposé :

$$\min_{x \geq 0, \ y} \ -b^\top y + c^\top x \quad \text{s.c.} \quad A^\top y \leq c, \quad Ax = b,$$

Ainsi,

$$\boxed{\text{(SD) est auto-dual.}}$$

4. **Lien avec la dualité forte.**

Comme vu en 3, le problème (SD) peut s'écrire comme la somme des valeurs optimales de (P) et (D) :

$$\text{(SD)} \ = \ p^\star - d^\star,$$

où $p^\star$ et $d^\star$ désignent respectivement les valeurs optimales des problèmes (P) et (D).

Or, (P) est le dual de (D) et réciproquement. Étant linéaires, ces deux programmes vérifient la dualité forte :

$$p^\star = d^\star$$

Ainsi, si $x^\star$ et $y^\star$ sont des solutions optimales de (P) et (D), on a

$$c^\top x^\star = b^\top y^\star.$$

En substituant dans (SD), la valeur optimale vaut donc

$$\boxed{c^\top x^\star - b^\top y^\star = 0.}$$

De plus, le couple $\left(x^\star, y^\star\right)$ est solution optimale du problème (SD).

Exercice 2

1. **Conjugué de $\|x\|_1$**

   On considère la fonction

$$f(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

Son conjugué de Fenchel est défini par

$$f^*(y) = \sup_x \big(y^\top x - f(x)\big) = \sup_x \sum_{i=1}^{n} (y_i x_i - |x_i|).$$

Analyse coordonnée par coordonnée. Si $y_i > 1$, on peut prendre $x_i = t > 0$, nulle ailleurs. Alors :

$$y_i x_i - |x_i| = t(y_i - 1) \longrightarrow +\infty \quad \text{quand } t \to +\infty.$$

Si $y_i < -1$, on prend $x_i = -t < 0$, et

$$y_i x_i - |x_i| = t(-y_i - 1) \longrightarrow +\infty \quad \text{quand } t \to +\infty.$$

Ainsi, pour que $f^*(y)$ soit fini, il faut nécessairement $|y_i| \le 1$ pour tout i.

Cas où $\|y\|_\infty \le 1$. Pour tout x, on a $|y_i x_i| \le |x_i|$, donc

$$y_i x_i - |x_i| \le 0 \quad \forall i, \qquad \text{et ainsi} \qquad y^\top x - \|x\|_1 \le 0.$$

L'égalité est atteinte pour $x = 0$, donc $f^*(y) = 0$.

$$\boxed{f^*(y) = \begin{cases} 0, & \text{si } \|y\|_\infty \le 1, \\ +\infty, & \text{sinon.} \end{cases}}$$

2. **Dual du problème (RLS)**

   On considère le problème

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

On introduit une variable y telle que $y = Ax - b$, soit équivalemment la contrainte

$$y - Ax + b = 0.$$

Le Lagrangien associé (multiplicateur dual $\mu \in \mathbb{R}^n$) est alors

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

La fonction duale s'écrit

$$g(\mu) = \inf_{x,y} L(x, y, \mu) = \mu^\top b + \inf_y (\|y\|_2^2 + \mu^\top y) + \inf_x (\|x\|_1 - \mu^\top Ax).$$

Minimisation en y. Pour chaque coordonnée,

$$\inf_{y_i} (y_i^2 + \mu_i y_i) = -\frac{\mu_i^2}{4},$$

donc

$$\inf_y (\|y\|_2^2 + \mu^\top y) = -\frac{1}{4}\|\mu\|_2^2.$$

Minimisation en x. On a

$$\inf_x(\|x\|_1 - \mu^\top Ax) = -\sup_x(\mu^\top Ax - \|x\|_1) = -f^*(A^\top\mu).$$

En utilisant le résultat de la question précédente :

$$f^*(A^\top\mu) = \begin{cases} 0, & \text{si } \|A^\top\mu\|_\infty \leq 1, \\ -\infty, & \text{sinon.} \end{cases}$$

Fonction duale. Ainsi,

$$g(\mu) = \begin{cases} \mu^\top b - \frac{1}{4}\|\mu\|_2^2, & \text{si } \|A^\top\mu\|_\infty \leq 1, \\ -\infty, & \text{sinon.} \end{cases}$$

Problème dual. On obtient donc :

$$\boxed{\max_\mu \ \mu^\top b - \frac{1}{4}\|\mu\|_2^2 \quad \text{s.c.} \quad \|A^\top\mu\|_\infty \leq 1.}$$

Exercice 3

1. **Équivalence entre (Sep.1) et (Sep.2).**
   On introduit $z_i = \max\big(0,\, 1 - y_i\,(\omega^\top x_i)\big)$.
   Le problème est alors équivalent (en divisant part τ) :

$$\min_{\omega,z} \frac{1}{n\tau}\, 1^\top z + \frac{1}{2}\|\omega\|_2^2 \quad \text{s.c.} \quad z_i = \max\big(0,\, 1 - y_i\,(\omega^\top x_i)\big)$$

   On remarque que rajouter la contrainte $z \geq 0$ ne change rien au problème. Or, pour $z \geq 0$, la fonction objective est croissante pour chacune des composantes de z. On peut donc reformuler le problème sous la forme :

$$\min_{\omega,z} \frac{1}{n\tau}\, 1^\top z + \frac{1}{2}\|\omega\|_2^2 \quad \text{s.c.} \quad z_i \geq 1 - y_i\,(\omega^\top x_i) \quad \forall i = 1,\dots,n,\ z \geq 0$$

2. **Dual de (Sep.2).**
   On introduit les multiplicateurs de Lagrange :

$$\lambda_i \geq 0 \text{ pour } 1 - y_i(\omega^\top x_i) - z_i \leq 0, \qquad \pi_i \geq 0 \text{ pour } -z_i \leq 0.$$

   Le Lagrangien est

$$L(\omega, z, \lambda, \pi) = \frac{1}{n\tau}\, 1^\top z + \frac{1}{2}\|\omega\|_2^2 + \sum_{i=1}^n \lambda_i\big(1 - y_i\,(\omega^\top x_i) - z_i\big) - \pi^\top z.$$

Infimum sur z. Les termes dépendant de z sont

$$\sum_i z_i\Big(\frac{1}{n\tau} - \lambda_i - \pi_i\Big).$$

Pour que l'infimum soit fini, il faut que les coefficients de z_i soient nuls :

$$\frac{1}{n\tau} - \lambda_i - \pi_i = 0 \quad \Rightarrow \quad \pi_i = \frac{1}{n\tau} - \lambda_i.$$

Comme $\pi_i \geq 0$, on obtient la contrainte

$$\lambda_i \leq \frac{1}{n\tau}.$$

Infimum sur ω. On isole les termes dépendant de ω :

$$L_\omega(\omega) = \frac{1}{2}\|\omega\|_2^2 - \sum_{i=1}^n \lambda_i\, y_i\, x_i^\top \omega = \frac{1}{2}\sum_{j=1}^d \omega_j^2 - \sum_{j=1}^d \omega_j \Big(\sum_{i=1}^n \lambda_i y_i x_{ij} \Big).$$

On peut donc minimiser séparément sur chaque coordonnée ω_j :

$$\inf_{\omega_j} \ \Big(\tfrac{1}{2}\omega_j^2 - \omega_j \sum_{i=1}^n \lambda_i y_i x_{ij} \Big).$$

Il s'agit d'une fonction quadratique convexe dont la dérivée s'annule en

$$\omega_j^\star = \sum_{i=1}^n \lambda_i y_i x_{ij}.$$

En réinjectant ce point stationnaire, la valeur minimale pour chaque j est

$$-\frac{1}{2}\Big(\sum_{i=1}^n \lambda_i y_i x_{ij} \Big)^2.$$

En sommant sur toutes les coordonnées $j = 1, \ldots, d$, on obtient

$$\inf_\omega L_\omega(\omega) = -\frac{1}{2}\sum_{j=1}^d \Big(\sum_{i=1}^n \lambda_i y_i x_{ij} \Big)^2 = -\frac{1}{2}\Big\| \sum_{i=1}^n \lambda_i y_i x_i \Big\|_2^2.$$

Fonction duale. En remplaçant dans le Lagrangien,

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

avec les contraintes $0 \le \lambda_i \le \frac{1}{n\tau}$.

Problème dual. Le dual de (Sep.2) est donc

$$\boxed{\max_{\lambda \in \mathbb{R}^n} \ \sum_{i=1}^n \lambda_i - \frac{1}{2}\Big\| \sum_{i=1}^n \lambda_i y_i x_i \Big\|_2^2 \quad \text{s.c.} \quad 0 \le \lambda_i \le \frac{1}{n\tau}.}$$