Méthodes d'optimisation des fonctions à plusieurs variables

1 Méthode analytique d'optimisation des fonctions à plusieurs variables

Exercice 1 (*). Calculer le gradient de la fonction f définie sur $\mathbb{R}^2$ par $f(x,y) = \left(3x^2+4xy\right)\cdot\sin(x+y^2)$, puis celui de la fonction g définie sur $\mathbb{R}^3$ par $g(x_1,x_2,x_3) = x_1^2\cdot e^{x_2x_3}+\cos(x_1+x_3)$.

Exercice 2 (*). Même exercice avec la fonction f définie sur $\mathbb{R}^2$ par $f(a,b) = \left((a+b)^2+(ab)^2\right)\sin((a+b)ab)$.

Exercice 3 (*). Étudier les extremums de la fonction f définie sur $\mathbb{R}^2$ par $f(x,y) = x^2y+xy^2$.

Exercice 4 (***). Soit $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ une matrice carrée de taille n et soit f la fonction définie sur Ω par $f(x) = \|Ax-y\|^2$, où $\|.\|$ désigne la norme euclidienne sur $\mathbb{R}^n$. Calculer le gradient et la hessienne de f.

2 Méthodes algorithmiques d'optimisation des fonctions à plusieurs variables

L'utilisation de Python est indispensable pour cette partie. Les fichiers `my_descent_td1.py` et `td1_squelette.py` sont disponibles sur Moodle.

Exercice 5 (**). Le but de cet exercice est d'écrire un premier algorithme de descente de gradient et de l'appliquer à une fonction simple d'une variable.

1. Compléter les fonctions `descent` et `update` dont les buts respectifs sont de retourner l'itéré correspondant au nombre d'itération maximal spécifié, et de calculer l'itéré suivant selon la formule étudiée en cours. Donner une valeur par défaut du pas égale à 0.01, et une valeur par défaut du nombre d'itération maximal égal à 1000.
2. Soit la fonction f définie sur $\mathbb{R}$ par $f : x \mapsto x^2+1$.
   (a) Lui appliquer l'algorithme de descente de gradient sur $k = 10$ itérations en partant du point $a_0 = 2$ et avec un pas fixe $\delta = 0.2$. Donner les coordonnées de a_k ainsi que son image par f.
   (b) Reprendre la question 1 en partant de $a_0 = -1.5$.
   (c) Reprendre la question 1 avec $\delta = 0.9$, $\delta = 1.1$ et $\delta = 0.05$. Commenter.

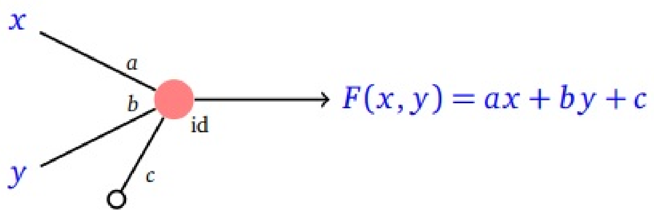
Exercice 6 (**). Soit la fonction f définie sur $\mathbb{R}$ par $f : x \mapsto x^4-2x^3+4$. Lui appliquer l'algorithme de descente de gradient sur $k = 1000$ itérations en partant du point $a_0 = -1$ et avec un pas fixe $\delta = 0.05$. Donner les coordonnées de a_k ainsi que son image par f. Commenter.

Exercice 7 (**).

1. Compléter la classe `GradientDescent` en intégrant le niveau d'erreur comme paramètre d'arrêt de l'algorithme et en lui ajoutant un attribut comptant le nombre d'itérations effectuées.
2. Soit la fonction f définie sur $\mathbb{R}^2$ par $f : (x,y) \mapsto x^2+3y^2$.
   (a) Lui appliquer l'algorithme de descente de gradient en partant du point $a_0 = (2,1)$, avec un pas fixe $\delta = 0.2$ et un taux d'erreur $\epsilon = 0.01$. Donner le nombre d'itérations k effectuées, les coordonnées de a_k ainsi que son image par f.
   (b) Reprendre la question 1 en partant d'un point a_0 dont les coordonnées sont tirées uniformément sur $[-1\ ;\ 1]$. Effectuer plusieurs simulations et commenter les résultats obtenus.
   (c) Reprendre la question 1 en partant du point $a_0 = (2,1)$ et en choisissant un pas qui décroît linéairement avec $\delta_0 = 0.2$. Commenter.

Exercice 8 (**). On considère un ensemble de $N = 5$ points du plan $A_i = (x_i, y_i)_{i\in\{1,\dots,5\}}$, pour lequel on souhaite trouver l'équation $y = ax + b$ de la droite qui approche au mieux tous ces points, en le sens suivant : la somme des carrés des distances entre la valeur observée y_i et celle donnée par l'équation de la droite doit être minimale. Leurs coordonnées sont $A_1 = (4,1)$, $A_2 = (7,3)$, $A_3 = (8,3)$, $A_4 = (10,6)$ et $A_5 = (12,7)$. Implémenter l'algorithme de descente de gradient stochastique, puis l'utiliser pour déterminer l'équation de la droite, en commençant par déterminer l'expression de la fonction que l'on cherche à minimiser. Les paramètres sont laissés au choix.

Exercice 9 (**). On considère le neurone suivant.



D'après [DM].

Étant données N données $(x_i, y_i, z_i)_{i\in\{1,\dots,N\}}$, il s'agit de trouver la valeur des poids a, b et c pour lesquels $F(x_i, y_i)$ approxime au mieux z_i, au sens des moindres carrés. On considère pour cela les 5 données suivantes :

$$(1,0,0), (0,1,5), (2,1,1), (1,2,0), (2,2,3).$$

Utiliser l'algorithme de descente de gradient stochastique pour déterminer la valeur des poids a, b et c. Les paramètres sont laissés au choix.

Exercice 10 (***). On considère un ensemble de N points du plan $A_i = (x_i, y_i)_{i\in\{1,\dots,N\}}$, pour lequel on souhaite trouver l'équation du cercle qui approche au mieux tous ces points.

1. Proposer l'expression d'une fonction erreur pertinente à minimiser.
2. **Application.** On considère les données suivantes :

   $$(-1,1), (0,-1), (0,3), (1,-2), (1,4), (2,-2), (2,4), (3,-2), (3,4), (4,-1), (4,3), (5,1).$$

   Utiliser l'algorithme de descente de gradient par lots pour déterminer l'équation du cercle qui minimise la fonction d'erreur choisie. La valeur initiale, le pas et la condition d'arrêt sont laissés au choix. Représenter l'évolution de l'erreur en fonction du nombre d'époques. Vous pouvez utiliser la méthode d'accélération de convergence basée sur le moment étudiée en cours.