

CC2 - Optimisation

Durée : 2h30.

Seuls le polycopié de cours et les notes personnelles de cours sont autorisés.

Exercice 1. QUESTIONS DE BASE

1. Soient $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$. Calculer le gradient et la hessienne de $f : x \mapsto \frac{1}{2} \|Ax - b\|_2^2$ en détaillant le calcul.
2. Quelle est la différence si $x \in \mathbb{C}^n$, $A \in \mathbb{C}^{m \times n}$ et $b \in \mathbb{C}^m$.
3. On rappelle que le simplexe en dimension n est défini par

$$\Delta_n = \{x \in \mathbb{R}^n, x_i \geq 0, \forall i \in \{1, \dots, n\}, \sum_{i=1}^n x_i = 1\}.$$

Dessiner le simplexe en dimension 2 et 3.

4. Déterminer le cône normal au simplexe pour $x \in \text{ri}(\Delta_n)$.
5. On souhaite projeter un point $x_0 \in \mathbb{R}^n$ sur Δ_n . Ecrire le problème d'optimisation associé au problème de projection.
6. Donner les conditions d'optimalité de ce problème.

Exercice 2. DESCENTES DE SOUS-GRADIENT

Le but de cet exercice est de mieux comprendre les propriétés intrinsèques des descentes de sous-gradient. On considère le problème suivant :

$$\min_{x \in \mathbb{R}^n} f(x)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe. On considère aussi la descente de sous-gradient suivante :

$$\begin{aligned} x_1 &\in \mathbb{R}^n \\ x_{k+1} &= x_k - \tau_k \frac{\eta_k}{\|\eta_k\|} \quad \text{avec } \eta_k \in \partial f(x_k) \\ \text{Si } 0 &\in \partial f(x_{k+1}) \text{ arrêter, sinon, itérer.} \end{aligned}$$

1. Soit $f(x) = |x|$ avec $x \in \mathbb{R}$. Rappeler l'expression du sous-différentiel de f .
2. Pour les choix suivants, dire si l'algorithme converge et si c'est le cas, donnez un taux de convergence approximatif.
 - (a) $x_1 = 1, \tau_k = 1, \forall k \in \mathbb{N}$.
 - (b) $x_1 = 1, \tau_k = 2, \forall k \in \mathbb{N}$.
 - (c) $x_1 = 10, \tau_k = \frac{1}{k^2}$.
 - (d) $x_1 = 2, \tau_k = \frac{1}{\sqrt{k}}$.

Ces exemples montrent qu'en général $(\tau_k)_{k \in \mathbb{N}}$ doit tendre vers 0 pour que la méthode converge. Cependant, il ne doit pas converger trop rapidement vers 0. Dans la suite de l'exercice, on s'intéresse aux propriétés géométriques de l'algorithme et on cherche une condition suffisante pour que la suite $(x_k)_{k \in \mathbb{N}}$ soit bornée.

3. On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = |x| + 2|y|$.
 - (a) Déterminer le sous-différentiel de f au point $(1, 0)$.
 - (b) Dessiner, 3 lignes de niveau de f , dont celle de niveau 1.
 - (c) Montrer sur le même schéma que les directions données par le sous-différentiel ne sont pas nécessairement des directions de descente (i.e. des directions suivant lesquelles la fonction décroît).
4. A partir de maintenant $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction convexe arbitraire. Une des étapes cruciales de la preuve de convergence d'un algorithme consiste à montrer que $(x_k)_{k \in \mathbb{N}}$ est bornée. Soit x^* un minimiseur de f et $r_k = \|x_k - x^*\|$. Comme les sous-gradient ne sont pas forcément des directions de descente, on étudie la distance au minimiseur r_k plutôt que $f(x_k) - f(x^*)$.
 - (a) Donner une expression de r_{k+1}^2 qui dépend de r_k , η_k , τ_k et du produit scalaire usuel.
 - (b) Montrer que $\langle \eta_k, x^* - x_k \rangle \leq f(x^*) - f(x_k)$.
 - (c) En combinant les inégalités précédentes, montrer que $r_{k+1}^2 \leq r_k^2 + \tau_k^2$.
 - (d) Soit $\tau_k = \frac{1}{k^\alpha}$. Dédurre des questions précédentes une condition suffisante sur α pour que la suite des itérées soit bornée.

Exercice 3. PROJECTIONS ALTERNÉES

Soient A et B des ensembles fermés, non vides, convexes de \mathbb{R}^n . On définit :

$$\forall x \in \mathbb{R}^n, d_A(x) = \min_{a \in A} \|x - a\|_2 \quad \text{et} \quad d_B(x) = \min_{b \in B} \|x - b\|_2.$$

$$\forall x \in \mathbb{R}^n, d_A^2(x) = \min_{a \in A} \|x - a\|_2^2 \quad \text{et} \quad d_B^2(x) = \min_{b \in B} \|x - b\|_2^2.$$

La projection d'un point $x \in \mathbb{R}^n$ sur A est notée $p_A(x)$.

Dans cet exercice, on considère le problème d'optimisation non contraint suivant. Minimiser :

$$f(x) = \frac{1}{2} \max(d_A^2(x), d_B^2(x)).$$

1. Comprendre les fonctions distances et f graphiquement.

- (a) On se concentre d'abord sur un problème dans \mathbb{R} . Soit $A = \{0\}$ et $B = \{1\}$. Tracer sur un même schéma le graphe de d_A^2 ainsi que ceux de d_B^2 et de f .
- (b) Est-ce que f est différentiable? Convexe?
- (c) Soient $A = [0, 1]$ et $B = [1, 2]$. Tracer les graphes de d_A^2 , de d_B^2 et de f .
- (d) On se place maintenant dans \mathbb{R}^2 . Soit $A = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$. Dessiner A ainsi que les lignes de niveau 1, 4 et 9 de d_A^2 .
- (e) **A partir de maintenant et jusqu'à la fin, on suppose que $A \cap B \neq \emptyset$.** Donner les valeurs de $f(x)$ pour $x \in A \cap B$, $x \in A$, $x \in B$. Dédurre les solutions globales ainsi que la valeur du minimum du problème : $\min_{x \in \mathbb{R}^n} f(x)$.

2. Propriétés analytiques des fonctions distances.

- (a) Montrer que d_A et d_B sont convexes.
- (b) Montrer que d_A^2 and d_B^2 sont convexes.
- (c) Montrer que f est convexe.
- (d) Montrer que d_A est Lipschitz de constante 1.
- (e) Montrer que pour tout $x \in \text{int}(A)$, d_A est différentiable et que $\nabla d_A(x) = 0$.

Dans la suite de l'exercice, on admet que d_A^2 est différentiable pour tout x et que :

$$\forall x \notin \text{int}(A), \nabla d_A^2(x) = 2(x - p_A(x)).$$

3. Un algorithme.

On s'intéresse maintenant à la résolution numérique de $\min_{x \in \mathbb{R}^n} f(x)$.

On considère l'algorithme suivant :

Initialisation : $x_0 \in A \setminus B$, $k := 0$;

Tant que $x_k \notin A \cup B$,

- i. Calculer la solution x_{k+1} du sous-problème :

$$(P_k) \quad \min_{x \in \mathbb{R}^n} f(x_k) + \langle g_k, x - x_k \rangle + \|x - x_k\|_2^2$$

où $g_k \in \partial f(x_k)$ est un sous-gradient de f au point x_k .

- ii. $k \leftarrow k + 1$;

- (a) Ecrire les conditions d'optimalité du premier ordre de (P_k) . En déduire l'expression de x_{k+1} en fonction de x_k et de g_k .
- (b) Donner l'expression du sous-différentiel $\partial f(x)$ de f au point x dans le cas où $x \in A$ et $x \notin B$. Faire de même dans le cas similaire où $x \in B$ et $x \notin A$.
- (c) Donner l'expression du sous-différentiel de f en tout point. Où se trouvent les singularités de f (les points de non différentiabilité de f) ?
- (d) Déduire de la question (3.b) l'expression de $g_0 \in \partial f(x_0)$ et x_1 . Calculer $g_1 \in \partial f(x_1)$, et x_2 . De façon plus générale, quelles sont les expressions de g_k et de x_{k+1} ?
- (e) Montrer la trajectoire de l'algorithme (i.e. les itérées $(x_k)_{k \in \mathbb{N}}$) quand $x_0 = (5, 5)$, A est la droite d'équation $y = x$, et B est la droite d'équation $y = 0$.

Cet algorithme est appelé l'algorithme des projections alternées. C'est un grand standard de l'optimisation.