Convex Optimization DM n°2

Gravier Thomas

Professeur: Alexandre d'Aspremont

Exercice n°1

Soit $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times d}$

(P) $\min_x c^T x$ st $Ax = b$, $x \geq 0$

(D) $\max_y b^T y$ st $A^T y \leq c$

1) (P) dans sa forme standart devient

$$\min_x c^T x \quad \text{st} \quad Ax - b = 0, \quad -x \leq 0$$

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

La fonction dual est $g(\lambda, \mu) = \inf_x L(x, \lambda, \mu)$, elle est concave car c'est l'infimum ponctuel de fonctions affines.

On peut donc chercher son gradient en 0.

$\nabla \mathcal{L}_x (x, \mu, \lambda) = c - \lambda + A^t \mu = 0$

Donc $c = \lambda - A^t \mu$ donc $g(\lambda, \mu) = -\mu^T b$

Si on factorise comme dans le cour on aurait eu le même résultat

Le problème dual est donc $\max g(\lambda, \mu) = \max_{\lambda, \mu} \begin{cases} -\mu^T b \text{ avec } c = \lambda - A^t \mu \\ \infty \text{ sinon} \end{cases}$

On peut le reformuler en $\max_{\lambda \mu} -\mu^T b$ st $\lambda \geq 0$, $c - \lambda - A^t \mu = 0$

2) On met D dans sa forme standart

$$\min_y -b^T y \quad \text{st} \quad A^T y \leq c \iff A^T y - c \leq 0$$

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

De la même façon qu'à la question 1

$\nabla_y \mathcal{L}(y, \lambda) = -b + A^* \lambda = 0 \implies b = A\lambda$.

$\implies g(\lambda) = \begin{cases} -\lambda^T c & \text{si } b = A\lambda \\ \infty & \text{sinon} \end{cases}$

Le problème dual est donc $\max_\lambda -\lambda^T c$ st $\lambda \geq 0$, $b - A\lambda = 0$

3) On veut calculer le dual de $\min_{x,y} c^T x - b^T y$ st $b - Ax = 0$, $-x \leq 0$, $A^T y - c \leq 0$

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

Donc $g(\mu, \lambda_1, \lambda_2) = \inf_{x,y} (c^T - \mu^T A - \lambda_1^T) x + (-b^T + \lambda_2^T A^T) y + \mu^T b - \lambda_2^T$

On a deux cas possible

- si $-b^T + \lambda_2^T A^T = 0$ et $c^T - \mu^T A - \lambda_1^T = 0$ alors $g(\mu, \lambda_1, \lambda_2) = \mu^T b - \lambda_2^T c$
- sinon $g(\mu, \lambda_1, \lambda_2) = -\infty$

Alors (SD) peut être réécrit en :

$$\max_{\mu,\lambda_1,\lambda_2} \mu^T b - \lambda_2^T c$$

st $A^T\mu + \lambda_1 = c$

$A\lambda_2 = b$

$\lambda_2 \geq 0$

$\lambda_1 \geq 0$

Avec $\lambda_2 = x$, $\mu = y$

Qui devient $\min_{x,y,\lambda_1} x^T c - y^T b$

$Ax = b$

$A^T y + \lambda_1 = c$

$\lambda_1, x \geq 0$

Sachant que dans $\mathbb{R}$ la transposé est égal à elle même pour une fonction à valeurs dans $\mathbb{R}$. De plus le problème ne dépend pas de λ_1

Le problème devient $\min_{x,y} c^T x - b^T y$

st $Ax = b$

$\begin{cases} A^T y + \lambda_1 = c \\ \lambda_1 \geq 0 \end{cases} \iff \{ A^T y \leq c$

$x \geq 0$

On a donc prouvé que le dual de (SD) est (SD)

4) Le self dual est $\min_{x,y} c^T x - b^T y$

st $Ax = b$

$x \geq 0$

$A^T y \leq c$

On se rend compte que ce problème minimise une variable à la fois sous des contraintes indépendantes

On peut réécrire le problème SD en :

$$(P)\ \min_x c^T x \text{ st } Ax = b,\ x \geq 0 \quad + \quad \max_y b^T y,\ A^T y \leq c\ (D)$$

Donc si x^* est l'optimale de P il le sera aussi pour (SD)

La même chose est valable pour D avec y^*

Le dual de $\min_x c^T x$ st $Ax = b$, $x \geq 0$ d'après la partie 1 est $\max_{\lambda,\mu} -\mu^T b$ st $\lambda \geq 0$, $c - \lambda + A^T\mu = 0$

qui est équivalent à $\max_y y^T b$ $A^T y \leq c$ $\quad$ $\max -b^T z$ st $-A^T z \leq c$

Donc $\min_{x,y} c^T x - b^T y$ st $Ax = b$, $x \geq 0$, $A^T y \leq c$ $= \underbrace{\min_x c^T x \text{ St } Ax = b,\ x \geq 0}_{P} - \underbrace{\min b^T y\ A^T y \leq c}_{P^*}$ si on pose $z = -y$

~~Avec la dualité forte~~

* les contraintes de P sont linéaire, pareil pour les inégalités.
* (P) est ~~faisable~~ faisable
* $\mathrm{dom}(c^T x)$ ~~est~~ $= \mathbb{R}^d$ qui est un ouvert

Avec les conditions de Slater si on définit

* p^* comme solution de (P)
* d^* comme solution de (P^*)

Alors par dualité forte $p^* = d^*$

Donc $\min c^T x - b^T y = p^* - d^* = 0$

st $Ax = b$

$x \geq 0$

$A^T y \leq c$

Exercice n°2 :

1) Soit $x, y \in \mathbb{R}^d$, on définit $f(x) := \|x\|_1$

Alors $f^*(y) = \sup_{x \in \text{dom} f} y^T x - f(x) = \sup_{x \in \text{dom} f} \sum_{i=1}^{d} (y_i x_i - |x_i|)$

On suppose $\exists j \in [\![1, d]\!]$ tel que $y_j > 1$. On choisit $x_j = t > 0$ et $x_i = 0$ pour $i \neq j$. On a $y^T x - f(x) = (y_j - 1)t \xrightarrow[t \to +\infty]{} +\infty$

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

De la même façon avec $y_j < -1$, $t < 0$ on aurait $f^*(y) = +\infty$

Supposons que $\|y\|_\infty \leq 1$:

$y^T x - f(x) \leq \sum_{i=1}^{d} |y_i x_i| - \sum_{i=1}^{d} |x_i|$

$= \sum_{i=1}^{d} \underbrace{|x_i|}_{\geq 0} \underbrace{(|y_i| - 1)}_{\leq 0} \leq 0$ car $\|y\|_\infty \leq 1$

l'inégalité est toujours valable avec $x = 0_d$ donc $f^*(y) = 0$ si $\|y\|_\infty \leq 1$

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

2) On réécrit le problème. Soit $y \in \mathbb{R}^n$

$\min_x \|y\|_2^2 + \|x\|_1$

st $y + b - Ax = 0$

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

$g(\mu) = \inf_{x, y} \|y\|_2^2 + \mu^T y + \mu^T b - \mu^T A x + \|x\|_1$

$= \inf_y \|y\|_2^2 + \mu^T y + \inf_x \|x\|_1 - \mu^T A x + \mu^T b$

* Le premier problème de l'égalité peut être minimisé car convexe et différentiable, on le note $h(y) = \|y\|_2^2 + \mu^T y$

$\nabla_y h(y) = 2y + \mu = 0 \iff y = -\mu/2$

Le minimum en y est donc $\frac{1}{4}\|\mu\|_2^2 - \frac{1}{2}\|\mu\|_2^2 = -\frac{1}{4}\|\mu\|_2^2$

* La deuxième partie du problème peut se résoudre avec la question 1

$\sup_x (A^T\mu)^T x - \|x\|_1 = f^*(A^T\mu) = \begin{cases} +\infty & \text{si } \|A^T\mu\|_\infty > 1 \\ 0 & \text{sinon} \end{cases}$

$\Rightarrow g(\mu) = \mu^T b - \frac{1}{4}\|\mu\|_2^2 - f^*(A^T\mu)$

Donc le problème dual peut être écrit :

$\max_\mu \mu^T b - \frac{1}{4}\|\mu\|_2^2$

st $\|A^T \mu\|_\infty \leq 1$

Exercice n°3 :

Le problème est $\min_{\omega} \frac{1}{n} \sum_{i=1}^{n} \ell(\omega, x_i, y_i) + \frac{\tau}{2} \|\omega\|_2^2$ (sep 1)

où $\ell(\omega, x_i, y_i) = \max\{0, 1 - y_i(\omega^T x_i)\}$

1) $\min_{\omega, z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|\omega\|_2^2$ (Sep 2)

$$\text{st } z_i \geq 1 - y_i(\omega^T x_i) \ \forall i \in [\![1, n]\!]$$
$$z \geq 0$$

Avec τ un paramètre de régularisation.

On réécrit (Sep 2) en : $\underbrace{\min_{\omega} \frac{1}{2} \|\omega\|_2^2}_{A} + \underbrace{\min_{z} \frac{1}{n\tau} \mathbb{1}^T z}_{B}$

st $\forall i \in [\![1, n]\!]$ $z_i \geq 1 - y_i(\omega^T x_i)$
$z \geq 0$

Le problème B est immédiat car $z \geq 0$ et on a une contrainte sur chaque z_i qui doit rester positif

donc B devient $\frac{1}{\tau} \sum_{i=1} \max\{0; 1 - y_i(\omega^T x_i)\}$

Donc (sep 2) peut se réécrire $\min_{\omega} \frac{1}{n\tau} \sum_{i=1}^{n} \max\{0; 1 - y_i(\omega^T x_i)\} + \frac{1}{2} \|\omega\|_2^2$

Avec $\tau > 0$ et d'après la définition de la fonction de perte on a (sep2) équivalent à (sep 1).

2) On commence par réécrire (Sep2) dans sa forme standard

$$\min_{\omega, z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|\omega\|_2^2$$

st $1 - z_i - y_i(\omega^T x_i) \leq 0 \quad \forall i \in [\![1, n]\!] \quad (\lambda_i)$

$-z \leq 0 \quad (\pi)$

On peut ensuite définir $(\lambda_i \in \mathbb{R}, i \in [\![1, n]\!], \pi \in \mathbb{R}^n)$

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

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

Donc $g(\lambda, \pi) = \begin{cases} \inf_{\omega} \frac{1}{2} \|\omega\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (\omega^T x_i) + \lambda^T \mathbb{1} & \text{si } \frac{1}{n\tau} \mathbb{1} - \pi - \lambda = 0 \\ -\infty & \text{sinon} \end{cases}$

De plus $f(\omega) := \frac{1}{2} \|\omega\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (\omega^T x_i) + \lambda^T \mathbb{1}$ est convexe et différentiable

On calcule $\nabla_\omega f(\omega) = \omega - \sum_{i=1}^{n} \lambda_i y_i x_i = 0 \Rightarrow \omega = \sum_{i=1}^{n} \lambda_i y_i x_i$

Donc $\inf_{\omega} f(\omega) = \frac{1}{2} \left(\sum_{i=1}^{n} \lambda_i y_i x_i \right)^T \left(\sum_{i=1}^{n} \lambda_i y_i x_i \right) - \sum_{i=1}^{n} \lambda_i y_i \left(\sum_{j=1}^{n} \lambda_j y_j x_j \right)^T x_i$

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

$$= \frac{1}{2} \sum_{1 \leq i, j \leq n} \lambda_i \lambda_j y_i y_j x_i^T x_j + \lambda^T \mathbb{1}$$

Donc $g(\lambda, \pi) = \begin{cases} \lambda^T \mathbb{1} - \frac{1}{2} \sum_{1 \leq i, j \leq n} \lambda_i \lambda_j y_i y_j x_i^T x_j & \text{si } \frac{1}{n\tau} \mathbb{1} - \pi - \lambda = 0 \\ -\infty & \text{sinon} \end{cases}$

Le dual de (Sep 2) est $\max_{\lambda, \pi} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{1 \leq i, j \leq n} \lambda_i \lambda_j y_i y_j x_i^T x_j$

$$\text{st } \frac{1}{n\tau} \mathbb{1} - \lambda - \pi = 0$$
$$\lambda \geq 0$$
$$\pi \geq 0$$

On le simplifie en $\max_{\lambda} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{1 \leq i, j \leq n} \lambda_i \lambda_j y_i y_j x_i^T x_j$

$$\text{st } 0 \leq \lambda \leq \frac{1}{n\tau} \mathbb{1}$$