DM2 - Optimisation Convexe

Coërchon Colin

4 Novembre 2024

1 Exercice 1

Pour $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$, et $A \in \mathbb{R}^{n\times d}$, on considère les deux problèmes suivants d'optimisation.

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

et

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

1. Le Lagrangien de la fonction $x \longmapsto c^T x$ est donné par :

$$L(x, \lambda, \nu) = c^T + \nu^T(Ax - b) - \lambda^T x = -b^T\nu + \left(c + A^T\nu - \lambda\right)^T x$$

La fonction duale lagrangienne est alors donnée par :

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

Le problème dual s'écrit alors :

$$\begin{aligned} \max_{\lambda,\nu} \quad & -b^T\nu \\ \text{s.c.} \quad & \lambda \succeq 0 \end{aligned} \tag{P}$$

Et donc, on a donc **notre problème dual de (P)** :

$$\begin{aligned} \max_\nu \quad & -b^T\nu \\ \text{s.c.} \quad & A^T\nu + c \succeq 0 \end{aligned}$$

On pose $y = -\nu \in \mathbb{R}^n$ pour alléger la notation.

$$\implies \boxed{\begin{aligned} \max_y \quad & b^T y \\ \text{s.c.} \quad & A^T y \preceq c \end{aligned}} \tag{DP}$$

2. Le Lagrangien de la fonction $y \longmapsto b^T y$ est donné par :

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

Ce qui peut se réécrire comme :

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

La fonction duale est obtenue en maximisant $L(y, \lambda)$ par rapport à y, ce qui donne :

$$g(\lambda) = \sup_y L(y, \lambda) = \sup_y \left(\lambda^T c + y^T (b - A\lambda)\right)$$

Pour que cette expression soit bornée, il faut que $b - A\lambda = 0$, sinon la fonction tendrait vers $+\infty$. Ainsi,

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

Le problème dual consiste alors à minimiser $g(\lambda)$, ce qui se traduit par :

$$\begin{array}{ll} \min\limits_{\lambda} & \lambda^T c \\ \text{s.c.} & A\lambda = b \\ & \lambda \succeq 0 \end{array} \tag{DD}$$

Ainsi, **le problème dual de (D)**, que nous avons noté (DD), est le suivant (j'ai juste remplacé λ par x) :

$$\boxed{\begin{array}{ll} \min\limits_{x} & x^T c \\ \text{s.c.} & Ax = b \\ & x \succeq 0 \end{array}}$$

Remarque :

C'est sympathique de remarquer que le dual du problème (P) correspond au problème (D) initial. Et **inversement**.

3. Voici le problème (Self-Dual). Nous allons prouver que son problème dual correspond au problème primal.

$$\begin{array}{ll} \min\limits_{x,y} & c^T x - b^T y \\ \text{s.c.} & Ax = b \\ & x \succeq 0 \\ & A^T y \preceq c \end{array} \tag{Self-Dual}$$

Étape 1 : Formulation du Lagrangien

Soit $\lambda \succeq 0$ le multiplicateur associé à la contrainte $x \succeq 0$, $\mu \succeq 0$ le multiplicateur associé à la contrainte $A^T y \preceq c$, et $\nu \in \mathbb{R}^n$ le multiplicateur associé à la contrainte d'égalité $Ax = b$. Le Lagrangien est donné par :

$$\begin{aligned} L(x, y, \lambda, \mu, \nu) &= c^T x - b^T y + \nu^T (b - Ax) - \lambda^T x + \mu^T (A^T y - c) \\ &= \nu^T b - \mu^T c + \left(c^T - \nu^T A - \lambda^T\right) x + \left(-b^T + \mu^T A^T\right) y \\ &= \nu^T b - \mu^T c + \left(c - A^T \nu - \lambda\right)^T x + (A\mu - b)^T y \end{aligned}$$

Étape 2 : Fonction duale lagrangienne

La fonction duale est obtenue en minimisant le Lagrangien par rapport aux variables x et y.

- Minimisation par rapport à x : Pour que l'expression en x soit bornée, il faut que :

$$c - A^T\nu - \lambda = 0$$

- Minimisation par rapport à y : Pour que l'expression en y soit bornée, il faut que :

$$A\mu - b = 0$$

En utilisant ces deux conditions, nous obtenons la fonction duale :

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

Étape 3 : Formulation du problème dual

Le problème dual consiste alors à maximiser $g(\lambda, \mu, \nu)$ par rapport aux variables ν, λ, et μ, en respectant les contraintes obtenues. Le problème dual s'écrit donc :

$$\begin{array}{ll} \max\limits_{\mu,\nu} & \nu^T b - \mu^T c \\ \text{s.c.} & \lambda = c - A^T\nu \\ & A\mu = b \\ & \lambda \succeq 0, \quad \mu \succeq 0 \end{array}$$

Or, on remarque que $\nu^T b$ et $\mu^T c$ sont simplement des produits scalaires. Donc on peut les écrire respectivement $b^T\nu$ et $c^T\mu$. De plus, on a facilement :

$$b^T\nu - c^T\mu = -(c^T\mu - b^T\nu)$$

On peut alors transformer le **maximum** en un **minimum**, ce qui donne :

$$\begin{array}{ll} \min\limits_{\mu,\nu} & c^T\mu - b^T\nu \\ \text{s.c.} & \lambda = c - A^T\nu \\ & A\mu = b \\ & \lambda \succeq 0, \quad \mu \succeq 0 \end{array}$$

Puis, on peut simplifier la contrainte $\lambda = c - A^T\nu$ puisqu'on a également la contrainte $\lambda \succeq 0$ et nous retrouvons alors :

$$\boxed{\begin{array}{ll} \min\limits_{\mu,\nu} & c^T\mu - b^T\nu \\ \text{s.c.} & A\mu = b \\ & \mu \succeq 0 \\ & A^T\nu \preceq c \end{array}}$$

Ainsi, nous retrouvons le problème primal (Self-Dual). Nous avons donc prouvé que le problème est **auto-dual (self-dual)**, car **le dual obtenu est identique au problème primal initial**.

4. On suppose que ce problème est faisable et borné, et que (x^*, y^*) est une solution optimale.

   Puisque (Self-Dual) est faisable et borné, le **théorème de dualité forte** assure l'existence d'une solution optimale (x^*, y^*) pour laquelle les conditions de complémentarité sont satisfaites entre le problème primal et son dual. En utilisant **le théorème des écarts complémentaires**, les conditions de faisabilité et de complémentarité pour les solutions optimales (x^*, y^*) sont :

   - **Conditions de faisabilité** :
     - $Ax^* = b$
     - $x^* \geq 0$
     - $A^T y^* \leq c$
   - **Conditions de complémentarité** :
     - $x_i^* \cdot (c - A^T y^*)_i = 0$ pour chaque composante i de x.
     - $(b - Ax^*)_j \cdot y_j^* = 0$ pour chaque composante j de y.

   Puisque x^* et y^* satisfont les contraintes, il est aisé de remarque que la deuxième condition de complémentarité implique que : $y^* = 0$. Il en découle alors rapidement avec la première contrainte que $x^* = 0$.

   Le théorème des écarts complémentaires nous a ainsi permi de prouver que :

   $$\boxed{(x^*, y^*) = (0, 0)}$$

   Cela montre que la valeur optimale de (Self-Dual) est exactement 0.

   De plus, d'après la question 2, nous avons observé que le dual du problème (P) est le problème (D) initial, et **inversement**, le dual de (D) est le problème (P) initial. Par conséquent, les deux problèmes (P) et (D) partagent la même valeur optimale lorsque les deux problèmes sont faisables et bornés.

   Le problème (Self-Dual) (qui est auto-dual) correspond simplement au problème [**(P) - (D)**], il en résulte que :

   - x^* est une solution faisable et optimale pour (P),
   - y^* est une solution faisable et optimale pour (D),
   - et la valeur optimale $c^T x^* - b^T y^* = 0$ montre que la dualité forte entre (P) et (D) est vérifiée.

   Conclusion

   - La valeur optimale de (Self-Dual) est exactement 0.
   - Le couple (x^*, y^*) peut être obtenu en résolvant séparément les problèmes (P) et (D).

2 Exercice 2

Pour une matrice $A \in \mathbb{R}^{n\times d}$ et un vecteur $b \in \mathbb{R}^n$, considérons le problème suivant :

$$\min_x \left(\|Ax-b\|_2^2 + \|x\|_1\right) \tag{RLS}$$

1. Pour calculer le conjugué de la norme ℓ_1, notée $\|x\|_1$, nous devons déterminer sa fonction conjuguée $\|x\|_1^*$, définie par :

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

La norme $\|x\|_1$ est définie comme $\|x\|_1 = \sum_{i=1}^d |x_i|$, et sa fonction conjuguée est la fonction indicatrice de la boule duale de ℓ_∞, c'est-à-dire l'ensemble $\{y \in \mathbb{R}^d \mid \|y\|_\infty \leq 1\}$. Ainsi :

$$\boxed{\|x\|_1^*(y) = \begin{cases} 0 & \text{si } \|y\|_\infty \leq 1 \\ +\infty & \text{sinon.} \end{cases}}$$

2. Le problème primal (RLS) s'écrit comme ceci :

$$\min_x \left(\|Ax-b\|_2^2 + \|x\|_1\right)$$

Pour simplifier, introduisons une variable auxiliaire $u = Ax - b$ et réécrivons le problème en termes de u et x :

$$\min_{x,u} \left(\|u\|_2^2 + \|x\|_1\right) \quad \text{s.c.} \quad u = Ax - b$$

En introduisant un multiplicateur de Lagrange $\nu \in \mathbb{R}^n$ pour la contrainte $u = Ax - b$, nous pouvons formuler le Lagrangien associé :

$$L(x,u,\nu) = \|u\|_2^2 + \|x\|_1 + \nu^T(Ax - b - u)$$

Pour obtenir la fonction duale, nous allons minimiser ce Lagrangien par rapport aux variables x et u.

- **Minimisation par rapport à u :**
La minimisation de $L(x,u,\nu)$ par rapport à u implique de résoudre :

$$\min_u \left(\|u\|_2^2 - \nu^T u\right)$$

En observant cette expression, nous remarquons qu'il s'agit d'un problème quadratique plutôt simple. La fonction objective peut être réécrite comme ceci :

$$\|u\|_2^2 - \nu^T u = \|u\|_2^2 - u^T\nu.$$

On minimise en prenant la dérivée par rapport à u :

$$\frac{\partial}{\partial u}\left(\|u\|_2^2 - u^T\nu\right) = 2u - \nu = 0.$$

Cela donne $\boxed{u = \frac{\nu}{2}}$.

En substituant cette valeur de u dans le Lagrangien, on obtient :

$$\begin{aligned}
L(x, u, \nu) &= \left\| \frac{\nu}{2} \right\|_2^2 + \|x\|_1 + \nu^T Ax - \nu^T b - \nu^T \left(\frac{\nu}{2} \right) \\
&= \frac{1}{4}\|\nu\|_2^2 + \|x\|_1 + \nu^T Ax - \nu^T b - \frac{1}{2}\|\nu\|_2^2 \\
&= -\frac{1}{4}\|\nu\|_2^2 + \|x\|_1 + \nu^T Ax - \nu^T b
\end{aligned}$$

- **Minimisation par rapport à x :**
  La minimisation par rapport à x implique de résoudre :

  $$\min_x \left(\|x\|_1 + \nu^T Ax \right).$$

  Nous avons déjà calculé dans la première question le conjugué de la norme $\|x\|_1$:

  $$\|x\|_1^*(y) = \begin{cases} 0 & \text{si } \|y\|_\infty \leq 1, \\ +\infty & \text{sinon} \end{cases}$$

  Ainsi, pour que cette minimisation soit bornée, il faut que ν satisfasse la contrainte suivante :

  $$\boxed{\|A^T \nu\|_\infty \leq 1}$$

  Lorsque cette contrainte est satisfaite, la minimisation de $\|x\|_1 + \nu^T Ax$ par rapport à x est égale à zéro, sinon elle tend vers $+\infty$.

Expression finale du dual :

En combinant les résultats des minimisations par rapport à x et u, nous obtenons la fonction duale :

$$g(\nu) = -\frac{1}{4}\|\nu\|_2^2 - \nu^T b$$

Ainsi, le problème dual devient :

$$\boxed{\max_\nu \left(-\frac{1}{4}\|\nu\|_2^2 - \nu^T b \right) \quad \text{s.c.} \quad \|A^T \nu\|_\infty \leq 1}$$

3 Exercice 3

1.
   - La variable z représente la valeur de la perte pour chaque point de données x_i, c'est-à-dire $z_i = \max\{0, 1 - y_i(\omega^T x_i)\}$. En ajoutant la contrainte $z_i \geq 1 - y_i(\omega^T x_i)$ pour chaque i, on garantit que z_i prend la valeur de la perte de classification.
   - La minimisation de $\frac{1}{n\tau}\mathbf{1}^T z$ correspond alors à la minimisation de la moyenne des pertes sur les n points, ce qui est équivalent au premier terme de l'objectif dans (Sep. 1).
   - Le terme de régularisation $\frac{1}{2}\|\omega\|_2^2$ est conservé dans (Sep. 2) de la même manière que dans (Sep. 1), ce qui garantit que la régularisation sur ω est identique dans les deux formulations.

   Ainsi, le problème (Sep. 2) reformule le problème initial en utilisant des variables auxiliaires pour linéariser la fonction de perte, mais il reste équivalent à (Sep. 1) en minimisant à la fois la perte et le terme de régularisation.