Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

Master $1^{\text{ère}}$ année, MMAS, 2025-2026
Optimisation

Partiel du 12/11/2025
Durée 1h30 ou 2h pour les 1/3 temps.

Exercice 1 (4.5pt)

On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = -2y^3 + 4y^2 + 5x^2y - 10xy + 6$.

1. Justifier brièvement la régularité de f, puis déterminer son gradient ainsi que ses points critiques.
2. La fonction f admet-elle un maximum local ?

Exercice 2 (4pt)

Soit $n \in \mathbb{N}^*$. On considère une fonction $g : \mathbb{R}^n \to \mathbb{R}$ supposée $\mathcal{C}^2$, convexe et minorée sur $\mathbb{R}^n$. On pose pour tout $x \in \mathbb{R}^n$, $f(x) = \frac{1}{2\tau} \|x - y\|_2^2 + g(x)$ où $y \in \mathbb{R}^n$ et $\tau > 0$ sont fixés.

On s'intéresse au problème d'optimisation suivant

$$(\mathcal{P}_1) \qquad \inf_{x \in \mathbb{R}^n} f(x).$$

1. Montrer que le problème $(\mathcal{P}_1)$ est bien posé.
2. Montrer que f admet un unique minimum global sur $\mathbb{R}^n$.

Exercice 3 (11.5pt)

Soit $M, n \in \mathbb{N} \setminus \{0, 1\}$. Dans la suite, on considère $\mathbb{R}^M$ et $\mathbb{R}^n$ munis de leurs produits scalaires canoniques respectifs que l'on note $\langle \cdot, \cdot \rangle$ indépendamment de la dimension. Associés à ces produits scalaires, on utilise la notation commune $\|\cdot\|$ pour les normes euclidiennes.

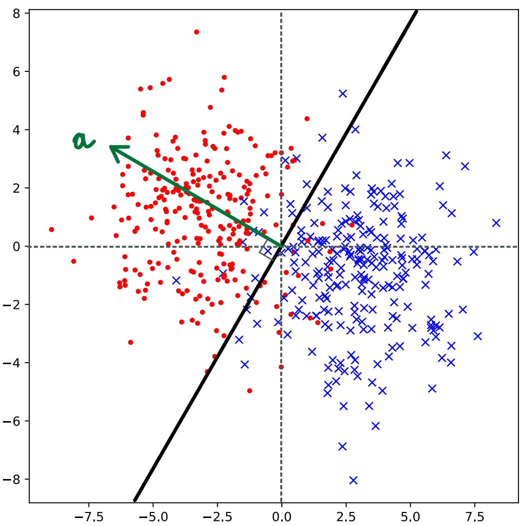

Figure 1. Exemple de points $x_i \in \mathbb{R}^2$. Ceux représentés par une croix, respectivement par un point, appartiennent à la classe -1, resp. 1. La droite représentée, notée $\mathcal{D}$, est $\{x \in \mathbb{R}^2 \,/\, \langle x, a \rangle = 0\}$. Vu le sens de a, tous les $x \in \mathbb{R}^2$ à gauche de $\mathcal{D}$ satisfont $\langle x, a \rangle > 0$, et à droite de $\mathcal{D}$, $\langle x, a \rangle < 0$. On associe à $f_a : x \in \mathbb{R}^n \mapsto \langle x, a \rangle$ (définissant en particulier $\mathcal{D}$), entièrement caractérisée par $a \in \mathbb{R}^2$, un coût $F(a)$. L'idée est de trouver un a de plus faible coût $F(a)$, en espérant que le F définie par le modèle permettent d'aboutir à un f_a effectuant une "bonne" classification.

On suppose donnée une séquence de M points $x_1, x_2, \ldots, x_M$ de $\mathbb{R}^n$, et pour chacun d'entre eux est associé une classe $y_i \in \{-1, 1\}$ pour $i \in \{1, \ldots, M\}$. Voir la figure 1 pour une illustration. On dit que la famille des couples, éléments de $\mathbb{R}^n \times \{-1, 1\}$, $((x_i, y_i))_{1 \leqslant i \leqslant M}$ forment les *données d'entraînement.*

On souhaite trouver un hyperplan (ou une droite dans le cas $n = 2$ comme sur la figure 1) passant par 0 et séparant, du mieux possible, ces points x_i entre ceux qui sont de la classe représentée par -1 et ceux de la classe représentée par 1.

Pour cela on considère la fonction suivante, pour $\lambda > 0$

$$F : a \in \mathbb{R}^n \mapsto \frac{1}{2M} \sum_{i=1}^{M} (1 - y_i \langle x_i,\, a \rangle)^2 + \frac{\lambda}{2} \|a\|^2 .$$

Et on s'intéresse au problème d'optimisation suivant

$$(\mathcal{P}_2) \qquad \inf_{a \in \mathbb{R}^n} F(a).$$

Partie I (7.5pt)

1. Soit $a \in \mathbb{R}^n$ tel que la quantité $\frac{1}{2M} \sum_{i=1}^{M} (1 - y_i \langle x_i,\, a \rangle)^2$ soit "petite". Pourquoi peut-on alors avoir confiance que la fonction $f_a : x \in \mathbb{R}^n \mapsto \langle x,\, a \rangle$ permette de séparer "correctement" les points x_i en fonction de leurs classes y_i ?
2. Justifier brièvement la régularité de F.
3. Montrer que le problème $(\mathcal{P}_2)$ est bien posé.
4. a) Soit $a \in \mathbb{R}^n$. Après avoir déterminer la différentielle de F en a, montrer que

$$\nabla F(a) = \frac{1}{M} \sum_{i=1}^{M} (y_i \langle x_i,\, a \rangle - 1) y_i x_i + \lambda a.$$

Indication : on pourra introduire, à la convenance, d'éventuelles fonctions intermédiaires.

   b) En déduire pour tout $a \in \mathbb{R}^n$, $\nabla^2 F(a)$.

5. a) Montrer que F est strictement convexe.

   b) En déduire que le problème $(\mathcal{P}_2)$ admet une unique solution, que l'on notera dans la suite $a_\lambda^\star \in \mathbb{R}^n$.

Partie II (4pt)

On va maintenant revisiter les résultats précédents en adoptant un point de vue matriciel. On définit le vecteur $\mathbb{1} = (1, \ldots, 1) \in \mathbb{R}^M$ et $Z \in \mathcal{M}_{n,M}(\mathbb{R})$ la matrice dont la i-ème colonne est donnée par le vecteur $y_i x_i = (y_i x_{i,1}, \ldots, y_i x_{i,n}) \in \mathbb{R}^n$.

6. a) Montrer que pour tout $a \in \mathbb{R}^n$, $F(a) = \frac{1}{2M} \left\| \mathbb{1} - Z^T a \right\|^2 + \frac{\lambda}{2} \|a\|^2$.

   b) En déduire que F est une fonctionnelle quadratique.

7. Donner alors, en utilisant les notations introduites dans cette partie, une expression de $\nabla F(a)$ et $\nabla^2 F(a)$, pout tout $a \in \mathbb{R}^n$.
8. En déduire une expression de $a_\lambda^\star$, puis déterminer $\lim\limits_{\lambda \to +\infty} a_\lambda^\star$. Pourquoi est-ce logique d'obtenir cette limite sachant la manière dont est construite la fonction objectif F ?