Optimisation et projection sur un convexe.

• Soit m un entier naturel et $\mathbb{R}^m$ l'espace euclidien usuel muni du produit scalaire noté $(\bullet\,,\,\bullet)$. On désigne par $||\bullet||$ la norme associée. Soit $K \subset \mathbb{R}^m$ un convexe fermé non vide et $P : \mathbb{R}^m \longrightarrow K$ la projection sur K : pour x appartenant à $\mathbb{R}^m$, Px est l'unique point de K tel que $||\, x - Px \,|| \leq ||\, x - z \,||$ pour tout z appartenant à K.

1) Démontrer que si x et y appartiennent à $\mathbb{R}^m$, on a :

(1) $\qquad ||\, Py - Px \,|| \leq ||\, y - x \,||$

(2) $\qquad 0 \leq (x - Px\,,\, Px - Py) \leq ||\, y - x \,||^2\,;$

on pourra utiliser la caractérisation suivante de Px :

$$w = Px \quad \Longleftrightarrow \quad w \in K \quad \text{et} \quad \forall\, z \in K\,,\ (x - w\,,\, z - w) \leq 0\,.$$

Les résultats de cette question pourront être admis pour traiter la suite de l'exercice.

• Pour $x \in \mathbb{R}^m$, on pose $J(x) = \frac{1}{2}\, ||\, x - Px \,||^2$.

2) Démontrer que J est différentiable et que

(3) $\qquad \mathrm{d}J(x)\bullet h = (x - Px\,,\, h)\,, \qquad \forall\, h \in \mathbb{R}^m\,, \qquad \text{soit } J'(x) = x - Px\,.$

• On se donne pour la suite de l'exercice un nombre réel $g > 0$ et on pose $K = \{x \in \mathbb{R}^m\,,\ |\, x_i \,| \leq g\,, \forall\, i,\ 1 \leq i \leq m\}$. On définit la fonction $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ par

(4) $$\psi(x) = \begin{cases} x & \text{si } |\, x \,| \leq g \\ g\,\dfrac{x}{|\, x \,|} & \text{sinon.} \end{cases}$$

3) Démontrer que

(5) $\qquad (Px)_i = \psi(x_i).$

• On introduit pour toute la suite un entier n, une matrice A d'ordre $n \times m$, de rang n, et un vecteur $b \in \mathbb{R}^n$ fixé. On définit l'ensemble T par la relation $T = \{y \in \mathbb{R}^m,\ A\,y = b\}$.

4) Démontrer que le problème qui consiste à chercher x tel que

(6) $$\begin{cases} x \in T \\ J(x) \leq J(y)\,, \qquad \forall\, y \in T \end{cases}$$

a au moins une solution.

5) Soit $z \in \mathbb{R}^m$ donné. On pose $||\, z \,||_1 \equiv \sum_i |\, z_i \,|$. Démontrer que

(7) $$\inf\{\, J(y) - (z\,,\, y)\,,\ y \in \mathbb{R}^m \,\} = -\frac{1}{2}\, ||\, z \,||^2 - g\, ||\, z \,||_1\ .$$

- On introduit le Lagrangien $\mathcal{L}$ suivant

(8) $\qquad \mathcal{L}(y,\, q) \,=\, J(y) \,+\, (q \,,\, b - A\, y) \,, \qquad y \in \mathbb{R}^m \,, \quad q \in \mathbb{R}^n \,.$

6) Démontrer que

(9) $\qquad \sup \{ \mathcal{L}(y,\, q),\; q \in \mathbb{R}^n \} \,=\, J(y) \quad \text{si } y \in T \qquad \text{et } +\infty \text{ sinon.}$

7) On introduit la fonctionnelle auxiliaire

(10) $\qquad G(q) \,=\, -\dfrac{1}{2} \,||\, A^{\rm t} q \,||^2 \,-g\, ||\, A^{\rm t} q \,||_1 \,+ (b \,,\, q)$

et le problème d'optimisation suivant : chercher $p^* \in \mathbb{R}^n$ de sorte que

(11) $\qquad G(p^*) \,\geq\, G(q) \,, \qquad \forall\, q \in \mathbb{R}^n \,.$

Montrer que le problème ci-dessus a une solution unique et proposer un algorithme pour résoudre le problème (10).

8) Le Lagrangien défini à la relation (8) admet-il un point selle $(x,\, p)$? Ecrire le système d'équations vérifié par le couple $(x,\, p)$. Comment p et p^* sont-ils reliés ?

Bertrand Mercier, mai 1997 ; FD, juin 2002.

Optimisation et projection sur un convexe.
Proposition de corrigé.

1) Rappelons que la projection sur le convexe K est caractérisée par les conditions suivantes :

(S1) $\quad Px \in K\,, \qquad (x - Px,\, z - Px) \,\leq\, 0\,, \quad \forall\, z \in K.$

Donc $||\, Py - Px \,||^2 = \,(Py - Px\,,\, Py - y + x - Px + y - x) \,\leq\, (Py - Px\,,\, y - x)$ et la relation (1) est une conséquence de l'inégalité de Cauchy-Schwarz. La condition $(x - Px\,,\, Px - Py) \,\geq\, 0$ est une conséquence directe de la relation (S1) et on a par ailleurs

$$\begin{aligned}(x - Px\,,\, Px - Py) &= \\ &= (x - y\,,\, Px - Py) + (y - Py\,,\, Px - Py) + (Py - Px\,,\, Px - Py) \\ &\leq (x - y\,,\, Px - Py) \,\leq ||\, x - y \,||\; ||\, Px - Py \,|| \leq ||\, x - y \,||^2\end{aligned}$$

ce qui achève de montrer la relation (2).

2) Pour x et $y \in \mathbb{R}^m$, on a

$$\begin{aligned}J(y) - J(x) &= \tfrac{1}{2}\, ||\, x - Px + (y - x + Py - Px) \,||^2 \,-\tfrac{1}{2}\, ||\, x - Px \,||^2 \\ &= (x - Px\,,\, y - x) + J\,(x - Px\,,\, Py - Px) \,+\, \tfrac{1}{2}\, ||\, y - x + Px - Py \,||^2\end{aligned}$$

donc compte tenu des relations (1) et (2),

$$||\, J(y) - J(x) - (x - Px\,,\, y - x) \,|| \leq ||\, y - x \,||^2 \,+ 2\, ||\, y - x \,||^2$$

ce qui établit que J est différentiable et que $\mathrm{d}J(x)$ est représenté par le vecteur $J'(x) = x - Px$, ce qu'exprime la relation (3).

3) On se donne un indice i dans l'intervalle $[1, m]$. Si on pose $\xi_i \,=\, P(x)_i\,$, la relation (S1) écrite avec un vecteur $z \in K$ de la forme $z \,=\, (0, \ldots, \zeta\,, \ldots, 0)$ entraîne

(S2) $\quad (x_i - \xi_i)\,(\zeta - \xi_i) \,\leq\, 0\,, \qquad \forall \zeta$ tel que $|\,\zeta\,| \leq\, g.$

Si x_i est tel que $x_i \,\geq\, g$, alors on a nécessairement $-g \,\leq\, \xi_i \,\leq\, g$ car le projeté de x appartient au convexe K. Donc $x_i \geq g \geq \xi_i$ et la condition (S2) implique que $\xi_i \geq \zeta$ pour tout ζ tel que $|\,\zeta\,| \leq\, g$, donc $\xi_i \geq g$. Comme $|\,\xi_i\,| \leq\, g$ on a nécessairement $\xi_i = g$. Le raisonnement est analogue pour dans le cas où $x_i \leq -g$ et conduit à la relation $\xi_i = -g$. Dans la situation intermédiaire où $-g < x_i < g$, on ne peut pas avoir $x_i < \xi_i \leq g$ car alors la condition (S2) est niée par le choix $\zeta = x_i$ et on ne peut pas avoir non plus $-g \leq \xi_i < x_i$ qui est mis en défaut grâce au même choix $\zeta = x_i$. Donc $\xi_i = x_i$ et la relation (5) est établie.

4) L'ensemble T est une variété affine non vide car la matrice A est de rang n. Donc c'est un fermé non vide de $\mathbb{R}^m$. Il suffit d'utiliser la continuité de J et de démontrer que J est infinie à l'infini pour appliquer le théorème 3.2.1 du cours qui permet de conclure à l'existence d'un point de minimum. Mais vu l'inégalité triangulaire, on a $\| x - Px \| \geq \| x \| - \| Px \| \geq \| x \| - \sqrt{m}\, g$ compte tenu du choix du convexe K ; donc la fonctionnelle J est infinie à l'infini et l'existence d'un point de minimum qui permet de résoudre le problème (6) est démontrée.

5) Une condition nécessaire d'extremum pour le problème (7) s'écrit $J'(x) = z$, soit $x - Px = z$, et pour la i ème coordonnée $x_i - \psi(x_i) = z_i$. Si $z_i = 0$, la condition $\psi(x_i) = x_i$ exprime que $| x_i | \leq g$. Si $z_i > 0$, la relation d'extremum s'écrit $x_i - g = z_i$ et le i ème contributeur de $J(x) - (z, x)$ vaut $\frac{1}{2} z_i^2 - z_i (g + z_i) = -\frac{1}{2} z_i^2 - g\, z_i$. Si $z_i < 0$, on a maintenant $x_i + g = z_i$ et le i ème contributeur de $J(x) - (z, x)$ prend la forme $\frac{1}{2} z_i^2 - z_i (-g + z_i) = -\frac{1}{2} z_i^2 + g\, z_i$. La relation (7) est donc une conséquence du fait que $J(\bullet)$ étant minorée par une fonction quadratique à l'infini, l'expression $J(y) - (z, y)$ est infinie à l'infini.

6) Si $y \in T$, on a $\mathcal{L}(y, q) = J(y)$ qui ne dépend pas de q. Si y n'appartient pas à T, alors la borne supérieure de $\mathcal{L}(y, q)$ si q est arbiraire dans $\mathbb{R}^n$ est égale à $+\infty$. On en déduit la relation (9).

7) Le problème (11) a une solution unique car la fonctionnelle $G(\bullet)$ est alpha-convexe puisque la matrice A est de rang n donc A^t est injective. L'algorithme du gradient simple à pas constant peut être utilisé pour calculer la solution p^* du problème (11).

8) Pour établir l'existence du point selle (x, p) pour le Lagrangien défini à la relation (8), on remarque d'abord que les contraintes "$x \in T$" sont qualifiées puisqu'elles sont affines et que le convexe T n'est pas vide. On établit ensuite que J est convexe. On a :

$$
\begin{aligned}
< \mathrm{d}J(y) - \mathrm{d}J(x)\,,\, y - x > &= (y - Py - x + Px\,,\, y - x) \\
&= (Px - Py\,,\, y - x) + \| y - x \|^2 \\
&\geq -\| Py - Px \| \; \| y - x \| + \| y - x \|^2 \;\; \geq \; 0
\end{aligned}
$$

compte tenu de (1), donc J est convexe. On utilise ensuite le théorème de Kuhn et Tucker qui garantit l'existence d'un point selle dans les conditions précédentes. La fonction J ainsi que les contraintes sont dérivables. Une condition nécessaire pour un point selle (u, p) du Lagrangien $\mathcal{L}$ est de satisfaire les conditions d'annulation des dérivées premières, c'est à dire ici

(S3) $\qquad x - Px - (p\,,\, A\,x) = 0\,, \qquad b - A\,x = 0\,.$

Le problème (11) est exactement le problème dual associé au Lagrangien (8) donc on a simplement $p = p^*$.

FD, juillet 2002.