Maths 5 2015–2016

Rappels de calcul différentiel

Notions de base

Notations

- L'espace ambient est $\mathbb{R}^n$
- Un point de $\mathbb{R}^n$ est désigné par x ou y, etc.
- $x = (x_1, \ldots, x_n)^T$
- $\cdot$ est "le" produit scalaire dans $\mathbb{R}^n$:

$$x \cdot y := x^T y = \sum_{j=1}^{n} x_j y_j$$

- $|\ |$ est "la" norme euclidienne dans $\mathbb{R}^n$: $|x| = \sqrt{x \cdot x}$
- Si $r > 0$, alors

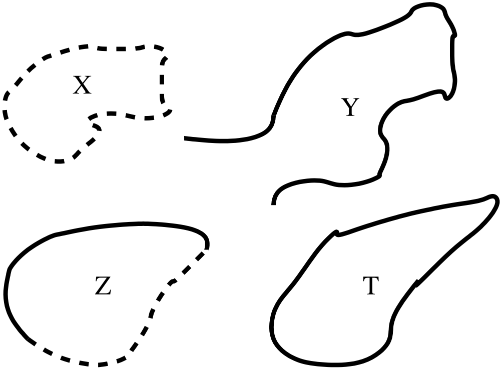
$$B(x, r) = \{y\ ;\ |y - x| < r\},\ \ \overline{B}(x, r) = \{y\ ;\ |y - x| \leq r\}$$

Ensembles

Définitions

- $\Omega \subset \mathbb{R}^n$ est un ouvert si, pour tout $x \in \Omega$, il existe un $r > 0$ (qui dépend de x) tel que $B(x, r) \subset \Omega$
- $F \subset \mathbb{R}^n$ est fermé si $\mathbb{R}^n \setminus F$ est ouvert
- $B \subset \mathbb{R}^n$ est borné s'il existe $M > 0$ tel que $B \subset \overline{B}(0, M)$
  Ou encore : il existe $M > 0$ tel que $|x| \leq M$, $\forall\, x \in B$
- $K \subset \mathbb{R}^n$ est compact si K est à la fois fermé et borné

Exemples



X est ouvert. Y est fermé, mais pas compact.

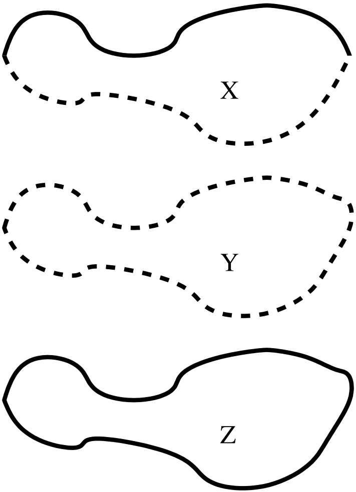
Z n'est ni ouvert, ni fermé. T est compact.

Ensembles

Définitions

- L'intérieur d'une partie X de $\mathbb{R}^n$ est le plus grand ouvert de $\mathbb{R}^n$ contenu dans X. L'intérieur est noté $\overset{\circ}{X}$
  Ou encore : $\overset{\circ}{X}$ est l'union de tous les ouverts contenus dans X
- L'adhérence d'une partie X de $\mathbb{R}^n$ est le plus petit fermé de $\mathbb{R}^n$ contenant X. L'adhérence est notée $\overline{X}$
  Ou encore : $\overline{X}$ est l'intersection de tous les fermés de $\mathbb{R}^n$ contenant X

Exemple



Y est l'intérieur de X. Z est l'adhérence de X

Suites

Notations

- Une suite de points de $\mathbb{R}^n$ est notée $x^0, x^1, \ldots$, ou $(x^k)_{k \geq 0}$
- Abus de notation : on écrit $(x^k) \subset \mathbb{R}^n$, et non pas $(x^k) \subset (\mathbb{R}^n)^{\mathbb{N}}$
- Par abus de notation, une suite extraite de (x^k) et notée (x^{k_l})
  On appelle encore (x^{k_l}) sous-suite de (x^k)
  Notation "officielle" : une suite extraite s'écrit $(x^{\varphi(k)})$, avec $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante

Exemples

Définitions

Une suite $(x^k) \subset \mathbb{R}^n$:

- converge vers un point $x \in \mathbb{R}^n$ si et seulement si $x_j^k \to x_j$ (quand $k \to \infty$) pour chaque $j = 1, \ldots, n$
- est bornée si et seulement s'il existe $M > 0$ tel que $|x^k| \leq M$ pour tout k

Exemples

- La suite de terme général $x^k = (1 + 1/(k+1), e^{-k})^T$ converge vers $(1, 0)^T \in \mathbb{R}^2$
- La suite de terme général $x^k = (1, (-1)^k)^T$ ne converge pas
- La suite de terme général $x^k = (1, (-1)^k)^T$ est bornée
- La suite de terme général $x^k = (0, k)^T$ n'est pas bornée

Théorème de Bolzano-Weierstrass

Théorème (Première forme)

Hypothèses

(i) $(x^k) \subset \mathbb{R}^n$

(ii) (x^k) bornée

Conclusion

(x^k) contient une sous-suite convergente : il existe $x \in \mathbb{R}^n$ et une sous-suite (x^{k_l}) tels que $x^{k_l} \to x$ (quand $l \to \infty$)

Théorème (Deuxième forme)

Hypothèses

(i) $K \subset \mathbb{R}^n$ compact

(ii) $(x^k) \subset K$

Conclusion

(x^k) contient une sous-suite convergente vers un point de K : il existe $x \in K$ et une sous-suite (x^{k_l}) tels que $x^{k_l} \to x$

Fonctions continues

La continuité expliquée “avec les mains”

Une fonction f est continue si, pour x suffisamment proche de y, $f(x)$ est arbitrairement proche de $f(y)$

Définition

Soient $X \subset \mathbb{R}^n$ et $f : X \to \mathbb{R}$. Alors f est continue si et seulement si :
pour tout $x \in X$ et pour toute suite $(x^k) \subset X$ telle que $x^k \to x$, on a $f(x^k) \to f(x)$
Plus généralement, $f : X \to \mathbb{R}^m$ est continue si et seulement si chaque coordonnée de f est continue

Opérations avec les fonctions continues

Exemple fondamental

Un monôme (de n variables) est une expression de la forme $c_{j_1,\dots,j_n} x_1^{j_1} \cdot \ldots \cdot x_n^{j_n}$, avec $c_{j_1,\dots,j_n}$ constante et $j_1, \dots, j_n$ entiers
Un polynôme (de n variables) est une somme finie de monômes
Un polynôme est une fonction continue
Mieux : un polynôme est une fonction C^∞

Proposition

Les opérations usuelles préservent la continuité :
la somme, le produit, la composition de fonctions continues est une fonction continue

Exemple

La fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(x^2 - y) - \sqrt{x^2 + y^4}$, est continue

Démonstration.

- $(x, y) \mapsto x^2 - y$ est continue (polynôme)
- $(x, y) \mapsto \sin(x^2 - y)$ est continue (composée de fonctions continues)
- $(x, y) \mapsto x^2 + y^4$ est continue (polynôme)
- $(x, y) \mapsto \sqrt{x^2 + y^4}$ est continue (composée de fonctions continues)
- f est continue (différence de fonctions continues) $\square$

Fonctions continues et ensembles

Proposition

Hypothèse

$f : \mathbb{R}^n \to \mathbb{R}$ est continue

Conclusions

- Les ensembles de la forme

  $$\{x \in \mathbb{R}^n \,;\; f(x) > a\} \quad (\text{ou } f(x) < a)$$

  (avec $a \in \mathbb{R}$ une constante)   sont ouverts
- Les ensembles de la forme

  $$\{x \in \mathbb{R}^n \,;\; f(x) \geq a\} \quad (\text{ou } f(x) \leq a \text{ ou } f(x) = a)$$

  (avec $a \in \mathbb{R}$ une constante)   sont fermés

Fonctions continues et ensembles

Proposition (Généralisation de la proposition précédente)

Hypothèses

- $f_1, \ldots, f_m, g_1, \ldots, g_k, h_1, \ldots, h_l : \mathbb{R}^n \to \mathbb{R}$ sont continues
- $a_1, \ldots, a_m, b_1, \ldots, b_k, c_1, \ldots, c_l$ sont des constantes

Conclusions

- L'ensemble

$$\{x \in \mathbb{R}^n \; ; \; f_j(x) < a_j, g_j(x) > b_j, \forall\, j\}$$

  est ouvert
- L'ensemble

$$\{x \in \mathbb{R}^n \; ; \; f_j(x) \leq a_j, g_j(x) \geq b_j, h_j(x) = c_j, \forall\, j\}$$

  est fermé

Fonctions continues et ensembles

Les ensembles expliqués “avec les mains”

- Les inégalités strictes donnent des ouverts
- Les inégalités larges et les égalités donnent des fermés

Fonctions de classe C^1

! C^1 demande des ensembles particuliers

On se donne $f : X \to \mathbb{R}$

- On peut parler de la continuité de f quel que soit X
- Mais on va parler du caractère C^1 de f seulement dans deux cas :
  - X est ouvert
  - X est l'adhérence d'un ouvert

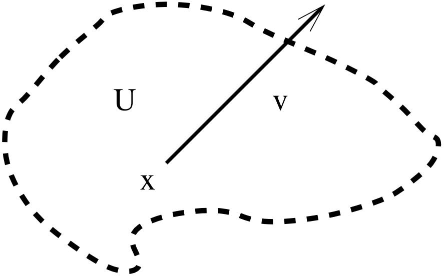
Dérivées directionnelles

- Cadre : $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert
- On fixe :
  - $x \in U$
  - $v \in \mathbb{R}^n$
- Pour $h \in \mathbb{R}$ proche de 0, on a $x + hv \in U$
- On pose alors :

$$\frac{\partial f}{\partial v}(x) := \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h}$$

- C'est la dérivée de f au point x dans la direction du vecteur v

Dérivées directionnelles



x + un petit multiple de v est dans U

Dérivées directionnelles

Exemple

Si $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$, alors

$$\frac{\partial f}{\partial (1, 2)}(1, 1) = 3$$

Lien avec la dérivée usuelle

Dans $\mathbb{R}$ ($n = 1$),

$$f'(x) = \frac{\partial f}{\partial 1}(x)$$

Dérivées partielles

Définition

Les dérivées partielles (du premier ordre) de $f : U \to \mathbb{R}$ sont

$$\frac{\partial f}{\partial x_j}(x) := \frac{\partial f}{\partial e_j}(x), \quad x \in U, j = 1, \ldots, n$$

Autres notations

$$\frac{\partial f}{\partial x_j}(x) = \partial_j f(x) = f_{x_j}(x)$$

Fonctions de classe C^1

Définition

Soit $f : U \to \mathbb{R}$, avec U ouvert
f est de classe C^1 si

- f admet des dérivées partielles (du premier ordre)
- les dérivées partielles $\partial_j f$ sont continues

Opérations avec les fonctions de classe C^1

Exemple fondamental

Un polynôme est de classe C^1

Proposition

Les opérations usuelles préservent le caractère C^1 :
la somme, le produit, la composition de fonctions de classe C^1 est une fonction de classe C^1

Exemple

La fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(x^2 - y) - \sqrt{x^2 + y^4}$, est de classe C^1 sur $\mathbb{R}^2 \setminus \{0\}$

Démonstration.

- $(x, y) \mapsto x^2 - y$ est C^1 (polynôme)
- $(x, y) \mapsto \sin(x^2 - y)$ est C^1 (composée de fonctions C^1)
- $(x, y) \mapsto x^2 + y^4$ est C^1 (polynôme)
- $t \mapsto \sqrt{t}$ est C^1 là où $t > 0$
- $(x, y) \mapsto \sqrt{x^2 + y^4}$ est C^1 là où $x^2 + y^4 > 0$ (composée de fonctions C^1), c'est-à-dire sur $\mathbb{R}^2 \setminus \{0\}$
- f est C^1 (différence de fonctions C^1) sur $\mathbb{R}^2 \setminus \{0\}$ $\square$

Fonctions de classe C^1

Les fonctions de classe C^1 expliquées “avec les mains”

Une fonction $f : U \to \mathbb{R}$ est de classe C^1 si
la graphe de f admet un (hyper)plan tangent en tout point
le plan tangent dépend continûment de $x \in U$

Fonctions de classe C^k

Définitions

On considère $f : U \to \mathbb{R}$, avec U ouvert

- f est de classe C^2 si les dérivées partielles (du premier ordre) de f sont de classe C^1. Ou encore : si les dérivées partielles des dérivées premières (existe et) sont continues. Ou encore : s'il existe $\partial_j f$, $j = 1, \ldots, n$ et s'il existe et sont continues $\partial_k \partial_j f$, $j, k = 1, \ldots, n$
- Les dérivées partielles $\partial_k \partial_j f$ sont les dérivées du second ordre
  On définit de même les dérivées partielles d'ordre m, pour $m = 2, 3, \ldots$
  Par exemple : $\partial_3 \partial_2 \partial_2 \partial_1 f$ est une dérivée du quatrième ordre
- Pour $k = 2, 3, \ldots$, f est une fonction de classe C^k si : ses dérivées partielles jusqu'à l'ordre k existe et les dérivées d'ordre (pile) k sont continues

Fonctions de classe C^k

Théorème

Hypothèse

f est de classe C^k

Conclusion

f et les dérivées partielles de f d'ordre $1, 2, \ldots, k-1$ sont continues

Théorème (de Schwarz)

Hypothèse

f est de classe C^k (pour un $k \geq 2$)

Conclusion

Une dérivée partielle à l'ordre k ne dépend pas de l'ordre dans lequel on dérive par rapport aux k variables

Exemple

- On considère $f \in C^4$
- Dans la dérivée partielle $\partial_2\partial_3\partial_2\partial_1 f$ (qui est du quatrième ordre) , on dérive par rapport aux variables x_1, x_2 (deux fois) et x_3
- Le théorème de Schwarz donne

$$\partial_2\partial_3\partial_2\partial_1 f = \partial_2\partial_2\partial_1\partial_3 f = \partial_3\partial_2\partial_2\partial_1 f = \text{etc}$$

Dérivées et classes de fonctions

Notations

- $\nabla f(x) = \vec{\nabla} f(x) := (\partial_1 f(x), \ldots, \partial_n f(x))^T$
- ∇f est le gradient de f
- $H_x(f) = \nabla^2 f(x) := \left(\partial_k \partial_j f(x)\right)_{j=1,\ldots,n}^{k=1,\ldots,n}$
- $H(f)$ est la matrice hessienne de f ou l'hessien(ne) de f
- $f \in C^k$ ou $f \in C^k(U)$ ou $f \in C^k(U, \mathbb{R})$ sont des notations pour : f de classe C^k sur l'ouvert U
- $f \in C^\infty$ (ou encore f est indéfiniment différentiable) veut dire $f \in C^k$, $k = 1, 2, \ldots$

Propriétés des fonctions de classe C^k

Exemple fondamental

Un polynôme est de classe C^∞ (donc C^k pour tout k)

Proposition

Les opérations usuelles préservent le caractère C^k :
la somme, le produit, la composition de fonctions de classe C^k est une fonction de classe C^k

Propriétés des fonctions de classe C^k

Proposition (Conséquence du théorème de Schwarz)

Hypothèse
$f \in C^2$
Conclusion
$H_x(f)$ est une matrice symétrique

Proposition

Hypothèse
$f \in C^1(U, \mathbb{R})$
Conclusion
f admet des dérivées directionnelles $\dfrac{\partial f}{\partial v}(x)$ en tout point $x \in U$ et en toute direction $v \in \mathbb{R}^n$ et

$$\frac{\partial f}{\partial v}(x) = \nabla f(x) \cdot v, \;\; \forall\, x \in U, \forall\, v \in \mathbb{R}^n$$

Exemples

Existe-t-il une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ telle que

- $\nabla f(x, y) = (x\, e^y, y\, e^{-x})^T$?
- $\nabla f(x, y) = (e^x \sin y - 2xy, e^x \cos y - x^2)^T$?

Notation de Landau

Notation

$o(h)$ dénote une quantité (nombre, vecteur, matrice, selon le contexte) dépendant de h (nombre ou vecteur) de sorte que

$$\lim_{|h|\to 0} \frac{|o(h)|}{|h|} = 0,$$

où $|\ |$ désigne le module, la longueur, la norme, selon le contexte

Exemples

- $\begin{pmatrix} h^2 & 0 \\ 0 & h^3 \end{pmatrix} = o(h)$
- $h \neq o(h)$

Segments

Notations

Si $x, y \in \mathbb{R}^n$, alors

- $[x, y]$ est le segment fermé de x à y :

  $$[x, y] = \{(1-t)x + ty \ ; \ t \in [0, 1]\} = \{x + t(y-x) \ ; \ t \in [0, 1]\}$$

- Si $x \neq y$, alors (x, y) est le segment ouvert de x à y :

  $$(x, y) = \{(1-t)x + ty \ ; \ t \in]0, 1[\} = \{x + t(y-x) \ ; \ t \in]0, 1[\}$$

Formules de Taylor dans $\mathbb{R}$

Théorème (Formules de Taylor à l'ordre 1)

Hypothèse

$f \in C^1(I)$, avec $I \subset \mathbb{R}$ intervalle ; $x, y \in I$

Conclusions

- (formule de Taylor avec reste)
  $f(y) = f(x) + f'(x)(y - x) + o(y - x)$
- (formule de Taylor avec point intermédiaire ou théorème de Lagrange) il existe $z \in [x, y]$ (ou $z \in (x, y)$ si $x \neq y$) tel que $f(y) = f(x) + f'(z)(y - x)$
- (théorème des accroissements finis)

$$|f(y) - f(x)| \leq |y - x| \max_{z \in [x,y]} |f'(z)|$$

Formules de Taylor dans $\mathbb{R}$

Théorème (Formules de Taylor à l'ordre 2)

Hypothèse

$f \in C^2(I)$, avec $I \subset \mathbb{R}$ intervalle ; $x, y \in I$

Conclusions

- (formule de Taylor avec reste)
  $f(y) = f(x) + f'(x)(y-x) + f''(x)(y-x)^2/2 + o((y-x)^2)$
- (formule de Taylor avec point intermédiaire) il existe $z \in [x, y]$ (ou $z \in (x, y)$ si $x \neq y$) tel que
  $f(y) = f(x) + f'(x)(y-x) + f''(z)(y-x)^2/2$
- (théorème des accroissements finis)
  $|f(y) - f(x) - f'(x)(y-x)| \leq (y-x)^2/2 \max_{z\in[x,y]} |f''(z)|$

Formules de Taylor dans $\mathbb{R}^n$

! Taylor demande des points particuliers

Les formules de Taylor se généralisent à $\mathbb{R}^n$
Mais : la formule de Taylor avec point intermédiaire et le théorème des accroissements finis demandent plus que $x, y \in U$: on doit avoir $[x, y] \subset U$
Cette condition est automatiquement satisfaite si U est convexe, mais pas en général

Formules de Taylor dans $\mathbb{R}^n$

Théorème (Formules de Taylor à l'ordre 1)

Hypothèse

$f \in C^1(U)$, avec $U \subset \mathbb{R}^n$ ouvert ; $[x, y] \subset U$

Conclusions

- (formule de Taylor avec reste)
  $f(y) = f(x) + \nabla f(x) \cdot (y - x) + o(y - x)$
- (formule de Taylor avec point intermédiaire ou théorème de Lagrange) il existe $z \in [x, y]$ (ou $z \in (x, y)$ si $x \neq y$) tel que $f(y) = f(x) + \nabla f(z) \cdot (y - x)$
- (théorème des accroissements finis)

$$|f(y) - f(x)| \leq |y - x| \max_{z \in [x,y]} |\nabla f(z)|$$

Formules de Taylor dans $\mathbb{R}^n$

Théorème (Formules de Taylor à l'ordre 2)

Hypothèse

$f \in C^1(U)$, avec $U \subset \mathbb{R}^n$ ouvert ; $[x, y] \subset U$

Conclusions

- (formule de Taylor avec reste)
  $f(y) =$
  $f(x) + \nabla f(x) \cdot (y - x) + (H_x f(y - x)) \cdot (y - x)/2 + o(|y - x|^2)$
- (formule de Taylor avec point intermédiaire) il existe $z \in [x, y]$ (ou $z \in (x, y)$ si $x \neq y$) tel que
  $f(y) = f(x) + \nabla f(x) \cdot (y - x) + (H_z f(y - x)) \cdot (y - x)/2$
- (théorème des accroissements finis)
  $|f(y) - f(x) - \nabla f(x) \cdot (y - x)| \leq |y - x|^2/2 \max_{z \in [x,y]} \|H_z f\|$

Formules de Taylor dans $\mathbb{R}^n$

Explication

Dans la formule précédente, $\|H_z f\|$ représente la norme matricielle de la matrice $H_z f$ en tant qu'application linéaire de $\mathbb{R}^n$ vers $\mathbb{R}^n$ (muni de la norme euclidienne standard $|\ |$)
La matrice $H_z f$ étant symétrique (par le théorème de Schwarz), on peut montrer que

$$\|H_z f\| = \max\{|\lambda|\ ;\ \lambda \text{ valeur propre de } H_z(f)\}$$

Exemple

Si $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 2x_2^2$, alors $H_z f = \begin{pmatrix} 2 & 2 \\ 2 & -4 \end{pmatrix}$, les valeurs propres de $H_z f$ sont $-1 \pm \sqrt{13}$ et $\|H_z f\| = 1 + \sqrt{13}$

Formules de Taylor dans $\mathbb{R}^n$

Idée des preuves

- On se donne $x, y \in U$ tels que $[x, y] \in U$ et $f \in C^k$ $(k = 1$ ou $2)$
- On ramène chaque formule dans $\mathbb{R}^n$ à une formule sur $I = [0, 1]$
- Pour ce faire, on introduit la fonction auxiliaire
$$g(t) = f((1-t)x + ty), \quad t \in [0, 1]$$
de sorte que :
    - $g \in C^k$
    - $g(0) = f(x)$, $g(1) = f(y)$
- On applique la formule de Taylor correspondante à g, avec
    - $I \rightsquigarrow [0, 1]$
    - $x \rightsquigarrow 0$
    - $y \rightsquigarrow 1$
- En calculant les dérivées de g en fonction de celles de f, on retrouve la formule de Taylor correspondante

Formules utiles

Proposition

On se donne : $f \in C^2(U, \mathbb{R})$, $x, y \in U$ tels que $[x, y] \subset U$
On pose $g(t) = g_{x,y}(t) = f((1-t)x + ty)$, $t \in [0, 1]$ et
$z = z(x, y, t) = (1-t)x + ty$
Alors

- $g'(t) = (\nabla f)(z) \cdot (y - x)$
- En particulier, $g'(0) = \nabla f(x) \cdot (y - x)$
- $g''(t) = (H_z(f)(y - x)) \cdot (y - x)$
- En particulier, $g''(0) = (H_x(f)(y - x)) \cdot (y - x)$

Convexité

Définitions

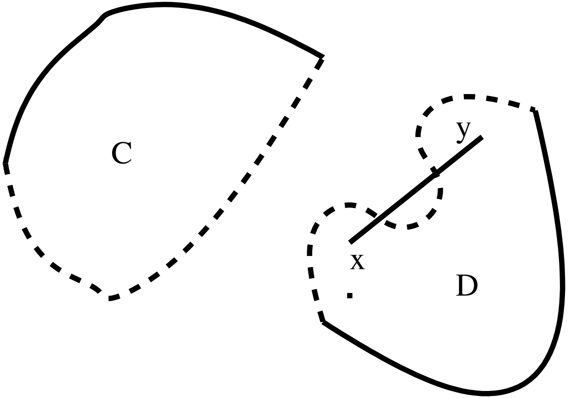
- $C \subset \mathbb{R}^n$ est convexe si et seulement si :

$$\text{pour tout } x, y \in C, \text{ on a } [x, y] \subset C$$

- $f : C \to \mathbb{R}$, avec $C \subset \mathbb{R}^n$, et convexe si et seulement si C est convexe et :

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y), \quad x, y \in C, t \in [0, 1]$$

Exemples



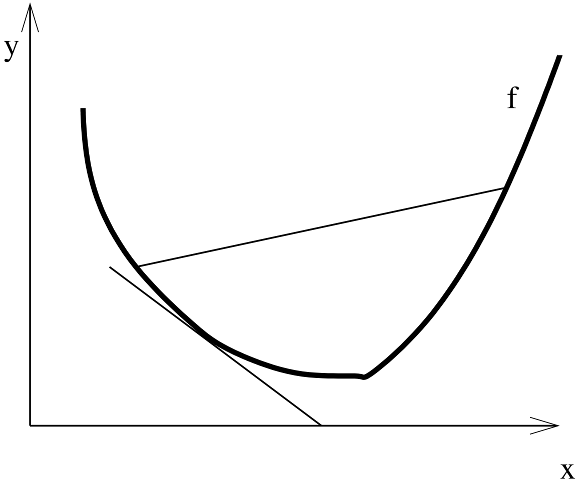
C est convexe. D n’est pas convexe

Convexité

La convexité expliquée “avec les mains”

Une fonction est convexe si son graphe retient l’eau

Fonctions convexes dans $\mathbb{R}$



Si f est convexe, alors le graphe de f se trouve

au–dessus des tangentes et en dessous des sécantes

Fonctions convexes dans $\mathbb{R}$

Théorème (Convexité et dérivées)

Soit $f : I \to \mathbb{R}$, avec $I \subset \mathbb{R}$ intervalle et $f \in C^2$
Les propriétés suivantes sont équivalentes :

- f convexe
- f' croissante
- $f'' \geq 0$
- $f(y) \geq f(x) + f'(x)(y - x), \forall\, x, y \in I$

La dernière propriété exprime le fait que le graphe de f se trouve au-dessus de la tangente en x

Fonctions convexes dans $\mathbb{R}^n$

Théorème (Convexité et dérivées partielles)

Soit $f : C \to \mathbb{R}$, avec $C \subset \mathbb{R}^n$ ouvert convexe et $f \in C^2$
Les propriétés suivantes sont équivalentes :

- f convexe
- $(\nabla f(y) - \nabla f(x)) \cdot (y - x) \geq 0, \forall\, x, y \in C$
- $H_x(f) \geq 0, \forall\, x \in C$
  (C'est-à-dire, les valeurs propres de la matrice symétrique $H_x(f)$ sont ≥ 0, pour tout $x \in C$)
- $f(y) \geq f(x) + \nabla f(x) \cdot (y - x), \forall\, x, y \in C$

La dernière propriété exprime le fait que le graphe de f se trouve au-dessus de l'hyperplan tangent en x

Exemples

- Une forme quadratique :

$$f(x) = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k, \quad a_{jk} = a_{kj} \in \mathbb{R}, x \in \mathbb{R}^n$$

  est convexe si et seulement si la matrice symétrique $A = (a_{jk})_{j=1,\dots,n}^{k=1,\dots,n}$ est ≥ 0
- $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^3$ n'est pas convexe

Problèmes d'optimisation

Définition

- Un problème de minimisation est un problème de la forme

$$(P) = (P_{f,V}) = \min_{v \in V} f(v)$$

  avec
  - $V \subset \mathbb{R}^n$
  - $f : V \to \mathbb{R}$
- Les points $v \in V$ sont des solutions admissibles (pour (P))
- f est la fonction objectif
- Une solution de (P) (si elle existe) est un point $v^0 \in V$ tel que $f(v^0) \leq f(v), \forall\, v \in V$
- Un tel v^0 est une solution optimale de (P)
- Le minimum de f sur V (c'est-à-dire la valeur de f en v^0) est le coût du problème (P)

Exemples

- Le problème (P) $\min_{x \in \mathbb{R}} e^x$ n'a pas de solution optimale
- Pour le problème $\min_{x \in \mathbb{R}} (e^x - x)$, la solution optimale est $x^0 = 0$, alors que le coût est 1

! Minimum $\not\Longrightarrow$ dérivée nulle

Pour le problème (P) $\min_{x \in [0,1]} f(x)$, avec $f(x) \equiv x$, la solution optimale est $x^0 = 0$, mais $f'(x^0) \neq 0$

Condition nécessaire d'optimalité

Définition

Si $f \in C^1(U)$, avec $U \subset \mathbb{R}^n$ ouvert, alors $v^0 \in \mathbb{R}^n$ est un point critique de f si et seulement si $\nabla f(v^0) = 0$

Théorème (de Fermat)

Hypothèses

- v^0 solution optimale de (P)
- $v^0 \in \overset{\circ}{V}$
- $f \in C^1(\overset{\circ}{V})$

Conclusion

v^0 est un point critique de f, c'est-à-dire $\nabla f(v^0) = 0$

Condition nécessaire d'optimalité

! Point critique $\not\Longrightarrow$ solution optimale

La réciproque est fausse
Pour la fonction $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, $x^0 = 0$ est point critique, mais n'est pas un point de minimum (ni de maximum)

Extrema locaux

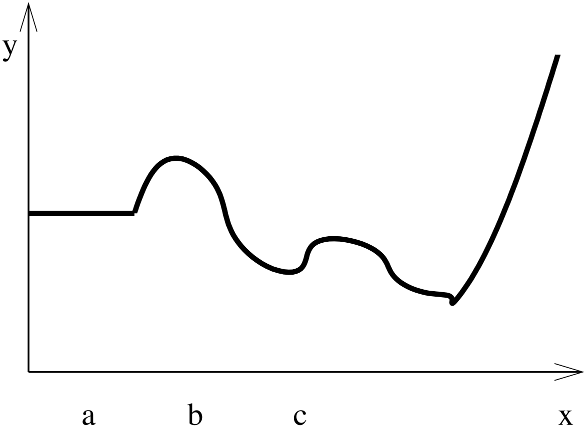
Définitions

Soit $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert. Un point $v^0 \in U$ est un

- minimum local s'il existe $r > 0$ tel que $f(v^0) \leq f(v)$, $\forall\ v \in B(v^0, r)$
- minimum local strict s'il existe $r > 0$ tel que $f(v^0) < f(v)$, $\forall\ v \in B(v^0, r) \setminus \{v^0\}$
- maximum local s'il existe $r > 0$ tel que $f(v^0) \geq f(v)$, $\forall\ v \in B(v^0, r)$
- maximum local strict s'il existe $r > 0$ tel que $f(v^0) > f(v)$, $\forall\ v \in B(v^0, r) \setminus \{v^0\}$

Dans tous ces cas, v^0 est un extremum local de f

Exemples



a est un minimum local et un maximum local. b est un maximum local strict. c est un minimum local strict

Théorème de Fermat

Théorème (de Fermat)

Hypothèses

- $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert
- $f \in C^1$
- v^0 extremum local de f

Conclusion

v^0 est point critique de f, c'est-à-dire $\nabla f(v^0) = 0$

! Point critique $\not\Longrightarrow$ extrémum local

La réciproque est fausse : prendre $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$

Nature locale des points critiques

Théorème

Hypothèses

- $f \in C^2(U, \mathbb{R})$, avec $U \subset \mathbb{R}^n$ ouvert
- v^0 point critique de f

Conclusions

(i) Si les valeurs propres de $H_{v^0}(f)$ sont strictement positives, alors v^0 est un minimum local strict

(ii) Si les valeurs propres de $H_{v^0}(f)$ sont strictement négatives, alors v^0 est un maximum local strict

(iii) Si $H_{v^0}(f)$ a à la fois des valeurs propres strictement positives et strictement négatives, alors v^0 n'est pas un extremum local (v^0 est un point-selle)

(iv) Dans les autres cas, on ne peut conclure

Comment reconnaître les matrices > 0 ?

Théorème (Critère de Laplace)

Une matrice symétrique $A = (a_{jk})_{j=1,\ldots,n}^{k=1,\ldots,n}$ a les valeurs propres strictement positives si et seulement si

$$\det(a_{jk})_{j=1,\ldots,l}^{k=1,\ldots,l} > 0, \; l = 1, \ldots, n$$

Une matrice symétrique $A = (a_{jk})_{j=1,\ldots,n}^{k=1,\ldots,n}$ a les valeurs propres strictement négatives si et seulement si

$$\operatorname{sgn}\det(a_{jk})_{j=1,\ldots,l}^{k=1,\ldots,l} = \operatorname{sgn}\,(-1)^l, \; l = 1, \ldots, n$$

Exemples

- La fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^3 + y^3 - 3xy$, a comme points critiques : $(0, 0)^T$, qui est un point-selle, et $(1, 1)^T$, qui est un minimum local strict
- La fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^4 + y^4$, a comme point critique $(0, 0)^T$, qui est un minimum local strict, mais ceci ne résulte pas du théorème précédent
- La fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^4 - y^4$, a comme point critique $(0, 0)^T$, qui n'est pas un extremum local, mais ceci ne résulte pas du théorème précédent

Quand la condition nécessaire devient suffisante

Théorème (réciproque au théorème de Fermat)

Hypothèses

- $f \in C^2(U, \mathbb{R})$, avec $U \subset \mathbb{R}^n$ ouvert convexe
- f convexe
- v^0 point critique de f

Conclusion

v^0 point de minimum de f

Problèmes d'optimisation

Le cadre

On considère le problème

$$(P)\ \min_{v \in V} f(v)$$

avec $V \subset \mathbb{R}^n$ fermé

Existence d'une solution optimale

Théorème (des bornes)

Hypothèses

- V compact
- f continue

Conclusion

(P) admet une solution optimale

Existence d'une solution optimale

Définition

$f : V \to \mathbb{R}$ est coerci(ti)ve si et seulement si $\lim\limits_{|v| \to \infty} f(v) = \infty$

Théorème

Hypothèses

f coercive et continue

Conclusion

(P) admet une solution optimale

Exemples

- $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2 - xy$, est coercive
- $f(x) = e^x$ est coercive sur $V = [0, +\infty[$, mais pas sur $\mathbb{R}$
- Si $f \in C^2(\mathbb{R}^n, \mathbb{R})$ est telle que $H_x(f) - I_n \geq 0$, $\forall\ x \in \mathbb{R}^n$, alors f est coercive

Exemples

- $\displaystyle \max_{x^2+y^2\leq 1} xy = \frac{1}{2}$
- $\displaystyle \min_{y\geq\sqrt{2}x^2} (x^2+y^2-2x) = -\frac{5}{8}$

Matrice jacobienne

$f : U \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$, avec $f_j : U \to \mathbb{R}$, $j = 1, \ldots, m$

$$J_f(x) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix} = \text{la matrice jacobienne de } f$$

Uniquement si $m = n$:
$\det J_f(x) =$ le déterminant jacobien de f en x

Exemples

Y a-t-il un écoulement de fluide

- donné en coordonnées lagrangiennes par $X = (1 + \cos t)\, x,\ Y = tx + e^t\, y$?
- incompressible donné en coordonnées eulériennes par $u = x + y,\ v = x - 2y$?
- incompressible donné en coordonnées lagrangiennes par $X = e^t\, x,\ Y = e^{-t}\, y$?

Donner une solution approchée du système

$$\begin{cases} e^{x+y} = 1,001 \\ \sin(x - y) = 0,003 \end{cases}$$

Problème

Si $f \in C^2(\mathbb{R}^n)$ vérifie $H_x f \geq a\,\mathrm{I}_n$, $\forall\, x \in \mathbb{R}^n$ (avec $a > 0$ fixé), alors $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ est bijective

Indications

- Appliquer le TAF à $x \mapsto (\nabla f(x) - \nabla f(y)) \cdot (x - y)$ (y fixé)
- En déduire que $x \mapsto \nabla f(x)$ est coerci(ti)ve
- À $z \in \mathbb{R}^n$ fixé, considérer le problème de minimisation $\min\limits_{x \in \mathbb{R}^n} |\nabla f(x) - z|^2$

Pour apprendre plus

Exercice

La fonction d'onde d'une particule quantique est donnée par

$$\psi(x,t) = \frac{1}{(\sqrt{2\pi})^3}\, e^{-\imath t - |x|^2/2},\ x \in \mathbb{R}^3,\ t \in \mathbb{R}$$

Est-ce (quantiquement parlant) une fonction lisse du temps ?