Partiel du 26 Mars 2015—Corrigé
"Optimisation et programmation dynamique"

Master mention Mathématiques appliquées 1ère année
Université Paris Dauphine

Dans tout le partiel, on note $\Pi_K(x)$ la projection orthogonale d'un point x de $\mathbb{R}^n$ sur un convexe fermé non vide K de $\mathbb{R}^n$. Si $y = (y_i)_{i=1,\ldots,m}$ et $z = (z_i)_{i=1,\ldots,m}$ sont deux vecteurs de $\mathbb{R}^m$, on écrit $y \leq z$, si et seulement si, $y_i \leq z_i$ pour tout $i = 1, \ldots, m$.

Exercice 1. On cherche à résoudre le problème

$$(\mathcal{P}) \qquad \min_{(x,y,z)\in K} x^2 + y^2 + z^2 \qquad \text{où} \qquad K = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1 \text{ et } x^2 + y^2 \leq 1\} .$$

1. Montrer que le problème admet une unique solution.
2. Montrer que la contrainte est qualifiée en tout point.
3. Ecrire les conditions nécessaires d'optimalité du problème et en déduire la solution du problème $(\mathcal{P})$.
4. Retrouver la solution du problème $(\mathcal{P})$ par la méthode de dualité.

Solution :

1. *On remarque que la contrainte K est fermée et le critère est coercif (c'est la norme euclidienne). Donc, par théorème de cours, le problème admet au moins une solution. Ensuite on note que le critère est strictement convexe (puisque $f(x, y, z) := x^2 + y^2 + z^2$ vérifie $Hess_f(x, y, z) = 2I_3$ qui est une matrice définie positive) et la contrainte est convexe (car la fonctions $h : (x, y, z) \to x + y + z - 1$ est affine et la fonction $g : (x, y, z) \to x^2 + y^2 - 1$ est convexe). Donc le problème admet une et une seule solution.*
2. *La contrainte est convexe, mais elle a un intérieur vide. Il faut revenir aux conditions de qualification usuelles. Posons $g(x, y, z) = x^2 + y^2 - 1$ et $h(x, y, z) = x + y + z - 1$. On note que ce sont des fonctions de classe C^1 (en fait C^∞).*
   Si $(x, y, z) \in K$ vérifie $g(x, y, z) < 0$, il suffit de montrer que la famille $\{\nabla h(x, y, z) = (1, 1, 1)\}$ est libre, ce qui est évident puisqu'elle est constituée d'un seul vecteur non nul. Si $(x, y, z) \in K$ vérifie $g(x, y, z) = 0$, alors il faut vérifier que la famille $\{\nabla h(x, y, z) = (1, 1, 1)\}$ est libre et qu'il existe un vecteur $v \in \mathbb{R}^3$ tel que
   $$\langle v, \nabla h(x, y, z)\rangle = 0 \text{ et } \langle v, \nabla g(x, y, z)\rangle < 0.$$
   La première condition est évidente. Pour la seconde, on peut prendre par exemple $v = (-x, -y, x + y)$, qui vérifie bien
   $$\langle v, \nabla h(x, y, z)\rangle = -x - y + (x + y) = 0 \text{ et } \langle v, \nabla g(x, y, z)\rangle = -2x^2 - 2y^2 = -2 < 0$$
   puisque $g(x, y, z) = x^2 + y^2 - 1 = 0$.
3. *Comme le critère et les fonctions définissant les contraintes sont de classe C^1 et comme la contrainte est qualifiée en tout point, les conditions nécessaires d'optimalité s'écrivent : si $(x, y, z) \in K$ est la solution du problème, alors il existe $\lambda \geq 0$ et $\mu \in \mathbb{R}$ tels que*
   $$\nabla f(x, y, z) + \lambda\nabla g(x, y, z) + \mu\nabla h(x, y, z) = 0$$

avec la condition d'exclusion $\lambda g(x,y,z) = 0$. Considérons deux cas : si $g(x,y,z) < 0$, alors $\lambda = 0$ et le système se réduit à

$$\begin{cases} 2x + \mu = 0 \\ 2y + \mu = 0 \\ 2z + \mu = 0 \\ x + y + z = 1 \end{cases}$$

qui ne possède qu'une solution solution : $x = y = z = 1/3$. Notons que $g(1/3, 1/3, 1/3) = 2/9 < 1$. Donc le point $(1/3, 1/3, 1/3)$ vérifie les conditions nécessaires d'optimalité. Comme la contrainte est qualifiée et que le problème est convexe, on en déduit que $(1/3, 1/3, 1/3)$ est un minimum du problème. Comme ce minium est unique, c'est la seule solution.
On pourrait étudier le cas où $g(x,y,z) = 0$ (même si c'est inutile). Le système devient :

$$\begin{cases} 2x + 2\lambda x + \mu = 0 \\ 2y + 2\lambda y + \mu = 0 \\ 2z + \mu = 0 \\ x + y + z = 1 \\ x^2 + y^2 = 1 \end{cases}$$

Comme $\lambda \geq 0$, on trouve en utilisant les deux premières égalités que $x = y$, ce qui implique par la saturation de la contrainte d'inégalité, que $x = y = \epsilon\sqrt{2}/2$ où $\epsilon = \pm 1$. Donc $z = 1 - \epsilon\sqrt{2}$. Alors $\mu = -2z = -2 + \epsilon 2\sqrt{2}$ et $\lambda = -1 - \mu/(2x) = -3 + 2\epsilon\sqrt{2}$. Comme $2\sqrt{2} < 3$, on a $\lambda < 0$, ce qui est impossible. Donc la contrainte d'inégalité n'est pas saturée à l'optimum.

4. *On pose*

$$L(x,y,z,\lambda,\mu) = f(x,y,z) + \lambda g(x,y,z) + \mu h(x,y,z) = (1+\lambda)(x^2+y^2) + \mu(x+y) + z^2 + \mu z - \lambda - \mu.$$

On cherche à minimiser L par rapport à (x,y,z) pour $\lambda \geq 0$. On trouve comme système d'optimalité (qui est aussi une condition suffisante puisque L est convexe en (x,y,z)) :

$$\begin{cases} 2(1+\lambda)x + \mu = 0 \\ 2(1+\lambda)y + \mu = 0 \\ 2z + \mu = 0 \end{cases}$$

soit $x = y = -\mu/((2(1+\lambda))$ et $z = -\mu/2$. D'où

$$d(\lambda,\mu) = \min_{(x,y,z)\in\mathbb{R}^3} L(x,y,z,\lambda,\mu) = -\frac{\mu^2}{2(1+\lambda)} - \frac{\mu^2}{4} - \lambda - \mu.$$

On cherche maintenant le maximum de d sur $\mathbb{R}_+ \times \mathbb{R}$. Comme d est concave (cf. le cours) et que la contrainte $\{\lambda \geq 0\}$ est convexe et qualifiée puisqu'affine, les conditions nécessaires sont suffisantes : il suffit de trouver un point vérifiant les conditions nécessaires d'optimalité, soit $\nabla d(\lambda,\mu) + (\theta, 0) = 0$ où $\theta \geq 0$. Or $\nabla d(\lambda,\mu) - (\theta, 0) = 0$ équivaut au système

$$\begin{cases} \frac{\mu^2}{2(1+\lambda)^2} - 1 - \theta = 0 \\ -\frac{\mu}{(1+\lambda)} - \frac{\mu}{2} - 1 = 0 \end{cases}$$

On note que $\lambda = 0$ et $\mu = -2/3$ est solution (avec $\theta = 7/9$). On peut s'arrêter là, puisque cela fournit une (et donc la) solution du problème : $(x,y,z) = (1/3, 1/3, 1/3)$.
Si on veut étudier ce qui se passe si la contrainte d'inégalité n'est pas saturée, alors on peut dire que $\theta = 0$, $\mu = \sqrt{2}\epsilon(1+\lambda)$ (où $\epsilon = \pm 1$) et donc $\lambda = -\sqrt{2}\epsilon(3\sqrt{2}\epsilon/2 + 1) < 0$ pour $\epsilon = 1$ et $\epsilon = -1$. Donc cette solution est impossible et ce cas est exclus.

Exercice 2. 1. Soient $n, m \in \mathbb{N}^*$, C une matrice réelle de taille $m \times n$ et $d \in \mathbb{R}^m$. On suppose que l'ensemble $K := \{x \in \mathbb{R}^n,\ Cx \leq d\}$ est non vide.
Soit $x, y \in \mathbb{R}^n$. Montrer que $y = \Pi_K(x)$, si et seulement si, il existe $\lambda \in \mathbb{R}^m$ avec $\lambda \geq 0$, $y - x + C^T\lambda = 0$ et $\lambda^T(Cy - d) = 0$ (où C^T est la transposée de la matrice C).
(Indication : Ecrire les conditions nécessaires du problème d'optimisation satisfait par $\Pi(x)$).

2. (Erreur d'énoncé)

Solution :

1. *La projection sur K est la solution du problème d'optimisation*

$$\min_{y \in K} \frac{1}{2}\|y - x\|^2$$

Comme il s'agit d'un problème convexe, avec une contrainte qualifiée puisqu'affine, les conditions nécessaires sont suffisantes. Or

$$L(y, \lambda) = \frac{1}{2}\|y - x\|^2 + \sum_{i=1}^{m} \lambda_i((Cx)_i - d_i) = \frac{1}{2}\|y - x\|^2 + \lambda^T Cx.$$

On en déduit que y est solution du problème, i.e., $y = \Pi_K(x)$, si et seulement si, il existe $\lambda \in \mathbb{R}^m$ avec $\lambda \geq 0$ et $\nabla_y L(y, \lambda) = 0$ et $\lambda^T(Cy - d) = 0$. Or $\nabla_y L(y, \lambda) = (y - x) + C^T\lambda$, ce qui prouve l'assertion.

Exercice 3. Soient C_1 et C_2 deux convexes fermés de $\mathbb{R}^n$ d'intersection non vide. On suppose qu'on sait effectuer numériquement la projection Π_{C_1} et Π_{C_2} sur les ensembles C_1 et C_2. On cherche à trouver un point de $C_1 \cap C_2$.

Pour $\tau \in]0, 2[$ un paramètre fixé et $\bar{x} \in \mathbb{R}^n$ une position initiale donnée, on définit l'algorithme

$$\begin{aligned} & x_0 = \bar{x} \\ & x_{n+1/2} = x_n + \tau\left(\Pi_{C_1}(x_n) - x_n\right) \qquad (n \in \mathbb{N}) \\ & x_{n+1} = x_{n+1/2} + \tau\left(\Pi_{C_2}(x_{n+1/2}) - x_{n+1/2}\right) \qquad (n \in \mathbb{N}) \end{aligned}$$

1. Montrer que, si C est un convexe fermé non vide, $x \in C$ et $y \in \mathbb{R}^n$, alors :

$$\|y + \tau\left(\Pi_C(y) - y\right) - x\|^2 \leq \|y - x\|^2 - \tau(2 - \tau)\left\|\Pi_C(y) - y\right\|^2.$$

2. En déduire que, pour tout $x \in C_1 \cap C_2$, on a :

$$\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \tau(2 - \tau)\left(\left\|\Pi_{C_2}(x_{n+1/2}) - x_{n+1/2}\right\|^2 + \|\Pi_{C_1}(x_n) - x_n\|^2\right).$$

3. Montrer alors que, pour tout $x \in C_1 \cap C_2$, la suite $(\|x_n - x\|^2)$ converge et vérifier que la suite (x_n) est bornée.

4. Montrer que toute valeur d'adhérence de la suite (x_n) appartient à $C_1 \cap C_2$.

Solution :

1. *On a :*

$$\|y + \tau\left(\Pi_C(y) - y\right) - x\|^2 = \|y - x\|^2 + 2\tau\langle y - x, \Pi_C(y) - y\rangle + \tau^2\|\Pi_C(y) - y\|^2$$

Or
$$\langle y-x, \Pi_C(y)-y\rangle = \langle y-\Pi_C(y), \Pi_C(y)-y\rangle + \langle \Pi_C(y)-x, \Pi_C(y)-y\rangle$$
où
$$\langle \Pi_C(y)-x, \Pi_C(y)-y\rangle \leq 0$$
par caractérisation de la projection. Donc
$$\langle y-x, \Pi_C(y)-y\rangle \leq \langle y-\Pi_C(y), \Pi_C(y)-y\rangle = -\|y-\Pi_C(y)\|^2,$$
ce qui montre que
$$\|y+\tau\left(\Pi_C(y)-y\right)-x\|^2 \leq \|y-x\|^2 - \tau(2-\tau)\|\Pi_C(y)-x_n\|^2$$
ce qui est le résultat voulu.

2. *On applique le résultat d'abord le résultat précédent entre* x_{n+1} *et* $x_{n+1/2}$ *: on trouve*
$$\|x_{n+1}-x\|^2 \leq \|x_{n+1/2}-x\|^2 - (2\tau-\tau^2)\|\Pi_{C_2}(x_{n+1/2})-x_{n+1/2}\|^2.$$
Puis on applique le résultat de la question précédente entre $x_{n+1/2}$ *et* x_n *pour obtenir l'inégalité désirée :*
$$\|x_{n+1}-x\|^2 \leq \|x_n-x\|^2 - \tau(2-\tau)\left(\left\|\Pi_{C_2}(x_{n+1/2})-x_{n+1/2}\right\|^2 + \|\Pi_{C_1}(x_n)-x_n\|^2\right).$$

3. *Comme* $\tau \in]0,2[$*, on a* $2\tau-\tau^2>0$*. Donc la suite* $(\|x_n-x\|^2)$ *est décroissante, minorée : elle admet une limite. En particulier, cette suite est bornée. Comme*
$$\|x_n\| \leq \|x_n-x\| + \|x\|$$
la suite (x_n) *est également bornée.*

4. *Soit* $\bar{x}$ *une valeur d'adhérence de la suite* (x_n) *et* (x_{n_k}) *une sous-suite qui converge vers* $\bar{x}$*. Notons d'abord que, d'après la question 2, on a*
$$\left\|\Pi_{C_2}(x_{n+1/2})-x_{n+1/2}\right\|^2 + \|\Pi_{C_1}(x_n)-x_n\|^2 \leq (2\tau-\tau^2)^{-1}\left(\|x_n-x\|^2 - \|x_{n+1}-x\|^2\right)$$
où la suite à droite tend vers 0 puisque la suite $(\|x_n-x\|^2)$ *converge. On en déduit que les suites* $(\|\Pi_{C_1}(x_n)-x_n\|)$ *et* $\left(\left\|\Pi_{C_2}(x_{n+1/2})-x_{n+1/2}\right\|\right)$ *tendent vers* 0*. Appliquées à la sous-suite* (x_{n_k})*, la première convergence implique que* $\Pi_{C_1}(\bar{x})=\bar{x}$*. En particulier,* $\bar{x}\in C_1$ *et la suite* $(x_{n_k+1/2})$ *converge également vers* $\bar{x}$*. Alors la second convergence entraîne de même que* $\Pi_{C_2}(\bar{x})=\bar{x}$*, ce qui signifie que* $\bar{x}\in C_2$*. Cela prouve finalement que* $\bar{x}\in C_1\cap C_2$*.*