GRANGETTE
Paule

$\boxed{\text{DM 2 : Convex Optimization}}$

Exercice 1 :

1) Le probleme (P) sous forme standard est : $\min_x c^T x$
$\text{st: } Ax-b=0, -x\leq 0$

Calculons son Lagrangien associé :

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

D'où :

$$g(\lambda,\nu) = \inf_x \mathcal{L}(x,\lambda,\nu) = \begin{cases} b^T\nu & \text{si } c-\lambda-A^T\nu = 0 \\ -\infty & \text{sinon} \end{cases}$$

g est concave car on prend la borne inférieure du Lagrangien qui est affiné en x.

Le probleme dual de (P) est alors :

$$\max_\nu \; b^T\nu$$
$$\text{st. } c-A^T\nu \geq 0$$

2) Le probleme (D) sous forme standard est : $\min_y -b^T y$
$\text{st. } A^T y - c \leq 0$

Son Lagrangien est :

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

D'où :

$$g(\lambda) = \inf_y \mathcal{L}(y,\lambda) = \begin{cases} -c^T\lambda & \text{si } A\lambda - b = 0 \\ -\infty & \text{sinon} \end{cases}$$

Le probleme dual de (D) est alors :

$$\max -c^T\lambda \qquad \Longleftrightarrow (P)$$
$$\text{st. } \lambda \geq 0 \text{ et } A\lambda - b = 0$$

3) Le Lagrangien du problème (SD) est :

$$\mathcal{L}(x, y, \lambda, \mu, \nu) = c^T x - b^T y + \lambda^T (Ax - b) - \mu x + \nu^T (A^T y - c)$$
$$= \left[c^T x + \lambda^T (Ax + b) - \mu x\right] + \left[-b^T y + \nu^T (A^T y - c)\right]$$
$$= \mathcal{L}_{(P)}(x, \lambda, \mu) + \mathcal{L}_{(D)}(y, \nu)$$

donc :

$$g(\lambda, \mu, \nu) = \inf_{x,y} \mathcal{L}(x, y, \lambda, \mu, \nu) = \inf_x \mathcal{L}_{(P)}(x, \lambda, \mu) + \inf_y \mathcal{L}_{(D)}(y, \nu)$$
$$= \begin{cases} b^T \mu - c^T \nu & \text{si } c - \lambda - A^T \mu = 0 \text{ et } A\lambda - b = 0 \\ -\infty & \text{sinon} \end{cases}$$

Donc le problème dual est :

$$\max_{\mu, \nu} \; b^T \mu - c^T \nu \quad \text{s.t.} \quad \begin{cases} c \geq A^T \mu \\ A\lambda = b \\ \lambda \geq 0 \end{cases}$$

qui est exactement le problème primal.

4) • Comme il y a dualité forte, la solution optimale du problème dual est égale à celle du problème primal.
Or, (D) est le dual de (P)
donc $[x^*, y^*]$ peut être obtenu en résolvant (P) et/ou (D).

• 0 est clairement dans l'épigraphe du problème (SD)
or, (SD) est son propre dual et l'intersection des épigraphes des problèmes primal et dual est exactement égal à la solution optimale.
Donc la solution optimale de (SD) est exactement 0.

Exercice 2 :

1) Soit $f(x) = \|x\|_1$

$$f^*(x) = \sup_{y \in \mathbb{R}^d} \left(x^T y - f(y)\right) = \sup_{y \in \mathbb{R}^d} \left(x^T y - \|y\|_1\right)$$

On considère la norme duale définie par :

$$\|x\|_* = \sup_{\|u\| \leq 1} u^T x$$

- Si $\|x\|_* > 1$, alors il existe $z \in \mathbb{R}^d$ tel que $x^T z > 1$

Soit $t \in \mathbb{R}$ tel que $x = tz$; on a :

$$x^T y - \|y\|_1 = t\left(z^T y - \|z\|_1\right) \xrightarrow[t \to \infty]{} \infty$$

Donc $f^*(x) = \infty$ quand $\|x\|_* > 1$

- Si $\|x\|_* \leq 1$, on a par l'inégalité de Cauchy-Schwarz

$$x^T y \leq \|x\|_* \|y\|_1 \quad \text{donc} \quad x^T y - \|y\|_1 \leq 0$$

De plus, si $y = 0$, la borne est atteinte.

En conclusion :

$$f^*(x) = \begin{cases} 0 & \text{si } \|x\|_* \leq 1 \\ \infty & \text{si } \|x\|_* > 1 \end{cases}$$

2) On a :

$$\min_x \|Ax - b\|_2^2 + \|x\|_1 \iff \min_{x,y} \; y^T y + \|x\|_1 \quad \text{s.t. } y = Ax - b$$

Le Lagrangien du problème est :

$$\mathcal{L}(x,y,\nu) = y^T y + \|x\|_1 + \nu^T y - \nu^T A x + \nu^T b$$
$$= \left(y^T y + \nu^T y\right) + \left(\|x\|_1 - \nu^T A x\right) + \nu^T b$$

or, $\inf_x \left(\|x\|_1 - \nu^T A x\right) = \begin{cases} 0 & \text{si } \|A^T \nu\|_* \leq 1 \\ -\infty & \text{sinon} \end{cases}$ (d'après 1))

et $\mathcal{L}_y(y) = y^T y + \nu^T y$

$\nabla \mathcal{L}_y(y) = 2y^T + \nu^T$

donc $\nabla \mathcal{L}_y(y) = 0 \iff y = -\frac{1}{2}\nu$

D'où $\inf_y \left(y^T y + \nu^T y \right) = -\frac{1}{4} \nu^T \nu$

D'où $g(\nu) = \inf_{x,y} \mathcal{L}(x,y,\nu) = \begin{cases} -\frac{1}{4}\nu^T\nu + \nu^T b & \text{si } \|A^T\nu\|_* \leq 1 \\ -\infty & \text{sinon} \end{cases}$

Donc le problème dual de (RLS) est :

$$\max_\nu \; -\frac{1}{4} \nu^T \nu + \nu^T b$$
$$\text{s.t. } \|A^T\nu\|_* \leq 1$$

Exercice 3 :

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

on pose $z_i = \max(0, 1-y_i(w^T x_i))$

on a $z_i \geq 0$ et $z_i \geq 1-y_i(w^T x_i)$

z_i est un majorant de 0 et $1-y_i(w^T x_i)$

mais il atteint une de ces 2 valeurs car il est défini comme un maximum.

Ainsi, "z_i" est le plus petit des majorants de $\{0, 1-y_i(w^T x_i)\}$. ie : $z_i = \min_{\substack{\gamma \geq 0 \\ \gamma \geq 1-y_i(w^T x_i)}} \gamma$

Soit $z = (z_i)_{1 \leq i \leq n} \in \mathbb{R}^n$

$$\sum_{i=1}^{n} \max(0, 1-y_i(w^T x_i)) = \mathbb{1}^T z$$

De plus, τ étant constante >0, on peut diviser l'expression par τ, le minimum reste inchangé.

Donc (Sep 1) est équivalent à :

$$\min_{w,z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 \qquad \text{(Sep 2)}$$
$$\text{st.} \begin{cases} z_i \geq 1-y_i(w^T x_i) \\ z_i \geq 0 \end{cases}$$

HW 2

2) Calculons le Lagrangien associé à (Sep 2).

$$\mathcal{L}(w,z,\lambda,\pi) = \frac{1}{n\tau}\mathbb{1}^T z + \frac{1}{2} w^T w + \sum_{i=1}^{n} \lambda_i (1 - y_i (\underbrace{w^T x_i}_{x_i^T w}) - z_i) - \sum_{i=1}^{n} \pi_i z_i$$

$$= \frac{1}{n\tau}\mathbb{1}^T z - \lambda^T z - \pi^T z + \mathbb{1}^T \lambda + \frac{1}{2n}\sum_{i=1}^{n} w^T w - 2n y_i \lambda_i x_i^T w$$

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

or, d'après la question 1 de l'exercice 2,

$$\inf_w \left(w^T w - 2n y_i \lambda_i x_i^T w\right) = \begin{cases} 0 & \text{si } \|x_i\|_* \leq \frac{1}{2n y_i \lambda_i} \\ -\infty & \text{sinon} \end{cases}$$

et

$$\inf_z \left(\frac{1}{n\tau}(\mathbb{1} - \lambda - \pi)^T z + \mathbb{1}^T \lambda\right) = \begin{cases} \mathbb{1}^T \lambda & \text{si } \frac{1}{n\tau}\mathbb{1} - \lambda - \pi = 0 \\ -\infty & \text{sinon} \end{cases}$$

car la fonction à minimiser est affine.

Ainsi :

$$g(\lambda,\pi) = \begin{cases} \mathbb{1}^T \lambda & \text{si } \|x_i\|_* \leq \frac{1}{2n y_i \lambda_i}, \forall i \in \{1,..,n\} \text{ et } \frac{1}{n\tau}\mathbb{1} - \lambda - \pi = 0 \\ -\infty & \text{sinon} \end{cases}$$

Donc le problème dual est :

$$\max_\lambda \ \mathbb{1}^T \lambda$$

$$\text{s.t.} \begin{cases} \|x_i\|_* \leq \frac{1}{2n y_i \lambda_i}, \forall i \in \{1,..,n\} \\ \frac{1}{n\tau}\mathbb{1} - \lambda - \pi = 0 \\ \lambda \geq 0 \\ \pi \geq 0 \end{cases}$$

Exercice 5 :

1) Le Lagrangien associé au problème est :

$$\mathcal{L}(x,\lambda,\mu) = c^T x + \lambda^T(Ax-b) - \sum_{i=1}^{n} \mu_i x_i (1-x_i)$$

$$= (A^T\lambda + c)^T x - \lambda^T b \quad x^T \mathrm{diag}(\mu_i)(\mathbb{1}-x)$$

NB : $\mathrm{diag}(\mu_i) = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$ et $\mu = (\mu_i)_{1\le i\le n} \in \mathbb{R}^n$

$$= (A^T\lambda + c - \mu)^T x - \lambda^T b + x^T \mathrm{diag}(\mu) x$$

$$\nabla_x \mathcal{L}(x,\lambda,\mu) = A^T\lambda + c - \mu \quad - 2\, \mathrm{diag}(\mu)^T x$$

$\mathrm{diag}(\mu)$

$$\nabla_x \mathcal{L}(x,\lambda,\mu) = 0 \iff x = \frac{1}{2}\mathrm{diag}(1/\mu_i)(A^T\lambda + c - \mu) \triangleq x_0$$

si $\mu_i > 0, \forall i$

D'où :

$$g(\lambda,\mu) = \inf_{x} \mathcal{L}(x,\lambda,\mu) = \mathcal{L}(x_0;\lambda,\mu)$$

$$= \begin{cases} -\lambda^T b + \frac{1}{4}(A^T\lambda + c - \mu)^T \mathrm{diag}(1/\mu_i)\,\mathrm{diag}(\mu_i)\,\mathrm{diag}(1/\mu)\times(A^T\lambda + c - \mu) \\ \quad - 1/2\,(A^T\lambda + c-\mu)^T (A^T\lambda + c - \mu) & \text{si } \mu \geq 0 \\ -\infty & \text{sinon} \end{cases}$$

$a_i = i^e$ colonne de A

$$= \begin{cases} -\lambda^T b - 1/4 \sum_{i=1}^{n} \frac{(a_i^T\lambda + c_i + \mu_i)^2}{\mu_i} & \text{si } \mu \geq 0 \\ -\infty & \text{sinon} \end{cases}$$

Le problème dual est alors :

$$\sup_{\lambda,\mu} -\lambda^T b - 1/4 \sum_{i=1}^{n} \frac{(a_i^T\lambda + c_i - \mu_i)^2}{\mu_i}$$

s.t. $\mu \geq 0, \; , \lambda \geq 0$

Or, d'après l'indice, $\sup_{\mu_i \geq 0} \frac{(a_i^T\lambda + c_i - \mu_i)^2}{\mu_i} = \max\{0, a_i^T\lambda + c_i\}$

On obtient finalement le problème dual suivant :

$$\begin{cases} \sup_{\lambda} \; -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n} \max(0, a_i^T\lambda + c_i) \\ \text{s.t.} \quad \lambda \succeq 0 \end{cases}$$

2) Le Lagrangien du problème LP relaxation est :

$$\mathcal{L}(x, \lambda, \mu, \nu) = c^Tx + \lambda^T(Ax-b) + \sum_{i=1}^{n} \nu_i(x_i - 1) - \mu_i x_i$$

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

Le Lagrangien est affine, alors :

$$g(\lambda, \mu, \nu) = \inf_x \mathcal{L}(x, \lambda, \mu, \nu) = \begin{cases} -\lambda^T b - \nu^T\mathbb{1} & \text{si } A^T\lambda + c + \nu - \mu = 0 \\ -\infty & \text{sinon} \end{cases}$$

donc le dual est :

$$\begin{aligned} \sup_{\lambda, \nu} \; & -\lambda^T b - \nu^T \mathbb{1} \\ \text{s.t.} \; & A^T\lambda + c + \nu - \mu = 0 \\ & \lambda \succeq 0, \; \mu \succeq 0, \; \nu \succeq 0 \end{aligned}$$

$$\iff \begin{aligned} \sup_{\lambda, \nu} \; & -\lambda^T b - \nu^T \mathbb{1} \\ \text{s.t.} \; & A^T\lambda + c + \nu \succeq 0 \\ & \lambda \succeq 0, \; \nu \succeq 0 \end{aligned}$$

Ce problème est équivalent au dual précédent. Donc, les bornes inférieures des deux problèmes de relaxations sont identiques.

Exercice 4

Considérons le problème :

$$\sup_{a\in P} a^T x \leq b \iff \sup_a \; a^T x - b \leq 0 \quad \text{s.t.} \; C^T a - d \preceq 0$$

$$\iff \inf_a \; -a^T x + b \geq 0 \quad \text{s.t.} \; C^T a - d \preceq 0$$

Trouvons le dual de $\inf_a -a^T x + b$ s.t. $C^T a - d \preceq 0$

Le Lagrangien est alors :

$$\mathcal{L}(a,\lambda) = -a^T x + b + \lambda^T (C^T a - d)$$

$$= (C\lambda - x)^T a + b - \lambda^T d$$

donc :

$$g(\lambda) = \inf_a \mathcal{L}(a,\lambda) = \begin{cases} b - \lambda^T d & \text{si } C\lambda = x \\ -\infty & \text{sinon} \end{cases}$$

car le Lagrangien est une fonction affine de a.

donc le problème dual est :

$$\sup_\lambda \; b - \lambda^T d \quad \text{s.t.} \; \begin{cases} C\lambda = x \\ \lambda \succeq 0 \end{cases}$$

Le problème dual et primal ont le même optimum car il s'agit d'un problème LP.

Ainsi le problème initial :

$$\min_x c^T x \quad \text{s.t.} \; \sup_{a\in P} a^T x \leq b$$

équivaut à :

$$\min_x c^T x \quad \text{s.t.} \; \begin{cases} b \geq \lambda^T d \\ C\lambda = x \\ \lambda \succeq 0 \end{cases}$$