Devoir maison n$^{\text{o}}$ 2 convex optimization

Maria Cherifa

1$^{\text{er}}$ novembre 2020

Exercice 1 :

Dans cet exercice nous avons deux problèmes d'optimisation linéaire avec $c \in \mathbf{R}^d, b \in \mathbf{R}^n, A \in \mathbf{R}^{n\times d}$:

$$\begin{aligned} &\min_x c^T x \\ &\text{s.t } Ax = b \text{ et } x \geq 0 \end{aligned} \tag{P}$$

$$\begin{aligned} &\max_y b^T y \\ &\text{s.t } A^T y \leq c \end{aligned} \tag{D}$$

1. Dans cette question on veut calculer le dual du problème (P), commençons par calculer $L(x, \lambda, v)$ le Lagrangien du problème (P) :

$$\begin{aligned} L(x, \lambda, v) &= c^T x - \lambda^T x + v^T (A\,x - b) \\ &= (c - \lambda + A^T v)^T x - v^T b \end{aligned}$$

   On voit bien que L est une fonction affine de x, alors la fonction duale est donc donnée par :

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

   Le problème dual de (P) est défini alors

$$\begin{aligned} &\max g(\lambda, v) \\ &\text{s.t } \lambda \geq 0 \end{aligned}$$

   Ce qui est équivalent à

$$\begin{aligned} &\max_v v^T b \\ &\text{s.t } A^T v \leq c \end{aligned}$$

   On constate que le dual de (P) est en fin de compte (D).

2. On veut calculer le problème dual de (D), commençons par calculer le Lagrangien du problème, avant cela on peut transformer le problème (D) en un problème de minimisation comme suit :

$$\begin{aligned} &-\min_x \; - b^T y \\ &\text{s.t } A^T y \leq c \end{aligned} \tag{1}$$

Et donc le Lagrangien est donné par :

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

On voit que L est une fonction affine de y, alors la fonction dual est donnée par :

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

Le problème dual de (D) est défini alors par :

$$\begin{aligned} -\max\, & g(\lambda) \\ \text{s.t } & \lambda \geq 0 \end{aligned}$$

Donc :

$$\begin{aligned} -\max_\lambda\ & -\lambda^T\,c \\ \text{s.t } & \lambda \geq 0 \\ \text{et } & A\,\lambda = b \end{aligned}$$

Ce qui est équivalent à

$$\begin{aligned} \min_\lambda\ & \lambda^T\,c \\ \text{s.t } & \lambda \geq 0 \\ \text{et } & A\,\lambda = b \end{aligned}$$

On constate que le dual de (D) est en fin de compte (P).

3. Un problème est dit self-dual si et seulement le problème dual est égal au problème primal, on veut alors prouver cette propriété pour le problème suivant :

$$\begin{aligned} \min_{x,y}\ & c^T\,x - b^T\,y \\ \text{s.t } & A^T\,y \leq c \\ & A\,x = b \\ & x \geq 0 \end{aligned} \tag{Self-Dual}$$

$$\text{(Self-Dual)} = \begin{cases} \min\limits_{x,y} c^T\,x + \min\limits_{x,y}\ -b^T\,y \\ \text{s.t } A^T\,y \leq c \\ A\,x = b \\ x \geq 0 \end{cases}$$

Si il existe un y tel que $A^T\,y \leq c$ et il existe un x tel que $A\,x = b$ et $x \geq 0$, alors (Self-Dual) peut s'écrire sous la forme de :

$$\text{(Self-Dual)} = \begin{cases} \min\limits_x c^T\,x \\ \text{s.t } A\,x = b \\ x \geq 0 \end{cases} + \begin{cases} \min\limits_y\ -b^T\,y \\ \text{s.t } A^T\,y \leq c \end{cases}$$

Ce qui est équivalent à :

$$(\text{Self-Dual}) = \begin{cases} \min_x c^T x \\ \text{s.t } A\,x = b \\ x \geq 0 \end{cases} + \begin{cases} -\max_y b^T y \\ \text{s.t } A^T y \leq c \end{cases}$$

Donc :

$$(\text{Self-Dual}) = (\text{P}) - (\text{D})$$

Alors si on passe au dual :

$$(\text{Self-Dual})^* = (\text{P})^* - (\text{D})^*$$

Par les questions précédentes :

$$(\text{Self-Dual})^* = (\text{P})^* - (\text{D})^* = (\text{D}) - (\text{P}) = -(\text{Self-Dual})$$

On a bien alors la propriété de la Self-Dualité qui est conservée.

4. On veut prouver que la solution $[x^*, y^*]$ du problème (Self-Dual) peut être obtenue en résolvant (P) et (D). En effet, si le problème (Self-Dual) est faisable et bornée et sa solution optimale est donnée par $[x^*, y^*]$, alors forcément les problèmes (P) et (D) sont faisables et bornées aussi puisque $(\text{Self-Dual}) = (\text{P}) - (\text{D})$, de plus comme $(\text{Self-Dual})^* = (\text{P})^* - (\text{D})^*$ alors ça implique que la solution $[x^*, y^*]$ peut être obtenu en résolvant (P) et (D) séparément, tel que x^* est la solution de (P) et y^* est la solution de (D).

5. Par les conditions de Slater sur les problèmes (P) et (D), nous avons dualité forte et donc nous avons égalité entre la valeur optimale pour (P) et (D) et la solution de leurs problèmes duaux. De plus comme $(\text{Self-Dual})^* = (\text{P})^* - (\text{D})^* = (\text{D}) - (\text{P})$ alors on déduit que la solution de (Self-Dual) est nulle.

Exercice 2 :

Soit $A \in \mathbf{R}^{n \times d}$ et $b \in \mathbf{R}^n$, on considère le problème d'optimisation suivant,

$$\min_x ||A\,x - b||_2^2 + ||x||_1 \tag{RLS}$$

1. On veut calculer la fonction conjuguée de $||x||_1$, En effet par définition de la fonction conjuguée nous avons :

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

De plus nous pouvons écrire :

$$||x||_1 = \sup_{||p||_\infty \leq 1} x^T p$$

Donc :

$$f^*(y) = \sup_x (y^T x - \sup_{||p||_\infty \leq 1} x^T p)$$

Ce qui implique que :

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

2. Calculons le dual du problème (RLS), tout d'abord nous pouvons voir que :

$$(\text{RLS}) \iff \left\{ \min_{y=A\,x-b} y^T\,y + ||x||_1 \right.$$

Calculons le Lagrangien de (RLS) :

$$\begin{aligned} L(x,y,v) &= y^T\,y + ||x||_1 + v^T\,y - v^T\,A\,x + v^T\,b \\ &= v^T\,b + y^T\,y + v^T\,y - (v^T\,A\,x - ||x||_1) \end{aligned}$$

en posant $f(y) = y^T\,x - ||x||_1$, alors le Lagrangien s'écrit comme suit :

$$\begin{aligned} L(x,y,v) &= v^T\,b + y^T\,y + v^T\,y - (v^T\,A\,x - ||x||_1) \\ &= v^T\,b + y^T\,y + v^T\,y - f(A^T\,v) \end{aligned}$$

et donc la fonction duale est donnée par :

$$g(v) = \inf_{x,y} L(x,y,v) = \begin{cases} v^T\,b + \inf\limits_{y}(y^T\,y + v^T\,y) & \text{si } ||A^T\,v||_\infty \leq 1 \\ -\infty & \text{sinon} \end{cases}$$

calculons $\inf\limits_{y}(y^T\,y + v^T\,y)$, en effet nous avons $0 \leq ||y + \frac{v}{2}||_2^2 = y^T\,y + v^T\,y + \frac{v^T\,v}{4}$, donc $y^T\,y + v^T\,y \geq \frac{-v^T\,v}{4}$, donc la fonction duale est donnée par :

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

Alors le problème duale est donné par :

$$\begin{cases} \max\limits_{v} \quad v^T\,b - \frac{v^T\,v}{4} \\ \text{s.t} \quad ||A^T\,v||_\infty \leq 1 \end{cases}$$

Exercice 3 :

Dans cet exercice nous traitons le problème du "support vector machine" connu sous SVM. La forumation du problème d'optimisation se fait grâce à la hinge loss fonction, en effet on veut trouver un hyperplan qui sépare les x_i selon leurs labels y_i, cela revient à résoudre le problème suivant :

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

avec τ un paramètre de régularisation. nous remarquons que (Sep.1) est équivalent à :

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

1. Dans cette question nous voulons prouver que résoudre :

$$\begin{aligned} \min_{w,z} \frac{1}{n\,\tau} \mathbf{1}^T\,z + \frac{\tau}{2}||w||_2^2 \\ \text{avec} \quad z_i \geq 1 - y_i(w^T\,x_i) \quad \forall i = 1, ...n \\ z \geq 0 \end{aligned} \tag{Sep.2}$$

revient à résoudre (Sep.1). En effet, (Sep.2) peut s'écrire sous la forme de :

$$\begin{aligned} &\min_{w,z} \frac{1}{n\,\tau}\mathbf{1}^T z + \frac{\tau}{2}||w||_2^2 \\ \text{avec} \quad &1 - z_i \leq y_i(w^T x_i) \quad \forall i = 1, ...n \\ &z \geq 0 \end{aligned} \tag{Sep.2}$$

Donc si $1 \leq y_i(w^T x_i)$ alors $z_i = 0$ et si $1 > y_i(w^T x_i)$ alors $z_i = 1 - y_i(w^T x_i)$ ce qui nous permet de dire que $z_i = \max(0, 1 - y_i(w^T x_i))$ et donc si on prend $z_i = \max(0, 1 - y_i(w^T x_i))$ le problème (Sep.2) devient un problème sans contrainte et il sera égal à :

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

On deduit alors que (Sep.1) $\Longleftrightarrow$(Sep.2).

2. On veut calculer le problème dual de (Sep.2), le Lagrangien est défini par :

$$\begin{aligned} L(w, z, \lambda, \pi) &= \frac{1}{n\,\tau}\mathbf{1}^T z + \frac{1}{2}||w||_2^2 + \sum_{i=1}^{n} \lambda_i\,(1 - z_i - y_i(w^T x_i)) - \sum_{i=1}^{n} \pi_i\, z_i \\ &= \frac{1}{n\,\tau}\mathbf{1}^T z + \frac{w^T w}{2} + \mathbf{1}^T \lambda - \lambda^T z - \sum_{i=1}^{n} \lambda_i\, y_i(w^T x_i) - \pi^T z \\ &= (\frac{1}{n\,\tau}\mathbf{1} - \lambda - \pi)^T z + \mathbf{1}^T \lambda + \frac{w^T w}{2} - \sum_{i=1}^{n} \lambda_i\, y_i(w^T x_i) \end{aligned}$$

Calculons la fonction duale :

$$g(\lambda, \pi) = \inf_{w,z} L(w, z, \lambda, \pi) = \begin{cases} \mathbf{1}^T \lambda + \inf_w \frac{w^T w}{2} - \sum_{i=1}^{n} \lambda_i\, y_i(w^T x_i) & \text{si } \frac{1}{n\,\tau}\mathbf{1}^T - \lambda^T - \pi^T = 0 \\ -\infty & \text{sinon} \end{cases}$$

Comme on a $0 \leq ||w - \sum_{i=1}^{n} \lambda_i\, y_i\, x_i||_2^2 = ||w||_2^2 - w^T(2\sum_{i=1}^{n} \lambda_i\, y_i\, x_i)) + ||\sum_{i=1}^{n} \lambda_i\, y_i\, x_i)||_2^2$ cela implique que : $\frac{||w||_2^2}{2} - \sum_{i=1}^{n} \lambda_i\, y_i w^T\, x_i) \geq -\frac{||\sum_{i=1}^{n} \lambda_i\, y_i\, x_i||_2^2}{2}$.

Et donc le problème dual est de la forme :

$$\begin{aligned} &\max_{\lambda,\pi} \mathbf{1}^T \lambda - \frac{||\sum_{i=1}^{n} \lambda_i\, y_i\, x_i||_2^2}{2} \\ \text{s.t} \;\; &\frac{1}{n\,\tau}\mathbf{1}^T - \lambda^T - \pi^T = 0 \;\; \text{et} \lambda_i \geq 0 \end{aligned}$$

Si on veut réduire la dimension, on peut supprimer la variable π, en effet comme nous avons : $\frac{1}{n\,\tau}\mathbf{1}^T - \lambda^T - \pi^T = 0$ et $\lambda_i \geq 0$ ceci est équivalent à $0 \leq \lambda_i \leq \frac{1}{n\,\tau}$ et donc le problème dual devient :

$$\begin{aligned} &\max_{\lambda} \mathbf{1}^T \lambda - \frac{||\sum_{i=1}^{n} \lambda_i\, y_i\, x_i||_2^2}{2} \\ \text{s.t} \;\; &0 \leq \lambda_i \leq \frac{1}{n\,\tau} \quad \forall i = 1....n \end{aligned}$$

Exercice 4 :

Cet exercice traite le robust linear programming. soit le programme linéaire robuste définit par :

$$\begin{aligned} \min_x \; & c^T x \\ \text{s.t} \sup_{a \in \mathcal{P}} \; & a^T x \leq b \end{aligned}$$

avec $x \in \mathbf{R}^n$ et $\mathcal{P} = \{a \mid C^T a \leq d\}$ est un polyhèdre non vide.

1. Tout d'abord le problème précédent est équivalent à :

$$\begin{aligned} \min \; & c^T x \\ \text{s.t} \; & f(x) \leq b \end{aligned} \tag{A}$$

Si on définit $f(x)$ comme la valeur optimale du programme linéaire précédent alors on aura :

$$(\text{A}) \Longleftrightarrow \begin{cases} \max \;\; x^T a \\ \text{s.t} \;\; C\, a \leq d \end{cases}$$

Ici a est la variable et x est considéré comme un paramètre du problème, alors le Lagrangien du problème précédent est donnée par :

$$L(a, z) = -x^T a + z^T C\, a - z^T d$$

La fonction duale est égale à :

$$g(z) = \inf_a L(a, z) = \begin{cases} -z^T d & \text{pour } \; z^T C = x^T \text{et } \; z \geq 0 \\ -\infty & \text{sinon} \end{cases}$$

Le problème duale est alors donné par :

$$\begin{aligned} -\max_z \;\; & -z^T d \\ \text{s.t} \;\; & C^T z = x \\ & z \geq 0 \end{aligned}$$

Ce qui équivalent à :

$$\begin{aligned} \min_z \;\; & z^T d \\ \text{s.t} \;\; & C^T z = x \\ & z \geq 0 \end{aligned}$$

Sans oublier que $f(x)$ est aussi solution du programme linéaire, alors $f(x) \leq b$, si et seulement si il existe un z qui vérifie :

$$d^T z \leq b \qquad C^T z = x \qquad z \geq 0$$

Donc le programme linéaire initiale est équivalent à :

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

Exercice 5 :

Un programme linéaire booléen est un problème de la forme :

$$\begin{aligned} \min\ & c^T x \\ \text{s.t}\ & A\,x \leq b \\ x_i \in \{0,1\}\ & i = 1, ...n \end{aligned}$$

En général ce problème est difficile à résoudre. Considèrons la relaxation du problème linéaire :

$$\begin{aligned} \min\ & c^T x \\ \text{s.t}\ & A\,x \leq b \\ 0 \leq x_i \leq 1\ & i = 1, ...n \end{aligned} \tag{R.L}$$

qui est plus facile à résoudre.

1. Le LP booléen peut être reformulé comme suit :

$$\begin{aligned} \min\ & c^T x \\ \text{s.t}\ & A\,x \leq b \\ x_i\,(1-x_i) = 0\ & i = 1, ...n \end{aligned} \tag{LP.R}$$

On cherche le formulation duale de ce problème, en effet le Lagrangien de ce problème est égale à :

$$\begin{aligned} L(x,\mu,v) &= c^T x + \mu^T (A\,x - b) - v^T x + x^T \mathbf{diag}(v)\,x \\ &= x^T \mathbf{diag}(v)\,x + (c^T + A^T \mu - v)^T x - b^T \mu \end{aligned}$$

Et donc la fonction duale est donnée par :

$$g(\mu, v) = \inf_x L(x,\mu,v) = \begin{cases} -b^T \mu + \inf\limits_x (x^T \mathbf{diag}(v)\,x + (c^T + A^T \mu - v)^T x) & \\ -\infty & \text{sinon} \end{cases}$$

Afin de trouver $\inf\limits_x (x^T \mathbf{diag}(v)\,x + (c^T + A^T \mu - v)^T x)$ nous utilisons :

$$0 \leq ||x + \frac{c^T + A^T \mu - v}{2}||_2^2 = x^T x + (c^T + A^T \mu - v)^T x + \frac{||c^T + A^T \mu - v||^2}{4}$$

Et donc on $x^T \mathbf{diag}(v)\,x + (c^T + A^T \mu - v)^T x \geq -\frac{||c^T + A^T \mu - v||^2}{4\,v}$.
Alors la fonction duale est égale :

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

Ou a_i est la ième colonne de A, nous considèrons de plus que $a^2/0 = \infty$ si $a \neq 0$ et $a^2/0 = 0$ si $a = 0$.
Donc le problème dual est donné par :

$$\begin{cases} \max\ -b^T \mu - \frac{1}{4}\sum_{i=1}^n (c_i^T + a_i^T \mu - v_i)^2 / v_i \\ \text{s.t } v \geq 0 \end{cases}$$

Simplifions l'expression précédente : nous allons utiliser l'aide donnée dans l'énoncé, en effet :

$$\begin{aligned} \max_{v_i \geq 0} \left(\frac{-(c_i^T + a_i^T \mu - v_i)^2}{v_i} \right) &= \begin{cases} 4(c_i + a_i^T \mu) & c_i + a_i^T \mu \leq 0 \\ 0 & c_i + a_i^T \mu \geq 0 \end{cases} \\ &= 4 \min(0, c_i + a_i^T \mu) \end{aligned}$$

Donc le problème devient :

$$\begin{cases} \max\ -b^T \mu + \sum_{i=1}^n \min(0, c_i + a_i^T \mu) \\ \text{s.t } \mu \geq 0 \end{cases}$$

2. Utilisons l'aide donnée dans l'énoncé. Le Lagrangien du problème (R.L) est donné par :

$$\begin{aligned} L(x,u,v,w) &= c^T\,x + u^T(A\,x-b) - v^T\,x + w^T\,(x-1) \\ &= (c + A^T\,u - v + w)^T\,x - b^T\,u - \mathbf{1}^T\,w \end{aligned}$$

La fonction duale est égale à :

$$g(u,v,w) = \inf_x L(x,u,v,w) = \begin{cases} -b^T\,u - \mathbf{1}^T\,w & \text{si } \ c + A^T\,u - v + w = 0 \\ -\infty & \text{sinon} \end{cases}$$

Et donc le problème dual est donné par :

$$\begin{aligned} \max \ & -b^T\,u - \mathbf{1}^T\,w \\ \text{s.t } & c + A^T\,u - v + w = 0 \\ & u \geq 0, v \geq 0, w \geq 0 \end{aligned}$$

Ceci est équivalent au Lagrange relaxation donné par (LP.R), donc les deux problèmes sont équivalents et donnent la même solution.