

Fonctions de plusieurs variables

Hervé Hocquard

Université de Bordeaux, France

16 février 2022

université
de **BORDEAUX**

LABORATOIRE
BORDELAIS
DE RECHERCHE
EN INFORMATIQUE

LaBRI

Définition

On note $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ fois}} = \{X = (x_1, \cdots, x_n) \mid x_i \in \mathbb{R}, \forall i \in \llbracket 1, n \rrbracket\}$

\mathbb{R}^n est un \mathbb{R} -espace vectoriel de dimension n .

Définition

On note $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ fois}} = \{X = (x_1, \cdots, x_n) \mid x_i \in \mathbb{R}, \forall i \in [|1, n|]\}$

\mathbb{R}^n est un \mathbb{R} -espace vectoriel de dimension n .

Introduction

On s'intéresse aux fonctions $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. Il faut d'abord étudier la structure du domaine D car le domaine est aussi important que la fonction. Pour cela on va définir une notion de distance.

Définition : distance

Soit E un ensemble non vide. On dit qu'une application $d : E \times E \rightarrow \mathbb{R}_+$, $d : (x, y) \mapsto d(x, y)$ est une distance sur E si elle vérifie les trois axiomes suivants :

D1 (séparation) $\forall (x, y) \in E \times E, \{x = y\} \Leftrightarrow \{d(x, y) = 0\}$;

D2 (symétrie) $\forall (x, y) \in E \times E, d(x, y) = d(y, x)$;

D3 (inégalité triangulaire)

$\forall (x, y, z) \in E \times E \times E, d(x, y) \leq d(x, z) + d(z, y)$.

Définition : distance

Soit E un ensemble non vide. On dit qu'une application $d : E \times E \rightarrow \mathbb{R}_+$, $d : (x, y) \mapsto d(x, y)$ est une distance sur E si elle vérifie les trois axiomes suivants :

D1 (séparation) $\forall (x, y) \in E \times E, \{x = y\} \Leftrightarrow \{d(x, y) = 0\}$;

D2 (symétrie) $\forall (x, y) \in E \times E, d(x, y) = d(y, x)$;

D3 (inégalité triangulaire)

$$\forall (x, y, z) \in E \times E \times E, d(x, y) \leq d(x, z) + d(z, y).$$

Définition

On appelle espace métrique tout couple (E, d) où $E \neq \emptyset$ est un espace vectoriel et d est une distance.

Définition : distance

Soit E un ensemble non vide. On dit qu'une application $d : E \times E \rightarrow \mathbb{R}_+$, $d : (x, y) \mapsto d(x, y)$ est une distance sur E si elle vérifie les trois axiomes suivants :

D1 (séparation) $\forall (x, y) \in E \times E, \{x = y\} \Leftrightarrow \{d(x, y) = 0\}$;

D2 (symétrie) $\forall (x, y) \in E \times E, d(x, y) = d(y, x)$;

D3 (inégalité triangulaire)

$$\forall (x, y, z) \in E \times E \times E, d(x, y) \leq d(x, z) + d(z, y).$$

Définition

On appelle espace métrique tout couple (E, d) où $E \neq \emptyset$ est un espace vectoriel et d est une distance.

Exemple

$$E = \mathbb{R}, d(x, y) = |x - y|$$

Définition : produit scalaire

Soient $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$. On appelle produit scalaire de X et Y , noté $X.Y$ ou $\langle X, Y \rangle$ le réel :

$$X.Y = x_1y_1 + x_2y_2 + \dots + x_ny_n$$

Définition : produit scalaire

Soient $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$. On appelle produit scalaire de X et Y , noté $X.Y$ ou $\langle X, Y \rangle$ le réel :

$$X.Y = x_1y_1 + x_2y_2 + \dots + x_ny_n$$

Définition

Soit E un espace vectoriel sur \mathbb{R} . On appelle norme sur E une application de E dans \mathbb{R}_+ qui à $x \mapsto \|x\| \in \mathbb{R}_+$, et vérifie

N1 (séparation) $\forall x \in E, \|x\| = 0 \Leftrightarrow x = 0$

N2 (homogénéité positive) $\forall \lambda \in \mathbb{R}, \forall x \in E, \|\lambda x\| = |\lambda| \cdot \|x\|$

N3 (inégalité triangulaire) $\forall x, y \in E, \|x + y\| \leq \|x\| + \|y\|$.

Un espace vectoriel sur \mathbb{R} muni d'une norme est appelé espace vectoriel normé (e.v.n.).

Définition

Soit E un e.v.n. L'application $d : E \times E \rightarrow \mathbb{R}_+$ qui au couple (x, y) associe $d(x, y) = \|x - y\|$ est une distance sur E .

On l'appelle distance induite sur E par la norme. Elle possède les propriétés suivantes :

- $\forall x \in E, d(0, x) = \|x\|$
- $\forall \lambda \in \mathbb{R}, \forall (x, y) \in E^2, d(\lambda x, \lambda y) = |\lambda|d(x, y)$
- $\forall (x, y, z) \in E \times E \times E, d(x + z, y + z) = d(x, y)$.

Définition

Soit E un e.v.n. L'application $d : E \times E \rightarrow \mathbb{R}_+$ qui au couple (x, y) associe $d(x, y) = \|x - y\|$ est une distance sur E .

On l'appelle distance induite sur E par la norme. Elle possède les propriétés suivantes :

- $\forall x \in E, d(0, x) = \|x\|$
- $\forall \lambda \in \mathbb{R}, \forall (x, y) \in E^2, d(\lambda x, \lambda y) = |\lambda|d(x, y)$
- $\forall (x, y, z) \in E \times E \times E, d(x + z, y + z) = d(x, y)$.

Remarque

Toute norme induit une distance, par contre toutes les distances ne proviennent pas d'une norme.

Exemple de normes sur \mathbb{R}^n .

Soit $x \in \mathbb{R}^n$, $X = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}, \forall i \in \llbracket 1, n \rrbracket$. Alors

$$\|X\|_1 = \sum_{i=1}^n |x_i| \quad (\text{norme de Manhattan})$$

$$\|X\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \quad (\text{norme euclidienne})$$

$$\|X\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

$$\|X\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

sont des normes sur \mathbb{R}^n .

Remarque

Dans toute la suite du cours :

$$d(X, Y) = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{1/2}$$

$$\| X \| = \| X \|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$$

Définition

Soit a un point de \mathbb{R}^n et $r > 0$ un nombre réel.

- 1 $\bar{B}(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) \leq r\}$ est appelée boule fermée de centre a et de rayon r .
- 2 Une boule ouverte de centre a et de rayon r est $B(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) < r\}$
- 3 Une sphère de centre a et de rayon r est $S(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) = r\}$

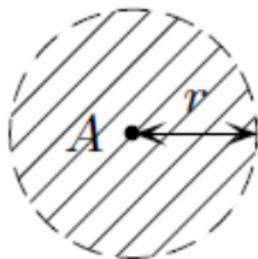
Définition

Soit a un point de \mathbb{R}^n et $r > 0$ un nombre réel.

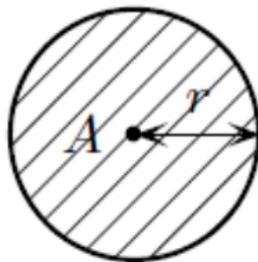
- 1 $\bar{B}(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) \leq r\}$ est appelée boule fermée de centre a et de rayon r .
- 2 Une boule ouverte de centre a et de rayon r est $B(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) < r\}$
- 3 Une sphère de centre a et de rayon r est $S(a, r) = \{x \in \mathbb{R}^n \mid d(a, x) = r\}$

Remarque

On obtient des boules de formes différentes pour des espaces métriques différents...



Boule ouverte sur \mathbb{R}^2



Boule fermée sur \mathbb{R}^2

Définition

Une partie bornée P de \mathbb{R}^n est une partie de \mathbb{R}^n pour laquelle on peut trouver une boule (ouverte ou fermée) qui contient tous les points de P .

Définition

Une partie ouverte (ou un ouvert) de \mathbb{R}^n est une partie Ω t.q.
 $\forall u \in \Omega, \exists r > 0$ tel que $B(u,r) \subset \Omega$ i.e. tout point de Ω est le
centre d'une boule ouverte, de rayon non nul, incluse dans Ω .

Une partie fermée (ou un fermé) de \mathbb{R}^n est une partie telle que
son complémentaire Ω dans \mathbb{R}^n est un ouvert.

Remarque

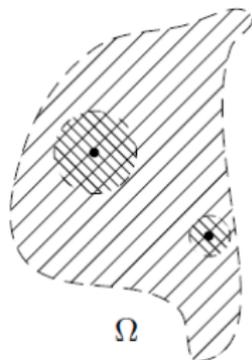
Un ouvert de \mathbb{R}^2 est une partie de \mathbb{R}^2 dans laquelle tous les points ont de la place autour d'eux. On peut faire un tour autour de chaque point sans sortir de l'ouvert.

On peut se représenter les ouverts en pensant que ce sont des taches qui ne contiennent pas leur frontière.

Remarque

Un ouvert de \mathbb{R}^2 est une partie de \mathbb{R}^2 dans laquelle tous les points ont de la place autour d'eux. On peut faire un tour autour de chaque point sans sortir de l'ouvert.

On peut se représenter les ouverts en pensant que ce sont des taches qui ne contiennent pas leur frontière.



Exemples

- \mathbb{R}^2 est une partie ouverte de \mathbb{R}^2 .
- L'ensemble vide est une partie ouverte de \mathbb{R}^2 .
- Une boule ouverte est une partie ouverte de \mathbb{R}^2 .
- Un demi-plan ouvert (i.e. bord exclu) est une partie ouverte.
- Une réunion quelconque de parties ouvertes est une partie ouverte.
- Une intersection finie de parties ouvertes est une partie ouverte.
- Tout produit de deux intervalles ouverts est un ouvert.
- Une boule fermée n'est pas une partie ouverte de \mathbb{R}^2 .
- Le complémentaire de toute boule fermée est un ouvert.

Définition

Soit $A \subset \mathbb{R}^n$.

- 1 On dit que a est intérieur à A si on peut trouver un ouvert $\Omega \subset \mathbb{R}^n$ tel que $a \in \Omega$ et $\Omega \subset A$. L'intérieur de A , noté $\overset{\circ}{A}$, est le plus grand ouvert inclus dans A .
- 2 On dit que a est un point frontière de A si tout ouvert $\Omega \subset \mathbb{R}^n$ contenant a rencontre à la fois A et le complémentaire de A .
- 3 On dit que a est adhérent à A si tout ouvert $\Omega \subset \mathbb{R}^n$ contenant a rencontre A .
- 4 L'adhérence de A , notée \overline{A} , est le plus petit fermé qui contient A .

Voisinages d'un point dans \mathbb{R}^2

Définition générale

On dit qu'une partie V de \mathbb{R}^n est un voisinage de $x \in \mathbb{R}^n$ si V contient un ouvert contenant x .

Définition

Soit $a \in \mathbb{R}^2$, on appelle voisinage de a , toute partie de \mathbb{R}^2 contenant une boule ouverte de centre a .

Remarque

Dire qu'une partie de \mathbb{R}^2 est ouverte, c'est exactement dire qu'elle est voisinage de chacun de ses points.

Définition

Soit X et Y deux éléments de \mathbb{R}^n . On définit le segment fermé (resp. ouvert) $[X, Y]$ (resp. $]X, Y[$) par :

$$[X, Y] = \{Z = \lambda X + (1 - \lambda)Y \text{ où } \lambda \in [0, 1] \text{ (resp. } \lambda \in]0, 1[)\}$$

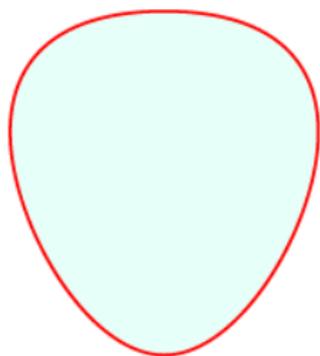
Définition

Soit X et Y deux éléments de \mathbb{R}^n . On définit le segment fermé (resp. ouvert) $[X, Y]$ (resp. $]X, Y[$) par :

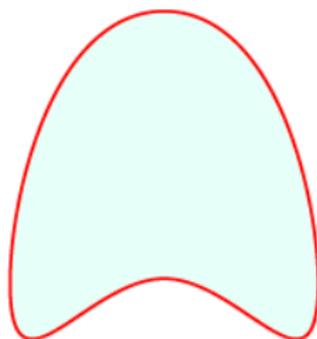
$$[X, Y] = \{Z = \lambda X + (1 - \lambda)Y \text{ où } \lambda \in [0, 1] \text{ (resp. } \lambda \in]0, 1[)\}$$

Définition

Un sous-ensemble non vide de \mathbb{R}^n est dit convexe si chaque fois qu'il contient deux points X et Y , il contient tout le segment $[X, Y]$.



Convexe



Non convexe

Remarques

- Toute boule (ouverte ou fermée) est convexe.
- Tout demi-espace est convexe.
- Toute intersection non vide de convexes est un convexe.

Définition

On s'intéresse maintenant aux fonctions $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. On distingue des **fonctions scalaires** : $\mathbb{R}^n \rightarrow \mathbb{R}$ et des **fonctions vectorielles** : $\mathbb{R}^n \rightarrow \mathbb{R}^p$, $p > 1$.

Définition

On s'intéresse maintenant aux fonctions $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. On distingue des **fonctions scalaires** : $\mathbb{R}^n \rightarrow \mathbb{R}$ et des **fonctions vectorielles** : $\mathbb{R}^n \rightarrow \mathbb{R}^p$, $p > 1$.

Introduction

On va commencer par l'étude des fonctions de deux variables. Une fonction définie sur une partie D de \mathbb{R}^2 et à valeurs réelles fait correspondre à tout point $X = (x, y)$ de D , (appelé le domaine de définition de f) un réel unique $f(X)$.

Introduction

Les fonctions utilisées dans les modèles économiques sont, le plus souvent, des fonctions de plusieurs variables.

Exemple

Les fonctions de production, les fonctions d'utilité.

Introduction

On considère un système de production produisant un certain produit "output" à partir de n autres biens, appelés facteurs de production ou "input".

Introduction

On considère un système de production produisant un certain produit "output" à partir de n autres biens, appelés facteurs de production ou "input".

La fonction de production associe à un n -uplet (x_1, \dots, x_n) de nombres positifs la quantité maximale q d'output que l'on peut produire.

Introduction

On considère un système de production produisant un certain produit "output" à partir de n autres biens, appelés facteurs de production ou "input".

La fonction de production associe à un n -uplet (x_1, \dots, x_n) de nombres positifs la quantité maximale q d'output que l'on peut produire.

Dans certains modèles, on se limite à deux facteurs de production : le **capital** et le **travail**. Dans ce cas, on s'intéresse à des fonctions de la forme : $q = f(K, L)$ où K représente le capital et L le travail. On a alors une fonction de deux variables, associant à un couple (K, L) de nombres réels (ici positifs) un nombre $f(K, L)$.

Exemple

Les fonctions suivantes apparaissent souvent en théorie de production :

- $f(K, L) = aK + bL$ (fonction linéaire)
- $f(K, L) = \min(K/a, L/b)$, avec $a > 0$ et $b > 0$ (fonction à facteurs complémentaires)
- $f(K, L) = kK^\alpha L^\beta$, avec $k \geq 0$, $\alpha \geq 0$, $\beta \geq 0$ (fonction de Cobb-Douglas).

Définition

Soit $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^2$.

- 1 L'ensemble des points de \mathbb{R}^3

$$S_f = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, z = f(x, y)\}.$$

est appelé la surface représentative de f . S est aussi appelé le graphe de la fonction f .

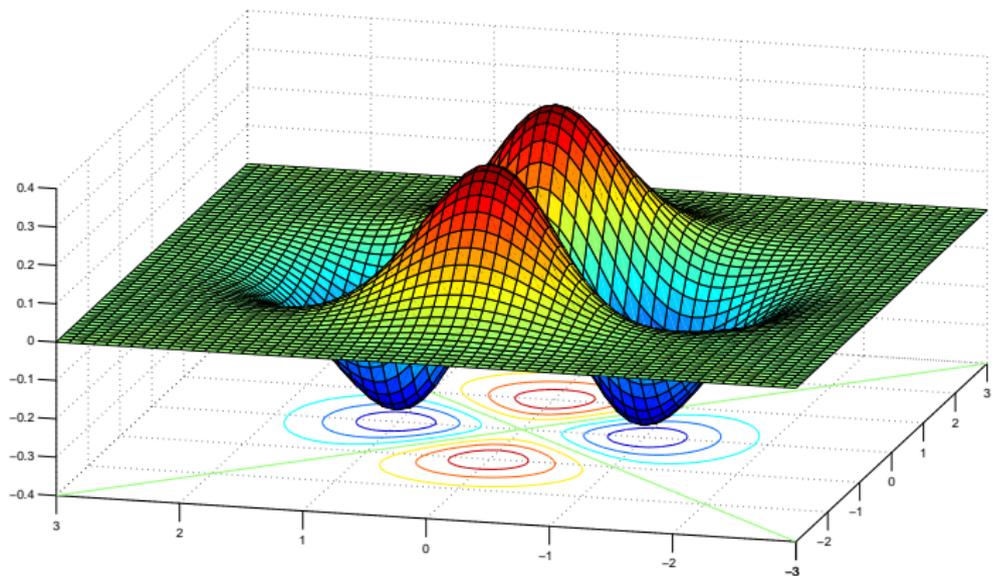
- 2 Soit $A = (a, b)$ un point intérieur de D . Les fonctions $x \mapsto f(x, b)$ et $y \mapsto f(a, y)$ définies sur des intervalles ouverts, contenant respectivement b et a sont appelées les fonctions partielles associées à f au point A .
- 3 Soit $k \in \mathbb{R}$. L'ensemble $C_k = \{(x, y) \in D \text{ et } f(x, y) = k\}$ est la courbe de niveau k (ou ligne de niveau) de la fonction f .

Remarque

Pour les fonctions de trois variables, la notion analogue à la courbe de niveau est celle de surface de niveau.

Les lignes de niveau et les fonctions partielles sont utiles pour dessiner les graphes des fonctions.

Fonctions de plusieurs variables : exemple



$$f(x,y) = (x^2 - y^2) \cdot \exp(-x^2 - y^2)$$

Une fois qu'on a les normes et les voisinages, la définition de limite est la même que dans \mathbb{R} :

Définition

Soit $(X_n)_{n \in \mathbb{N}}$ une suite d'éléments de \mathbb{R}^n et $A \in \mathbb{R}^n$.

On dit que $\lim_{n \rightarrow +\infty} X_n = A$ ssi $\forall V$ voisinage de A , $\exists N_V \in \mathbb{N}$ tel que $n \geq N_V \Rightarrow X_n \in V$.

C'est-à-dire que :

$\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N}$ tel que $n \geq N_\varepsilon \Rightarrow \|X_n - A\| \leq \varepsilon$.

Propriété

Soit $(X_n)_{n \in \mathbb{N}} = ((x_{n_1}, \dots, x_{n_n}))_{n \in \mathbb{N}}$ une suite de \mathbb{R}^n et $A = (a_1, \dots, a_n) \in \mathbb{R}^n$, alors $\lim_{n \rightarrow +\infty} X_n = A$ ssi $\forall i \in [1, n]$,
 $\lim_{n \rightarrow +\infty} x_{n_i} = a_i$.

Définition

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $A \in D$.

On dit que f a une limite $L \in \mathbb{R}^p$ en A ssi $\forall (X_n)_{n \in \mathbb{N}}$ suite de D telle que $\lim_{n \rightarrow +\infty} X_n = A$, on a $\lim_{n \rightarrow +\infty} f(X_n) = L$.

Il y a une autre définition de la limite d'une fonction qui est équivalente à la définition précédente.

Définition

Soit $f : D \rightarrow \mathbb{R}^p$ une fonction définie sur une partie D de \mathbb{R}^n et A un point adhérent à D , L un point de \mathbb{R}^p .

On dit que f a pour limite L lorsque $X \rightarrow A$ si :

$$(\forall \varepsilon > 0, \exists \eta > 0 ; \|X - A\| \leq \eta, X \in D) \implies (\|f(X) - L\| \leq \varepsilon).$$

Il y a une autre définition de la limite d'une fonction qui est équivalente à la définition précédente.

Définition

Soit $f: D \rightarrow \mathbb{R}^p$ une fonction définie sur une partie D de \mathbb{R}^n et A un point adhérent à D , L un point de \mathbb{R}^p .

On dit que f a pour limite L lorsque $X \rightarrow A$ si :

$$(\forall \varepsilon > 0, \exists \eta > 0 ; \|X - A\| \leq \eta, X \in D) \implies (\|f(X) - L\| \leq \varepsilon).$$

Remarques

- 1 La notion de limite ne dépend pas des normes utilisées.
- 2 La limite, si elle existe, est unique.

Définition

Une fonction f est continue en un point $A \in D$ si la limite de f en ce point existe et est égale à la valeur de la fonction en A .
La fonction est continue sur D si elle est continue en tout point de D .

Définition

Une fonction f est continue en un point $A \in D$ si la limite de f en ce point existe et est égale à la valeur de la fonction en A .
La fonction est continue sur D si elle est continue en tout point de D .

Ou bien on peut reformuler cette définition à l'aide des suites :

Définition

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $A \in D$.

On dit que f est continue en A ssi $\forall (X_n)_{n \in \mathbb{N}}$ suite de D telle que $\lim_{n \rightarrow +\infty} X_n = A$, on a $\lim_{n \rightarrow +\infty} f(X_n) = f(A)$.

Propriétés

La somme, le produit et le quotient (là où le dénominateur ne s'annule pas) des fonctions continues sont continus. La composée de fonctions continues est continue.

Propriétés

La somme, le produit et le quotient (là où le dénominateur ne s'annule pas) des fonctions continues sont continus. La composée de fonctions continues est continue.

Remarque

Toute fonction obtenue à l'aide de fonctions continues élémentaires de variables (x_1, \dots, x_n) en utilisant les opérations algébriques et la composition est continue dans son domaine naturel de définition.

Propriétés

La somme, le produit et le quotient (là où le dénominateur ne s'annule pas) des fonctions continues sont continus. La composée de fonctions continues est continue.

Remarque

Toute fonction obtenue à l'aide de fonctions continues élémentaires de variables (x_1, \dots, x_n) en utilisant les opérations algébriques et la composition est continue dans son domaine naturel de définition.

Exemples

Les fonctions polynômes $x^k y^n$, exponentielles e^{2x+xy} , sont continues sur \mathbb{R}^2 .

Il peut être pratique de fixer toutes les composantes sauf une :

Définition

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$. Soit $X_0 = (x_0^1, \dots, x_0^n) \in D$. Pour $i \in \llbracket 1, n \rrbracket$, on appelle i-ème fonction partielle de f en X_0 la fonction :

$$f_{X_0, i} : \begin{cases} D_i \subset \mathbb{R} \rightarrow \mathbb{R}^p \\ x \mapsto f(x_0^1, \dots, x_0^{i-1}, x, x_0^{i+1}, \dots, x_0^n) \end{cases}$$

où x est à la i -ème place, et D_i est tel que pour $x \in D_i$, $(x_0^1, \dots, x, \dots, x_0^n) \in D$.

Proposition

Si f est continue en $x_0 = (x_0^1, \dots, x_0^n)$ alors $\forall i \in [1, n]$, la fonction partielle $f_{x_0, i}$ est continue en x_0^i .

Remarque

La réciproque est fausse !

Exemple

On considère une fonction $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ définie de la façon suivante

$$f: (x,y) \mapsto \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Exemple : correction

Ses 2 fonctions partielles en $(0,0)$ sont

$$f_{(0,0),1} : x \mapsto \begin{cases} \frac{x \cdot 0}{x^2 + 0} & \text{si } x \neq 0, \\ 0 & \text{si } x = 0, \end{cases}$$

et

$$f_{(0,0),2} : y \mapsto \begin{cases} \frac{0 \cdot y}{0 + y^2} & \text{si } y \neq 0, \\ 0 & \text{si } y = 0. \end{cases}$$

Elles sont donc continues.

Pourtant f n'est pas continue en $(0,0) \dots$

Exemple : correction

Soient $x_n = y_n = 1/n$. On a $\lim_{n \rightarrow +\infty} x_n = \lim_{n \rightarrow +\infty} y_n = 0$, mais

$$f(x_n, y_n) = \frac{1/n^2}{2/n^2} = \frac{1}{2}. \text{ Donc } \lim_{n \rightarrow +\infty} f(x_n, y_n) \neq 0 = f(0, 0).$$

Exemple : correction

Une autre démonstration du fait que f n'est pas continue en $(0,0)$: prenons une restriction de f sur la droite D_1 définie par l'équation $y = x$.

$$\lim_{(x,y) \rightarrow (0,0)} \left(f(x,y) \Big|_{D_1} \right) = \lim_{x \rightarrow 0} \frac{xx}{x^2 + x^2} = \frac{1}{2}.$$

Donc la fonction f restreinte à un sous-ensemble D_1 de \mathbb{R}^2 n'a pas la même limite que la même fonction restreinte à deux autres sous-ensembles de \mathbb{R}^2 . (Les fonctions partielles $f_{(0,0),1}$ et $f_{(0,0),2}$ sont des restrictions de f aux droites $y = 0$ et $x = 0$ respectivement). Or la limite, si elle existe, doit être unique, donc la limite n'existe pas.

Définition

Une partie compacte (un compact) de \mathbb{R}^n est une partie fermée et bornée de \mathbb{R}^n .

Définition

Une partie compacte (un compact) de \mathbb{R}^n est une partie fermée et bornée de \mathbb{R}^n .

Il existe au moins deux façons différentes de définir un compact dans un espace normé, mais dans \mathbb{R}^n elles sont équivalentes à celle qu'on donne ici.

Exemple

Dans \mathbb{R} un intervalle fermé, et dans \mathbb{R}^n les boules fermées sont des exemples de compacts.

Théorème

Soit $f : D \rightarrow \mathbb{R}^p$ une fonction continue sur une partie $D \subset \mathbb{R}^n$ et K une partie compacte de \mathbb{R}^n contenue dans D .
Alors, $f(K)$ est une partie compacte de \mathbb{R}^p .

Théorème

Soit $f : D \rightarrow \mathbb{R}^p$ une fonction continue sur une partie $D \subset \mathbb{R}^n$ et K une partie compacte de \mathbb{R}^n contenue dans D .
Alors, $f(K)$ est une partie compacte de \mathbb{R}^p .

Corollaire

Une fonction continue sur un compact est bornée et atteint ses bornes.

Théorème

Soit $f : D \rightarrow \mathbb{R}^p$ une fonction continue sur une partie $D \subset \mathbb{R}^n$ et K une partie compacte de \mathbb{R}^n contenue dans D .
Alors, $f(K)$ est une partie compacte de \mathbb{R}^p .

Corollaire

Une fonction continue sur un compact est bornée et atteint ses bornes.

Interprétation

Cela signifie que sur un compact $K \subset \mathbb{R}^n$ il existe au moins un point $X_m \in K$ et au moins un point $X_M \in K$ tels que pour tout $X \in K$ on ait

$$\|f(X_m)\| \leq \|f(X)\| \leq \|f(X_M)\|.$$

Rappels

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur un intervalle I de \mathbb{R} .
La dérivée de f au point $a \in I$ est :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

Rappels

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur un intervalle I de \mathbb{R} .
La dérivée de f au point $a \in I$ est :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

Approximation (équation de la tangente)

$$f(x) \simeq f(a) + f'(a)(x - a)$$

Rappels

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur un intervalle I de \mathbb{R} .
La dérivée de f au point $a \in I$ est :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

Approximation (équation de la tangente)

$$f(x) \simeq f(a) + f'(a)(x - a)$$

Élasticité

$$\varepsilon_f(a) \simeq \frac{\frac{\Delta f}{f}}{\frac{\Delta x}{x}} \simeq a \frac{f'(a)}{f(a)}$$

Introduction

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $A \in D$.

Une expression du type “ $\lim_{X \rightarrow A} \frac{f(X) - f(A)}{X - A}$ ” n’est pas bien définie parce que diviser par $X - A$, qui est un vecteur de \mathbb{R}^n , n’a aucun sens !

Néanmoins, si on fixe toutes les composantes de X sauf une, on peut définir des dérivées partielles.

Définition générale

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ et $A \in D$.

Pour $i = \llbracket 1, n \rrbracket$, on appelle dérivée partielle par rapport à x_i de f en $A = (a_1, \dots, a_n)$, et on note $\frac{\partial f}{\partial x_i}(A)$, ou bien $f'_{x_i}(A)$, la dérivée de la fonction partielle $f_{A,i}$ prise en a_i :

$$\frac{\partial f}{\partial x_i}(A) = f'_{A,i}(a_i) = \lim_{x_i \rightarrow a_i} \frac{f(a_1, \dots, x_i, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)}{x_i - a_i}.$$

Définition

Pour une fonction de deux variables $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ en un point $A = (a, b) \in D$. Les dérivées partielles de $f(x, y)$ en (a, b) sont les dérivées des fonctions partielles $f(x, b)$ et $f(a, y)$ qui se calculent alors :

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h}$$

$$\frac{\partial f}{\partial y}(a, b) = \lim_{k \rightarrow 0} \frac{f(a, b + k) - f(a, b)}{k}$$

Parfois, on les note aussi $f'_x(a, b)$ et $f'_y(a, b)$.

Exercice

Soit $f(x,y) = 2x^2 - 3xy + 4y^2$.

Calculer les dérivées partielles au point $(1,2)$.

Exercice

Soit $f(x,y) = 2x^2 - 3xy + 4y^2$.

Calculer les dérivées partielles au point $(1,2)$.

En considérant y constant et en dérivant par rapport à x on a :

$$\frac{\partial f}{\partial x}(x,y) = 4x - 3y \quad \text{et} \quad \frac{\partial f}{\partial x}(1,2) = -2$$

Exercice

Soit $f(x,y) = 2x^2 - 3xy + 4y^2$.

Calculer les dérivées partielles au point $(1,2)$.

En considérant y constant et en dérivant par rapport à x on a :

$$\frac{\partial f}{\partial x}(x,y) = 4x - 3y \quad \text{et} \quad \frac{\partial f}{\partial x}(1,2) = -2$$

En considérant x constant et en dérivant par rapport à y on a :

$$\frac{\partial f}{\partial y}(x,y) = -3x + 8y \quad \text{et} \quad \frac{\partial f}{\partial y}(1,2) = 13$$

Exemple

Déterminer les dérivées partielles premières de la fonction $f(x, y) = 5x^2y^3$.

Exemple

Déterminer les dérivées partielles premières de la fonction $f(x, y) = 5x^2y^3$.

$$\frac{\partial f}{\partial x}(x, y) = 10xy^3$$

$$\frac{\partial f}{\partial y}(x, y) = 15x^2y^2$$

Exemple

Déterminer les dérivées partielles premières de la fonction $f(x, y) = 5x^2y^3$.

$$\frac{\partial f}{\partial x}(x, y) = 10xy^3$$

$$\frac{\partial f}{\partial y}(x, y) = 15x^2y^2$$

Exercice

Déterminer les dérivées partielles premières de la fonction $f(x, y) = 3x^2y^2 + 4xy^3 + 7y$ puis celles de la fonction de production de Cobb-Douglas $g(K, L) = 4K^{\frac{3}{4}}L^{\frac{1}{4}}$.

Définition

Si $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ a dérivées partielles au voisinage de $A = (a, b) \in D$ et si ces dérivées partielles sont continues en (a, b) , alors S_f admet au point $A(a, b, f(a, b))$ un plan tangent dont une équation est :

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

Définition

Si $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ a dérivées partielles au voisinage de $A = (a, b) \in D$ et si ces dérivées partielles sont continues en (a, b) , alors S_f admet au point $A(a, b, f(a, b))$ un plan tangent dont une équation est :

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

Remarques

- ① $\nabla f(a, b) = \left(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b) \right) = \overrightarrow{\text{grad}} f(a, b)$ s'appelle le vecteur gradient de f en (a, b) .
- ② Si h et k sont petits ((x, y) proche de (a, b))

$$f(a + h, b + k) \simeq f(a, b) + \frac{\partial f}{\partial x}(a, b).h + \frac{\partial f}{\partial y}(a, b).k$$

Exercice

Soit $f : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ définie par $f(x, y) = 5x^{\frac{1}{2}}y^{\frac{1}{3}}$.

- 1 Donner les expressions des dérivées partielles de f en x et en y en tout point de $\mathbb{R}_*^+ \times \mathbb{R}_*^+$.
- 2 Déterminer une équation du plan tangent à la surface au point $(1; 1; 5)$.
- 3 En déduire une valeur approchée de $f(1,04; 0,97)$.

Exemple introductif

La fonction de production d'une entreprise est donnée par

$$Q(L, K) = 5KL - K^2 - 3L^2.$$

Le niveau de production obtenu pour $L = 100$ et pour $K = 200$ est $Q(100, 200) = 30000$.

Exemple introductif

La fonction de production d'une entreprise est donnée par $Q(L, K) = 5KL - K^2 - 3L^2$.

Le niveau de production obtenu pour $L = 100$ et pour $K = 200$ est $Q(100, 200) = 30000$.

Si on diminue le niveau de travail de 100 à 99 ($\Delta L = -1$) de combien doit-on faire varier le capital pour maintenir un niveau de production égal à 30000 ?

On peut faire le calcul exact et chercher la valeur de K solution de $Q(99, K) = 30000$ et proche de 200.

Exemple introductif

On doit donc résoudre l'équation $5 \times 99K - K^2 - 3 \times 99^2 = 30000$
ou encore $K^2 - 495K + 30000 + 29403 = 0$
soit $K^2 - 495K + 59403 = 0$ de discriminant 7413.

Exemple introductif

On doit donc résoudre l'équation $5 \times 99K - K^2 - 3 \times 99^2 = 30000$
ou encore $K^2 - 495K + 30000 + 29403 = 0$

soit $K^2 - 495K + 59403 = 0$ de discriminant 7413.

Les racines sont en valeur arrondie 290,55 et 204,45 et la solution proche de 200 est 204,45.

Il faut donc compenser la perte d'une unité de travail par l'augmentation de 4,45 unités de capital ($\Delta K = 4,45$) pour conserver le même niveau de production qu'avec 100 unités de travail et 200 unités de capital.

Exemple introductif

On doit donc résoudre l'équation $5 \times 99K - K^2 - 3 \times 99^2 = 30000$
ou encore $K^2 - 495K + 30000 + 29403 = 0$
soit $K^2 - 495K + 59403 = 0$ de discriminant 7413.

Les racines sont en valeur arrondie 290,55 et 204,45 et la solution proche de 200 est 204,45.

Il faut donc compenser la perte d'une unité de travail par l'augmentation de 4,45 unités de capital ($\Delta K = 4,45$) pour conserver le même niveau de production qu'avec 100 unités de travail et 200 unités de capital.

On va construire une approximation de ce calcul exact grâce aux dérivées partielles ou fonctions marginales.

Exemple introductif

Soit un couple de valeurs (x,y) appartenant à la courbe de niveau k de f c'est à dire $f(x,y) = k$; on fait varier x de x à $x + \Delta x$ et on cherche la variation à imposer alors à y (y passe de y à $y + \Delta y$) pour que le couple $(x + \Delta x, y + \Delta y)$ soit aussi sur la courbe de niveau k ; on a donc :

Exemple introductif

$f(x + \Delta x, y + \Delta y) = f(x, y) = k$ et pour $(\Delta x, \Delta y)$ “faible” on peut utiliser la valeur approchée de $f(x + \Delta x, y + \Delta y)$ donnée par les dérivées partielles en (x, y) et écrire

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + \frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y$$

Exemple introductif

$f(x + \Delta x, y + \Delta y) = f(x, y) = k$ et pour $(\Delta x, \Delta y)$ “faible” on peut utiliser la valeur approchée de $f(x + \Delta x, y + \Delta y)$ donnée par les dérivées partielles en (x, y) et écrire

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + \frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y$$

ce qui impose la relation

$$\frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y = 0$$

$$\text{ou } -\frac{\Delta y}{\Delta x} = \frac{\frac{\partial f}{\partial x}(x, y)}{\frac{\partial f}{\partial y}(x, y)} = \frac{f_{m_x}(x, y)}{f_{m_y}(x, y)}$$

Exemple introductif

$f(x + \Delta x, y + \Delta y) = f(x, y) = k$ et pour $(\Delta x, \Delta y)$ “faible” on peut utiliser la valeur approchée de $f(x + \Delta x, y + \Delta y)$ donnée par les dérivées partielles en (x, y) et écrire

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + \frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y$$

ce qui impose la relation

$$\frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y = 0$$

$$\text{ou } -\frac{\Delta y}{\Delta x} = \frac{\frac{\partial f}{\partial x}(x, y)}{\frac{\partial f}{\partial y}(x, y)} = \frac{f_{m_x}(x, y)}{f_{m_y}(x, y)}$$

cette expression s'appelle le taux marginal de substitution technique de x en y et est notée $\text{TMST}_{x,y}$; elle est positive si les fonctions marginales en (x, y) sont positives et dans ce cas Δx et Δy sont de signe contraire.

À retenir

$$\text{TMST}_{x,y} = -\frac{\Delta y}{\Delta x} = \frac{f_{m_x}(x,y)}{f_{m_y}(x,y)}$$

À retenir

$$\text{TMST}_{x,y} = -\frac{\Delta y}{\Delta x} = \frac{f_{m_x}(x,y)}{f_{m_y}(x,y)}$$

Retour à l'exemple

Le taux marginal de substitution de L en K est

$$\text{TMST}_{L,K} = -\frac{\Delta K}{\Delta L} = \frac{P_{mL}}{P_{mK}} = \frac{5K - 6L}{5L - 2K} \text{ et donc pour } L=100 \text{ et } K=200 \text{ ce taux marginal de substitution de L en K est égal à}$$

$$\frac{5 \times 200 - 6 \times 100}{5 \times 100 - 2 \times 200} = \frac{400}{100} = 4.$$

Retour à l'exemple

Ce qui signifie que si on baisse L de 100 à 99 il faut augmenter K de 200 à $200+4=204$ qui est une valeur approchée de 204,45 et si on calcule

$Q(99, 204) = 5 \times 204 \times 99 - 204^2 - 3 \times 99^2 = 29961$ le niveau 30000 est approximativement maintenu (erreur de $39/30000$)

(pour 204,45 le calcul donne

$Q(99, 204.45) = 5 \times 204,45 \times 99 - 204,45^2 - 3 \times 99^2 = 29999.95$).

Retour à l'exemple

Ce qui signifie que si on baisse L de 100 à 99 il faut augmenter K de 200 à $200+4=204$ qui est une valeur approchée de 204,45 et si on calcule

$Q(99, 204) = 5 \times 204 \times 99 - 204^2 - 3 \times 99^2 = 29961$ le niveau 30000 est approximativement maintenu (erreur de 39/30000)

(pour 204,45 le calcul donne

$Q(99, 204.45) = 5 \times 204,45 \times 99 - 204,45^2 - 3 \times 99^2 = 29999.95$).

L'avantage est que ce calcul est valable aussi pour L passant de 100 à 101 et donc K doit passer de 200 à 196 de même pour d'autres variations faibles de L autour de 100 induisant des variations de K autour de 200 pour rester au même niveau de production.

Introduction

Si on connaît les variations relatives de chacune des variables x et y on peut se demander quelle variation relative il en résulte pour $f(x,y)$.

Introduction

Si on connaît les variations relatives de chacune des variables x et y on peut se demander quelle variation relative il en résulte pour $f(x, y)$. Puisque

$$f(x + \Delta x, y + \Delta y) \approx f(x, y) + \frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y$$

on en déduit

Introduction

$$\frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{f(x, y)} \approx \frac{\frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y}{f(x, y)}$$

Introduction

$$\frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{f(x, y)} \approx \frac{\frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y}{f(x, y)}$$

$$\text{ou } \frac{\Delta f}{f} \approx \frac{\frac{\partial f}{\partial x}(x, y)}{f(x, y)} \Delta x + \frac{\frac{\partial f}{\partial y}(x, y)}{f(x, y)} \Delta y$$

Introduction

$$\frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{f(x, y)} \approx \frac{\frac{\partial f}{\partial x}(x, y) \times \Delta x + \frac{\partial f}{\partial y}(x, y) \times \Delta y}{f(x, y)}$$

$$\text{ou } \frac{\Delta f}{f} \approx \frac{\frac{\partial f}{\partial x}(x, y)}{f(x, y)} \Delta x + \frac{\frac{\partial f}{\partial y}(x, y)}{f(x, y)} \Delta y$$

$$\text{ou } \frac{\Delta f}{f} \approx x \frac{\frac{\partial f}{\partial x}(x, y)}{f(x, y)} \frac{\Delta x}{x} + y \frac{\frac{\partial f}{\partial y}(x, y)}{f(x, y)} \frac{\Delta y}{y}$$

Introduction

En définissant les élasticité partielles de f par rapport à x et

par rapport à y comme $x \frac{\frac{\partial f}{\partial x}(x,y)}{f(x,y)}$ notée $\epsilon_{f/x}$ et $y \frac{\frac{\partial f}{\partial y}(x,y)}{f(x,y)}$

notée $\epsilon_{f/y}$ on peut en déduire :

Introduction

En définissant les élasticités partielles de f par rapport à x et

par rapport à y comme $x \frac{\frac{\partial f}{\partial x}(x,y)}{f(x,y)}$ notée $\varepsilon_{f/x}$ et $y \frac{\frac{\partial f}{\partial y}(x,y)}{f(x,y)}$

notée $\varepsilon_{f/y}$ on peut en déduire :

$$\frac{\Delta f}{f} \approx \varepsilon_{f/x} \frac{\Delta x}{x} + \varepsilon_{f/y} \frac{\Delta y}{y}$$

Exercice

Soit f définie sur $\mathbb{R}^+ \times \mathbb{R}^+$ par $f(x, y) = 3x^{\frac{1}{4}}y^{\frac{1}{3}}$.

À partir d'une valeur (x_0, y_0) , on augmente x de 3% et on diminue y de 2%.

Quelle est la variation en pourcentage de f ?

Exercice

Soit f définie sur $\mathbb{R}^+ \times \mathbb{R}^+$ par $f(x, y) = 3x^{\frac{1}{4}}y^{\frac{1}{3}}$.

À partir d'une valeur (x_0, y_0) , on augmente x de 3% et on diminue y de 2%.

Quelle est la variation en pourcentage de f ?

Correction

$$\varepsilon_{f/x} = \frac{1}{4} \text{ et } \varepsilon_{f/y} = \frac{1}{3}$$

$$\frac{\Delta f}{f} \approx \varepsilon_{f/x} \frac{\Delta x}{x} + \varepsilon_{f/y} \frac{\Delta y}{y}$$

$$\frac{\Delta f}{f} \approx \frac{1}{4} \times \frac{3}{100} + \frac{1}{3} \times \left(-\frac{2}{100} \right)$$

Exercice

Soit f définie sur $\mathbb{R}^+ \times \mathbb{R}^+$ par $f(x,y) = 3x^{\frac{1}{4}}y^{\frac{1}{3}}$.

À partir d'une valeur (x_0, y_0) , on augmente x de 3% et on diminue y de 2%.

Quelle est la variation en pourcentage de f ?

Correction

$$\varepsilon_{f/x} = \frac{1}{4} \text{ et } \varepsilon_{f/y} = \frac{1}{3}$$

$$\frac{\Delta f}{f} \approx \varepsilon_{f/x} \frac{\Delta x}{x} + \varepsilon_{f/y} \frac{\Delta y}{y}$$

$$\frac{\Delta f}{f} \approx \frac{1}{4} \times \frac{3}{100} + \frac{1}{3} \times \left(-\frac{2}{100}\right)$$

$$\frac{\Delta f}{f} \approx \frac{1}{1200} \text{ soit } \frac{\Delta f}{f} \approx 0,08\%$$

On définit de même les dérivées partielles des fonctions de n variables de la façon suivant :

Proposition

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de n variables qui admet des dérivées partielles dans un voisinage de $X^* = (x_1^*, \dots, x_n^*)$, continues en X^* .

Alors si $X = (x_1, \dots, x_n)$ est assez près de X^* , on a :

$$f(X) = f(X^*) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X^*) \cdot (X - X^*)_i + \|X - X^*\| \varepsilon(X)$$

où $\lim_{X \rightarrow X^*} \varepsilon(X) = 0$.

Ou encore...

Proposition

$$f(x_1^* + h_1, \dots, x_n^* + h_n) = f(X^*) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X^*) h_i + \|(h_1, \dots, h_n)\| \varepsilon(h)$$

où $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

$$x_{n+1} = f(x_1, \dots, x_n) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X^*) \cdot (X - X^*)$$

est l'équation de l'hyperplan tangent à la surface.

Définition

Soit $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe C^1 . Son gradient, pris en tout point de D définit une fonction à valeurs vectorielles $\vec{\text{grad}}f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$, noté aussi :

$$\vec{\nabla}f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) (x,y).$$

Théorème

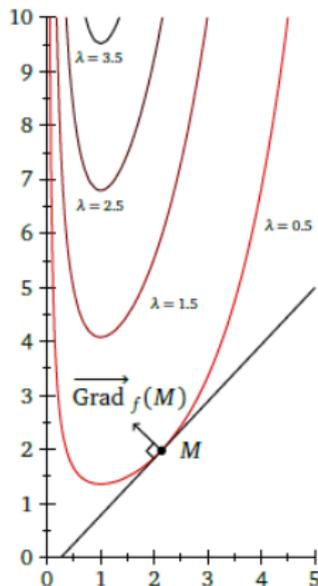
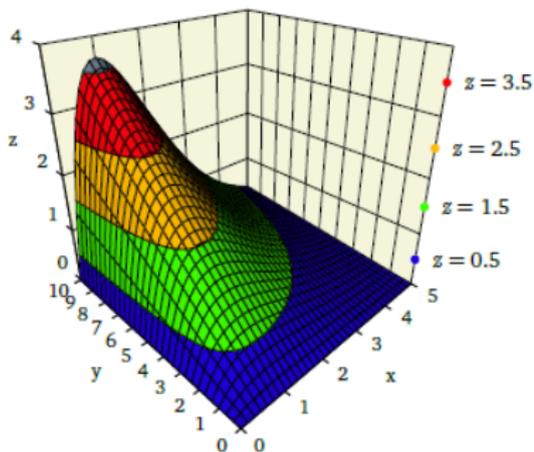
Soient $D \subset \mathbb{R}^2$, $f : D \rightarrow \mathbb{R}$ et $(x, y) \in D$, alors si $f(x, y) = \lambda$, (x, y) appartient à la courbe de niveau $C_\lambda(f)$.

Le vecteur gradient $\vec{\nabla} f(x, y)$ est normal à la courbe $C_\lambda(f)$ au point (x, y) .

Théorème

Soient $D \subset \mathbb{R}^2$, $f: D \rightarrow \mathbb{R}$ et $(x,y) \in D$, alors si $f(x,y) = \lambda$, (x,y) appartient à la courbe de niveau $C_\lambda(f)$.

Le vecteur gradient $\vec{\nabla}f(x,y)$ est normal à la courbe $C_\lambda(f)$ au point (x,y) .



Exemple 1

$f(x,y) = x^2 + y^2$. $C_\lambda(f) = C((0,0), \sqrt{\lambda})$ - cercle de centre $(0,0)$ et de rayon $\sqrt{\lambda}$. $\frac{\partial f}{\partial x}(x,y) = 2x$, $\frac{\partial f}{\partial y}(x,y) = 2y$. On remarque que $(2x, 2y) = 2(x,y)$ est 2 fois le vecteur radial qui est en effet orthogonal au cercle.

Exemple 2

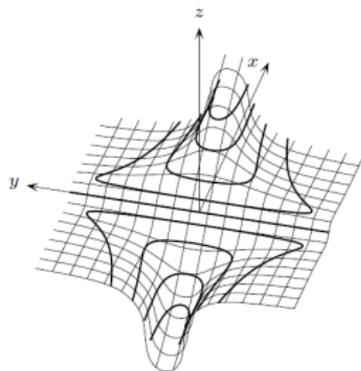
Soit la courbe d'équation $x^2 - y = 0$. Pour calculer la normale en chaque point de cette courbe, on la voit comme une ligne de niveau 0 de la fonction $f(x,y) = x^2 - y$. La normale est donc donnée par son gradient : $\vec{\nabla} f(x,y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix}$.

Théorème

Le gradient en (x,y) indique la direction de plus grande pente sur Γ_f (à partir du point (x,y)).

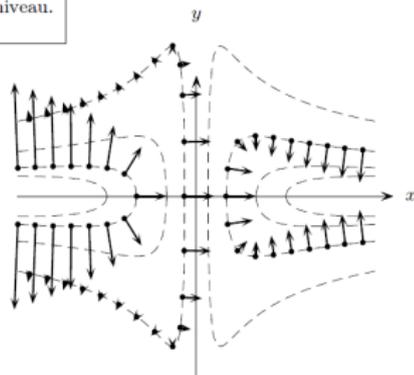
Théorème

Le gradient en (x,y) indique la direction de plus grande pente sur Γ_f (à partir du point (x,y)).



Le gradient est orthogonal aux lignes de niveau.

On voit de plus que le gradient est dirigé dans le sens des pentes croissantes, et que plus la pente est forte, plus le gradient est grand en norme.



Preuve

$$f((x, y) + \vec{v}) - f(x, y) = \vec{\nabla} f(x, y) \cdot \vec{v} + o(\|\vec{v}\|)$$

Le produit scalaire $\vec{\nabla} f(x, y) \cdot \vec{v}$ vaut $\|\vec{\nabla} f(x, y)\| \times \|\vec{v}\| \times \cos \theta$, où θ est l'angle entre les deux vecteurs. L'accroissement de la fonction atteint le maximum quand $\cos \theta = 1$, alors \vec{v} doit être parallèle à $\vec{\nabla} f(x, y)$.

Remarque

En suivant la ligne de plus grande pente dans D on a, sur le graphe, le chemin le plus court à parcourir pour obtenir une variation donnée de f . Autrement dit, si on veut passer le plus vite possible du niveau a au niveau b à partir d'un point (x,y) donné de niveau $a = f(x,y)$, il faut suivre le gradient.

Définition

Soit $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Les dérivées partielles définissent n nouvelles fonctions

$$f'_{x_i}(x_1, \dots, x_n) = \frac{\partial f}{\partial x_i}(x_1, \dots, x_n).$$

Définition

Soit $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Les dérivées partielles définissent n nouvelles fonctions

$$f'_{x_i}(x_1, \dots, x_n) = \frac{\partial f}{\partial x_i}(x_1, \dots, x_n).$$

On peut regarder les dérivées partielles de chacune de ces nouvelles fonctions. Cela nous donne les dérivées partielles d'ordre 2 (aussi appelées les dérivées partielles secondes) et à leur tour on peut regarder les dérivées partielles des dérivées partielles d'ordre 2, etc. Cela s'écrit par exemple :

$$f''_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)$$

Définition

Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^k sur un ouvert Ω de \mathbb{R}^n est une fonction dont toutes les dérivées partielles jusqu'à l'ordre k existent et sont continues sur cet ouvert Ω .

Une fonction est dite de classe C^∞ si elle est de classe C^k pour tout $k \in \mathbb{N}$.

Theorème de Schwarz

Soit $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 sur D . Les fonctions de dérivées partielles d'ordre 2,

$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)$ et $\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$ sont égales en tout point de D .

Remarque

Le théorème de Schwarz implique que les dérivées partielles d'ordre k , $k \geq 2$, d'une fonction de classe C^k , $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ne dépendent pas de l'ordre dans lequel les dérivées partielles sont prises. Par exemple, pour une fonction de deux variables $f(x, y)$

de classe C^3 , on a :
$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y}.$$

Exercice

Soit $f(x,y) = 3x^{\frac{1}{3}}y^{\frac{1}{4}}$, $x > 0$, $y > 0$.

Déterminer les dérivées partielles d'ordre 2 de f .

Théorème

Soit f de classe C^1 autour de (x_0, y_0) .

On suppose que $\overrightarrow{\text{grad}f}(x_0, y_0) \neq (0, 0)$.

Alors la courbe de niveau de f passant par (x_0, y_0) a une tangente (T) d'équation :

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0$$

En particulier, $\overrightarrow{\text{grad}f}(x_0, y_0)$ est normal à cette tangente.

Théorème

Soit f de classe C^1 autour de (x_0, y_0) .

On suppose que $\overrightarrow{\text{grad}f}(x_0, y_0) \neq (0, 0)$.

Alors la courbe de niveau de f passant par (x_0, y_0) a une tangente (T) d'équation :

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0$$

En particulier, $\overrightarrow{\text{grad}f}(x_0, y_0)$ est normal à cette tangente.

Définition

Un point où $\overrightarrow{\text{grad}f}(x_0, y_0) \neq (0, 0)$ s'appelle un point régulier.

Si $\overrightarrow{\text{grad}f}(x_0, y_0) = (0, 0)$ alors (x_0, y_0) est un point critique (ou stationnaire).

Définition

Soit f une fonction de $\mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 sur un ouvert Ω .
On appelle matrice Hessienne de f en X la matrice symétrique

$$H_f(X) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(X) \right)$$

Définition

Soit f une fonction de $\mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 sur un ouvert Ω .
On appelle matrice Hessienne de f en X la matrice symétrique

$$H_f(X) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(X) \right)$$

Rappel

Soit S une matrice symétrique.

Alors $(x_1, x_2, \dots, x_n) \mapsto (x_1 \ x_2 \ \dots \ x_n) S^t (x_1 \ x_2 \ \dots \ x_n) = X S^t X$ est une forme quadratique (on dit que c'est la forme quadratique associée à S).

Exercice

On considère une fonction de Cobb-Douglas de deux variables définie par

$$f(x,y) = x^\alpha y^\beta \quad \text{pour } x > 0 \text{ et } y > 0$$

- 1 Calculer la matrice hessienne de f au point (a,b) .

On la note $H_f(a,b)$.

- 2 Montrer que la forme quadratique définie par $H_f(a,b)$ est :

$$q(x,y) = a^{\alpha-2} b^{\beta-2} [\alpha(\alpha-1)b^2 x^2 + \beta(\beta-1)a^2 y^2 + 2\alpha\beta abxy]$$

Définition : petit o...

Soient f et g deux fonctions d'une variable à valeurs réelles.
On dit que $g = o(f)$ au point a si :

$$\lim_{x \rightarrow a} \frac{g(x)}{f(x)} = 0$$

Définition : petit o...

Soient f et g deux fonctions d'une variable à valeurs réelles.
On dit que $g = o(f)$ au point a si :

$$\lim_{x \rightarrow a} \frac{g(x)}{f(x)} = 0$$

Exemple

$u(x) = x^3, v(x) = x^2 + 2x$. En $a = 0$ on a $u(x) = o(v(x))$ et en $a = +\infty$ on a $v(x) = o(u(x))$.

Théorème : Formule de Taylor à l'ordre 2

Soit f une fonction de classe C^2 au voisinage de $X_0 \in \mathbb{R}^n$. Alors, pour $Y \in \mathbb{R}^n$ de norme assez petite, on a :

$$f(X_0 + Y) = f(X_0) + Df_{X_0}(Y) + \frac{1}{2} Y H_f(X_0) Y + \|Y\|^2 \varepsilon(Y)$$

$$\text{où } \lim_{Y \rightarrow 0} \varepsilon(Y) = 0 \text{ et } Df_{X_0}(Y) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X_0) \times y_i = \vec{\nabla} f(X_0) \cdot Y$$

Théorème : Formule de Taylor à l'ordre 2

Soit f une fonction de classe C^2 au voisinage de $X_0 \in \mathbb{R}^n$. Alors, pour $Y \in \mathbb{R}^n$ de norme assez petite, on a :

$$f(X_0 + Y) = f(X_0) + Df_{X_0}(Y) + \frac{1}{2} Y H_f(X_0) Y + \|Y\|^2 \varepsilon(Y)$$

$$\text{où } \lim_{Y \rightarrow 0} \varepsilon(Y) = 0 \text{ et } Df_{X_0}(Y) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X_0) \times y_i = \vec{\nabla} f(X_0) \cdot Y$$

Remarque

Cette formule de Taylor donne l'approximation de f au voisinage de X_0 par un polynôme de degré 2. Ce qui suit permet d'approcher f au voisinage de X_0 par des polynômes de degré supérieur.

Définition

Soit f une fonction de \mathbb{R}^n dans \mathbb{R} de classe C^k sur un ouvert U de \mathbb{R}^n . Pour $l \leq k$, on pose :

$$\begin{aligned} D^l f_X(Y, \dots, Y) &= \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_l=1}^n \frac{\partial^l f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_l}}(X) \cdot y_{i_1} y_{i_2} \dots y_{i_l} \\ &= \sum_{1 \leq i_1, i_2, \dots, i_l \leq n} \frac{\partial^l f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_l}}(X) \cdot y_{i_1} y_{i_2} \dots y_{i_l} \end{aligned}$$

où $Y = (y_1, \dots, y_n)$.

On peut montrer que :

$$D^l f_X(Y, \dots, Y) = \sum_{\substack{\alpha_1, \dots, \alpha_n \in \mathbb{N} \\ \alpha_1 + \alpha_2 + \dots + \alpha_n = l}} \frac{l!}{\alpha_1! \alpha_2! \dots \alpha_n!} \frac{\partial^l f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(X) \cdot y_1^{\alpha_1} y_2^{\alpha_2} \dots y_n^{\alpha_n}$$

On peut montrer que :

$$D^1 f_X(Y, \dots, Y) = \sum_{\substack{\alpha_1, \dots, \alpha_n \in \mathbb{N} \\ \alpha_1 + \alpha_2 + \dots + \alpha_n = 1}} \frac{1!}{\alpha_1! \alpha_2! \dots \alpha_n!} \frac{\partial^1 f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(X) \cdot y_1^{\alpha_1} y_2^{\alpha_2} \dots y_n^{\alpha_n}$$

Exemple

$$\begin{aligned} D^2 f_X(Y, Y) &= \sum_{l_1 + l_2 + \dots + l_n = 2} \frac{\partial^2 f}{\partial x_1^{l_1} \partial x_2^{l_2} \dots \partial x_n^{l_n}}(X) \cdot y_1^{l_1} y_2^{l_2} \dots y_n^{l_n} \\ &= \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X) y_i y_j \\ &= {}^t Y H_f(X) Y \end{aligned}$$

Proposition

Soit f une fonction de \mathbb{R}^n dans \mathbb{R} de classe C^k sur un voisinage de X_0 . Alors, pour Y de norme assez petite, on a :

$$f(X_0 + Y) = f(X_0) + Df_{X_0}(Y) + \frac{1}{2}D^2f_{X_0}(Y, Y) + \dots + \frac{1}{k!}D^kf_{X_0}(Y, \dots, Y) + \|Y\|^k \varepsilon(Y)$$

où $\lim_{Y \rightarrow 0} \varepsilon(Y) = 0$.

Une fonction F de \mathbb{R}^n dans \mathbb{R}^m sera notée :

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$X = (x_1, \dots, x_n) \rightarrow F(X) = (f_1(X), \dots, f_m(X))$$

Où les f_i sont des fonctions de \mathbb{R}^n dans \mathbb{R} que l'on appelle les fonctions coordonnées de F .

Définition

On dit qu'une fonction F de \mathbb{R}^n dans \mathbb{R}^m , définie sur un voisinage de X_0 est différentiable en X_0 si ses fonctions coordonnées sont différentiables en X_0 .

Définition

On dit qu'une fonction F de \mathbb{R}^n dans \mathbb{R}^m , définie sur un voisinage de X_0 est différentiable en X_0 si ses fonctions coordonnées sont différentiables en X_0 .

On peut donc écrire, pour tout $i = 1, \dots, m$ et pour X assez proche de X_0 :

$$f_i(X) = f_i(X_0) + Df_{iX_0}(X - X_0) + \|X - X_0\| \varepsilon_i(X)$$
$$\text{où } \lim_{X \rightarrow X_0} \varepsilon_i(X) = 0$$

On peut alors définir une application linéaire de \mathbb{R}^n dans \mathbb{R}^m , notée DF_{X_0} par :

$$DF_{X_0}(Y) = (Df_{1X_0}(Y), \dots, Df_{mX_0}(Y))$$

DF_{X_0} s'appelle la différentielle de F en X_0 . On peut alors écrire :

$$F(X) = F(X_0) + DF_{X_0}(X - X_0) + \|X - X_0\| \varepsilon(X)$$

où $\lim_{X \rightarrow X_0} \varepsilon(X) = 0$

On peut alors définir une application linéaire de \mathbb{R}^n dans \mathbb{R}^m , notée DF_{X_0} par :

$$DF_{X_0}(Y) = (Df_{1X_0}(Y), \dots, Df_{mX_0}(Y))$$

DF_{X_0} s'appelle la différentielle de F en X_0 . On peut alors écrire :

$$F(X) = F(X_0) + DF_{X_0}(X - X_0) + \|X - X_0\| \varepsilon(X)$$

où $\lim_{X \rightarrow X_0} \varepsilon(X) = 0$

Cette égalité donne encore une approximation de F au voisinage de X_0 par une application linéaire. Le fait de travailler en dimension finie va permettre de traduire cette égalité matriciellement.

Supposons que les $\frac{\partial f_i}{\partial x_j}(X_0)$ existent et soient continues en X_0 pour tout $i = 1, \dots, m$ et tout $j = 1, \dots, n$ (on dit que F est de classe C^1). On sait alors que la différentielle Df_{iX_0} est définie par :

Supposons que les $\frac{\partial f_i}{\partial x_j}(X_0)$ existent et soient continues en X_0 pour tout $i = 1, \dots, m$ et tout $j = 1, \dots, n$ (on dit que F est de classe C^1). On sait alors que la différentielle Df_{iX_0} est définie par :

$$Df_{iX_0}(y_1, \dots, y_n) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(X_0) y_j$$

Donc chaque composante de $DF_{X_0}(Y)$ dans la base canonique de \mathbb{R}^m sera : $\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(X_0)y_j$, ou encore que le vecteur colonne de ses coordonnées dans la base canonique est :

Donc chaque composante de $DF_{X_0}(Y)$ dans la base canonique de \mathbb{R}^m sera : $\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(X_0)y_j$, ou encore que le vecteur colonne de ses coordonnées dans la base canonique est :

$$DF_{X_0}(Y) \begin{pmatrix} \sum_{j=1}^n \frac{\partial f_1}{\partial x_j}(X_0)y_j \\ \vdots \\ \sum_{j=1}^n \frac{\partial f_m}{\partial x_j}(X_0)y_j \end{pmatrix}$$

On a alors l'égalité matricielle :

$$\begin{pmatrix} \sum_{j=1}^n \frac{\partial f_1}{\partial x_j}(X_0) y_j \\ \vdots \\ \sum_{j=1}^n \frac{\partial f_m}{\partial x_j}(X_0) y_j \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X_0) & \cdots & \frac{\partial f_1}{\partial x_n}(X_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(X_0) & \cdots & \frac{\partial f_m}{\partial x_n}(X_0) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Ce qui prouve que la matrice de DF_{X_0} dans les bases canoniques de \mathbb{R}^n et \mathbb{R}^m est :

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X_0) & \cdots & \frac{\partial f_1}{\partial x_n}(X_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(X_0) & \cdots & \frac{\partial f_m}{\partial x_n}(X_0) \end{pmatrix}$$

Ce qui prouve que la matrice de DF_{X_0} dans les bases canoniques de \mathbb{R}^n et \mathbb{R}^m est :

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X_0) & \cdots & \frac{\partial f_1}{\partial x_n}(X_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(X_0) & \cdots & \frac{\partial f_m}{\partial x_n}(X_0) \end{pmatrix}$$

Cette matrice s'appelle la jacobienne de F en X_0 et se note en général $J(F)_{X_0}$. C'est une matrice $m \times n$. Si $m = n$, son déterminant s'appelle le jacobien de F en X_0 (noté $|J(F)_{X_0}|$).

Introduction

Le jacobien sert de test à la dépendance fonctionnelle d'un système d'équations linéaires et non linéaires. C'est ainsi, si $|J| = 0$ alors les équations sont fonctionnellement dépendantes ; si $|J| \neq 0$ alors les équations sont fonctionnellement indépendantes.

Exemple

Soit le système linéaire suivant :

$$(S) \begin{cases} f_1(x_1, x_2) = 5x_1 + 3x_2 \\ f_2(x_1, x_2) = 25x_1^2 + 30x_1x_2 + 9x_2^2 \end{cases}$$

Testons la dépendance fonctionnelle du système linéaire S.

Exemple

Soit le système linéaire suivant :

$$(S) \begin{cases} f_1(x_1, x_2) = 5x_1 + 3x_2 \\ f_2(x_1, x_2) = 25x_1^2 + 30x_1x_2 + 9x_2^2 \end{cases}$$

Testons la dépendance fonctionnelle du système linéaire S.
Le calcul du jacobien donne :

$$|J| = \begin{vmatrix} 5 & 3 \\ 50x_1 + 30x_2 & 30x_1 + 18x_2 \end{vmatrix}$$

Exemple

Soit le système linéaire suivant :

$$(S) \begin{cases} f_1(x_1, x_2) = 5x_1 + 3x_2 \\ f_2(x_1, x_2) = 25x_1^2 + 30x_1x_2 + 9x_2^2 \end{cases}$$

Testons la dépendance fonctionnelle du système linéaire S.
Le calcul du jacobien donne :

$$|J| = \begin{vmatrix} 5 & 3 \\ 50x_1 + 30x_2 & 30x_1 + 18x_2 \end{vmatrix}$$

$$|J| = 150x_1 + 90x_2 - 150x_1 - 90x_2 = 0$$

Exemple

Soit le système linéaire suivant :

$$(S) \begin{cases} f_1(x_1, x_2) = 5x_1 + 3x_2 \\ f_2(x_1, x_2) = 25x_1^2 + 30x_1x_2 + 9x_2^2 \end{cases}$$

Testons la dépendance fonctionnelle du système linéaire S.
Le calcul du jacobien donne :

$$|J| = \begin{vmatrix} 5 & 3 \\ 50x_1 + 30x_2 & 30x_1 + 18x_2 \end{vmatrix}$$

$$|J| = 150x_1 + 90x_2 - 150x_1 - 90x_2 = 0$$

Comme la valeur de $|J| = 0$, il y a une dépendance fonctionnelle entre f_1 et f_2 , en effet : $(5x_1 + 3x_2)^2 = f_2$.

Théorème (Cas de deux variables)

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 sur un ouvert contenant $X_0 = (x_0, y_0)$. On pose $k = f(X_0)$. On suppose que $\frac{\partial f}{\partial y}(X_0) \neq 0$. Il existe alors un voisinage de x_0 ($]x_0 - \alpha, x_0 + \alpha[$, $\alpha > 0$) et une fonction $x \rightarrow y(x)$ de \mathbb{R} dans \mathbb{R} , de classe C^1 sur ce voisinage,

$$\text{tels que : } \left\{ \begin{array}{l} \blacklozenge \forall x \in]x_0 - \alpha, x_0 + \alpha[, f(x, y(x)) = k \\ \blacklozenge y(x_0) = y_0 \\ \blacklozenge y'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)} \end{array} \right.$$

Les fonctions implicites

Ce théorème signifie que la courbe de niveau k de f peut être considérée localement (c'est à dire pour X proche de X_0) comme la courbe représentative d'une fonction d'une variable et la dernière formule permet de calculer la dérivée de cette fonction en x_0 .

Fonction réciproque

Le théorème suivant s'obtient en appliquant le théorème des fonctions implicites à la fonction $f(x) - y$.

Le théorème suivant s'obtient en appliquant le théorème des fonctions implicites à la fonction $f(x) - y$.

Théorème

Soit f une fonction de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, de classe C^1 sur un voisinage de x_0 , et dont la différentielle en x_0 , Df_{x_0} , est bijective. Il existe alors un voisinage ouvert U de x_0 et un voisinage ouvert V de $y_0 = f(x_0)$ tels que f soit une bijection de U sur V . Si on note f^{-1} la bijection réciproque (définie sur V), f^{-1} est de classe C^1 sur V et $Df_y^{-1} = (Df_x)^{-1} \quad \forall x \in U \text{ et } y = f(x)$.

Fonctions homogènes

Définition

Soit $f : (\mathbb{R}_+^*)^n \rightarrow \mathbb{R}$. Soit $r \in \mathbb{R}$.

On dit que f est homogène de degré r si :

$$\forall t \in \mathbb{R}_+^*, f(tX) = t^r f(X), \text{ pour tout } X \in (\mathbb{R}_+^*)^n$$

Définition

Soit $f : (\mathbb{R}_+^*)^n \rightarrow \mathbb{R}$. Soit $r \in \mathbb{R}$.

On dit que f est homogène de degré r si :

$$\forall t \in \mathbb{R}_+^*, f(tX) = t^r f(X), \text{ pour tout } X \in (\mathbb{R}_+^*)^n$$

Exemple

$$f(x, y, z) = kx^\alpha y^\beta z^\gamma \text{ (Cobb-Douglas à 3 variables)}$$

$$f(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz \text{ (quadratique)}$$

Fonctions homogènes

Définition

Soit $f : (\mathbb{R}_+^*)^n \rightarrow \mathbb{R}$. Soit $r \in \mathbb{R}$.

On dit que f est homogène de degré r si :

$$\forall t \in \mathbb{R}_+^*, f(tX) = t^r f(X), \text{ pour tout } X \in (\mathbb{R}_+^*)^n$$

Exemple

$f(x, y, z) = kx^\alpha y^\beta z^\gamma$ (Cobb-Douglas à 3 variables)

$f(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$ (quadratique)



Attention

$$f(x, y) = \sqrt{25 - x^2 - y^2}$$

f n'est pas homogène.

Définition

Si f est une fonction de production de n variables : $f(x_1, \dots, x_n)$
 n inputs pour 1 output

Définition

Si f est une fonction de production de n variables : $f(x_1, \dots, x_n)$
 n inputs pour 1 output

- ① Si f est homogène de degré 1 alors

$$f(tx_1, \dots, tx_n) = tf(x_1, \dots, x_n).$$

Si on double tous les inputs alors la production double.
On dit qu'une telle fonction est à rendements d'échelle constants.

Définition

Si f est une fonction de production de n variables : $f(x_1, \dots, x_n)$
 n inputs pour 1 output

- 1 Si f est homogène de degré 1 alors
 $f(tx_1, \dots, tx_n) = tf(x_1, \dots, x_n)$.
Si on double tous les inputs alors la production double.
On dit qu'une telle fonction est à rendements d'échelle constants.
- 2 Si f est homogène de degré r avec $0 < r < 1$ alors f est dite à rendements d'échelle décroissants.

Définition

Si f est une fonction de production de n variables : $f(x_1, \dots, x_n)$
 n inputs pour 1 output

- 1 Si f est homogène de degré 1 alors
 $f(tx_1, \dots, tx_n) = tf(x_1, \dots, x_n)$.
Si on double tous les inputs alors la production double.
On dit qu'une telle fonction est à rendements d'échelle constants.
- 2 Si f est homogène de degré r avec $0 < r < 1$ alors f est dite à rendements d'échelle décroissants.
- 3 Si f est homogène de degré r avec $r > 1$ alors f est dite à rendements d'échelle croissants.

Propriété

Une fonction de Cobb-Douglas $f(x_1, \dots, x_n) = kx_1^{\alpha_1} \dots x_n^{\alpha_n}$ est homogène de degré $\alpha_1 + \dots + \alpha_n$.

Propriété

Une fonction de Cobb-Douglas $f(x_1, \dots, x_n) = kx_1^{\alpha_1} \dots x_n^{\alpha_n}$ est homogène de degré $\alpha_1 + \dots + \alpha_n$.

Théorème

Si f est de classe C^1 et si f est homogène de degré r , alors ses dérivées partielles sont homogènes de degré $r - 1$.

Propriété

Une fonction de Cobb-Douglas $f(x_1, \dots, x_n) = kx_1^{\alpha_1} \dots x_n^{\alpha_n}$ est homogène de degré $\alpha_1 + \dots + \alpha_n$.

Théorème

Si f est de classe C^1 et si f est homogène de degré r , alors ses dérivées partielles sont homogènes de degré $r - 1$.

Exercice

Preuves de la propriété et du théorème...

Théorème d'Euler

Soit f une fonction de classe C^1 sur $(\mathbb{R}_+^*)^n$ et homogène de degré r .

Alors, pour tout $X \in (\mathbb{R}_+^*)^n$:

$$x_1 \frac{\partial f}{\partial x_1}(X) + \cdots + x_n \frac{\partial f}{\partial x_n}(X) = r \times f(X) = X \cdot \nabla f(X)$$

Ce qui s'écrit :

$$\sum_{i=1}^n \varepsilon_{f/x_i}(X) = r.$$

Théorème d'Euler

Soit f une fonction de classe C^1 sur $(\mathbb{R}_+^*)^n$ et homogène de degré r .

Alors, pour tout $X \in (\mathbb{R}_+^*)^n$:

$$x_1 \frac{\partial f}{\partial x_1}(X) + \cdots + x_n \frac{\partial f}{\partial x_n}(X) = r \times f(X) = X \cdot \nabla f(X)$$

Ce qui s'écrit :

$$\sum_{i=1}^n \varepsilon_{f/x_i}(X) = r.$$

Exercice

Preuve du théorème...

Exercice

Soit f la fonction de production définie sur $D' = (\mathbb{R}_+^*)^2$ par

$$f(x, y) = x^{1/3} \sqrt{y}$$

- 1 Montrer que f est homogène et déterminer son degré d'homogénéité.
- 2 Si on double les quantités x et y , comment varie la production de f ?
- 3 Montrer que f vérifie l'égalité d'Euler.

Réciproque du Théorème d'Euler

Si f est une fonction de classe C^1 sur $(\mathbb{R}_+^*)^n$ et si

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(\mathbf{X}) = r \times f(\mathbf{X})$$

pour tout $\mathbf{X} \in (\mathbb{R}_+^*)^n$. Alors, f est homogène de degré r .

Définition

La fonction f est dite convexe si

$$\forall (x,y) \in I^2, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Elle est dite concave si $-f$ est convexe.

Définition

La fonction f est dite convexe si

$$\forall (x,y) \in I^2, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Elle est dite concave si $-f$ est convexe.

Définition bis

On dit que f est convexe (resp. concave) sur un intervalle I si pour tous points A et B de la courbe représentant f , l'arc de courbe est situé au-dessous (au-dessus) du segment $[AB]$.

Définition

La fonction f est dite convexe si

$$\forall (x, y) \in I^2, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Elle est dite concave si $-f$ est convexe.

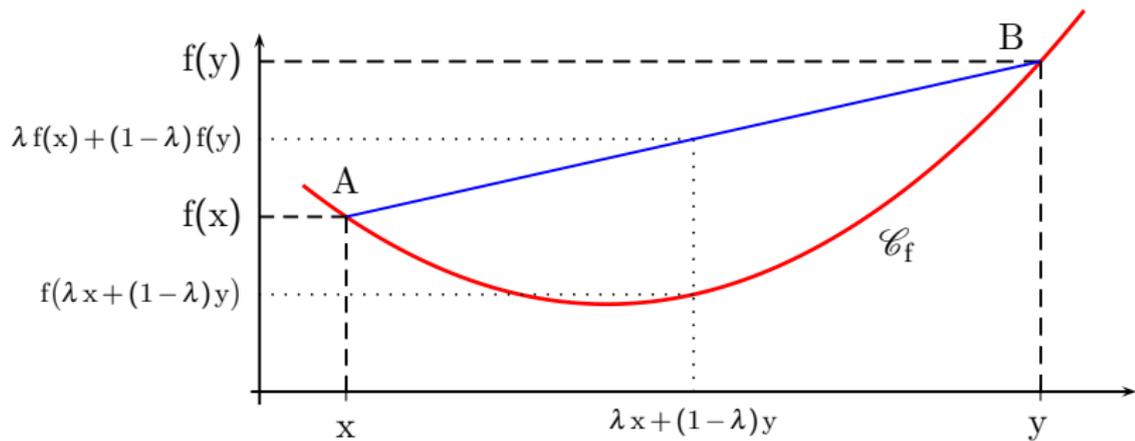
Définition bis

On dit que f est convexe (resp. concave) sur un intervalle I si pour tous points A et B de la courbe représentant f , l'arc de courbe est situé au-dessous (au-dessus) du segment $[AB]$.

Exemples

Les fonctions $x \mapsto x^2$, $x \mapsto |x|$ sont convexes.

Fonctions d'une variable : Rappels



Théorème

Soit f une fonction de classe C^1 sur un intervalle I .

Alors f est convexe (resp. concave) sur I si et seulement si :

$$f(y) - f(x) \geq f'(x)(y - x), \quad \forall (x, y) \in I^2$$

$$\text{(resp. } f(y) - f(x) \leq f'(x)(y - x), \quad \forall (x, y) \in I^2)$$

Théorème

Soit f une fonction de classe C^2 sur un intervalle I . Alors :

- f est convexe sur $I \iff f''(x) \geq 0$ pour tout $x \in I$
- f est concave sur $I \iff f''(x) \leq 0$ pour tout $x \in I$

Théorème

Soit f une fonction de classe C^2 sur un intervalle I . Alors :

- f est convexe sur $I \iff f''(x) \geq 0$ pour tout $x \in I$
- f est concave sur $I \iff f''(x) \leq 0$ pour tout $x \in I$

Inégalités de convexité

Si f est dérivable et convexe sur I alors sa courbe représentative est au-dessus de chacune de ses tangentes :

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

Pour les fonctions concaves l'inégalité est inversée.

Fonctions d'une variable : Rappels

L'intérêt des fonctions convexes ou concaves dans les problèmes d'optimisation s'explique par le résultat suivant :

L'intérêt des fonctions convexes ou concaves dans les problèmes d'optimisation s'explique par le résultat suivant :

Théorème

Soit f une fonction concave (resp. convexe) sur intervalle ouvert I . Si x_0 est un point critique pour f , alors f présente en x_0 un maximum (resp. minimum) global sur I .

Définition

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie sur un ensemble convexe U .
On dit que f est concave (resp. convexe) sur U si :

$$\forall \lambda \in [0, 1], \forall X, Y \in U$$

$$f(\lambda X + (1 - \lambda)Y) \geq \lambda f(X) + (1 - \lambda)f(Y)$$

$$(\text{resp. } f(\lambda X + (1 - \lambda)Y) \leq \lambda f(X) + (1 - \lambda)f(Y))$$

Théorème

Soit f de classe C^1 définie sur un convexe U .

Alors f est concave (resp. convexe) sur U si et seulement :

$$\forall X, Y \in U$$

$$f(Y) - f(X) \leq \nabla f(X) \cdot (Y - X)$$

ou encore $f(y_1, \dots, y_n) - f(x_1, \dots, x_n) \leq \sum_{i=1}^n \frac{\partial f}{\partial x_i} (y_i - x_i)$

$$\left(\text{resp. } f(y_1, \dots, y_n) - f(x_1, \dots, x_n) \geq \sum_{i=1}^n \frac{\partial f}{\partial x_i} (y_i - x_i) \right)$$

Définition : Rappels

Une forme quadratique définie sur \mathbb{R}^n est une fonction à valeurs réelles qui peut s'écrire sous la forme :

$$q(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{i < j} a_{ij}x_i x_j$$

où $A = (a_{ij})$ est l'unique matrice symétrique correspondant à l'expression de q .

Réciproquement, à toute matrice symétrique A , peut être associée une fonction à valeurs réelles $q(x_1, \dots, x_n) = X A^t X$ avec $X = (x_1 \ x_2 \ \dots \ x_n)$, identique à celle définie précédemment. Cette fonction est une forme quadratique.

Exercice

Déterminer les formes quadratiques associées aux matrices suivantes :

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & -3 \\ 2 & -3 & 4 \end{pmatrix}$$

Définition

Soit A une matrice symétrique d'ordre n .

Alors A est dite :

Définition

Soit A une matrice symétrique d'ordre n .

Alors A est dite :

- ① définie positive (resp. semi-définie positive) si
 $q(x_1, \dots, x_n) = X A^t X > 0$ pour tout $X \in \mathbb{R}^n$ et $X \neq O_{\mathbb{R}^n}$
(resp. $q(x_1, \dots, x_n) = X A^t X \geq 0$)

Définition

Soit A une matrice symétrique d'ordre n .

Alors A est dite :

- 1 définie positive (resp. semi-définie positive) si
 $q(x_1, \dots, x_n) = X A^t X > 0$ pour tout $X \in \mathbb{R}^n$ et $X \neq O_{\mathbb{R}^n}$
(resp. $q(x_1, \dots, x_n) = X A^t X \geq 0$)
- 2 définie négative (resp. semi-définie négative) si
 $q(x_1, \dots, x_n) = X A^t X < 0$ pour tout $X \in \mathbb{R}^n$ et $X \neq O_{\mathbb{R}^n}$
(resp. $q(x_1, \dots, x_n) = X A^t X \leq 0$)

Définition

Soit A une matrice symétrique d'ordre n .

Alors A est dite :

- 1 définie positive (resp. semi-définie positive) si
 $q(x_1, \dots, x_n) = X A^t X > 0$ pour tout $X \in \mathbb{R}^n$ et $X \neq O_{\mathbb{R}^n}$
(resp. $q(x_1, \dots, x_n) = X A^t X \geq 0$)
- 2 définie négative (resp. semi-définie négative) si
 $q(x_1, \dots, x_n) = X A^t X < 0$ pour tout $X \in \mathbb{R}^n$ et $X \neq O_{\mathbb{R}^n}$
(resp. $q(x_1, \dots, x_n) = X A^t X \leq 0$)
- 3 indéfinie s'il existe des vecteurs X de \mathbb{R}^n tel que $X A^t X > 0$
et d'autres tel que $X A^t X < 0$.

Théorème

Toute matrice symétrique à coefficients réels est diagonalisable.
Toute matrice hessienne à coefficients réels est diagonalisable.

Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^n . Alors, f est une fonction concave sur U si et seulement si la matrice hessienne de f est semi-définie négative pour tout X de U .
La fonction f est convexe sur U si et seulement si la matrice hessienne de f est semi-définie positive pour tout X de U .

Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^n . Alors, f est une fonction concave sur U si et seulement si la matrice hessienne de f a toutes ses valeurs propres négatives ou nulles. La fonction f est convexe sur U si et seulement si la matrice hessienne de f a toutes ses valeurs propres positives ou nulles.



Remarque

Dans \mathbb{R}^2 , on utilisera une propriété pratique...car décomposer des formes quadratiques en somme de carrés...ce n'est pas gagné



Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^2 .

Alors :

Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^2 .

Alors :

- 1 f est concave sur U si et seulement si $\forall (x,y) \in U$,
 $\text{tr}(H_f(x,y)) < 0$ et $\det(H_f(x,y)) \geq 0$.

Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^2 .

Alors :

- 1 f est concave sur U si et seulement si $\forall (x,y) \in U$,
 $\text{tr}(H_f(x,y)) < 0$ et $\det(H_f(x,y)) \geq 0$.
- 2 f est convexe sur U si et seulement si $\forall (x,y) \in U$,
 $\text{tr}(H_f(x,y)) > 0$ et $\det(H_f(x,y)) \geq 0$.

Théorème

Soit f de classe C^2 sur un sous-ensemble ouvert U de \mathbb{R}^2 .

Alors :

- 1 f est concave sur U si et seulement si $\forall (x,y) \in U$,
 $\text{tr}(H_f(x,y)) < 0$ et $\det(H_f(x,y)) \geq 0$.
- 2 f est convexe sur U si et seulement si $\forall (x,y) \in U$,
 $\text{tr}(H_f(x,y)) > 0$ et $\det(H_f(x,y)) \geq 0$.
- 3 sinon, on ne peut rien dire...

Exercice

Les fonctions f et g suivantes sont-elles convexes ou concaves sur \mathbb{R}^2 ?

$$f(x,y) = x^4 + x^2y^2 + y^4 - 3x - 8y \quad , \quad g(x,y) = xy$$

Exercice

On considère une fonction de Cobb-Douglas de deux variables définie par

$$f(x,y) = x^\alpha y^\beta \quad \text{pour } x > 0 \text{ et } y > 0$$

Montrer que f est strictement concave sur $(\mathbb{R}_+^*)^2$ si et seulement

$$\text{si : } \begin{cases} \alpha > 0 \\ \beta > 0 \\ \alpha + \beta < 1 \end{cases} .$$

Remarque

Une fonction de production de Cobb-Douglas définie sur $(\mathbb{R}_+^*)^2$ est concave si et seulement si ses rendements d'échelle sont constants ou décroissants.

Exercice

Soit f la fonction la fonction de production définie sur $D' = (\mathbb{R}_+^*)^2$ par

$$f(x, y) = x^{1/3} \sqrt{y}$$

- 1 Calculer la hessienne de f au point $(a, b) \in D'$.
- 2 Montrer que la fonction f est concave sur D' .

Complément

Etude du signe d'une forme quadratique

Proposition

On peut décomposer une forme quadratique sur \mathbb{R}^n comme combinaison linéaire de carrés de formes linéaires indépendantes. Cette décomposition n'est pas unique, mais les nombres de coefficients strictement positifs et strictement négatifs de ces CL est fixe. Le couple (m, p) de ces deux nombres s'appelle la signature de la forme quadratique.

Etude du signe d'une forme quadratique

Proposition

On peut décomposer une forme quadratique sur \mathbb{R}^n comme combinaison linéaire de carrés de formes linéaires indépendantes. Cette décomposition n'est pas unique, mais les nombres de coefficients strictement positifs et strictement négatifs de ces CL est fixe. Le couple (m, p) de ces deux nombres s'appelle la signature de la forme quadratique.

Définition

Soit A une matrice carrée d'ordre n . On appelle mineur principal d'ordre k ($k = 1, 2, \dots, n$) le déterminant de la matrice obtenue en éliminant de A les mêmes $n - k$ colonnes et $n - k$ lignes. On appelle mineur diagonal principal d'ordre k ($k = 1, 2, \dots, n$) le déterminant de la matrice obtenue en éliminant de A les $n - k$ dernières colonnes et les $n - k$ dernières lignes.

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs principaux de A .

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs principaux de A.

Trois mineurs principaux d'ordre un :

$$M_{11} = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11} \quad , \quad M_{12} = \begin{vmatrix} a_{22} \end{vmatrix} = a_{22} \quad , \quad M_{13} = \begin{vmatrix} a_{33} \end{vmatrix} = a_{33}$$

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs principaux de A.

Trois mineurs principaux d'ordre un :

$$M_{11} = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11} \quad , \quad M_{22} = \begin{vmatrix} a_{22} \end{vmatrix} = a_{22} \quad , \quad M_{33} = \begin{vmatrix} a_{33} \end{vmatrix} = a_{33}$$

Trois mineurs principaux d'ordre deux :

$$M_{21} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad , \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \quad , \quad M_{23} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

Exercice

$$\text{Soit } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R}).$$

Déterminer les mineurs principaux de A.

Trois mineurs principaux d'ordre un :

$$M_{11} = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11} \quad , \quad M_{12} = \begin{vmatrix} a_{22} \end{vmatrix} = a_{22} \quad , \quad M_{13} = \begin{vmatrix} a_{33} \end{vmatrix} = a_{33}$$

Trois mineurs principaux d'ordre deux :

$$M_{21} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad , \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \quad , \quad M_{23} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

Un mineur principal d'ordre trois : $M_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs diagonaux principaux de A.

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs diagonaux principaux de A .

$$M_1 = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11}$$

Exercice

Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R})$.

Déterminer les mineurs diagonaux principaux de A .

$$M_1 = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11}$$

$$M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Exercice

$$\text{Soit } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{M}_{(3,3)}(\mathbb{R}).$$

Déterminer les mineurs diagonaux principaux de A .

$$M_1 = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11}$$

$$M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$M_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Proposition

Soit q une forme quadratique sur \mathbb{R}^n et A la matrice symétrique associée. Alors :

- ◆ q est définie positive ssi tous ses n mineurs diagonaux principaux sont strictement positifs.
- ◆ q est définie négative ssi tous ses n mineurs diagonaux principaux d'ordre k ($k = 1, \dots, n$) M_k vérifient : $(-1)^k M_k > 0$.
- ◆ Si l'un des mineurs diagonaux principaux est non nul et ne vérifie pas l'une des deux règles de signes ci-dessus, alors q est de signe variable (q est indéfinie).

Proposition

Soit q une forme quadratique sur \mathbb{R}^n et A la matrice symétrique associée. Alors :

- ◆ q est semi-définie positive ssi tous ses $2^n - 1$ mineurs principaux sont positifs ou nuls.
- ◆ q est semi-définie négative ssi tous ses $2^n - 1$ mineurs principaux d'ordre k ($k = 1, \dots, n$) M_k vérifient : $(-1)^k M_k \geq 0$.
- ◆ Sinon q est de signe variable (q est indéfinie).

Optimisation des fonctions de plusieurs variables

Définition

On étudie le comportement d'une fonction de plusieurs variables à valeurs réelles. Une telle fonction peut avoir des valeurs extrémales : des minima (des valeurs les plus petites) ou des maxima (des valeurs les plus grandes) sur tout le domaine de définition ou bien sur une certaine partie.

On les appelle des extrema.

Définition

Soit $f : D \rightarrow \mathbb{R}$ une fonction définie sur une partie $D \subset \mathbb{R}^n$.

Définition

Soit $f : D \rightarrow \mathbb{R}$ une fonction définie sur une partie $D \subset \mathbb{R}^n$.

- 1 On dit que f admet un maximum (resp. minimum) global au point $A \in D$ si pour tout $X \in D$ on a $f(X) \leq f(A)$ (resp. $f(X) \geq f(A)$). Le maximum (resp. minimum) est appelé strict si $f(X) < f(A)$ (resp. $f(X) > f(A)$).

Définition

Soit $f : D \rightarrow \mathbb{R}$ une fonction définie sur une partie $D \subset \mathbb{R}^n$.

- 1 On dit que f admet un maximum (resp. minimum) global au point $A \in D$ si pour tout $X \in D$ on a $f(X) \leq f(A)$ (resp. $f(X) \geq f(A)$). Le maximum (resp. minimum) est appelé strict si $f(X) < f(A)$ (resp. $f(X) > f(A)$).
- 2 On dit que f admet un maximum (resp. minimum) local au point $A \in D$ si on peut trouver un nombre $r > 0$ tel que $X \in D$ et $\|X - A\| < r$ entraîne $f(X) \leq f(A)$ (resp. $f(X) \geq f(A)$).

Définition

Soit $f : D \rightarrow \mathbb{R}$ une fonction définie sur une partie $D \subset \mathbb{R}^n$.

- 1 On dit que f admet un maximum (resp. minimum) global au point $A \in D$ si pour tout $X \in D$ on a $f(X) \leq f(A)$ (resp. $f(X) \geq f(A)$). Le maximum (resp. minimum) est appelé strict si $f(X) < f(A)$ (resp. $f(X) > f(A)$).
- 2 On dit que f admet un maximum (resp. minimum) local au point $A \in D$ si on peut trouver un nombre $r > 0$ tel que $X \in D$ et $\|X - A\| < r$ entraîne $f(X) \leq f(A)$ (resp. $f(X) \geq f(A)$).

Les extrema globaux sont appelés aussi extrema absolus.

Théorème des extrema sur un compact

Théorème

Soit $f : K \rightarrow \mathbb{R}$ une fonction continue sur un compact $K \subset \mathbb{R}^n$
(compact=fermé+borné).

Alors f admet un maximum global et un minimum global sur K .

Théorème des extrema sur un compact

Théorème

Soit $f : K \rightarrow \mathbb{R}$ une fonction continue sur un compact $K \subset \mathbb{R}^n$ (compact=fermé+borné).

Alors f admet un maximum global et un minimum global sur K .

Remarque

En dimension $n = 1$ la fonction a des points extrémaux sur un intervalle. Soit ils sont à l'intérieur de l'intervalle, auquel cas ils vérifient $f'(x) = 0$, soit ils sont au bord de l'intervalle (sur le bord, la condition $f'(x) = 0$ n'est pas forcément satisfaite).

Donc pour trouver les extrema on cherche d'abord des points critiques (où la dérivée s'annule), puis on compare la valeur des points critiques avec les valeurs sur le bord de l'intervalle. Les valeurs max et min se trouvent parmi ces valeurs-là.

Définition

Soit $f : D \rightarrow \mathbb{R}$ une fonction de classe C^1 sur une partie D de \mathbb{R}^n . On dit que $A \in D$ est un point critique de f si toutes les dérivées partielles s'annulent en A (équivalent à dire que le gradient de f est nul en A , équivalent à dire aussi que la différentielle de f est nulle en A).

Théorème

Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe C^2 définie sur un ouvert $U \subset \mathbb{R}^n$ admettant un maximum ou un minimum local au point $A \in U$. Alors A est un point critique de f .

Théorème

Soit $f : U \rightarrow \mathbb{R}$ une fonction de classe C^2 définie sur un ouvert $U \subset \mathbb{R}^n$ admettant un maximum ou un minimum local au point $A \in U$. Alors A est un point critique de f .



Attention

La réciproque est fausse !!!

Preuve du Théorème

Reprenons la formule de Taylor à l'ordre 2 en dimension 2.
La preuve se généralise sans problème aux dimensions supérieures.

Preuve du Théorème

Reprenons la formule de Taylor à l'ordre 2 en dimension 2.
La preuve se généralise sans problème aux dimensions supérieures.

$$\begin{aligned} f(a+h, b+k) - f(a, b) &= h \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b) \\ &+ \frac{1}{2} \left(h^2 \frac{\partial^2 f}{\partial x^2}(a, b) + 2hk \frac{\partial^2 f}{\partial x \partial y}(a, b) + k^2 \frac{\partial^2 f}{\partial y^2}(a, b) \right) \\ &+ o(\|(h, k)\|^2) \end{aligned}$$

Preuve du Théorème

Si on a un maximum local en A , alors $f(a+h, b+k) - f(a, b) \leq 0$ pour tout (h, k) suffisamment petit. La valeur de la fonction linéaire de deux variables $h \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b)$, si elle n'est pas 0, est grande par rapport aux termes suivants. Donc cette valeur, si elle n'est pas égale à 0, doit être négative.

Preuve du Théorème

Si on a un maximum local en A , alors $f(a+h, b+k) - f(a, b) \leq 0$ pour tout (h, k) suffisamment petit. La valeur de la fonction linéaire de deux variables $h \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b)$, si elle n'est pas 0, est grande par rapport aux termes suivants. Donc cette valeur, si elle n'est pas égale à 0, doit être négative.

Pourtant pour h, k positifs il faut que les constantes

$\frac{\partial f}{\partial x_i}(a, b) \leq 0$, $i = 1, 2$ et pour h, k négatifs il faut que les mêmes

valeurs $\frac{\partial f}{\partial x_i}(a, b) \geq 0$, $i = 1, 2$, d'où $\frac{\partial f}{\partial x_i} \equiv 0$, $i = 1, 2$.

On peut refaire le même raisonnement pour un min local.

Exemple

Soit $f : (x, y) \mapsto x^2 - 2x + xy + y^2$, définie et de classe C^1 sur \mathbb{R}^2 .

Montrer que f admet un minimum local en $\left(\frac{4}{3}, -\frac{2}{3}\right)$.

Exemple

Soit $f : (x, y) \mapsto x^2 - 2x + xy + y^2$, définie et de classe C^1 sur \mathbb{R}^2 .
Montrer que f admet un minimum local en $\left(\frac{4}{3}, -\frac{2}{3}\right)$.

Exemple

Si f admet un extremum local en $(x, y) \in \mathbb{R}^2$, alors

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0 \text{ ou encore } \begin{cases} 2x + y = 2 \\ x + 2y = 0 \end{cases}.$$

On trouve alors après calculs que $(x, y) = \left(\frac{4}{3}, -\frac{2}{3}\right)$ est l'unique point critique de f sur \mathbb{R}^2 .

Extremum local : cas pratique

Exemple

Soit $f : (x, y) \mapsto x^2 - 2x + xy + y^2$, définie et de classe C^1 sur \mathbb{R}^2 .
Montrer que f admet un minimum local en $\left(\frac{4}{3}, -\frac{2}{3}\right)$.

Exemple

Si f admet un extremum local en $(x, y) \in \mathbb{R}^2$, alors

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0 \text{ ou encore } \begin{cases} 2x + y = 2 \\ x + 2y = 0 \end{cases}.$$

On trouve alors après calculs que $(x, y) = \left(\frac{4}{3}, -\frac{2}{3}\right)$ est l'unique point critique de f sur \mathbb{R}^2 .

Est-il vraiment un extremum local ?

Exemple

Il faut comparer $f(x,y)$ et $f\left(\frac{4}{3}, -\frac{2}{3}\right)$ pour (x,y) voisin de $\left(\frac{4}{3}, -\frac{2}{3}\right)$. Ce qui revient à étudier le signe de

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) \text{ pour } (h,k) \text{ voisin de } (0,0)$$

Exemple

Il faut comparer $f(x,y)$ et $f\left(\frac{4}{3}, -\frac{2}{3}\right)$ pour (x,y) voisin de $\left(\frac{4}{3}, -\frac{2}{3}\right)$. Ce qui revient à étudier le signe de

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) \text{ pour } (h,k) \text{ voisin de } (0,0)$$

D'après la formule de Taylor à l'ordre 2, on a pour tout $(h,k) \in \mathbb{R}^2$:

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) - f\left(\frac{4}{3}, -\frac{2}{3}\right) = h^2 + hk + k^2 = \left(h + \frac{k}{2}\right)^2 + \frac{3k^2}{4}$$

Exemple

Il faut comparer $f(x,y)$ et $f\left(\frac{4}{3}, -\frac{2}{3}\right)$ pour (x,y) voisin de $\left(\frac{4}{3}, -\frac{2}{3}\right)$. Ce qui revient à étudier le signe de

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) \text{ pour } (h,k) \text{ voisin de } (0,0)$$

D'après la formule de Taylor à l'ordre 2, on a pour tout $(h,k) \in \mathbb{R}^2$:

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) - f\left(\frac{4}{3}, -\frac{2}{3}\right) = h^2 + hk + k^2 = \left(h + \frac{k}{2}\right)^2 + \frac{3k^2}{4}$$

$$f\left(\frac{4}{3} + h, -\frac{2}{3} + k\right) - f\left(\frac{4}{3}, -\frac{2}{3}\right) \geq 0$$

Extremum local : cas pratique

f admet un minimum local en $\left(\frac{4}{3}, -\frac{2}{3}\right)$ qui vaut

$$f\left(\frac{4}{3}, -\frac{2}{3}\right) = -\frac{4}{3}.$$

Extremum local : cas pratique

f admet un minimum local en $\left(\frac{4}{3}, -\frac{2}{3}\right)$ qui vaut

$$f\left(\frac{4}{3}, -\frac{2}{3}\right) = -\frac{4}{3}.$$

Remarque

Tout ceci est long et fastidieux il faut étudier des conditions du second ordre en se servant de la matrice Hessienne.

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et $X_0 \in D$.

Rappels

Quand $n = 1$, pour savoir si un point critique X_0 est un maximum local ou un minimum local, on étudie la dérivée seconde (quand elle existe) :

- si $f''(X_0) > 0$, alors $f(X_0)$ est un minimum local,
- si $f''(X_0) < 0$, alors $f(X_0)$ est un maximum local,
- si $f''(X_0) = 0$, il faut faire des calculs supplémentaires de dérivées supérieures - ce peut être un point d'inflexion, un maximum ou un minimum.

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et $X_0 \in D$.

Rappels

Quand $n = 1$, pour savoir si un point critique X_0 est un maximum local ou un minimum local, on étudie la dérivée seconde (quand elle existe) :

- si $f''(X_0) > 0$, alors $f(X_0)$ est un minimum local,
- si $f''(X_0) < 0$, alors $f(X_0)$ est un maximum local,
- si $f''(X_0) = 0$, il faut faire des calculs supplémentaires de dérivées supérieures - ce peut être un point d'inflexion, un maximum ou un minimum.

Dans le cas des fonctions de plusieurs variables, à la place de f'' , on étudie la Hessienne.

Théorème

Soit $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et $X_0 \in D$ un point critique de f .

On suppose que la Hessienne $H_f(X_0)$ existe. Alors

- si $H_f(X_0)$ est définie positive (${}^t x H_f(X_0) x > 0$, pour tout $x \in \mathbb{R}^n$) alors $f(X_0)$ est un minimum local,
- si $H_f(X_0)$ est définie négative (${}^t x H_f(X_0) x < 0$, pour tout $x \in \mathbb{R}^n$) alors $f(X_0)$ est un maximum local,
- sinon il faut étudier des termes d'ordre supérieur dans la décomposition de Taylor en X_0 ...

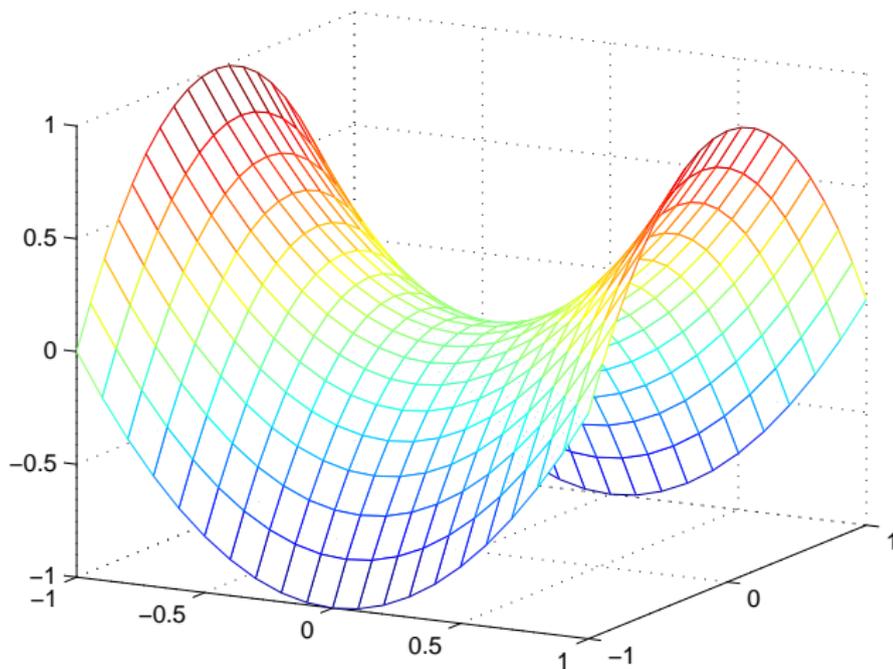
Théorème

Soit $f: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ et $X_0 \in D$ un point critique de f .

On suppose que la Hessienne $H_f(X_0)$ existe. Alors

- si $\det(H_f(X_0)) > 0$, f admet un extremum local en X_0
 - si $\text{tr}(H_f(X_0)) > 0$ alors $f(X_0)$ est un minimum local,
 - si $\text{tr}(H_f(X_0)) < 0$ alors $f(X_0)$ est un maximum local,
- si $\det(H_f(X_0)) < 0$ alors X_0 est un point selle (comme selle de cheval) ou point col (comme dans les montagnes)
- si $\det(H_f(X_0)) = 0$ alors on ne peut pas conclure, on doit regarder la formule de Taylor à l'ordre supérieur (à l'ordre 3 et parfois plus)...

Recherche des extrema : fonctions de deux variables



$$f(x,y) = x^2 - y^2$$

$(0;0)$ point selle

Exemple

- Déterminer des points où f n'est pas de classe C^1 et regarder les valeurs de f en ces points. Par exemple, la fonction $f(x,y) = 1 - \sqrt{x^2 + y^2}$ admet un maximum à l'origine mais on ne le trouve pas parmi les points critiques.
- Rechercher les points critiques.
- Etudier les points critiques.

Exemple

Extrema locaux et globaux de $f(x,y) = 2x^2y + 2x^2 + y^2$ sur \mathbb{R}^2 .

Points critiques :

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) = 4xy + 4x = 0 \\ \frac{\partial f}{\partial y}(x,y) = 2x^2 + 2y = 0 \end{cases} \Rightarrow \begin{cases} x(y+1) = 0 \\ x^2 + y = 0 \end{cases}$$

On trouve alors trois points critiques $(0,0)$, $(-1,-1)$ et $(1,-1)$.

Exemple

Points critiques	(0,0)	(-1,-1)	(1,-1)
$\frac{\partial^2 f}{\partial x^2}(x,y) = 4y + 4$	4	0	0
$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 4x$	0	-4	4
$\frac{\partial^2 f}{\partial y^2}(x,y) = 2$	2	2	2
$\det(H_f(X_0))$	8	-16	-16
Signe de $\text{tr}(H_f(X_0))$	> 0		
Nature du point critique :	min	point selle	point selle

Exemple

Pour les extrema globaux on voit que :

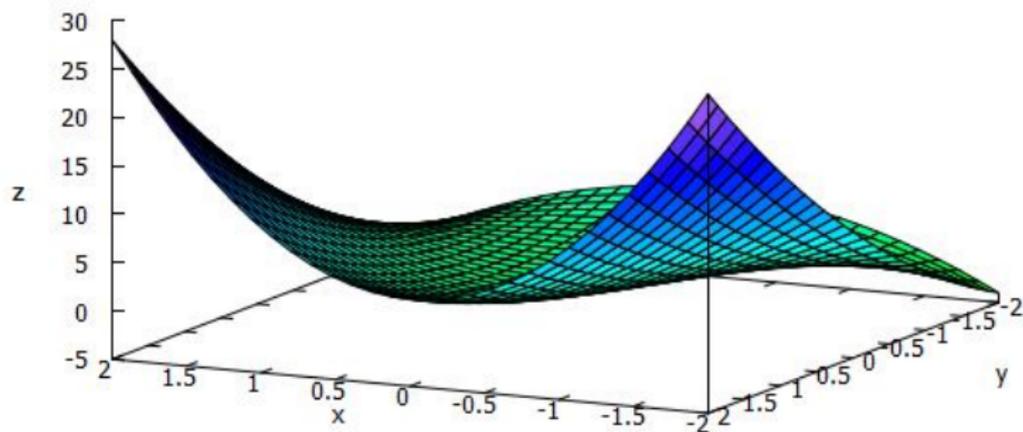
$$\lim_{x \rightarrow \pm\infty} f(x, 0) = \lim_{x \rightarrow \pm\infty} 2x^2 = +\infty$$

donc pas de maximum global.

Pas de minimum global non plus car

$$\lim_{x \rightarrow \pm\infty} f(x, -2) = \lim_{x \rightarrow \pm\infty} -2x^2 + 4 = -\infty$$

Recherche des extrema : fonctions de deux variables



$$f(x, y) = 2x^2y + 2x^2 + y^2$$

Théorème

Soit f une fonction convexe sur un ouvert convexe Ω de \mathbb{R}^n .

- Si f admet en $X_0 \in \Omega$ un minimum local, alors f admet en X_0 un minimum global.
- Si f est de classe C^1 sur Ω . Alors si $\text{grad}f(X_0) = 0$, f admet en X_0 un minimum global sur Ω .

On a des énoncés équivalents en remplaçant convexe par concave et minimum par maximum.

Introduction

Il s'agit d'optimiser une fonction f sur un domaine fermé borné Δ de $\mathbb{R}^2 \times \mathbb{R}^2$ délimité par des inégalités que doivent vérifier les variables ; en microéconomie elles seront du type $x \geq 0$ et $y \geq 0$ et $ax + by \leq c$ avec a, b, c constantes positives. La fonction f étant continue elle admet sur Δ un maximum (global) et un minimum (global).

Marche à suivre

- On recherche les éventuels points critiques de f à l'intérieur de Δ et on examine en ces points la condition du second ordre : on a ainsi optimisé la fonction à l'intérieur de Δ .
- On étudie localement la fonction sur le bord de Δ (en se ramenant à une fonction d'une variable à étudier sur un segment).
- On conclut en regroupant tous les points critiques à l'intérieur et sur le bord de Δ .

Exercice

Optimiser la fonction $f(x,y) = x^2 + y^2 - xy - x - y + 2$ sur l'ensemble $\Delta = \{ x \geq 0 \text{ et } y \geq 0 \text{ et } x + y \leq 3 \}$.