Optimisation algorithmique - feuille de TD 1

Exercice 1. Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. Soit $f : \mathbb{R}^n \to \mathbb{R}$ l'application $f : x \mapsto \frac{1}{2}\langle Ax , x \rangle$.

1. Montrer que f est différentiable sur $\mathbb{R}^n$ et déterminer $\nabla f(x)$ pour tout x.
2. Quelle est l'expression de ∇f si A est symétrique ?
3. Quel est le gradient de l'application $x \mapsto \frac{1}{2}\|x\|^2$?

Exercice 2. Un réel $\alpha > 0$ étant fixé, on définit

$$\forall x \in \mathbb{R}^n, \quad J_\alpha(x) = \sum_{i=2}^{n-1} N_\alpha(x_{i+1} + x_{i-1} - 2x_i), \quad \text{où} \quad N_\alpha(t) = \sqrt{\alpha + t^2}.$$

a) Prouver que J_α est différentiable et calculer, pour tout $x, h \in \mathbb{R}^n$ l'expression $DJ_\alpha(x).h$ de trois façons différentes:

- en calculant $\frac{\partial J_\alpha}{\partial x_j}$ pour tout j;
- en calculant les dérivées directionnelles de J_α
- en calculant directement la forme différentielle DJ_α.

b) Écrire une fonction `gradientJ(x,alpha)` qui prend en argument un vecteur colonne `x`, un réel strictement positif `alpha`, et renvoie le gradient de J_α en `x` sous forme de vecteur colonne.

Exercice 3. Soit f la fonction définie sur $\mathbb{R}^2$ par :

$$f(x_1, x_2) = x_1^3 + x_2^3 - 6(x_1^2 - x_2^2)$$

1. Montrer que f admet quatre points critiques.
2. Calculer $f(t, 0)$ et $f(0, t)$, et en déduire directement que f n'admet pas d'extremum en $(0, 0)$.
3. Écrire le développement de Taylor de f en $(4, 0)$. En déduire que f admet un minimum local en $(4, 0)$.
4. Calculer les valeurs propres de la matrice hessienne aux deux autres points critiques et en déduire directement leur type (minimum local, maximum local, ou point selle).

Exercice 4. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par :

$$f(x,y) = x^2 + y^3$$

Cette fonction admet-elle un extremum global sur $\mathbb{R}^2$? Admet-elle un extremum local ?

Exercice 5. Etudier les extrema locaux et globaux des fonctions suivantes de $\mathbb{R}^2$ dans $\mathbb{R}$:

1. $f_1(x,y) = xy(x+y-1)$
2. $f_2(x,y) = 4xy + \frac{1}{x} + \frac{1}{y}$
3. $f_3(x,y) = x^3 + y^3 - 9xy + 27$
4. $f_4(x,y) = x^2 - 3x^2y + 2x^4$
5. $f_5(x,y) = (y^2 - x^2)(y^2 - 2x^2)$
6. $f_6(x,y) = (x+y)^2 - (x^4 + y^4)$
7. $f_7(x,y) = x \ln y + y \ln x$

Exercice 6. Soit $p^1 = (p_1^1, p_2^1)$, $p^2 = (p_1^2, p_2^2)$, ..., $p^n = (p_1^n, p_2^n)$ n points dans $\mathbb{R}^2$. Le but de cet exercice est de trouver le point qui minimise la somme des distances au carré à tous les points p^i : ainsi, pour tout $x = (x_1, x_2)$ nous définissons la fonctionnelle suivante :

$$J(x) = \sum_{i=1}^{n} \|x - p^i\|^2.$$

1. Calculer $\frac{\partial J}{\partial x_1}(x)$ et $\frac{\partial J}{\partial x_2}(x)$, et en déduire que J admet un unique point critique x^* sur $\mathbb{R}^2$ à déterminer. Comment s'appelle ce point en termes géométriques ?
2. Montrer que x^* est un minimiseur local de J.
3. Expliquer pourquoi x^* est aussi l'unique minimiseur global de J.

Exercice 7. Régression linéaire simple Soit $(x_i, y_i)_{i=1}^n$ un n-uplet de points dans $\mathbb{R}^2$, avec $n \geq 2$. On suppose qu'au moins deux points x_i sont distincts. La régression linéaire simple consiste à trouver un relation affine $y = \alpha x + \beta$ qui s'adapte au mieux aux observations, ce qui s'obtient en minimisant la fonctionnelle suivante dans $\mathbb{R}^2$: $J(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha x_i - \beta)^2$

1. Montrer que cette fonctionnelle admet un minimiseur global unique et le calculer.

2. Écrire une fonction RegressionLineaire(x,y) qui calcule cette solution à partir de vecteurs x, y donnant les coordonnées des points. La fonction doit renvoyer deux variables alpha et beta correspondant aux coefficients calculés. Tester cette fonction pour

   ```
   x = np.random.rand(1,10)
   y = -5 + 12*x + np.random.randn(1,10)
   ```

   et afficher sur le même graphique les points et la droite de régression.

Exercice 8. Modèle linéaire Soient $(x_i, y_i)_{i=1}^n$ un n-uplet de points de $\mathbb{R}^2$. Nous cherchons à trouver une relation entre les variables x_i et y_i. On considère le modèle

$$f(x) = \sum_{j=0}^{k} \beta_j w_j(x),$$

où les w_j sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et les β_j sont des coefficients. On cherche les coefficients β_j qui s'adaptent le mieux au modèle $y_i = f(x_i)$ pour $i = 1, \cdots, n$ en minimisant

$$J(\beta_0, \cdots, \beta_k) = \sum_{i=1}^{n} \left(\sum_{j=0}^{k} \beta_j w_j(x_i) - y_i \right)^2 .$$

On supposera que $n \geq k + 1$.

1. Montrer que l'on peut écrire
   $$J(\beta) = \|M\beta - y\|^2,$$
   où $\beta = (\beta_0, \cdots, \beta_k)^T$, $y = (y_1, \cdots, y_n)^T$, et M est une matrice à définir.

2. Montrer que le vecteur β des coefficients optimaux satisfait $M^T M\beta = M^T y$.

3. Montrer que si M est de rang maximal alors la solution est unique.

4. On considère le cas de la régression polynomiale : $w_j(x) = x^j$. Écrire une fonction RegressionPolynomiale(x,y,k) qui calcule cette solution à partir de vecteurs x, y donnant les coordonnées des points, et de l'ordre k. La fonction doit renvoyer un vecteur beta dans $\mathbb{R}^{k+1}$ correspondant aux coefficients calculés. Tester cette fonction pour

   ```
   x = np.random.rand(10,1)
   y = -5 + 12*x + np.random.randn(10,1)
   k = 3
   ```

   et afficher sur le même graphique les points et la courbe de régression.