Techniques Algorithmiques pour le Calcul Scientifique (LI364)

TD-TME n°1 et 2 - Introduction à l'optimisation

1. TD

Exercice 1 (Caractérisation de la convexité)**.** Soit C un sous-ensemble ouvert convexe non vide de $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une application différentiable sur C. Montrer que les propositions suivantes sont équivalentes :

1. f est convexe sur C ;
2. pour tout $x, y \in C$, $f(y) \geqslant f(x) + \langle \nabla f(x), y - x\rangle$;
3. l'application ∇f est monotone sur C, c'est-à-dire

$$\forall x, y \in C, \qquad \langle \nabla f(y) - \nabla f(x), y - x\rangle \geqslant 0.$$

Montrer que si, de plus, f est deux fois différentiable sur $\mathbb{R}^n$, alors

$$f \text{ est convexe sur } \mathbb{R}^n \quad \Leftrightarrow \quad \forall x \in \mathbb{R}^n, \quad \nabla^2 f(x) \text{ est semi-définie positive.}$$

Exercice 2 (Fonctions convexes et extrema)**.** Soit f une fonction numérique convexe sur un ouvert convexe U de $\mathbb{R}^n$. Si f est différentiable en $a \in U$ et si $df(a) = 0$, montrer que f admet un minimun global en a sur U.

Exercice 3 (Fonctions coercives et extrema)**.** Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite coercive si

$$\lim_{\|x\| \to +\infty} f(x) = +\infty.$$

Montrer que si f est continue et coercive alors f admet au moins un minimum sur $\mathbb{R}^n$.

Exercice 4 (Fonction unimodale)**.** On dit qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est unimodale sur l'interval $[a, b]$ si elle a admet un unique minimum local sur $[a, b]$. Montrer qu'une fonction unimodale est strictement décroissante jusqu'au minimun et strictement croissante après le minimum.

Exercice 5 (Calcul d'extrema locaux)**.** **1.** Considérons l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^3 + y^3 - 3xy$. Admet-elle des minima locaux, des maxima locaux ?

2. Considérons maintenant l'application $g(x, y) = e^{1-x^2-y^2} + x^2 + y^2$. Admet-elle des minima locaux, des maxima locaux ? des minima ou des maxima globaux ?

Exercice 6 (Fonction quadratique)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ définie par $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ où A est une matrice symétrique de taille $n \times n$ et $b \in \mathbb{R}^n$.

1. Montrer que $\nabla f(x) = Ax - b$.
2. En déduire la matrice hessienne $H_f(x)$.
3. Proposer un algorithme d'optimisation pour résoudre un système linéaire.

Exercice 7 (Moindres carrés)**.** Étants données n points (x_i, y_i) de $\mathbb{R}^2$ avec x_i non tous égaux entre eux, montrer qu'il existe des nombres λ et μ, uniques, qui minimisent la somme

$$\sum_{i=1}^{n} (\lambda x_i + \mu - y_i)^2.$$

Exercice 8 (Inégalité d'Hadamard)**.** On munit l'espace $E = \mathbb{R}^n$ du produit scalaire usuel. On note $f(v_1,\dots,v_n)$ le déterminant de la matrice $n \times n$ de vecteurs colonnes $v_1,\dots,v_n \in E$.

1. Montrer que le maximum de f sur l'ensemble X défini par

$$\|v_1\| = \cdots = \|v_n\| = 1$$

est atteint et est strictement positif.

2. Montrer par le théorème des extrémas liés que si le maximum est atteint en $(v_1,\dots,v_n)$, les v_i forment une base orthonormale de E.

3. En déduire l'inégalité d'Hadamard,

$$|\det(v_1,\dots,v_n)| \leqslant \|v_1\| \cdots \|v_n\|,$$

pour tout vecteur $v_1,\dots,v_n \in \mathbb{R}^n$. Quant a-t-on égilité ?

Exercice 9 (Décomposition de Choleski)**.** Soit A une matrice symétrique définie positive de taille n. Notons $A = LL^T$ sa décomposition de Choleski (L étant triangulaire inférieure).

1. Pour $n = 3$, écrire a_{ij} pour $i, j = 1,2,3$ en fonction des coefficients de L

2. Remarquer que si l'on calcule colonne par colonne, on peut alors calculer dans l'ordre l_{11}, l_{21}, l_{31}, l_{22}, l_{32} et l_{33}. En déduire une fonction MATLAB calculant la décomposition de Choleski pour tout valeur de n.

3. Donner la décomposition de Choleski de la matrice

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

2. TME

Exercice 10 (Méthode de la section dorée)**. 1.** Écrire un programme MATLAB implémentant la méthode de la section dorée. Votre programme doit prendre en paramètre une fonction et un intervalle. La recherche doit continuer jusqu'à ce qu'on obtienne la borne d'erreur voulue mais on ne doit pas dépasser 100 itérations.

2. Écrire un programme MATLAB implémentant la méthode de Newton.

3. Tester vos algorithmes sur les exemples suivants. Vous pourrez comparer vos résultats avec ceux donnés par la fonction `fminbnd` de MATLAB.

a) $f(x) = \sin(x)$ sur $[0, \pi/2]$

b) $f(x) = (\arctan x)^2$ sur $[-1, 1]$

c) $f(x) = |\ln(x)|$ sur $[1/2, 4]$

d) $f(x) = |x|$ sur $[-1, 1]$

Exercice 11 (Fonction de Rosenbrock et méthode de Newton)**.** La *fonction de Rosenbrock* est une fonction non convexe de deux variables utilisée comme test pour des problèmes d'optimisation mathématique. Elle a été introduite par Rosenbrock en 1960. Elle est définie par

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

1. Calculer le gadient $g(x)$ et la Hessienne $H(x)$ de la fonction f (on utilisera la Symbolic Math Toolbox).

2. Vérifier que $x^* = [1,1]^T$ est un minimum local de f.

3. Calculer les 5 premiers itérés de la méthode de Newton pour minimiser f en commençant avec $x_0 = [-1,-2]^T$. Tracer les lignes de niveau de la fonction f en utilisant `ezcontour` dans le domaine `[-1.5;2;-3;3]`. Afficher les itérés sur le même graphe.

4. Calculer la norme de l'erreur $\|x - x^*\|$ à chaque itération et déterminer si le taux de convergence est quadratique.

Exercice 12 (Méthode de gradient à pas optimal et méthode de Wolfe)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables à valeur réelle. La méthode du gradient à pas constant consiste à calculer les itérations

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

avec α constant.

1. Implémenter la méthode du gradient avec $\alpha = 1$. Tester votre programme sur la fonction $f(x) = x_1^2 + 2x_2^2$ en partant de $x_0 = (-1,-1)$. Tester pour plusieurs pas de descente : par exemple $\alpha = 0.1$, $\alpha = 0.1$ et $\alpha = 0.5$ et $\alpha = 1$. Commenter.

2. Faire de même avec la fonction de Rosenbrock en commençant par exemple en $x_0 = (-1, 1.2)$ et $\alpha = 0.001$.

3. Dans la méthode du gradient à pas optimal, on cherche α_k tel que

$$\min_{\alpha_k \geqslant 0} f(x_k - \alpha_k \nabla f(x_k)).$$

Pour trouver ce α_k, on va utiliser la méthode de Wolfe. Soit donc $g(t) = f(x+td)$ où d est une direction de descente. Étant donné un $t \in \mathbb{R}^+$, la recherche linéaire de Wolfe consiste à rétrécir un intervalle de confiance $[t_g, t_d]$ dans lequel on choisit un t que l'on teste.
- Au départ, $t_g = 0$, $t_d = +\infty$ et $t = 1$, $m_1 = 0.1$, $m_2 = 0.9$
- si $g(t) \leqslant g(0) + m_1 t g'(0)$ et $g'(t) \geqslant m_2 g'(0)$ alors arrêt
- si $g(t) > g(0) + m_1 t g'(0)$ alors on pose $t_d = t$, $t_g = t_g$ et $t = (t_d + t_g)/2$ (si $t_d = +\infty$ alors $t = 10 t_g$)
- si $g(t) \leqslant g(0) + m_1 t g'(0)$ et $g'(t) < m_2 g'(0)$ alors $t_g = t$, $t_d = t_d$ et $t = (t_d + t_g)/2$ (si $t_d = +\infty$ alors $t = 10 t_g$)

Implanter la méthode du gradient à pas optimal avec la méthode de Wolfe. Tester votre implantation sur la fonction de Rosenbrock.