Université Paris-Dauphine
Département MIDO
Chargé de cours : F.-X. Vialard.
Chargés de TD/TP : A. Frouvelle, G. Jankowiak, F.-X. Vialard.
DUMI2E troisième année

Optimisation numérique

année 2013/2014

Séance de TD/TP n^{o}1

1. Formulation du problème d'optimisation

La formulation du problème d'optimisation (non-linéaire) est :
Soient une fonction $f : \mathbb{R}^n \to \mathbb{R}$ et S un sous-ensemble (non vide) de $\mathbb{R}^n$, $S \subset \mathbb{R}^n$, résoudre

$$\operatorname{argmin}_{x \in S} f(x) \,. \tag{1}$$

Dans ce cours, on propose des méthodes (numériques) pour résoudre ce problème. A priori, il n'y a pas unicité de la solution et on ne peut trouver des minimums globaux qu'à une précision donnée $\varepsilon > 0$. En pratique, on ne peut pas résoudre ce problème pour f ou S quelconques.

De plus, les méthodes développées s'appliquent à des classes de fonctions (et non à une fonction en particulier...). On cherche donc à évaluer la performance de la méthode sur une classe de fonctions.
Considérons par exemple le cas $S = [0, 1]^n$, et $C := \{f : S \to \mathbb{R} \,|\, f$ Lipschitzienne de constante $L\}$, i.e. $f : S \to \mathbb{R}$ appartient à C si $|f(x) - f(y)| \leq L|x - y|$ pour tout $x, y \in S$.

On considère la méthode d'optimisation suivante : et en prendre le minimum et retourner

1. Évaluer f sur $(\frac{i}{p})_{i=0,\ldots,p}$.
2. Sélectionner le minimum x_*.
3. Retourner $(x_*, f(x_*))$.

Afin d'évaluer la performance de cette méthode, on définit la complexité de cette méthode comme le nombre d'évaluations de la fonction nécessaires pour terminer dans le pire des cas. Terminer signifie trouver x_* tel que $f(x_*) \leq \min f + \varepsilon$.

1. Quelle est la complexité de la méthode pour $\varepsilon > 0$ donné ?
2. Généraliser en dimension n quelconque pour $S = [0, 1]^n$.
3. Peut-on espérer localiser un minimimum global à ε près ?
4. Implémenter cette méthode en dimension 1 et l'appliquer sur $x \to (x - \frac{1}{\sqrt{2}})^2$ sur $[0, 1]$.

Pour la question 3, définir une fonction MethodeSimple qui prend ε en paramètre et la fonction f à optimiser. MethodeSimple retourne le minimum trouvé x_* et sa valeur $f(x_*)$.

2. Approximation numérique des dérivées

Rappel : Soit $f : \mathbb{R} \mapsto \mathbb{R}$ une fonction dérivable. Si f admet un extremum local en un point $a \in \mathbb{R}$, alors $f'(a) = 0$. [en donner une preuve]

Cette condition nécessaire d'optimalité peut être vérifiée numériquement en approximant la dérivée au point considéré par une méthode de différence finies :

Exercice 1 TD

On considère une fonction $f : \mathbb{R}^n \mapsto \mathbb{R}$ de classe C^1.

1. Justifier l'approximation

$$\partial_i f(x) = f'(x)(e_i) \simeq \frac{f(x + \varepsilon e_i) - f(x)}{\varepsilon} \cdot \tag{2}$$

   Quel est l'ordre de cette approximation ?

2. On suppose la fonction f de classe C^2. Étudier le cas de la différence finie centrée :

$$\partial_i f(x) = f'(x)(e_i) \simeq \frac{f(x + \varepsilon e_i) - f(x - \varepsilon e_i)}{2\varepsilon} \cdot \tag{3}$$

3. Comparer les ordres d'approximation obtenus.

Exercice 2 TD

On s'intéresse au choix numérique de la valeur du paramètre ε. On rappelle que python (matlab...) représente les réels sous la forme de nombre à virgule flottante : en double précision, 64 bits sont alloués à un nombre donné : 1 bit est alloué au signe, 11 à l'exposant et les 52 bits restant sont alloués à la mantisse (chiffres significatifs). La conséquence directe de cette approximation est que les calculs sont ne sont pas nécessairement exacts. Par exemple, $1.0e30 + 1 - 1.0e30$ donne un résultat nul.

1. Pour choisir ε, il faut prendre en compte cet approximation : théoriquement, il faut faire tendre ε vers 0. Pourquoi choisir ε le plus petit possible n'est pas un bon choix pratique ?
2. On suppose que l'évaluation numérique de f est donnée par $\hat{f}$ vérifiant $\|f(x) - \hat{f}(x)\| \leq \eta \|f\|_\infty$ avec η de l'ordre de 10^{-p}. En utilisant une inégalité triangulaire, comparer la différence finie (2) appliquée à $\hat{f}$ et la valeur de la dérivée de f au point x.
3. En minimisant le majorant précédent en ε, proposer une valeur de ε offrant un bon compromis.
4. Mêmes questions (2 et 3) pour l'approximation par différences finies centrées (3).

Exercice 3 TP

On considère la fonction $f : \mathbb{R}^n \mapsto \mathbb{R}$ définie par

$$f(x) = \sum_{i=1}^{n} x_i^2 \,, \tag{4}$$

et le point $a = [i]_{i=1,\dots,n}$.

1. Calculer la dérivée partielle de f par rapport à x_1.
2. Utiliser la différence finie à droite (2) pour calculer une valeur approchée de la dérivée partielle par rapport à x_1 au point a, pour différentes valeurs de n et différentes valeurs de ε. On choisira $n = 10^i$ pour $3 \le i \le 8$.
3. Étudier le cas de la différence finie centrée :

$$f'(x) \simeq \frac{f(x + \varepsilon e_i) - f(x - \varepsilon e_i)}{2\varepsilon} \,. \tag{5}$$

4. Comparer et discuter les approximations obtenues. Comment expliquer ces résultats ?

3. Optimisation en dimension 1

Definition 1 *Une fonction $f : [0, T] \mapsto \mathbb{R}$ est dite unimodale sur $[0, T]$ s'il existe un minimum strict $t^* \in]0, T[$ tel que f est strictement décroissante sur $[0, t^*]$ et strictement croissante sur $[t^*, T]$.*

La minimisation en dimension 1 intervient dans les algorithmes de descente pour la minimisation de fonctionnelles sur $\mathbb{R}^n$. Lorsque le minimum est recherché exactement, on appelle ces méthodes *méthodes à pas optimal*. En pratique, Il est intéressant de proposer des méthodes faisant seulement appel à l'évaluation de la fonction à minimiser. La méthode la plus simple pour localiser le minimum à $M\varepsilon$ près d'une fonction lipschitzienne de constante M, $f : [a, b] \to \mathbb{R}$ est d'évaluer f sur une grille de pas ε. Vous pourrez l'utiliser en TP pour confirmer les résultats des autres méthodes. On rappelle que les méthodes sont souvent basée sur la méthode générique de dichotomie suivante :

```
Initialisation de l'intervalle[a = 0, b] pour b ∈ ℝ*₊.
Tant que (f(a) ≤ f(b)) faire
  | b ← b/2
Fin
Initialisation de c ← 2b
Tant que (f(b) > f(c)) faire
  | a ← b,
  | b ← c,
  | c ← 2c
Fin
Tant que ((b − a) > tolérance) faire
  | Application d'une méthode de réduction du triplet a, b, c.
Fin
Retourner c
```

Algorithme 1: Méthode de dichotomie

Exercice 4 TD/TP

1. Montrer que la recherche du triplet de point dans l'algorithme précédent termine pour une fonction unimodale.
2. Proposer une méthode de réduction du triplet basée sur le polynôme d'interpolation de Lagrange pour ce triplet de point.
3. Montrer la convergence de la méthode pour une fonction unimodale C^1 dont la dérivée s'annule uniquement au minimum.
4. (optionnel) Montrer que l'ordre de convergence est super-linéaire si cet algorithme converge si on suppose la fonction suffisamment régulière.
5. Implémenter l'algorithme en faisant bien attention aux cas de boucles infinies.

Exercice 5 TD/TP

Un véhicule doit se rendre d'un point $a \in \mathbb{R}^2$ à un point $b \in \mathbb{R}^2$ en un minimum de temps. Sa vitesse est v_1 dans un domaine D convexe contenant a et $v_2 > v_1$ sur D^c le complémentaire de D. On suppose que $D = \{(x, y) \,|\, y \leq f(x)\}$ pour f une fonction concave que vous choisirez.

1. Formuler ce probléme d'optimisation et le résoudre pour retrouver la loi de Snell dans le cas d'une fonction f constante.

2. Implémenter une méthode d'optimisation simple pour trouver numériquement l'optimum et tracer la trajectoire optimale pour différentes valeurs de v_1 et v_2.

Exercice 6 TD/TP

On cherche à appliquer les algorithmes codés précedemment au problème de la séparation de sources dans le cas particulier de deux enregistrements. Télécharger les données **micro1_1** et **micro1_2** sur `http://www.ceremade.dauphine.fr/~vialard/` et le fichier python OutilsSons.py. Pour chaque signal D_i, soustraire la moyenne et normaliser par la déviation standard. On notera $E[X]$ la moyenne du vecteur X.

1. On définit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ par $(x, y) \to E[(xD_1 + yD_2)^4]$. Comment se ramener à un problème d'optimisation unidimensionnel pour minimiser f sous la contrainte (*) $x^2 + 2c\,xy + y^2 = 1$ avec $c = E[D_1 D_2]$?
2. Tracer la fonction unidimensionnelle obtenue.
3. Utiliser une méthode d'optimisation implémentée précédemment et écouter le signal optimal obtenu pour l'optimum (x_0, y_0) ainsi que le signal associé à l'orthogonal de (x_0, y_0) pour la forme quadratique définissant la contrainte (*).

Exercice 7 TP

1. Implémenter la méthode de la section dorée.
2. Comparer sur différents exemples la robustesse des deux méthodes unidimensionnelles implémentées (interpolation quadratique et section dorée).
3. Proposer un algorithme plus robuste profitant de la convergence superlinéaire de l'algorithme d'interpolation quadratique.