TD 2 : problème d'optimisation et convexité

Exercice 1. Soient $p^1 = (p_1^1, p_2^1)$, $p^2 = (p_1^2, p_2^2)$, ..., $p^n = (p_1^n, p_2^n)$ 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 2. Régression linéaire simple

Soit $(x_i, y_i)_{i=1}^n$ un n-uplet de points dans $\mathbb{R}^2$, avec $n \geqslant 2$. 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 $\mathbb{R}^2$: $J(\alpha, \beta) = \sum_{i=1}^n (\alpha x_i + \beta - y_i)^2$

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(100)
y = -5 + 12*x + np.random.randn(100)
```

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

Exercice 3. 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 $n \geqslant k + 1$.

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(100)
y = -5 + 12*x - 3*x**2 + np.random.randn(100)
k = 3
```
et afficher sur le même graphique les points et la courbe de régression.

Exercice 4. Méthode par dichotomie

Soit $f : [a, b] \to \mathbb{R}$ une fonction continue strictement monotone et telle que $f(a)f(b) < 0$.

1. Justifier qu'il existe un unique $x^* \in]a, b[$ tel que $f(x^*) = 0$.
2. Justifier qu'il est possible de construire $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites respectivement croissante et décroissante telles que $a_0 = a$, $b_0 = b$, et pour tout $n \in \mathbb{N}$, $b_n - a_n = \frac{b-a}{2^n}$ et $x^* \in [a_n, b_n]$. Etudier la convergence de $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$.
3. Ecrire une fonction dichotomie qui prend en argument la fonction f, les bornes de l'interval, une tolérance $\varepsilon > 0$, et qui renvoie une approximation de x^* ainsi que la liste des itérés. *On prendra garde à évaluer la fonction f le moins possible de fois.*
4. On suppose que f est $\mathcal{C}^2$, avec $f'' > 0$, et f admet un extremum local sur $]a, b[$. Justifier que l'on peut utiliser l'algorithme de la question précédente pour approcher cet extremum local.
5. Illustrer la convergence de la méthode, grâce à la fonction plt.semilogy, sur un exemple.
6. Ecrire une fonction diffFiniesDroite, qui prend en argument une fonction f et qui renvoie une fonction approximant la dérivée de f par la formule des différences finies à droite :
$$f'(x) \simeq \frac{f(x+\varepsilon) - f(x)}{\varepsilon}.$$
On prendra $\varepsilon = 10^{-8}$.
7. Trouver de manière approchée les minimums locaux de la fonction $f(x) = x^2 + 5(1 - \cos(2\pi x))$ définie sur l'interval $[0.5, 2.5]$, en ne calculant pas explicitement la dérivée de f.
8. Comparer la convergence obtenue à la question précédente avec celle obtenue lorsque l'on utilise à la place l'expression littérale de la dérivée. Commenter.