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

Master 1ère année, MMA, 2022-2023
Optimisation

Examen du 11/01/2024

Durée 2h. Aucun document n'est autorisé. Les exercices sont indépendants.

Exercice 1

Soit $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. On pose $g : x \in \mathbb{R}^n \mapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c$ et $f = e^g$.

1. On suppose dans cette question que A est définie positive.
   a) Déterminer pour tout $x \in \mathbb{R}^n$, $\nabla f(x)$.
   b) Déterminer pour tout $x \in \mathbb{R}^n$, $\nabla^2 f(x)$.
   c) Montrer que f admet un minimum global sur $\mathbb{R}^n$.
   d) Déduire des précédentes questions qu'il existe $m > 0$ tel que f est m-fortement convexe sur $\mathbb{R}^n$.

2. On considère dans cette question que $n = 3$, $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $b = (1, -2, 1)$, $c = \frac{5}{2}$. Soit $C = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_3 = 0\}$ et $U = \mathrm{Vect}((1, 1, 0))$.
   a) Montrer que C et U sont des convexes fermés.
   b) Montrer que f est strictement convexe sur C. En déduire que f est strictement convexe sur U.
   c) Est-ce que f est convexe sur $\mathbb{R}^3$?
   d) En vous aidant des ensembles de niveau de f dans C, proposez un candidat pour le minimum global de f sur U.
   e) Montrer que c'est bien l'unique minimum global de f sur U.

Exercice 2

Soit $M, n \in \mathbb{N}^*$. On considère une collection de points $(x_i)_{1 \leqslant i \leqslant M}$ de $[0, 1]$ deux à deux distincts et des valeurs réelles associées $(y_i)_{1 \leqslant i \leqslant M}$. Une telle collection de couples (x_i, y_i) est représentée en exemple sur la figure 1 suivante.

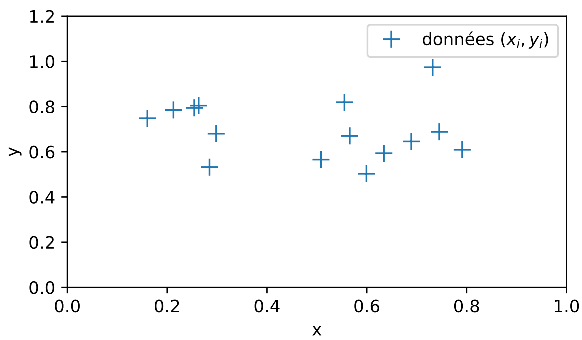

Figure 1. Représentation des points x_i de $[0, 1]$ et des valeurs y_i associées, pour $1 \leqslant i \leqslant M$.

On souhaite trouver une fonction $f : [0, 1] \to \mathbb{R}$ de la forme suivante

$$\forall x \in [0, 1], \quad f(x) = \sum_{k=1}^{n} a_k e^{-\frac{(x-p_k)^2}{2\sigma^2}}, \tag{1}$$

où

— $\sigma > 0$, $0 \leqslant p_1 < p_2 < \ldots < p_n \leqslant 1$ sont des données fixées du problème,

— les poids $a_1, \dots, a_n \in \mathbb{R}$ sont les inconnus du problème paramétrant la fonction f,

qui interpole le mieux possible les données précédentes et de manière simple. Plus précisément, on souhaite que f satisfasse les contraintes $f(x_i) \simeq y_i$ pour tout $i \in \{1, \dots, M\}$ avec le maximum de coefficients a_k nuls. Voir la figure 2 pour des représentations de fonctions du type donné dans (1).

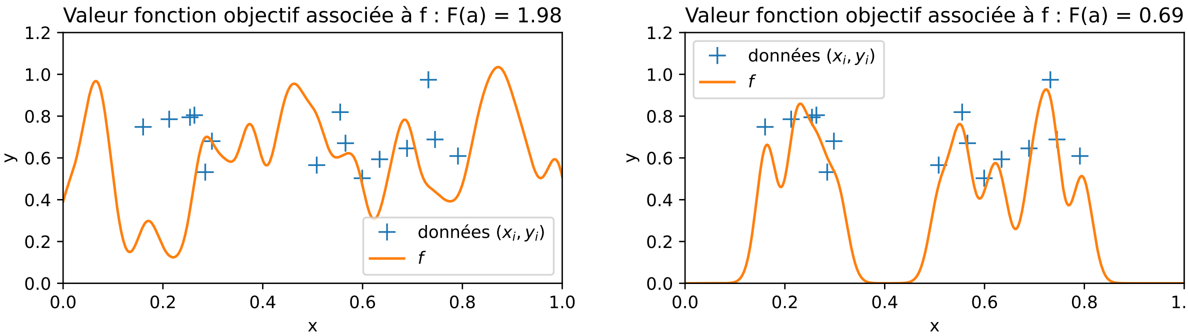

FIGURE 2. Représentation de deux fonctions de la forme donnée dans (1). A gauche les poids $a_1, a_2, \dots, a_n \in \mathbb{R}$ sont aléatoires, à droite les poids $a = (a_k)_{1 \leqslant k \leqslant n}$ sont bien choisis pour assurer une interpolation satisfaisant nos critères.

La fonction f de (1), qui est donc une somme pondérée par les poids $a_k \in \mathbb{R}$ de gaussiennes centrées en les points $p_k \in [0,1]$, peut se réécrire de la manière suivante

$$\forall x \in [0,1], \quad f(x) = \sum_{k=1}^{n} a_k e^{-\frac{(x-p_k)^2}{2\sigma^2}} = \sum_{k=1}^{n} a_k u_k(x) = \langle u(x),\, a\rangle\,,$$

où pour tout $k \in \{1, \dots, n\}$, $u_k : x \in [0,1] \mapsto e^{-\frac{(x-p_k)^2}{2\sigma^2}}$ est la gaussienne centrée en p_k et pour tout $x \in [0,1]$, $u(x) = (u_k(x))_{1 \leqslant k \leqslant n} \in \mathbb{R}^n$.

Pour trouver la meilleure fonction interpolatrice comme décrite précédemment, on va s'intéresser au problème d'optimisation suivant

$$\inf_{a \in \mathbb{R}^n} F(a), \tag{2}$$

où

$$\forall a \in \mathbb{R}^n, \quad F(a) = \frac{1}{2} \sum_{i=1}^{M} (\langle u(x_i),\, a\rangle - y_i)^2 + \lambda \sum_{k=1}^{n} N_\varepsilon(a_k). \tag{3}$$

avec $\lambda > 0$ et $N_\varepsilon : t \in \mathbb{R} \mapsto \sqrt{t^2 + \varepsilon}$, $\varepsilon > 0$.

<u>Résumé.</u>

— Données du problème :
  — les points à interpoler $(x_i, y_i) \in [0,1] \times \mathbb{R}$, $i \in \{1, \dots, M\}$,
  — $\sigma > 0$, $0 \leqslant p_1 < p_2 < \dots < p_n \leqslant 1$ les paramètres des gaussiennes.
  — $\lambda > 0$, $\varepsilon > 0$.

— Objectif : trouver le meilleur vecteur $a \in \mathbb{R}^n$ définissant une fonction interpolante $f : x \in [0,1] \mapsto \langle u(x),\, a\rangle$ en résolvant le problème (2).

Indication : on pourra utiliser les notations N_ε, N'_ε, N''_ε quand cela est pertinent afin d'alléger les calculs.

1. On suppose, <u>dans cette question uniquement</u>, que $n = 5$, $p_1 = 0$, $p_2 = 1/4$, $p_3 = 1/2$, $p_4 = 3/4$ et $p_5 = 1$. Représenter sur un même graphe l'allure des fonctions u_2 et $2u_4$ lorsque $\sigma = 0.1$.

2. On pose $G : a \in \mathbb{R}^n \mapsto \frac{1}{2}\sum_{i=1}^{M} (\langle u(x_i),\, a\rangle - y_i)^2$. Quel est l'intérêt, vis à vis d'une éventuelle solution de (2), d'avoir le terme $G(a)$ dans l'expression de F ?

3. On pose $H : a \in \mathbb{R}^n \mapsto \sum_{k=1}^{n} N_\varepsilon(a_k)$. Montrer que H est une fonction coercive.

4. Déterminer pour tout $a \in \mathbb{R}^n$, $\nabla F(a)$. *Indication : on pourra utiliser le fait que pour tout $k \in \{1, \dots, n\}$, $a_k = \langle e_k,\, a\rangle$, où e_k est le k-ème vecteur de la base canonique de $\mathbb{R}^n$.*

5. Déterminer pour tout $a \in \mathbb{R}^n$, $\nabla^2 F(a)$.

6. Montrer que F est strictement convexe sur $\mathbb{R}^n$.

7. Montrer que F admet un unique minimum global sur $\mathbb{R}^n$.

8. Montrer que les ensembles de sous-niveau de F, lorsqu'ils sont non vides, sont des ensembles convexes et compacts de $\mathbb{R}^n$.

9. Soit $a \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ telle que $\langle \nabla F(a), d \rangle < 0$.

   a) Montrer que d est une direction de descente pour F en a.

   b) Rappeler la définition de la condition d'Armijo en a dans la direction d pour F.

   c) Illustrer cette condition par un dessin.

10. Écrire en Python une fonction renvoyant la valeur du pas d'Armijo en $a \in \mathbb{R}^n$ dans la direction de descente $-\nabla F(a)$ pour F. Cette fonction prendra en argument les variables que vous jugerez nécessaires.

11. Montrer que la convergence de l'algorithme de descente de gradient à pas d'Armijo vers l'unique minimum global de F est assurée.

12. <u>Bonus.</u> Écrire en Python deux fonctions renvoyant respectivement $F(a)$ et $\nabla F(a)$ pour $a \in \mathbb{R}^n$ de sorte que leurs implémentations sont vectorielles, i.e. sans aucunes boucles for. <u>Bonus**2.</u> On suppose que $M > n$. Trouvez des implémentations qui minimisent le nombre d'opérations. Vous pourrez précalculer en amont de vos fonctions les quantités que vous jugerez nécessaires.

    Indication : vous pourrez écrire sur votre feuille les calculs que vous jugerez pertinents.