5MM71 : Méthodes du premier ordre pour l'optimisation non convexe et non lisse

Examen final du 9 janvier 2024 – Durée : 2 heures

Consignes – *Aucun document,* ***à l'exception de deux pages format A4 de notes manuscrites****, n'est autorisé. L'utilisation de calculatrices, téléphones portables (ainsi que tout autre appareil électronique) est interdite. Cet énoncé comporte* ***deux*** *exercices indépendants qui peuvent être traités dans l'ordre de votre choix. Les réponses doivent être soigneusement justifiées, sauf mention contraire. La qualité de la rédaction et la rigueur des raisonnements seront prises en compte dans la notation.* ***Le candidat est autorisé à rédiger en anglais.***

Soient $n, m \in \mathbb{N}^*$. On note $\langle \cdot, \cdot \rangle$ le produit scalaire associé à la base canonique de $\mathbb{R}^n$ (resp. $\mathbb{R}^m$) et $\|\cdot\|$ les normes euclidiennes associées.

Exercice 1

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $g : \mathbb{R}^m \to \mathbb{R}$ deux fonctions convexes, respectivement L_f- et L_g-régulières. Soit $R > 0$. On note

$$\mathcal{B} = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid \|x\|^2 + \|y\|^2 \leq R^2 \right\}$$

la boule fermée de $\mathbb{R}^n \times \mathbb{R}^m$, centrée en l'origine, de rayon R. On s'intéresse à la résolution du problème d'optimisation suivant :

$$\min_{(x,y)\in\mathcal{B}} f(x) + g(y) \tag{$\mathcal{P}$}$$

1. Justifier que le problème $(\mathcal{P})$ admet au moins une solution.
2. Réécrire le problème $(\mathcal{P})$ sous la forme d'un problème d'optimisation non contrainte, de fonction objectif $J : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$.
3. Écrire les itérations de l'algorithme *Forward-Backward Splitting* (FBS) de pas de temps $\tau > 0$ pour la minimisation de J. Sans démonstration (mais en justifiant votre réponse), indiquer à quelle condition sur le pas de temps τ cet algorithme converge vers une solution de $(\mathcal{P})$. *On explicitera, en fonction des données du problème, les valeurs que τ peut prendre pour assurer la convergence.*

Soient $\tau, \sigma > 0$. On considère l'algorithme suivant :

$$y_0 \in \mathcal{B}(0, R) \qquad \text{et} \qquad \forall\, k \in \mathbb{N}, \quad \begin{cases} x_{k+1} = \operatorname{proj}_{\mathcal{B}(0,\sqrt{R^2-\|y_k\|^2})}\big(x_k - \tau\,\nabla f(x_k)\big) \\ y_{k+1} = \operatorname{proj}_{\mathcal{B}(0,\sqrt{R^2-\|x_{k+1}\|^2})}\big(y_k - \sigma\,\nabla g(y_k)\big) \end{cases}$$

où $\mathcal{B}(0, r)$ désigne la boule fermée centrée en l'origine, de rayon r.

4. Justifier que l'algorithme considéré est bien défini.

5. Montrer que

$$\forall\, k \in \mathbb{N}^*, \qquad f(x_k) \geq f(x_{k+1}) + \frac{1}{2}\left(\frac{2}{\tau} - L_f\right)\|x_{k+1} - x_k\|^2$$

et $$\forall\, k \in \mathbb{N}^*, \qquad g(y_k) \geq g(y_{k+1}) + \frac{1}{2}\left(\frac{2}{\sigma} - L_g\right)\|y_{k+1} - y_k\|^2$$

En déduire que, si $0 < \tau < 2/L_f$ et que $0 < \sigma < 2/L_g$, alors $(f(x_k)+g(y_k))_{k\in\mathbb{N}}$ est décroissante et convergente.

6. *À partir de cette question, $0 < \tau < 2/L_f$ et que $0 < \sigma < 2/L_g$.* Montrer que les suites $(\|x_{k+1} - x_k\|)_{k\in\mathbb{N}}$ et $(\|y_{k+1} - y_k\|)_{k\in\mathbb{N}}$ convergent vers 0.

7. Soit $k \in \mathbb{N}$. En utilisant la caractérisation du point proximal, montrer que

$$q_{k+1} = \frac{x_k - x_{k+1}}{\tau} + \nabla f(x_{k+1}) - \nabla f(x_k) \in \partial_x J(x_{k+1}, y_k)$$

et $$p_{k+1} = \frac{y_k - y_{k+1}}{\tau} + \nabla g(y_{k+1}) - \nabla g(y_k) \in \partial_y J(x_{k+1}, y_{k+1})$$

($\star$) 8. *On suppose que la suite $((x_k, y_k))_{k\in\mathbb{N}}$ converge, de limite (x^*, y^*).* Montrer que

$$\forall\,(x,y) \in \mathbb{R}^n \times \mathbb{R}^m, \qquad J(x, y^*) \geq J(x^*, y^*) \quad \text{et} \quad J(x^*, y) \geq J(x^*, y^*)$$

En déduire que

$$\forall\,(x,y) \in \mathcal{B}, \qquad f(x) + g(y) \geq f(x_*) + g(y^*)$$

puis que (x^*, y^*) est solution de $(\mathcal{P})$. *On pourra considérer dans un premier temps le cas où (x,y) est à l'intérieur de $\mathcal{B}$.*

9. Dans le cas convergent, quel intérêt peut présenter cet algorithme par rapport à l'algorithme FBS ? *On pourra considérer le cas où $L_f = 1$ et $L_g = 100$.*

Exercice 2

Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, s.c.i. et propre. On pose

$$\forall\,(x,y) \in \mathbb{R}^n \times \mathbb{R}^n, \qquad \mathcal{L}(x;y) = \frac{1}{2}\|x\|^2 + \langle x, y\rangle - f^*(y)$$

On considère au problème de recherche de point-selle

$$\min_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^n} \mathcal{L}(x;y)$$

1. Quel est le problème primal associé à la fonction de couplage $\mathcal{L}$? Justifier qu'il admet une unique solution.
2. Quel est le problème dual associé à $\mathcal{L}$? Justifier qu'il admet une unique solution.
3. Caractériser la solution primale au premier ordre.
4. Montrer que la dualité forte est vérifiée pour $\mathcal{L}$.
5. En déduire l'expression du point-selle de $\mathcal{L}$.

Fin de l'énoncé

Exercice 1

Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $g : \mathbb{R}^m \to \mathbb{R}$ deux fonctions convexes, respectivement L_f- et L_g-régulières. Soit $R > 0$. On note

$$\mathcal{B} = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid \|x\|^2 + \|y\|^2 \leq R^2 \right\}$$

la boule fermée de $\mathbb{R}^n \times \mathbb{R}^m$, centrée en l'origine, de rayon R. On s'intéresse à la résolution du problème d'optimisation suivant :

$$\min_{(x,y)\in\mathcal{B}} f(x) + g(y) \qquad (\mathcal{P})$$

1. Justifier que le problème $(\mathcal{P})$ admet au moins une solution.

(1) Puisque B est compact et $h(x,y) = f(x) + g(y)$ est différentiable (donc continue), par le théorème de Weierstrass, on a que

$h(x,y)$ atteint son borne sur B.

autrement dit,

le problème (P) admet au moins une solution.

2. Réécrire le problème $(\mathcal{P})$ sous la forme d'un problème d'optimisation non contrainte, de fonction objectif $J : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$.

(2) En considérons

$$\chi_B : \begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^m & \longrightarrow & \mathbb{R} \cup \{+\infty\} \\ (x,y) & \longmapsto & \begin{cases} 0 & \text{si } (x,y) \in B \\ +\infty & \text{sinon} \end{cases} \end{array}$$

Le problème non contrainte est

$$\min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} f(x) + g(y) + \chi_B(x,y)$$

donc

$$J(x,y) = f(x) + g(y) + \chi_B(x,y)$$

3. Écrire les itérations de l'algorithme *Forward-Backward Splitting* (FBS) de pas de temps $\tau > 0$ pour la minimisation de J. Sans démonstration (mais en justifiant votre réponse), indiquer à quelle condition sur le pas de temps τ cet algorithme converge vers une solution de $(\mathcal{P})$. *On explicitera, en fonction des données du problème, les valeurs que τ peut prendre pour assurer la convergence.*

---

(3) La fonction J est la somme de la fonction différentiable $h(x,y) = f(x) + g(y)$ et la fonction simple $\chi_B(x,y)$.

Puisque $\operatorname{prox}_{\tau\chi_B} = \operatorname{prox}_{\chi_B} = \operatorname{proj}_B$ et que

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^m, \quad \nabla h(x,y) = \begin{bmatrix} \nabla f(x) \\ \nabla g(y) \end{bmatrix}$$

les itérations de l'algorithme FBS s'écrit

$$(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m \text{ et } \forall k \in \mathbb{N}, \quad \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \operatorname{proj}_B \begin{pmatrix} x_k - \tau \nabla f(x_k) \\ y_k - \tau \cdot \nabla g(y_k) \end{pmatrix}$$

Soit $(x,y), (x',y') \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned} \| h(x,y) - h(x',y') \|^2 &= \| f(x) + g(y) - f(x') - g(y') \|^2 \\ &\leq \| f(x) - f(x') \|^2 + \| g(y) - g(y') \|^2 \\ &\leq L_f \cdot \| x - x' \|^2 + L_g \cdot \| y - y' \|^2 \\ &\leq \max\{L_f, L_g\} \cdot (\| x - x' \|^2 + \| y - y' \|^2) \\ &= \max\{L_f, L_g\} \cdot \| (x,y) - (x',y') \|^2 \end{aligned}$$

donc h est $\max\{L_f, L_g\}$ − régulière fonction.

Et χ_B est convexe, propre et continue sur son domaine.

Si $\tau \in \left] 0, \frac{1}{\max\{L_f, L_g\}} \right]$, la suite $(x_k, y_k)_{k \in \mathbb{N}}$ converge vers le minimiseur de J, c-à-d une solution de $(\mathcal{P})$.

Soient $\tau, \sigma > 0$. On considère l'algorithme suivant :

$$y_0 \in \mathcal{B}(0, R) \qquad \text{et} \qquad \forall\, k \in \mathbb{N}, \quad \begin{cases} x_{k+1} = \operatorname{proj}_{\mathcal{B}(0,\sqrt{R^2 - \|y_k\|^2})}\left(x_k - \tau\, \nabla f(x_k)\right) \\ y_{k+1} = \operatorname{proj}_{\mathcal{B}(0,\sqrt{R^2 - \|x_{k+1}\|^2})}\left(y_k - \sigma\, \nabla g(y_k)\right) \end{cases}$$

où $\mathcal{B}(0, r)$ désigne la boule fermée centrée en l'origine, de rayon r.

4. Justifier que l'algorithme considéré est bien défini.

---

(4) Il suffit de montrer que les projections sur les ensembles $B(0, \sqrt{R^2 - \|y_k\|^2})$ et $B(0, \sqrt{R^2 - \|x_{k+1}\|^2})$ sont bien définies pour tout $k \in \mathbb{N}$.

(i) les ensembles considérés sont convexes

(ii) ils sont fermés par définitions

(iii) ils sont non-vide :

si $y_k \in B(0, R)$, alors $R^2 - \|y_k\|^2 \geqslant 0$.

le boule $B(0, \sqrt{R^2 - \|y_k\|^2})$ est bien définie.

ainsi, x_{k+1} est bien défini et appartient à $B(0, \sqrt{R^2 - \|y_k\|^2}) \subset B(0, R)$, de sorte que le même argument assure que y_{k+1} est bien défini.

---

5. Montrer que

$$\forall\, k \in \mathbb{N}^*, \qquad f(x_k) \geq f(x_{k+1}) + \frac{1}{2}\left(\frac{2}{\tau} - L_f\right)\|x_{k+1} - x_k\|^2$$

$$\text{et} \qquad \forall\, k \in \mathbb{N}^*, \qquad g(y_k) \geq g(y_{k+1}) + \frac{1}{2}\left(\frac{2}{\sigma} - L_g\right)\|y_{k+1} - y_k\|^2$$

En déduire que, si $0 < \tau < 2/L_f$ et que $0 < \sigma < 2/L_g$, alors $(f(x_k) + g(y_k))_{k \in \mathbb{N}}$ est décroissante et convergente.

日期: /

15) Soit $k \in \mathbb{N}^*$, puisque f est L_f − régulière, par le lemme de descent,

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{L_f}{2} \|x_{k+1} - x_k\|^2$$

Puisque $\mathrm{proj}_C = \mathrm{prox}_{\chi_C}$, alors

$$x_{k+1} = \mathrm{prox}_{\chi_{B(0,\sqrt{R^2 - \|y_k\|^2})}}(x_k - \tau \nabla f(x_k))$$

par la caractérisation du point proximal,

$$x_k - \tau \nabla f(x_k) - x_{k+1} \in \partial \chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_{k+1})$$

par la définition de la sous gradient,

$\forall x \in \mathbb{R}^n$, $\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_{k+1})$

$$\geq \chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x) + \langle x_k - \tau \nabla f(x_k) - x_{k+1}, x_{k+1} - x \rangle$$

alors $\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_{k+1})$

$$\geq \chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_k) + \langle x_k - \tau \nabla f(x_k) - x_{k+1}, x_{k+1} - x_k \rangle$$

par définition, x_{k+1} est projété de $B(0,\sqrt{R^2-\|y_k\|^2})$, donc

$x_{k+1} \in B(0,\sqrt{R^2-\|y_k\|^2})$, $\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_{k+1}) = 0$

et y_k est projété de $B(0,\sqrt{R^2-\|x_k\|^2})$, alors $y_k \in B(0,\sqrt{R^2-\|x_k\|^2})$

donc $\|y_k\|^2 \leq R^2 - \|x_k\|^2$

$$\|x_k\|^2 \leq R^2 - \|y_k\|^2$$

c-à-d $x_k \in B(0,\sqrt{R^2-\|y_k\|^2})$ donc $\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_k) = 0$

日期: /

f convexe
≥ 0

Alors

$$0 \geq \langle x_k - \tau \nabla f(x_k) - x_{k+1} , x_{k+1} - x_k \rangle$$
$$= -\tau \langle \nabla f(x_k), x_{k+1} - x_k \rangle - \|x_{k+1} - x_k\|^2$$

on en déduit que

??

$$\langle \nabla f(x_k), x_{k+1} - x_k \rangle = -\frac{1}{\tau} \|x_{k+1} - x_k\|^2$$

donc

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{\tau}\|x_{k+1} - x_k\|^2 + \frac{L_f}{2}\|x_{k+1} - x_k\|^2$$
$$f(x_k) \geq f(x_{k+1}) + \frac{1}{2}\left(\frac{2}{\tau} - L_f\right) \cdot \|x_{k+1} - x_k\|^2$$

De même

$$g(y_k) \geq g(y_{k+1}) + \frac{1}{2}\left(\frac{2}{\sigma} - L_g\right) \cdot \|y_{k+1} - y_k\|^2$$

Si $0 < \tau < \frac{2}{L_f}$ et $0 < \sigma < \frac{2}{L_g}$, on a

$$\frac{2}{\tau} - L_f > 0 \quad \text{et} \quad \frac{2}{\sigma} - L_g > 0$$

en sommant les deux inégalités,

$$f(x_k) + g(y_k) \geq f(x_{k+1}) + g(y_{k+1}) + \frac{1}{2}\left(\frac{2}{\tau} - L_f\right) \cdot \|x_{k+1} - x_k\|^2 + \frac{1}{2}\left(\frac{2}{\sigma} - L_g\right) \cdot \|y_{k+1} - y_k\|^2$$

alors

$$f(x_{k+1}) + g(y_{k+1}) - (f(x_k) + g(y_k))$$
$$\leq -\frac{1}{2}\left(\frac{2}{\tau} - L_f\right) \cdot \|x_{k+1} - x_k\|^2 - \frac{1}{2}\left(\frac{2}{\sigma} - L_g\right) \cdot \|y_{k+1} - y_k\|^2$$
$$\leq 0$$

donc $(f(x_k)+g(y_k))_{k\in\mathbb{N}}$ est décroissante.

Par construction, les x_k et y_k sont de norme inférieures que R. Ainsi, les (x_k, y_k) sont de norme inférieures que $\sqrt{2}R$.
Alors la suite $((x_k,y_k))_{k\in\mathbb{N}}$ est contenue dans la boule fermé de rayon $\sqrt{2}R$, c-à-d dans un compact.
Or $h(x,y)$ est continue, donc elle y atteint son borne. En particulier, $h(x,y)$ est minorée.
Donc $(f(x_k)+g(y_k))_{k\in\mathbb{N}}$ est minorée.
Et comme elle est décroissante, alors $(f(x_k)+g(y_k))_{k\in\mathbb{N}}$ est convergente.

6. *À partir de cette question,* $0 < \tau < 2/L_f$ *et que* $0 < \sigma < 2/L_g$. Montrer que les suites $(\|x_{k+1}-x_k\|)_{k\in\mathbb{N}}$ et $(\|y_{k+1}-y_k\|)_{k\in\mathbb{N}}$ convergent vers 0.

(6) D'après (5), on a

$$f(x_{k+1})+g(y_{k+1})-(f(x_k)+g(y_k))$$
$$\leq -\frac{1}{2}\left(\frac{2}{\tau}-L_f\right)\cdot\|x_{k+1}-x_k\|^2-\frac{1}{2}\left(\frac{2}{\sigma}-L_g\right)\cdot\|y_{k+1}-y_k\|^2$$

alors

$$f(x_k)+g(y_k)-(f(x_{k+1})+g(y_{k+1}))$$
$$\geq \frac{1}{2}\left(\frac{2}{\tau}-L_f\right)\cdot\|x_{k+1}-x_k\|^2+\frac{1}{2}\left(\frac{2}{\sigma}-L_g\right)\cdot\|y_{k+1}-y_k\|^2$$

Puisque $(f(x_k)+g(y_k))_{k\in\mathbb{N}}$ converge vers 0 quand k tends vers $+\infty$, alors

$$\frac{1}{2}\left(\frac{2}{\tau}-L_f\right)\cdot\|x_{k+1}-x_k\|^2+\frac{1}{2}\left(\frac{2}{\sigma}-L_g\right)\cdot\|y_{k+1}-y_k\|^2$$

converge vers 0 quand k tends vers $+\infty$.

Et comme $\left(\frac{2}{\tau}-L_f\right)\cdot\|x_{k+1}-x_k\|^2\geq 0$ et $\left(\frac{2}{\sigma}-L_g\right)\cdot\|y_{k+1}-y_k\|^2\geq 0$

donc $\|x_{k+1}-x_k\|$ et $\|y_{k+1}-y_k\|$ converge vers 0.

7. Soit $k\in\mathbb{N}$. En utilisant la caractérisation du point proximal, montrer que

$$q_{k+1}=\frac{x_k-x_{k+1}}{\tau}+\nabla f(x_{k+1})-\nabla f(x_k)\in\partial_x J(x_{k+1},y_k)$$

et
$$p_{k+1}=\frac{y_k-y_{k+1}}{\tau}+\nabla g(y_{k+1})-\nabla g(y_k)\in\partial_y J(x_{k+1},y_{k+1})$$

(7) Par la définition

$$J(x,y)=f(x)+g(y)+\chi_B(x,y)$$

on fixe $y=y_k$, alors

$$J(x,y_k)=f(x)+g(y_k)+\chi_B(x,y_k)$$
$$=f(x)+g(y_k)+\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x)$$

et

$$\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x)=\tau\cdot\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x)$$
$$=\tau\cdot J(x,y_k)-\tau\cdot f(x)-\tau\cdot g(y_k)$$

donc

$$\partial\chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x)=\tau\cdot\partial_x J(x,y_k)-\tau\cdot\nabla f(x)$$

日期: /

Par l'hypothèse,

$$x_{k+1} = \mathrm{prox}_{\chi_{B(0,\sqrt{R^2+\|y_k\|^2})}}(x_k - \tau\cdot\nabla f(x_k))$$

par la caractérisation du point proximal,

$$x_k - \tau\cdot\nabla f(x_k) - x_{k+1} \in \partial \chi_{B(0,\sqrt{R^2-\|y_k\|^2})}(x_{k+1})$$

$$x_k - \tau\cdot\nabla f(x_k) - x_{k+1} \in \tau\cdot\partial_x J(x_{k+1}, y_k) - \tau\cdot\nabla f(x_{k+1})$$

alors

$$\frac{x_k - x_{k+1}}{\tau} + \nabla f(x_{k+1}) - \nabla f(x_k) \in \partial_x J(x_{k+1}, y_k)$$

La même, on obtient que

$$\frac{y_k - y_{k+1}}{6} + \nabla g(y_{k+1}) - \nabla g(y_k) \in \partial_y J(x_{k+1}, y_{k+1})$$

($\star$) 8. *On suppose que la suite $((x_k, y_k))_{k\in\mathbb{N}}$ converge, de limite (x^*, y^*).* Montrer que

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^m, \qquad J(x, y^*) \geq J(x^*, y^*) \quad \text{et} \quad J(x^*, y) \geq J(x^*, y^*)$$

En déduire que

$$\forall (x,y) \in \mathcal{B}, \qquad f(x) + g(y) \geq f(x_*) + g(y^*)$$

puis que (x^*, y^*) est solution de $(\mathcal{P})$. *On pourra considérer dans un premier temps le cas où (x, y) est à l'intérieur de $\mathcal{B}$.*

(8) Par la définition de la sous gradient convexe,

$\forall x \in \mathbb{R}^n, \qquad J(x, y_k) \geq J(x_{k+1}, y_k) + \langle q_{k+1}, x - x_{k+1} \rangle$

puisque $(x_{k+1}, y_k) \in B$ et B est fermé, $(x^*, y^*) \in B$

alors $\quad \chi_B(x_{k+1}, y_k) = \chi_B(x^*, y^*) = 0$

puisque f et g sont continue,

$$\lim_{k\to+\infty} J(x_{k+1}, y_k) = \lim_{k\to+\infty} f(x_{k+1}) + g(x_k) = f(x^*) + g(x^*) = J(x^*, y^*)$$

Soit $x \in \mathbb{R}^n$. puisque

$$\begin{aligned}\|q_{k+1}\| &\leq \frac{\|x_k - x_{k+1}\|}{\tau} + \|\nabla f(x_{k+1}) - \nabla f(x_k)\| \\ &\leq \frac{1}{\tau}\cdot\|x_{k+1} - x_k\| + L_f \cdot \|x_{k+1} - x_k\| \\ &= \left(\frac{1}{\tau} + L_f\right)\cdot\|x_{k+1} - x_k\|\end{aligned}$$

alors

$$\begin{aligned}\langle q_{k+1}, x - x_{k+1}\rangle &\leq \|q_{k+1}\|\cdot\|x - x_{k+1}\| \\ &\leq \left(\frac{1}{\tau} + L_f\right)\cdot\|x_{k+1} - x_k\|\cdot\|x - x_{k+1}\|\end{aligned}$$

et comme $(\|x_{k+1} - x_k\|)_{k\in\mathbb{N}}$ converge vers 0, alors

$$\lim_{k\to+\infty} \langle q_{k+1}, x - x_{k+1}\rangle = 0$$

Si $(x, y^*) \notin B$, alors $J(x, y^*) = +\infty$, le resultat est immédiat, puisque

$$+\infty > J(x^*, y^*) + 0$$

Si $(x, y^*) \in B$, alors deux situations sont possibles.

① Il existe $k_0 \in \mathbb{N}$ tel que pour tout $k \geq k_0$, $(x, y_k) \in B$.
Puisque B est fermé, par passage à la limite, on obtient que $(x, y^*) \in B$.

Puisque J est continue sur son domaine, en passant à la limite, on obtient

$$J(x, y^*) \geqslant J(x^*, y^*)$$

② Pour tout $k \in \mathbb{N}$, $(x, y_k) \notin B$, c-à-d que

$$\|x\|^2 + \|y_k\|^2 > R^2$$

et on a $J(x, y_k) = +\infty$ mais $J(x, y^*) < +\infty$.

Soit $((u_j, y^*))_{j\in\mathbb{N}}$ une suite de $\mathrm{int}(B)$ de limite (x, y^*).

Soit $j \in \mathbb{N}$, puisque $((u_j, y_k))_{k\in\mathbb{N}}$ converge vers (u_j, y^*), il existe nécessairement $k_0 \in \mathbb{N}$ tel que $(u_j, y_k) \in B$ pour tout $k \geqslant k_0$.

D'après le point précédent, on en déduit que

$$\forall j \in \mathbb{N}, \quad J(u_j, y^*) \geqslant J(x^*, y^*)$$

Or, on a $J(u_j, y^*) = f(u_j) + g(y^*)$ et que f est continue, on peut passer à limite quand $j \to +\infty$ et obtient

$$J(x, y^*) \geqslant J(x^*, y^*)$$

La même, on obtient $J(x^*, y) \geqslant J(x^*, y^*)$

Alors on a $\quad f(x) + g(y^*) \geqslant f(x^*) + g(y^*)$

$\quad\quad\quad\quad f(x^*) + g(y) \geqslant f(x^*) + g(y^*)$

donc $f(x) \geq f(x^*)$ et $g(y) \geq g(y^*)$

alors $f(x) + g(y) \geq f(x^*) + g(y^*)$

Par définition de la solution du probléme d'optimisation sous contraint, alors

(x^*, y^*) est la solution de (P)

9. Dans le cas convergent, quel intérêt peut présenter cet algorithme par rapport à l'algorithme FBS ? *On pourra considérer le cas où $L_f = 1$ et $L_g = 100$.*

(9) Dans FBS, le pas de temps est majoré par $\frac{1}{\max\{L_f, L_g\}}$ qui vaut dans l'exemple $\frac{1}{100}$.

Dans l'algorithme proposé, les pas de temps s'adaptent à la régularité des deux termes réguliers. En particulier, le pas de temps sont majorés respectivement par 1 et $\frac{1}{100}$, ce qui signifie qu'il est 100 fois plus grand pour la mise-à-jour de x_k.

On peut donc espérer une convergence plus rapide comparé à celle de FBS.

Exercice 2

Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, s.c.i. et propre. On pose

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \qquad \mathcal{L}(x; y) = \frac{1}{2}\|x\|^2 + \langle x, y\rangle - f^*(y)$$

On considère au problème de recherche de point-selle

$$\min_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^n} \mathcal{L}(x; y)$$

1. Quel est le problème primal associé à la fonction de couplage $\mathcal{L}$? Justifier qu'il admet une unique solution.

(1) Par définition, le problème primal est

$$\min_{x\in\mathbb{R}^n} J(x)$$

avec

$$J(x) = \sup_{y\in\mathbb{R}^n} \mathcal{L}(x;y) = \frac{1}{2}\|x\|^2 + \sup_{y\in\mathbb{R}^n} \langle x, y\rangle - f^*(y)$$

Puisque f est convexe, s.c.i. et propre, on a $f^{**} = f$

$$\sup_{y\in\mathbb{R}^n} \langle x, y\rangle - f^*(y) = f^{**}(x) = f(x)$$

Donc

$$J(x) = \frac{1}{2}\|x\|^2 + f(x)$$

Et comme J est fortement convexe, le problème primal admet une unique solution.

2. Quel est le problème dual associé à $\mathcal{L}$? Justifier qu'il admet une unique solution.

(2) Le problème dual est

$$\max_{y\in\mathbb{R}^n} E(y)$$

avec $\forall y \in \mathbb{R}^n$, $E(y) = \inf_{x\in\mathbb{R}^n} \mathcal{L}(x;y) = \inf_{x\in\mathbb{R}^n}\left\{\frac{1}{2}\|x\|^2 + \langle x,y\rangle\right\} - f^*(y)$

car $x \longmapsto \frac{1}{2}\|x\|^2 + \langle x,y\rangle$ est fortement convexe, elle admet une unique solution, par la règle de Fermat

$$x^* + y = 0 \iff x^* = -y$$

donc $\quad E(y) = -\frac{1}{2}\|y\|^2 - f^*(y)$

Puisque $-E(y)$ est fortement convexe, elle admet une seule solution.

3. Caractériser la solution primale au premier ordre.

(3) $$\partial J(x) = x + \partial f(x)$$

Par la régle de Fermat

$$0 \in \partial J(x^*) = x^* + \partial f(x^*)$$

donc $\quad -x^* \in \partial f(x^*)$

4. Montrer que la dualité forte est vérifiée pour $\mathcal{L}$.

(4) On remarque que la solution primale est l'unique élément de $\mathrm{prox}_f(0)$. Et la solution duale est l'unique élément de $\mathrm{prox}_{f^*}(0)$.

Par l'identité de MOREAU,

$$\mathrm{prox}_f(0) + \mathrm{prox}_{f^*}(0) = 0$$

日期:　　/

Donc $y^* = -x^* \in \partial f(x^*)$

Par l'identité de LEGENDRE-FENCHEL,

$$f(x^*) + f^*(-x^*) = \langle -x^*, x^* \rangle = -\|x^*\|^2$$

Autrement dit

$$J(x^*) = \frac{1}{2}\|x^*\|^2 + f(x^*) = -\frac{1}{2}\|x^*\|^2 - f^*(-x^*) = E(-x^*)$$

donc le saut de dualité est nul.

La dualité forte est vérifié.

5. En déduire l'expression du point-selle de $\mathcal{L}$.

(5) Puisque

- $x^* = \text{prox}_f(0)$ est l'unique solution primale
- $-x^*$ est l'unique solution duale
- La dualité forte est vérifiée.

Alors

L'unique point-selle de $\mathcal{L}$ est $(x^*, -x^*)$

日期:　　/