OPTIMISATION

LICENCE MATHÉMATIQUES ET GESTION 2013-2014

Plan général du cours

1. **Modélisation**
2. **Programmation linéaire**
3. **Calcul différentiel**
4. **Analyse matricielle**
5. **Optimisation nonlinéaire**

Table des matières

3 Calcul différentiel **1**
- 3.1 Notions de base 1
- 3.2 Fonctions de classe C^k 3
- 3.3 Développements limités 9
- 3.4 Ensembles et fonctions convexes 12
- 3.5 Extrema libres et liés 14
- 3.6 Quelques problèmes simples d'optimisation 17

3 Calcul différentiel

3.1 Notions de base

3.1 Notations. a) L'espace ambient est $\mathbb{R}^n$

b) Un point de $\mathbb{R}^n$ est désigné par x ou y, etc.

c) $x = (x_1, \dots, x_n)^T$

d) $\cdot$ est "le" produit scalaire dans $\mathbb{R}^n$:

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

e) $| \ |$ est "la" norme euclidienne dans $\mathbb{R}^n$: $|x| = \sqrt{x \cdot x}$

f) Si $r > 0$, alors

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

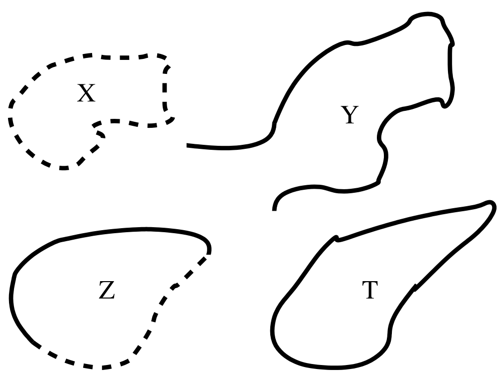
Ensembles

3.2 Définitions. a) $\Omega \subset \mathbb{R}^n$ est un ouvert si, pour tout $x \in \Omega$, il existe un $r > 0$ (qui dépend de x) tel que $B(x, r) \subset \Omega$

b) $F \subset \mathbb{R}^n$ est fermé si $\mathbb{R}^n \setminus F$ est ouvert

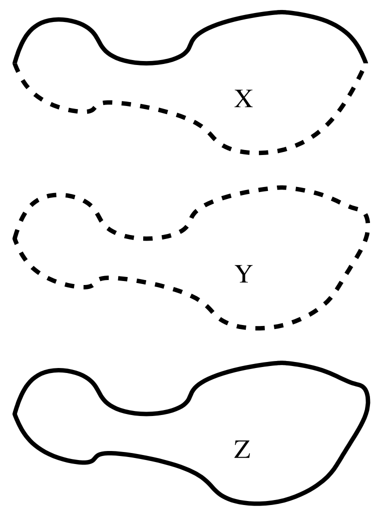
c) $B \subset \mathbb{R}^n$ est borné s'il existe $M > 0$ tel que $B \subset \overline{B}(0, M)$
Ou encore : il existe $M > 0$ tel que $|x| \leq M$, $\forall\ x \in B$

d) $K \subset \mathbb{R}^n$ est compact si K est à la fois fermé et borné



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

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



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

3.3 Définitions. a) 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 $\mathring{X}$
Ou encore : $\mathring{X}$ est l'union de tous les ouverts contenus dans X

b) 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

Suites

3.4 Notations. a) Une suite de points de $\mathbb{R}^n$ est notée $x^0, x^1, \ldots$, ou $(x^k)_{k \geq 0}$

b) Abus de notation : on écrit $(x^k) \subset \mathbb{R}^n$, et non pas $(x^k) \subset (\mathbb{R}^n)^{\mathbb{N}}$

c) Par abus de notation, une suite extraite de (x^k) est 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

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

a) 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 $j = 1, \dots, n$

b) est bornée si et seulement s'il existe $M > 0$ tel que $|x^k| \leq M$ pour tout k

Exemples

a) La suite de terme général $x^k = (1 + 1/(k+1), e^{-k})^T$ converge vers $(1,0)^T \in \mathbb{R}^2$

b) La suite de terme général $x^k = (1, (-1)^k)^T$ ne converge pas

c) La suite de terme général $x^k = (1, (-1)^k)^T$ est bornée

d) La suite de terme général $x^k = (0, k)^T$ n'est pas bornée

3.6 Théorème (Bolzano-Weierstrass ; 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 $l \to \infty$

3.7 Théorème (Bolzano-Weierstrass ; 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 $x^{k_l} \to x$ quand $l \to \infty$

3.2 Fonctions de classe C^k

La continuité expliquée à ma fille

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

3.8 Définition. Soient $X \subset \mathbb{R}^n$ et $f : X \to \mathbb{R}$. Alors f est continue si et seulement si :

$$x \in X,\ (x^k) \subset X,\ x^k \to x \implies f(x^k) \to f(x)$$

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

Exemples fondamentaux
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 $j_1, \ldots, j_n$ 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^∞

3.9 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. a) $(x, y) \mapsto x^2 - y$ est continue (polynôme)
b) $(x, y) \mapsto \sin(x^2 - y)$ est continue (composée de fonctions continues)
c) $(x, y) \mapsto x^2 + y^4$ est continue (polynôme)
d) $(x, y) \mapsto \sqrt{x^2 + y^4}$ est continue (composée de fonctions continues)
e) f est continue (différence de fonctions continues) □

Fonctions continues et ensembles

3.10 Proposition.
Hypothèse $f : \mathbb{R}^n \to \mathbb{R}$ est continue
Conclusions
a) 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
b) 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

3.11 Proposition (Généralisation de la proposition précédente).
Hypothèses

(i) $f_1, \ldots, f_m, g_1, \ldots, g_k, h_1, \ldots, h_l : \mathbb{R}^n \to \mathbb{R}$ sont continues

(ii) $a_1, \ldots, a_m, b_1, \ldots, b_k, c_1, \ldots, c_l$ sont des constantes

Conclusions

a) L'ensemble
$$\{x \in \mathbb{R}^n \ ; \ f_j(x) < a_j, g_j(x) > b_j, \forall \ j\}$$
est ouvert

b) 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

Les ensembles expliqués à ma fille

a) Les inégalités strictes donnent des ouverts

b) Les inégalités larges et les égalités donnent des fermés

Fonctions de classe C^1

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

a) On peut parler de la continuité de f quel que soit X

b) Mais on va parler du caractère C^1 de f seulement dans deux cas :
   (i) X est ouvert
   (ii) X est l'adhérence d'un ouvert

Dérivées directionnelles

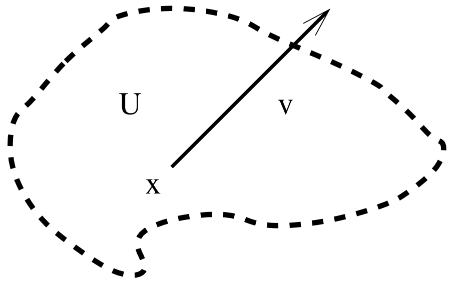
a) Cadre : $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert

b) On fixe :
   (i) $x \in U$
   (ii) $v \in \mathbb{R}^n$

c) Pour $h \in \mathbb{R}$ proche de 0, on a $x + hv \in U$

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

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



x + un petit multiple de v est dans U

Exemple
Si $f:\mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = xy$, alors $\dfrac{\partial f}{\partial(1,2)}(1,1) = 3$

Lien avec la dérivée usuelle
Dans $\mathbb{R}$ $(n=1)$, $f'(x) = \dfrac{\partial f}{\partial 1}(x)$

Dérivées partielles

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

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

Autres notations

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

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

a) f admet des dérivées partielles (du premier ordre)

b) 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

3.14 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. a) $(x,y)\mapsto x^2-y$ est C^1 (polynôme)

b) $(x,y)\mapsto\sin(x^2-y)$ est C^1 (composée de fonctions C^1)

c) $(x,y)\mapsto x^2+y^4$ est C^1 (polynôme)

d) $t\mapsto\sqrt{t}$ est C^1 là où $t>0$

e) $(x,y)\mapsto\sqrt{x^2+y^4}$ 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) f est C^1 (différence de fonctions C^1) sur $\mathbb{R}^2\setminus\{0\}$ $\square$

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

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

a) 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 $\partial_j f$, $j=1,\dots,n$ et s'il existe et sont continues $\partial_k\partial_j f$, $j,k=1,\dots,n$

b) 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 $m=2,3,\dots$

   Par exemple : $\partial_3\partial_2\partial_2\partial_1 f$ est une dérivée du quatrième ordre

c) Pour $k=2,3,\dots$, 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

3.16 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

3.17 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

a) On considère $f \in C^4$

b) 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

c) Le théorème de Schwarz donne

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

3.18 Notations. a) $\nabla f(x) = \vec{\nabla} f(x) := (\partial_1 f(x), \ldots, \partial_n f(x))^T$

b) ∇f est le gradient de f

c) $H_x(f) = \nabla^2 f(x) := \left(\partial_k \partial_j f(x)\right)_{j=1,\ldots,n}^{k=1,\ldots,n}$

d) $H(f)$ est la matrice hessienne de f ou l'hessien(ne) de f

e) $f \in C^k$ 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) $f \in C^\infty$ (ou encore f est indéfiniment différentiable) veut dire $f \in C^k$, $k = 1, 2, \ldots$

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

3.19 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

3.20 Proposition (Conséquence du théorème de Schwarz)**.**
Hypothèse $f \in C^2$
Conclusion $H_x(f)$ est une matrice symétrique

3.21 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

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

3.3 Développements limités

3.22 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

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

b) $h \neq o(h)$

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

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

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

b) 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 $\mathbb{R}$

3.24 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

a) (formule de Taylor avec reste) $f(y) = f(x) + f'(x)(y-x) + o(y-x)$

b) (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 $f(y) = f(x) + f'(z)(y-x)$

c) (théorème des accroissements finis)

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

3.25 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

a) (formule de Taylor avec reste)

$$f(y) = f(x) + f'(x)(y - x) + f''(x)(y - x)^2/2 + o((y - x)^2)$$

b) (formule de Taylor avec point intermédiaire) il existe $z \in [x, y]$ (ou $z \in (x, y)$ si $x \neq y$) tel que $f(y) = f(x) + f'(x)(y - x) + f''(z)(y - x)^2/2$

c) (théorème des accroissements finis)

$$|f(y) - f(x) - f'(x)(y - x)| \le (y - x)^2/2 \max_{z \in [x,y]} |f''(z)|$$

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

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

3.26 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

a) (formule de Taylor avec reste) $f(y) = f(x) + \nabla f(x) \cdot (y - x) + o(y - x)$

b) (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 $f(y) = f(x) + \nabla f(z) \cdot (y - x)$

c) (théorème des accroissements finis)

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

3.27 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

a) (formule de Taylor avec reste)

$$f(y) = f(x) + \nabla f(x) \cdot (y-x) + (H_x f(y-x)) \cdot (y-x)/2 + o(|y-x|^2)$$

b) (formule de Taylor avec point intermédiaire) il existe $z \in [x,y]$ (ou $z \in (x,y)$ si $x \neq y$) tel que

$$f(y) = f(x) + \nabla f(x) \cdot (y-x) + (H_z f(y-x)) \cdot (y-x)/2$$

c) (théorème des accroissements finis)

$$|f(y) - f(x) - \nabla f(x) \cdot (y-x)| \leq |y-x|^2/2 \max_{z \in [x,y]} \|H_z f\|$$

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 $\|H_z f\| = 1 + \sqrt{13}$

Idée des preuves

a) On se donne $x, y \in U$ tels que $[x,y] \in U$ et $f \in C^k$ ($k = 1$ ou 2)

b) On ramène chaque formule dans $\mathbb{R}^n$ à une formule sur $I = [0,1]$

c) Pour ce faire, on introduit la fonction auxiliaire

$$g(t) = f((1-t)x + ty), \quad t \in [0,1]$$

de sorte que :

(i) $g \in C^k$

(ii) $g(0) = f(x)$, $g(1) = f(y)$

d) On applique la formule de Taylor correspondante à g, avec

(i) $I \rightsquigarrow [0,1]$

(ii) $x \rightsquigarrow 0$

(iii) $y \rightsquigarrow 1$

e) En calculant les dérivées de g en fonction de celles de f, on retrouve la formule de Taylor correspondante

Formules utiles

3.28 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

a) $g'(t) = (\nabla f)(z) \cdot (y - x)$

b) En particulier, $g'(0) = \nabla f(x) \cdot (y - x)$

c) $g''(t) = (H_z(f)(y-x)) \cdot (y - x)$

d) En particulier, $g''(0) = (H_x(f)(y-x)) \cdot (y - x)$

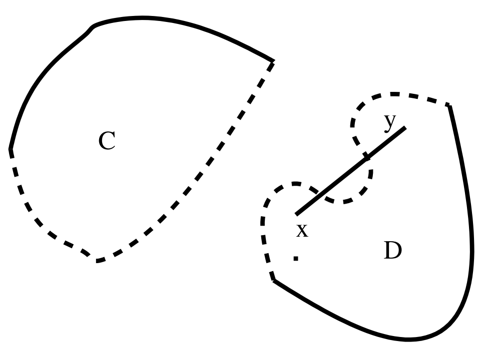
3.4 Ensembles et fonctions convexes

3.29 Définitions. a) $C \subset \mathbb{R}^n$ est convexe si et seulement si :

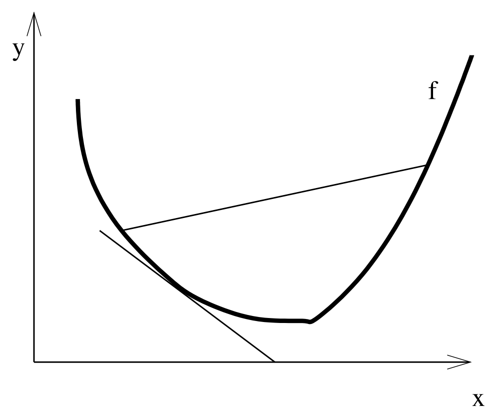
$$\text{pour tout } x, y \in C, \text{ on a } [x,y] \subset C$$

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

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



C est convexe. D n'est pas convexe



Si f est convexe, alors le graphe de f se trouve au−dessus des tangentes et en dessous des sécantes

La convexité expliquée à ma fille

Une fonction est convexe si son graphe retient l'eau

Fonctions convexes dans $\mathbb{R}$

3.30 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 :

a) f convexe

b) f' croissante

c) $f'' \geq 0$

d) $f(y) \geq f(x) + f'(x)(y - x)$, $\forall\ x, y \in I$

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

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

3.31 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 :

a) f convexe

b) $(\nabla f(y) - \nabla f(x)) \cdot (y - x) \geq 0$, $\forall\ x, y \in C$

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 $x \in C$)

d) $f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$, $\forall\ x, y \in C$

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

Exemples

a) 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 $A = (a_{jk})_{j=1,\ldots,n}^{k=1,\ldots,n}$ est ≥ 0

b) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^3$ n'est pas convexe

3.5 Extrema libres et liés

3.32 Définitions. a) Un problème de minimisation est un problème de la forme

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

avec

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

(ii) $f : V \to \mathbb{R}$

b) Les points $v \in V$ sont des solutions admissibles (pour (P))

c) f est la fonction objectif

d) Une solution de (P) (si elle existe) est un point $v^0 \in V$ tel que $f(v^0) \leq f(v)$, $\forall\, v \in V$

e) Un tel v^0 est une solution optimale de (P)

f) Le minimum de f sur V (c'est-à-dire la valeur de f en v^0) est le coût du problème (P)

Exemples

a) Le problème (P) $\min_{x \in \mathbb{R}} e^x$ n'a pas de solution optimale

b) Pour le problème $\min_{x \in \mathbb{R}} (e^x - x)$, la solution optimale est $x^0 = 0$, alors que le coût est 1

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

Points critiques

3.33 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 $\nabla f(v^0) = 0$

3.34 Théorème (de Fermat).
Hypothèses

(i) v^0 solution optimale de (P)

(ii) $v^0 \in \mathring{V}$

(iii) $f \in C^1(\mathring{V})$

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

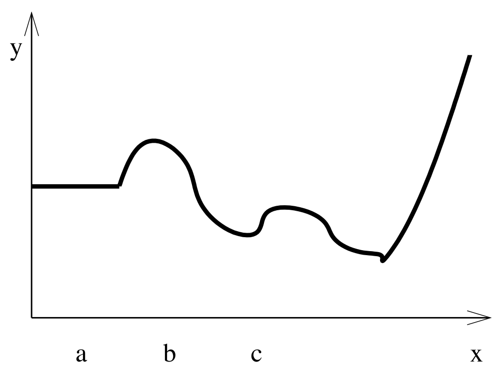
Attention !
La réciproque est fausse
Pour la fonction $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, $x^0 = 0$ est point critique, mais n'est pas un point de minimum (ni de maximum)

Extrema locaux

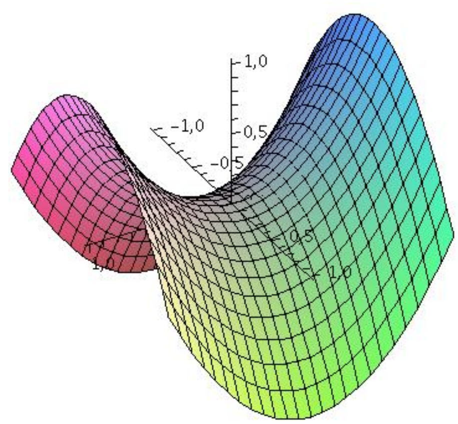
3.35 Définitions. Soit $f : U \to \mathbb{R}$, avec $U \subset \mathbb{R}^n$ ouvert. Un point $v^0 \in U$ est un

a) minimum local s'il existe $r > 0$ tel que $f(v^0) \le f(v)$, $\forall\ v \in B(v^0, r)$
b) minimum local strict s'il existe $r > 0$ tel que $f(v^0) < f(v)$, $\forall\ v \in B(v^0, r) \setminus \{v^0\}$
c) maximum local s'il existe $r > 0$ tel que $f(v^0) \ge f(v)$, $\forall\ v \in B(v^0, r)$
d) 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



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



$(0,0)^T$ est un point selle pour $f(x, y) = x^2 - y^2$

3.36 Théorème (de Fermat).
Hypothèses

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

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

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

Nature locale des points critiques

3.37 Théorème.
Hypothèses

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

Conclusions

a) Si les valeurs propres de $H_{v^0}(f)$ sont strictement positives, alors v^0 est un minimum local strict
b) Si les valeurs propres de $H_{v^0}(f)$ sont strictement négatives, alors v^0 est un maximum local strict
c) 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)
d) Dans les autres cas, on ne peut conclure

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

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

b) 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

a) 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
b) 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
c) 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

3.39 Théorème (réciproque au théorème de Fermat)**.**
Hypothèses

(i) $f \in C^2(U, \mathbb{R})$, avec $U \subset \mathbb{R}^n$ ouvert convexe

(ii) f convexe

(iii) v^0 point critique de f

Conclusion
v^0 point de minimum de f

3.6 Quelques problèmes simples d'optimisation

Le cadre
On considère le problème

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

avec $V \subset \mathbb{R}^n$ fermé
Notations :

a) $U = \overset{\circ}{V}$

b) $\Sigma = \partial V := V \setminus U$ (c'est le bord de V)

3.40 Théorème (des bornes)**.**
Hypothèses

(i) V compact

(ii) f continue

Conclusion
(P) admet une solution optimale

3.41 Définition. $f : V \to \mathbb{R}$ est coerci(ti)ve si et seulement si $\lim_{|v| \to \infty} f(v) = \infty$

3.42 Théorème.
Hypothèses
f coercive et continue
Conclusion
(P) admet une solution optimale

Exemples

a) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + y^2 - xy$, est coercive

b) $f(x) = e^x$ est coercive sur $V = [0, +\infty[$, mais pas sur $\mathbb{R}$

c) Si $f \in C^2(\mathbb{R}^n, \mathbb{R})$ est telle que $H_x(f) - aI_n \geq 0$, $\forall\ x \in \mathbb{R}^n$, où $a > 0$ est une constante, alors f est coercive
On dit alors que f est a-coercive

Problème du central téléphonique
Trois villes, A, B et C, ont la même population
Placez de manière optimale un central téléphonique desservant les trois villes

Contrainte régulière

Le cadre
On suppose
$$V = \{x \in \mathbb{R}^n \ ;\ h(x) \leq 0\}$$

3.43 Définition. $h : \mathbb{R}^n \to \mathbb{R}$ est une contrainte régulière si et seulement si :

a) $h \in C^2$

b) $\{x\ ; h(x) = 0\} \neq \emptyset$

c) pour tout x, $h(x) = 0 \Longrightarrow \nabla h(x) \neq 0$
Ou encore, le système $\begin{cases} \nabla h(x) = 0 \\ h(x) = 0 \end{cases}$ n'a pas de solution

Exemples

a) $h : \mathbb{R}^n \to \mathbb{R}$, $h(x) = |x|^2 - 1$ est une contrainte régulière

b) $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x,y) = xy$ n'est pas une contrainte régulière

c) $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x,y) = e^{x+y}$ n'est pas une contrainte régulière

3.44 Théorème (Fermé associé à une contrainte régulière).
Hypothèses

(i) h est une contrainte régulière

(ii) $V = \{x \in \mathbb{R}^n \ ;\ h(x) \leq 0\}$

Conclusions

a) $\overset{\circ}{V} = \{x \in \mathbb{R}^n \ ;\ h(x) < 0\}$

b) $\Sigma = V \setminus \overset{\circ}{V} = \{x \in \mathbb{R}^n \ ;\ h(x) = 0\}$

Minimisation sous contrainte régulière

Le cadre
On considère le problème (P) avec V donné par une contrainte régulière
On suppose la fonction objectif, f, de classe C^2
On suppose qu'il existe une solution optimale v^0 (par exemple, si f est continue et soit V est compact, soit f est coercive)
Alors :

a) ou bien $v^0 \in \overset{\circ}{V}$, et alors $\nabla f(v^0) = 0$ (théorème de Fermat)

b) ou bien $v^0 \in \Sigma$

3.45 Théorème (multiplicateur de Lagrange)**.**
Hypothèses

(i) $V = \{x \in \mathbb{R}^n \ ; h(x) \leq 0\}$

(ii) h régulière

(iii) $v^0 \in \Sigma$ solution optimale de (P)

Conclusion
Il existe λ réel négatif tel que $\nabla f(v^0) = \lambda \nabla h(v^0)$

Exemples

a) $\displaystyle\max_{x^2+y^2\leq 1} xy = \frac{1}{2}$

b) $\displaystyle\inf_{y>\sqrt{2}x^2} (x^2 + y^2 - 2x) = -\frac{5}{8}$