

TP 3: Projection sur un convexe

Dans ce TD, qui servira d'appui au TP suivant, le but est d'utiliser les théorèmes du cours (notamment le théorème de projection sur un convexe fermé) pour calculer des projections et écrire des conditions d'optimalité pour des problèmes d'optimisation sous contrainte.

1 Projection sur un demi-espace

Q1. Soit $v \in \mathbb{R}^n$ non nul et soit pose $K = \{x \in \mathbb{R}^n \mid \langle v|x \rangle \leq 0\}$. Démontrer que la projection d'un point y sur K est donnée par la formule

$$p_K(y) = \begin{cases} y & \text{si } \langle v|y \rangle \leq 0 \\ y - \langle v|y \rangle \frac{v}{\|v\|^2} & \text{sinon} \end{cases}$$

2 Minimisation avec contrainte de borne inférieure

Dans cette partie, on se donne $K = \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\}, x_i \geq f_i\}$ où $f \in \mathbb{R}^n$ est donné. On s'intéresse au problème de minimiser une fonction $J \in \mathcal{C}^1(\mathbb{R}^n)$ sur K . Dans la suite, on note e_i le i ème vecteur de la base canonique de \mathbb{R}^n .

Q1. Montrer que $p_K(x) = (\max(x_i, f_i))_{1 \leq i \leq n}$.

Q2. Soit $J \in \mathcal{C}^1(\mathbb{R}^n)$ et $x^* \in \arg \min_K J$.

1. Si $x_i^* > f_i$, démontrer que $x_\varepsilon := x^* + \varepsilon e_i$ appartient à K pour $\varepsilon \in [-\varepsilon_0, +\infty[$ avec $\varepsilon_0 = x_i^* - f_i$, puis que $J(x_\varepsilon) \geq J(x^*)$. En déduire que $\frac{\partial J}{\partial x_i}(x^*) = 0$.

2. Si $x_i^* = f_i$, démontrer que $\forall \varepsilon \geq 0, x_\varepsilon = x^* + \varepsilon e_i$ appartient à K . En déduire que $\frac{\partial J}{\partial x_i}(x^*) \geq 0$. Ainsi, $x^* \in \arg \min_K J \implies \nabla J(x^*) = \lambda \in \mathbb{R}_+^n$ où $\lambda_i(x_i^* - f_i) = 0$ pour tout $i \in \{1, \dots, n\}$.

Q3. On suppose que $J \in \mathcal{C}^1(\mathbb{R}^n)$ est convexe. Soit $\bar{x} \in K$ vérifiant $\nabla J(\bar{x}) = \lambda \in \mathbb{R}_+^n$ et $\lambda_i(\bar{x}_i - f_i) = 0$. En utilisant l'inégalité de convexité¹, démontrer que $\bar{x} \in \arg \min_K J$.

Q4. Dans cette question on considère

$$J : x \in \mathbb{R}^n \mapsto \frac{1}{2} \sum_{i=0}^n (x_{i+1} - x_i)^2$$

où l'on pose par convention $x_0 = x_{n+1} = 0$.

1. Calculer $\nabla J(x)$.

2. En utilisant les questions Q2 et Q3, démontrer que

$$\bar{x} \in \arg \min_K J \iff \bar{x} \in K \text{ et } \begin{cases} \bar{x}_i = \frac{1}{2}(\bar{x}_{i-1} + \bar{x}_{i+1}) & \text{si } \bar{x}_i > f_i \\ \bar{x}_i \geq \frac{1}{2}(\bar{x}_{i-1} + \bar{x}_{i+1}) & \text{si } \bar{x}_i = f_i \end{cases}$$

¹ $J(y) \geq J(x) + \langle \nabla J(x)|y - x \rangle$

3 Minimisation sur le simplexe unité

On note $\Delta = \{x \in \mathbb{R}_+^n \mid \sum_{1 \leq i \leq n} x_i = 1\}$, et on appelle cet ensemble Δ le simplexe unité. Cet ensemble apparaît naturellement dans de nombreux problèmes d'optimisation appliqués aux probabilités, en géométrie, ou en finance (allocation de portefeuilles: la contrainte $x_i \geq 0$ dit qu'on investit des sommes positives et la contrainte $\sum_i x_i = 1$ traduit la somme totale que l'on souhaite investir). Dans cet exercice, on cherche à calculer la projection de $y \in \mathbb{R}^n$ sur Δ , i.e.

$$\min_{x \in \Delta} \|x - y\|^2 \quad (1)$$

Pour $\lambda \in \mathbb{R}$, on définit $x^\lambda \in \mathbb{R}^n$ par

$$x_i^\lambda = \max(y_i - \lambda, 0).$$

Notre objectif est de démontrer que la solution de (1) est de la forme x^{λ^*} pour un certain $\lambda^* \in \mathbb{R}$.

Q1. Montrer que le problème (1) admet une unique solution.

Q2. On pose $g(\lambda) = \sum x_i^\lambda$. Montrer que $g(\lambda)$ est continue, que $\lim_{\lambda \rightarrow -\infty} g(\lambda) = +\infty$ et $\lim_{\lambda \rightarrow +\infty} g(\lambda) = 0$. En déduire qu'il existe λ^* tel que $g(\lambda^*) = 1$, puis que $x^{\lambda^*} \in \Delta$.

Q3. On va montrer que $\bar{x} := x^{\lambda^*}$ est solution de (1). Pour alléger les notations, on pose $\lambda = \lambda^*$.

1. Justifier qu'il suffit de démontrer que

$$\forall z \in \Delta, \langle y - \bar{x} | \bar{x} - z \rangle = \sum_{1 \leq i \leq n} (y_i - \bar{x}_i)(\bar{x}_i - z_i) \geq 0.$$

2. Soit $i \in \{1, \dots, n\}$ tel que $\bar{x}_i > 0$. Démontrer que $\bar{x}_i = y_i - \lambda$ puis que

$$(y_i - \bar{x}_i)(\bar{x}_i - z_i) = \lambda(\bar{x}_i - z_i)$$

3. Soit $i \in \{1, \dots, n\}$ tel que $\bar{x}_i = 0$. Démontrer alors que $y_i = 0$ puis que

$$(y_i - \bar{x}_i)(\bar{x}_i - z_i) \geq \lambda(\bar{x}_i - z_i).$$

4. En déduire que $\langle y - \bar{x} | \bar{x} - z \rangle \geq 0$ (en se rappelant $\bar{x}, z \in \Delta$), puis conclure: $p_\Delta(y) = x^\lambda$.

Q4. Soit $B = \{x \in \mathbb{R}^n \mid \sum_i |x_i| \leq 1\}$ la boule unité pour la norme ℓ^1 . On s'intéresse maintenant à la projection d'un point $y \in \mathbb{R}^n$ sur B .

1. Montrer que si $y \in \mathbb{R}_+^n$, alors $p_B(y) \in \mathbb{R}_+^n$.
2. En déduire dans ce cas ($y \in \mathbb{R}_+^n$), que si $y \notin B$, $p_B(y) = p_\Delta(y)$.
3. Montrer que si $y \in \mathbb{R}^n \setminus B$, et si l'on note $y_i = \varepsilon_i \tilde{y}_i$ où $\tilde{y} \in \mathbb{R}_+^n$ et $\varepsilon_i \in \{\pm 1\}$, alors

$$(p_B(y))_i = \varepsilon_i (p_\Delta(\tilde{y}))_i.$$