5MAS02 : Méthodes du premier ordre pour l'optimisation non convexe et non lisse
Examen final du 29 janvier 2021 – 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 largement 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.***

Exercice 1

Soit $(n, m) \in (\mathbb{N}^*)^2$, $A \in \mathcal{M}_{n,m}(\mathbb{R})$ et $b \in \mathbb{R}^m$. Considérons le problème d'optimisation suivant :

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|A\,x - b\|_2^2 + \|x\|_1$$

1. Appliquer l'algorithme *Forward-Backward Splitting* pour résoudre ce problème. Justifier **soigneusement** de la convergence de cette méthode dans le cas considéré.
2. Appliquer l'algorithme de Chambolle–Pock pour résoudre ce problème. Quelle est la formulation primale-duale du problème associée ? Justifier **soigneusement** de la convergence de cette méthode dans le cas considéré.

Exercice 2

Soit $n \in \mathbb{N}^*$. On note $\langle \cdot, \cdot \rangle$ le produit scalaire associé à la base canonique de $\mathbb{R}^n$ et $\|\cdot\|$ la norme euclidienne associée.

Introduction d'une norme. Soit $\Sigma \in \mathcal{M}_{n,n}(\mathbb{R})$ une matrice symétrique définie positive. On considère la norme suivante :

$$\|\cdot\|_\Sigma : \left\{ \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \sqrt{\langle x, \Sigma^{-1} x \rangle} \end{array} \right.$$

Problème étudié. Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est une fonction convexe, propre et continue sur son domaine, que l'on suppose fermé. On suppose que f est coercive.

On s'intéresse à la résolution du problème d'optimisation suivant :

$$\min_{x \in \mathbb{R}^n} f(x) \tag{$\mathcal{P}$}$$

Algorithme proposé. Pour résoudre $(\mathcal{P})$, on introduit la notation suivante :

$$\forall\, x^0 \in \mathbb{R}^n, \qquad (\mathrm{Id} + \Sigma\,\partial f)^{-1}(x^0) = \arg\min_{x\in\mathbb{R}^n} \left\{ f(x) + \frac{1}{2}\,\|x - x^0\|_\Sigma^2 \right\}$$

et on s'intéresse à l'algorithme suivant : pour $x_0 \in \mathbb{R}^n$, on définit pour tout $k \in \mathbb{N}$

$$x_{k+1} = (\mathrm{Id} + \Sigma\,\partial f)^{-1}(x_k)$$

1. Lorsque $\Sigma = \sigma\, I_n$, avec $\sigma > 0$, quel algorithme reconnaît-on ?
2. Justifier rapidement que $(\mathrm{Id} + \Sigma\,\partial f)^{-1}(x^0)$ contient exactement un élément.
3. À quoi correspondent les points fixes de cet algorithme ? (*On rappelle qu'il s'agit des points $\hat{x} \in \mathbb{R}^n$ pour lesquels $(Id + \Sigma\,\partial f)^{-1}(\hat{x}) = \hat{x}$*).
4. Montrer qu'il existe $C > 0$ tel que, pour tout $k \in \mathbb{N}$,
$$f(x_{k+1}) \leq f(x_{k+1}) + \frac{C}{2}\,\|x_{k+1} - x_k\|^2 \leq f(x_k)$$
5. Montrer que la suite $(f(x_k))_{k\in\mathbb{N}}$ est décroissante. En déduire qu'elle converge.
6. Montrer que la série $\sum \|x_{k+1} - x_k\|^2$ est convergente. En déduire que la suite des $x_{k+1} - x_k$ converge vers 0.
7. Soit $k \in \mathbb{N}^*$. À l'aide de la règle de Fermat, caractériser l'unique point de $(\mathrm{Id} + \Sigma\,\partial f)^{-1}(x^0)$. En déduire que $p_k = \Sigma^{-1}(x_k - x_{k+1})$ est un sous-gradient de f en x_k.
8. Justifier que la suite des x_k admet une sous-suite $(x_{k_j})_{j\in\mathbb{N}}$ qui converge. Qu'en est-il de la suite des $J(x_{k_j})$? *Indice : on pourra commencer par remarquer que, pour tout $k \in \mathbb{N}$, le point x_k appartient au domaine de f.* Justifier également que la suite $(p_{k_j})_{j\in\mathbb{N}}$ converge vers 0.
9. En appliquant la propriété de fermeture du sous-différentiel, montrer que les valeurs d'adhérence de $(x_k)_{k\in\mathbb{N}}$ sont des minimiseurs de f.
10. En déduire que, si f est strictement convexe, alors l'algorithme considéré converge vers une solution du problème $(\mathcal{P})$.

Résultats admis – Vous pouvez utiliser n'importe lequel des résultats suivants sans les justifier.

1. $\Sigma^\top = \Sigma$ et $\langle \Sigma\, x, x\rangle > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$.
2. Σ est inversible.
3. Σ^{-1} est symétrique définie positive.
4. $\|\cdot\|_\Sigma^2$ définit une norme et est différentiable de gradient
$$\forall\, x \in \mathbb{R}^n, \qquad \nabla(\|\cdot\|_\Sigma^2)(x) = 2\Sigma^{-1}x$$
5. $\|\cdot\|_\Sigma^2$ est fortement convexe.
6. Sur $\mathbb{R}^n$, toutes les normes sont équivalentes. Autrement dit, il existe $(a, b) \in (\mathbb{R}^+)^2$ tel que
$$\forall\, x \in \mathbb{R}^n, \qquad a\,\|x\| \leq \|x\|_\Sigma \leq b\,\|x\|$$

Fin de l'énoncé