Math ECG 2. 2022-2023

Mathématiques Approfondies - F. Gaunard
http://frederic.gaunard.com
H2B - Lycée Carnot, Paris 17e.

Chapitre III. Recherche d'*extrema* pour des fonctions réelles de plusieurs variables. Optimisation sous contrainte.

Avant-propos

Ce chapitre est la suite d'un premier chapitre d'introduction sur les fonctions de plusieurs variables. On en reprend toutes les notations. Plus précisément, dans tout le chapitre, on considère $\mathbb{R}^n$ muni de sa norme euclidienne, notée $||\cdot||$. Si f $\mathbb{R}^n \to \mathbb{R}$ admet des dérivées partielles d'ordre 1, on note ∇f le gradient de f

$$\nabla f = (\partial_1 f,\ \partial_2 f,\ ...,\ \partial_n f).$$

Durant les séances de travaux pratiques sous Python consacrés à ce premier chapitre, on a représenté graphiquement certaines fonctions de plusieurs variables (surfaces, lignes de niveaux, champ de gradient...) et observé parfois la présence de points où f semblait présenter un *extremum* (local) ou un *point selle*. Un des objectifs de ce chapitre est de donner un critère pour conclure quand à la nature d'un *point critique* (où le gradient s'annule).

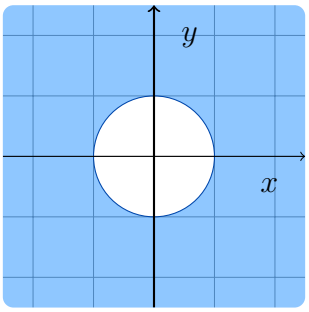
Par ailleurs, jusqu'à présent, les fonctions de plusieurs variables rencontrées étaient définies sur $\mathbb{R}^n$. Cependant, il est possible qu'elles ne soient pas définies partout. Il apparait alors nécessaire d'introduire quelques notions *topologiques* (ouverte ? fermée ? bornée ?) pour les parties de $\mathbb{R}^n$.

Exemple

Considérons la fonction $g:\ (x,y) \longmapsto \sqrt{x^2+y^2-1}$.
Cette fonction n'est pas définie sur $\mathbb{R}^2$ tout entier, elle est définie sur l'ensemble

$$\mathcal{D}_g = \{(x,y) \in \mathbb{R}^2 :\ x^2+y^2-1 \geq 0\} = \{(x,y) \in \mathbb{R}^2 :\ ||(x,y)|| \geq 1\},$$

que l'on représente (en bleu) ci-dessous



1 Rudiments de topologie sur $\mathbb{R}^n$

Tout comme on travaille sur des intervalles ouverts (ou fermés) lors de l'étude de fonctions d'une variable, on a besoin de notions analogues avec plusieurs variables. On commence donc par l'introduction de la notion de boules (ouvertes ou fermées) ci-dessous.

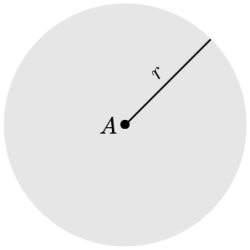
Définition

Soit r un réel strictement positif et a un point de $\mathbb{R}^n$.

- On appelle **boule ouverte de centre a et de rayon r** l'ensemble $\mathcal{B}(a,r)$ défini par
$$\mathcal{B}(a,r) = \{x \in \mathbb{R}^n \ : \ ||a-x|| < r\};$$
- On appelle **boule fermée de centre a et de rayon r** l'ensemble $\overline{\mathcal{B}}(a,r)$ défini par
$$\overline{\mathcal{B}}(a,r) = \{x \in \mathbb{R}^n \ : \ ||a-x|| \leq r\};$$

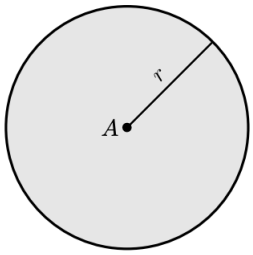
Illustration

Boule ouverte $\mathcal{B}(A,r)$



Illustration

Boule fermée $\overline{\mathcal{B}}(A,r)$



☞ La notion de boule de centre a est alors une *généralisation* de la notion d'intervalle centré en a (de la forme $]a-r;a+r[$ ou $[a-r;a+r]$ selon que l'intervalle est ouvert ou fermé).

Définition

Soit $a \in \mathbb{R}^n$. Une partie $\mathcal{V}$ de $\mathbb{R}^n$ est un **voisinage** de a si elle contient une boule ouverte de centre a et de rayon $r > 0$.

Définition

Soit U une partie non vide de $\mathbb{R}^n$.

- Une partie U de $\mathbb{R}^n$ est dite **ouverte** si et seulement si pour tout point $a \in U$, il existe $r > 0$ tel que $\mathcal{B}(a,r) \subset U$, soit si et seulement si, pour tout $a \in U$ il existe une boule ouverte de centre a incluse dans U.
On dit aussi que U est un ouvert de $\mathbb{R}^n$.

- Une partie F de $\mathbb{R}^n$ est dite **fermée** si et seulement si son complémentaire $\overline{F}$ est une partie ouverte.
On dit aussi que F est un fermé de $\mathbb{R}^n$.

☞ Certaines parties de $\mathbb{R}^n$ ne sont **ni ouvertes, ni fermées**!

☞ Une partie de $\mathbb{R}^n$ est donc ouverte si elle est voisinage de tous ses points.

Exercice 1. Soient $a \in \mathbb{R}^n$ et $r > 0$. Montrer que la boule ouverte $\mathcal{B}(a, r)$ est un ouvert de $\mathbb{R}^n$ et que la boule fermée $\overline{\mathcal{B}}(a, r)$ est un fermé de $\mathbb{R}^n$.

Proposition

- $\mathbb{R}^n$ et $\emptyset$ sont à la fois ouvertes ou fermées.
  On peut montrer (mais c'est difficile et dépasse de loin le cadre de ce cours) que ce sont les seules parties de $\mathbb{R}^n$ avec cette propriété.
- Tout singleton est une partie fermée de $\mathbb{R}^n$.
- Toute réunion d'ouverts de $\mathbb{R}^n$ est encore un ouvert de $\mathbb{R}^n$.
  Toute intersection d'un **nombre fini** d'ouverts de $\mathbb{R}^n$ est encore un ouvert.
  Attention, ce n'est plus vrai si l'intersection n'est pas finie.
- Toute intersection de fermés de $\mathbb{R}^n$ est encore un fermé de $\mathbb{R}^n$.
  Toute réunion d'un **nombre fini** de fermés de $\mathbb{R}^n$ est encore un fermé.
  Attention, ce n'est plus vrai si la réunion n'est pas finie.

Exercice 2.

(1) Montrer que toute partie finie de $\mathbb{R}^n$ est fermée.

(2) Soit $a \in \mathbb{R}^n$. Déterminer $\bigcap_{k=1}^{+\infty} \mathcal{B}\left(a, \frac{1}{k}\right)$. Sur quel point cet exemple permet-il d'insister?

Théorème

L'image réciproque $f^{-1}(I)$ d'un intervalle ouvert de $\mathbb{R}$ par une fonction f continue de $/R^n$ dans $\mathbb{R}$ est un ouvert de $\mathbb{R}^n$.

L'image réciproque $f^{-1}(S)$ d'un intervalle fermé de $\mathbb{R}$ par une fonction f continue de $/R^n$ dans $\mathbb{R}$ est un fermé de $\mathbb{R}^n$.

Corollaire

- Les ensembles de $\mathbb{R}^n$ de la forme
  $$\{(x_1, x_2, ..., x_n) : \ f(x_1, ..., x_n) < a\} \qquad \text{ou} \qquad \{(x_1, x_2, ..., x_n) : \ f(x_1, ..., x_n) > a\},$$
  où $a > 0$ et f continue de $\mathbb{R}^n$ dans $\mathbb{R}$, sont des **ouverts** de $\mathbb{R}^n$.
- Les ensembles de $\mathbb{R}^n$ de la forme
  $$\{(x_1, x_2, ..., x_n) : \ f(x_1, ..., x_n) \leq a\}, \qquad \{(x_1, x_2, ..., x_n) : \ f(x_1, ..., x_n) \geq a\}$$
  ou
  $$\{(x_1, x_2, ..., x_n) : \ f(x_1, ..., x_n) = a\}$$
  où $a > 0$ et f continue de $\mathbb{R}^n$ dans $\mathbb{R}$, sont des **fermés** de $\mathbb{R}^n$.

☞ Le programme officiel est formel; l'étude des fonctions réelles de n variables se limitera à des fonctions pouvant être définies sur des parties de $\mathbb{R}^n$ pouvant être définies simplement (réunion ou intersection finies, ..) à l'aide des ensembles ouverts et fermés précédent.

Exercice 3. Soient $a_1, a_2, ..., a_n$, $b_1, b_2, .., b_n$ des réels tels que $a_i < b_i$.

(1) Montrer que le produit cartésien

$$[a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$$

est un fermé de $\mathbb{R}^n$. Représenter graphiquement cet ensemble pour $n = 2$.

(2) Montrer que le produit cartésien

$$]a_1, b_1[\times]a_2, b_2[\times ... \times]a_n, b_n[$$

est un ouvert de $\mathbb{R}^n$. Représenter graphiquement cet ensemble pour $n = 2$.

Exercice 4. On considère la fonction φ définie par

$$\varphi(x, y) = \ln\left(y - x^2 + 5x - 6\right)$$

(1) Déterminer l'ensemble de définition $\mathcal{D}$ de φ.
(2) Le représenter graphiquement.
(3) Montrer que $\mathcal{D}$ est un ouvert de $\mathbb{R}^2$.

Définition

Une partie U de $\mathbb{R}^n$ est dite **bornée** s'il existe $M > 0$ telle que, pour tout $x \in U$, $||x|| \leq M$.

De manière équivalente, U est bornée si et seulement si il existe $M > 0$ tel que

$$U \subset \overline{\mathcal{B}}(0, M).$$

☞ L'inégalité triangulaire permet de dire qu'une partie est bornée si et seulement si elle est incluse dans une boule fermée (pas nécessairement centrée en 0).

☞ En pratique, pour montrer qu'un ensemble est borné, on montre souvent que chaque composante de ses éléments est bornée (par une même constante). En effet, si

$$\exists \kappa > 0, \quad \forall x = (x_1, x_2, ..., x_n) \in U, \quad \forall i \in [\![1, n]\!] : \ |x_i| \leq \kappa,$$

alors,

$$||x||^2 = \sum_{i=1}^{n} x_i^2 \leq n\kappa^2$$

et donc $U \subset \overline{\mathcal{B}}(0, \ \kappa\sqrt{n})$.

Exercice 5. Montrer que la partie U de $\mathbb{R}^2$ définie par $U =]-1, 2[\times[3, 4]$ est bornée.
Est-elle ouverte ? Est-elle fermée ?

2 Extension des notions vues sur $\mathbb{R}^n$

Ce paragraphe a pour (unique) but de voir que l'on peut étendre toutes les notions (continuité, existence des dérivées partielles, caractère $\mathcal{C}^1$...) introduites pour des fonctions définies sur $\mathbb{R}^n$ à des fonctions définies sur une partie (ouverte) de $\mathbb{R}^n$.

2.1 Continuité

Définition

Soient f une fonction définie sur une partie $\mathcal{D}$ et $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et $a \in D$.
On dit que f est **continue** en a, si

$$\forall \varepsilon > 0,\ \exists \eta > 0 :\ \forall x \in \mathcal{D},\ x \in \mathcal{B}(a,\eta) \Rightarrow |f(x) - f(a)| < \varepsilon.$$

On dit que f est continue sur $\mathcal{D}$ si elle est continue en tout point de $\mathcal{D}$.

Théorème

Toutes les propriétés (somme, produit, quotient, composition à gauche par une fonction d'une variable continue) vues pour les fonctions continues sur $\mathbb{R}^n$ sont valables pour des fonctions continues sur une partie $\mathcal{D}$ de $\mathbb{R}^n$.

Exercice 6. Justifier que la fonction φ de l'Exercice 4 est continue sur son ensemble de définition.

2.2 Calcul différentiel d'ordre 1

On rappelle que si f est une fonction définie au voisinage d'un point $a = (a_1, a_2, .., a_n) \in \mathbb{R}^n$, la $i-$ème fonction partielle de f en a est la fonction réelle définie par

$$f_{i,a}(t) = f(a_1, ..., a_{i-1}, \boldsymbol{t}, a_{i+1}, ..., a_n).$$

(Cette fonction est définie sur un voisinage de a_i.)

Définition

Soient U un ouvert non vide de $\mathbb{R}^n$ et $f :\ U \to \mathbb{R}$.

- On dit que f admet une dérivée partielle d'ordre 1 en $a \in U$ par rapport à la $i-$ème variable si la $i-$ème fonction partielle de f en a est dérivable en a_i. Auquel cas, on note
$$\partial_i f(a) = f'_{i,a}(a_i) = \lim_{t \to a_i} \frac{f(a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n) - f(a)}{t - a_i}.$$
- Si f admet une dérivée partielle en $a \in U$ par rapport à chaque variable, on appelle **gradient** de f en a, le vecteur de $\mathbb{R}^n$
$$\nabla f(a) = (\partial_1 f(a),\ \partial_2 f(a),\ ...,\ \partial_n f(a)).$$
- Si f admet une dérivée partielle par rapport à la $i-$ème variable en tout point a de U, on note $\partial_i f$ la fonction
$$\partial_i f :\ U \longrightarrow \mathbb{R}^n, \quad x \longmapsto \partial_i f(x).$$

Remarque

Attention, l'existence de dérivée(s) partielle(s) pour f en a ne garantit pas la continuité en a.

Définition

Soient U un ouvert de $\mathbb{R}^n$ et $f :\ U \to \mathbb{R}$. On dit que f est de classe $\mathcal{C}^1$ sur U si f admet une dérivée partielle d'ordre 1 par rapport à chaque variable en tout point de U et si, pour tout $i \in [\![1, n]\!]$, $\partial_i f$ est continue sur U.

Théorème

Toutes les propriétés (somme, produit, quotient, composition à gauche par une fonction d'une variable de classe $\mathcal{C}^1$) vues pour les fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$ sont valables pour des fonctions de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$.

Définition

Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$. On dit que f admet un développement limité d'ordre 1 en a s'il existe un réel α, un vecteur $u \in \mathbb{R}^n$ et une fonction $\varepsilon : U \to \mathbb{R}^n$ continue en 0 et vérifiant $\varepsilon(0) = 0$ tels que

$$\forall h \in \mathbb{R}^n \text{ tel que } a + h \in U, \quad f(a+h) = \alpha + \langle h, u\rangle + ||h||\varepsilon(h).$$

Remarque

On peut montrer (mais c'est hors programme) que l'existence d'un développement limité d'ordre 1 en a implique que f est continue en a, que f admet un gradient en a et que $\alpha = f(a)$, $u = \nabla f(a)$.
En particulier, en cas d'existence du DL, celui-ci est unique.

Théorème

Une fonction de classe $\mathcal{C}^1$ sur un ouvert admet en tout point un (et un seul) développement limité d'ordre 1.

Plus précisément, si $f : U \to \mathbb{R}^n$ est de classe $\mathcal{C}^1$ sur U, alors, pour tout $a \in U$, il existe une fonction $\varepsilon : U \to \mathbb{R}$ continue vérifiant $\varepsilon(0) = 0$ telle que

$$\forall h \in \mathbb{R}^n \text{ tel que } a + h \in U, \quad f(a+h) = f(a) + \langle h, \nabla f(a)\rangle + ||h||\varepsilon(h).$$

Ceci s'écrit aussi, pour x dans un voisinage de a,

$$f(x) = f(a) + \langle \nabla f(a), x - a\rangle + o\left(||x - a||\right).$$

☞ En particulier, une fonction de classe $\mathcal{C}^1$ sur un ouvert y est aussi continue.

Définition

Soient U un ouvert non vide de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ une fonction, $a \in U$ et u un vecteur non nul de $\mathbb{R}^n$. On dit que f admet une dérivée directionnelle en a dans la direction de u si la fonction

$$g : t \longmapsto f(a + tu),$$

définie au voisinage de 0, admet une dérivée. Auquel cas, la valeur de la dérivée directionnelle de f en a dans la direction de u est $g'(0)$.

☞ En notant $(e_1, e_2, .., e_n)$ les vecteurs de la base canonique de $\mathbb{R}^n$, dire que f admet une dérivée partielle en a par rapport à la $i-$ème variable revient à dire que f admet une dérivée directionnelle en a dans la direction de e_i.

Proposition

Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$, $a \in U$ et $u = (u_1, u_2, .., u_n) \in \mathbb{R}^n$. Alors, la fonction $g : \ t \mapsto f(a + tu)$, définie sur un voisinage V de 0, est dérivable sur V et vérifie, pour tout $t \in V$

$$g'(t) = \langle \nabla f(a + tu), u \rangle = \sum_{i=1}^{n} u_i \partial f_i(a + tu).$$

Corollaire

Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$ et $a \in U$. Alors, f admet une dérivée directionnelle en a dans la direction de tout vecteur non nul $u \in \mathbb{R}^n$ qui vaut $\langle \nabla f(a), u \rangle$.

Définition

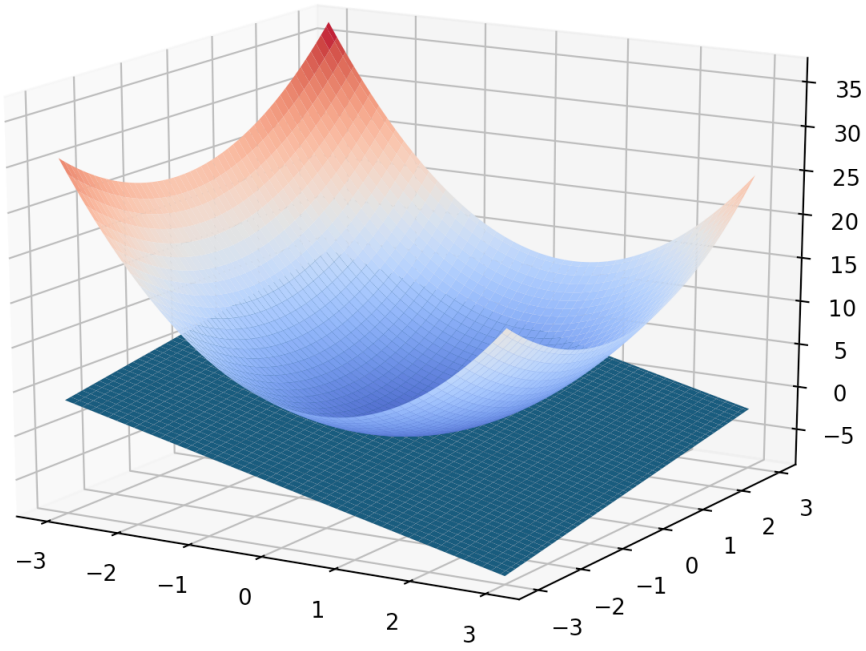
Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$ et $a \in U$.
L'application $x \in \mathbb{R}^n \longmapsto f(a) + \langle \nabla f(a), x - a \rangle$ s'appelle l'**approximation affine** de f au voisinage de a.
Le graphe de celle-ci, c'est à dire l'hyperplan affine de $\mathbb{R}^{n+1}$ d'équation $x_{n+1} = f(a) + \langle \nabla f(a), x - a \rangle$ s'appelle l'hyperplan affine tangent au graphe de f en a.

Exemple

Le code Python téléchargeable *ici* permet de visualiser un exemple; on a représenté l'hyperplan affine tangent au graphe de $f : \ x \mapsto 2(x - 1/2)^2 + (y + 1/2)^2$ en 0.

Cet hyperplan affine a pour équation

$$z = f(0) + \langle \nabla f(0), (x, y) \rangle = \frac{3}{4} - 2x + y$$



3 Calcul différentiel d'ordre 2 sur un ouvert de $\mathbb{R}^n$.

Définition

Soit f une fonction définie sur un ouvert non vide U de $\mathbb{R}^n$ et $a \in U$. Soit $(i,j) \in [\![1,n]\!]^2$.

- Si $\partial_i f$ est définie au voisinage de a et admet une dérivée partielle d'ordre 1 en a par rapport à la $j-$ième variable, on dit que f admet une dérivée partielle d'ordre 2 en a par rapport à la $i-$ième variable puis par rapport à la $j-$ième variable et sa valeur est notée
$$\partial^2_{j,i} f(a) = \partial_j\left(\partial_i f\right)(a)$$
- Si f admet en a une dérivée partielle d'ordre 2 par rapport à la $i-$ième variable puis par rapport à la $j-$ième variable pour tout couple (i,j), on appelle **matrice hessienne** (ou tout simplement Hessienne) de f en a la matrice
$$\nabla^2 f(a) = \left(\partial^2_{i,j}(a)\right)_{1\leq i,j\leq n}.$$
- On dit que f est de classe $\mathcal{C}^2$ sur U si f admet en tout point de U une dérivée partielle d'ordre 2 par rapport à la $i-$ième variable puis par rapport à la $j-$ième variable, pour tout couple (i,j) et que, pour tout couple (i,j) la fonction $\partial^2_{i,j} f$ est continue sur U.

Remarque

f est de classe $\mathcal{C}^2$ sur U si et seulement si toutes les dérivées partielles d'ordre 1 de f sont elles-mêmes de classe $\mathcal{C}^1$ sur U. En particulier, si f est $\mathcal{C}^2$ sur U, elle est aussi $\mathcal{C}^1$ sur U.

Exemple

Il est immédiat de montrer que les fonctions polynomiales sont de classe $\mathcal{C}^2$ sur $\mathbb{R}^n$.

Exercice 7. On considère la fonction $\varphi : x \in \mathbb{R}^n \longmapsto ||x||^2$. Justifier que φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^n$ et déterminer la Hessienne de φ en 0.

Théorème

Soit U un ouvert de $\mathbb{R}^n$.

- Tout combinaison linéaire de fonctions de classe $\mathcal{C}^2$ sur U est encore de classe $\mathcal{C}^2$ sur U.
- Le produit de fonctions de classe $\mathcal{C}^2$ sur U est encore de classe $\mathcal{C}^2$ sur U.
- Le quotient de fonctions de classe $\mathcal{C}^2$ sur U et dont le dénominateur de s'annule pas sur U est encore de classe $\mathcal{C}^2$ sur U.
- Si f est de classe $\mathcal{C}^2$ sur U et $g : V \to \mathbb{R}$ est une fonction définie et de classe $\mathcal{C}^2$ sur un ouvert V tel que $f(U) \subset V$, alors $g \circ f$ est encore de classe $\mathcal{C}^2$ sur U.

Exercice 8. Soit g la fonction de deux variables définie par $g(x,y) = \frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right)(1+x)(1+y)$.

(1) Déterminer le domaine de définition U de g. Vérifier que c'est un ouvert de $\mathbb{R}^2$.
(2) Justifier que g est de classe $\mathcal{C}^2$ sur U.
(3) Déterminer le gradient de g puis montrer que g n'admet qu'un seul point critique sur U.
(4) Déterminer l'expression de la matrice hessienne de g au point critique. Est-elle diagonalisable ?

Théorème

Théorème de Schwarz.

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert U de $\mathbb{R}^n$. Alors, pour tout couple $(i,j) \in [\![1,n]\!]$,

$$\partial^2_{i,j} f = \partial^2_{j,i} f.$$

Autrement dit, la Hessienne de f en tout point de U est symétrique.

☞ En particulier, la Hessienne en un point d'une fonction $\mathcal{C}^2$ est diagonalisable dans une base orthonormée.

Définition

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert U et $a \in U$. On appelle **forme quadratique** associée à la Hessienne de f en a la forme quadratique q_a définie par

$$q_a : \ h = (h_1, ..., h_n) \in \mathbb{R}^n \longmapsto \langle \nabla^2 f(a)h, h\rangle = \ {}^t H \nabla^2 f(a) H = \sum_{1\leq i,j\leq n} \partial^2_{i,j} f(a) h_i h_j$$

Définition

Soient U un ouvert de $\mathbb{R}^n$ et $f : \ U \to \mathbb{R}$. On dit que f admet un développement limité d'ordre 2 en a s'il existe un réel α, un vecteur $u \in \mathbb{R}^n$, une forme quadratique q sur $\mathbb{R}^n$ et une fonction $\varepsilon : U \to \mathbb{R}^n$ continue en 0 et vérifiant $\varepsilon(0) = 0$ tels que

$$\forall h \in \mathbb{R}^n \text{ tel que } a+h \in U, \quad f(a+h) = \alpha + \langle h, u\rangle + q(h) + ||h||^2 \varepsilon(h).$$

Théorème

Une fonction de classe $\mathcal{C}^2$ sur un ouvert admet en tout point un (et un seul) développement limité d'ordre 2.

Plus précisément, si $f : \ U \to \mathbb{R}^n$ est de classe $\mathcal{C}^2$ sur U, alors, pour tout $a \in U$, il existe une fonction $\varepsilon : \ U \to \mathbb{R}$ continue vérifiant $\varepsilon(0) = 0$ telle que

$$\forall h \in \mathbb{R}^n \text{ tel que } a+h \in U, \quad f(a+h) = f(a) + \langle h, \nabla f(a)\rangle + \frac{1}{2} q_a(h) + ||h||^2 \varepsilon(h).$$

Ceci s'écrit aussi, pour x dans un voisinage de a,

$$f(x) = f(a) + \langle \nabla f(a), x-a\rangle + \frac{1}{2} q_a(x-a) + o\left(||x-a||^2\right).$$

☞ Ce résultat est une généralisation de la formule de Taylor-Young à l'ordre 2.

Définition

Soient U un ouvert non vide de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ une fonction, $a \in U$ et u un vecteur non nul de $\mathbb{R}^n$. On dit que f admet une dérivée seconde en a dans la direction de u si la fonction

$$g : \ t \longmapsto f(a+tu),$$

définie au voisinage de 0, est dérivable deux fois. Auquel cas, la valeur de la dérivée seconde de f en a dans la direction de u est $g''(0)$ et se note $f''_u(a)$.

Proposition

Soient f une fonction de classe $\mathcal{C}^2$ sur un ouvert U non vide de $\mathbb{R}^n$, $a \in U$ et $u = (u_1, u_2, .., u_n) \in \mathbb{R}^n$. Alors,

- la fonction $g : t \mapsto f(a+tu)$, définie sur un voisinage V de 0, est dérivable deux fois sur V et vérifie, pour tout $t \in V$
$$g''(t) = \sum_{1 \leq i,j \leq n} \partial^2_{i,j} f(a+tu) u_i u_j = q_{a+tu}(u).$$
- f admet une dérivée seconde en a dans la direction de u et $f''_u(a) = q_a(u)$.

4 Recherche d'*extrema*

4.1 Notion d'extremum

Définition

Soit f une fonction définie sur un ensemble $\mathcal{D}$ de $\mathbb{R}^n$. Soit $a \in \mathcal{D}$.

- On dit que f admet un **minimum** local en a si il existe un voisinage $\mathcal{V}$ de a tel que
$$\forall x \in \mathcal{V}, \qquad f(x) \geq f(a).$$
- On dit que f admet un **maximum** local en a si il existe un voisinage $\mathcal{V}$ de a tel que
$$\forall x \in \mathcal{V}, \qquad f(x) \leq f(a).$$
- On appelle **extremum** un minimum ou un maximum de f. L'extremum est dit **global** si les inégalités ci-dessus ont lieu pour tout x dans $\mathcal{D}$. Il est dit **strict** si les inégalités ci-dessus sont strictes dès lors que $x \neq a$.

☞ En particulier, un extremum global est aussi un extremum local, mais naturellement le contraire n'est pas vrai.

4.2 Fonction continue sur un fermé borné

Théorème

Extrema sur un fermé borné.

Toute fonction continue sur un ensemble fermé et borné de $\mathbb{R}^n$ y est bornée est atteint ses bornes (et admet donc un minimum et un maximum).

4.3 Condition nécessaire de premier ordre

Théorème

Soit f une fonction de classe $\mathcal{C}^1$ sur un **ouvert** U de $\mathbb{R}^n$. Soit $a \in U$. Si f présente un extremum en a, alors nécessairement, le gradient de f en a est nul.

Définition

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert U. On dit que $a \in U$ est un **point critique** de f si $\nabla f(a) = 0$.
Toutes les dérivées directionnelles de f en a sont nulles.

À retenir!

Une fonction f de classe $\mathcal{C}^1$ ne peut donc présenter un extremum qu'en un point critique.
Attention, tout ce qui brille n'est pas de l'or; f ne présente pas toujours un extremum en chaque point critique.

Exercice 9. On considère la fonction f définie par $f(x,y) = (x+y)\left(\frac{1}{x}+\frac{1}{y}\right)$.

(1) Déterminer l'ensemble de définition $\mathcal{D}$ de f. Justifier que c'est un ouvert de $\mathbb{R}^2$.
(2) Montrer que, pour tout couple $(x,y) \in \mathcal{D}$

$$f(x,y) = 2 + \frac{y}{x} + \frac{x}{y} \quad \text{et} \quad f(x,y) = \frac{(x+y)^2}{xy}.$$

(3) Montrer que f est de classe $\mathcal{C}^2$ sur $\mathcal{D}$.
(4) Montrer que f possède une infinité de points critiques et les déterminer.
(5) (a) Comparer les réels $(x+y)^2$ et $4xy$.
    (b) En déduire que f admet sur $\mathcal{D}$ un minimum global en tous ses points critiques et donner sa valeur.

Définition

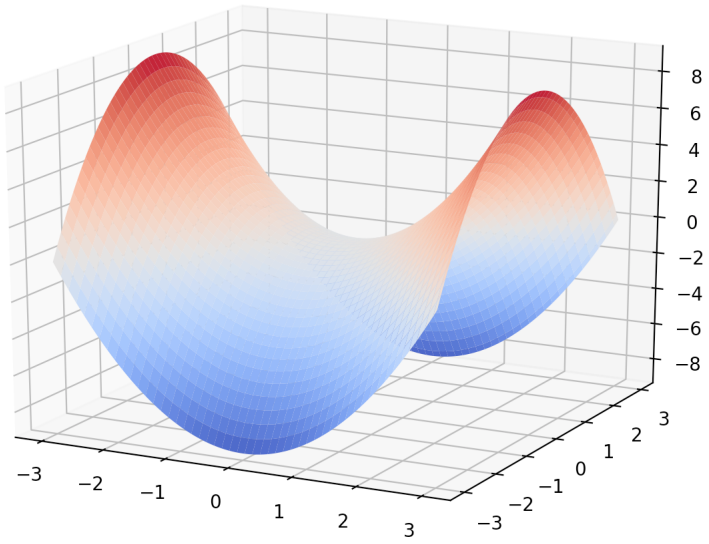
Soit f de classe $\mathcal{C}^1$ sur un ouvert U. On appelle **point selle** (ou **point col**) tout point critique de f où f ne présente pas d'extremum.

☞ Un col de montagne est le passage le plus bas entre deux montagnes mais il faut monter pour l'atteindre.

Exemple

Un bel exemple de point selle : Le *Pringles* (paraboloïde hyperbolique).

On a représenté le graphe de $f(x,y) = x^2 - y^2$ sur $[-3,3] \times [-3,3]$.



4.4 Condition suffisante locale de second ordre

Le développement limité (formule de Taylor-Young) à l'ordre 2 permet de donner un critère à la présence d'un extremum. En effet, si a est un point critique de f, pour h dans un voisinage de 0, on a

$$f(a+h) - f(a) = q_a(h) + ||h||^2 \varepsilon(h).$$

Théorème

Condition suffisante locale de second ordre.

Soient f une fonction de classe $\mathcal{C}^2$ sur un ouvert U et $a \in U$ un point critique de f.

- Si $\mathrm{Sp}(\nabla^2 f(a)) \subset \mathbb{R}_+^*$, (la forme quadratique q_a est définie positive et) f présente un minimum local (strict) en a.
- Si $\mathrm{Sp}(\nabla^2 f(a)) \subset \mathbb{R}_-^*$, (la forme quadratique q_a est définie négative et) f présente un maximum local (strict) en a.
- Si $\mathrm{Sp}(\nabla^2 f(a)) \cap \mathbb{R}_+^* \neq \emptyset$ **et** $\mathrm{Sp}(\nabla^2 f(a)) \cap \mathbb{R}_-^* \neq \emptyset$, (la forme quadratique q_a n'est pas de signe constant et) f présente un point selle en a.

Exercice 10. Soit $f(x,y) = 3xy - x^3 - y^3$. Justifier que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ puis, déterminer, si ils existent, les extremums locaux de f et préciser leurs natures.

Exercice 11. Soit la fonction f définie par : $f(x,y) = x\left((\ln(x))^2 + y^2\right)$.

(1) Préciser le domaine de définition de f et le représenter. Justifier qu'il s'agit d'un ouvert de $\mathbb{R}^2$.
(2) Montrer que f est de classe $\mathcal{C}^2$ sur ce domaine.
(3) Montrer que f admet un seul extremum local. Préciser sa nature et sa valeur.
(4) Cet extremum est-il global ?

Exercice 12. On désigne par n un entier naturel supérieur ou égal à 2. On considère une épreuve aléatoire pouvant aboutir à 3 résultats différents R_1 , R_2 et R_3 de probabilités respectives p_1, p_2 et p_3 (où on suppose que $0 < p_i < 1$, pour $1 \leq i \leq 3$).

On effectue n épreuves indépendantes du type de celle décrite ci-dessus. Pour tout i de $\{1,2,3\}$, on note X_i la variable aléatoire qui vaut 1 si le résultat numéro i n'est pas obtenu à l'issue des n épreuves et 0 sinon.

On désigne par X la variable égale au nombre de résultats qui n'ont pas été obtenus à l'issue des n épreuves.

(1) (a) Expliciter un lien entre X, X_1, X_2 et X_3.
(b) Donner la loi de X_i, pour tout $i \in [\![1,3]\!]$.
(c) En déduire l'espérance de X, notée $E(X)$.

On note f la fonction définie sur l'ouvert $]0,1[\times]0,1[$ par $f(x,y) = (1-x)^n + (1-y)^n + (x+y)^n$.

(2) (a) Justifier que f est une fonction de classe $\mathcal{C}^2$ sur $]0,1[\times]0,1[$.
(b) Déterminer les dérivées partielles d'ordre 1 de f.
(c) En déduire qu'il existe un seul point A de $]0;1[\times]0;1[$ où f est susceptible de présenter un extremum et l'expliciter.
(d) Démontrer que f présente un minimum local en ce point.

(3) Vérifier que $E(X) = f(p_1,p_2)$. En déduire les valeurs de p_1, p_2 et p_3 qui rendent $E(X)$ minimale et préciser la valeur de ce minimum.

4.5 Recherche d'*extrema* globaux

On propose ci-après un critère d'existence d'un extremum global sur un ouvert. L'énoncé de celui-ci nécessite une propriété géométrique sur l'ouvert où l'on cherche l'extremum. On introduit donc cette propriété en insistant néanmoins sur le fait que le programme stipule clairement que la vérification de cette propriété ne fait pas partie des attentes.

Définition

Un ouvert U non vide de $\mathbb{R}^n$ est dit **convexe** si, pour tous $x, y \in U$ et pour tout $t \in [0;1]$,

$$tx + (1-t)y \in U.$$

(C'est à dire qu'un segment reliant deux points de U est encore inclus dans U.)

☞ En particulier si U est un ouvert convexe et $a \in U$, alors si $||h||$ est assez petite, $a + h \in U$ (c'est le caractère ouvert) mais alors par convexité, pour tout $t \in [0;1]$, $a + th = t(a+h) + (1-t)a \in U$.

Théorème

Extremum global sur un ouvert convexe.

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert convexe U de $\mathbb{R}^n$. Si, pour tout $x \in U$, q_x est positive (resp. négative) alors f présente un minimum (resp. maximum) global en tout point critique.

Exercice 13. Montrer que, pour tout $(x, y) \in \mathbb{R}^2$,

$$2x^2 + y^2 + 12x + 4y + 2xy + 24 \geq 4.$$

4.6 Recherche d'*extrema* sur un fermé borné

Soit K un ensemble fermé et borné de $\mathbb{R}^n$.
On appelle *intérieur* de K le plus grand ouvert contenu dans K, que l'on note $\overset{\circ}{K}$. C'est donc un ouvert, et il est borné.
On appelle *frontière* (ou bord) de K l'ensemble (fermé) défini par $\partial K = K \setminus \overset{\circ}{K}$. C'est encore un fermé borné.

Par exemple, si $K = \overline{\mathcal{B}}(a, r)$, alors $\overset{\circ}{K} = \mathcal{B}(a, r)$ et $\partial K = \{x \in \mathbb{R}^n : \ ||a - x|| = r\}$.

À retenir!

Soient K un fermé borné de $\mathbb{R}^n$ et f une fonction continue sur K et de classe $\mathcal{C}^2$ sur $\overset{\circ}{K}$. On **sait** donc (par théorème) que f admet minimum et maximum sur K.

- On recherche les points critiques de f sur $\overset{\circ}{K}$ car ce sont les seuls points intérieurs à K où f peut présenter un extremum.
- On étudie les *extrema* de f sur le bord ∂K.
- Si f admet des *extrema* sur $\overset{\circ}{K}$, on les compare avec ceux sur ∂K. Si f n'admet aucun extremum sur $\overset{\circ}{K}$, c'est nécessairement sur ∂K.

Exercice 14. Soit T est l'ensemble des couples $(x, y) \in \mathbb{R}^2$ solutions du système d'inéquations

$$x \geq \frac{1}{4}, \quad y \geq \frac{1}{4}, \quad x + y \leq \frac{3}{4}.$$

On note $\overset{\circ}{T}$ l'*intérieur* de T, à savoir l'ensemble des couples $(x,y) \in \mathbb{R}^2$ solutions du système d'inéquations

$$x > \frac{1}{4}, \quad y > \frac{1}{4}, \quad x+y < \frac{3}{4}.$$

Soit f la fonction définie sur T par

$$f(x,y) = \frac{1}{x} + \frac{1}{y} - \frac{2}{x+y}.$$

(1) Représenter, sur un même graphique, les domaines T et $\overset{\circ}{T}$. Justifier qu'ils sont tous les deux bornés.

(2) (a) Justifier que f est de classe $\mathcal{C}^2$ sur $\overset{\circ}{T}$ et continue sur T.

(b) Déterminer les dérivées partielles d'ordre 1 sur $\overset{\circ}{T}$ de la fonction f.

(c) Montrer que f n'admet pas de point critique sur $\overset{\circ}{T}$.

(3) Montrer alors que f admet un maximum et un minimum sur T, atteint sur ∂T et dont on précisera les valeurs.

5 Recherche d'*extrema* sous contrainte linéaire

On s'intéresse dans ce dernier paragraphe à la recherche d'*extrema* sous une contrainte linéaire.

Prenons un exemple. On dispose de 2α mètres de clôture que l'on cherche à poser de sorte à former un terrain rectangulaire d'aire maximale. On cherche donc à *optimiser* la valeur de $f(x,y) = xy$, où $x > 0$ et $y > 0$ sont les longueurs du rectangle sous la contrainte linéaire $x + y = 2\alpha$. Les éléments introduits ci-dessous permettront (entre autres) de répondre à cette question.

Définition

On considère alors $g_1, g_2, .., g_p$ des éléments linéairement indépendants de $\mathcal{L}(\mathbb{R}^n,\ \mathbb{R})$ (ce sont des formes linéaires) et $b_1, ..., b_p$ des réels.

L'ensemble des solutions

$$\mathcal{C} = \{x \in \mathbb{R}^n :\ \forall i \in [\![1,p]\!],\ g_i(x) = b_i\}$$

est appelé **contrainte d'égalité linéaire**.
On lui associé un ensemble de solution du système homogène

$$\mathcal{H} = \{x \in \mathbb{R}^n :\ \forall i \in [\![1,p]\!],\ g_i(x) = 0\} = \bigcap_{i=1}^{p} \mathrm{Ker}(g_i).$$

Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert non vide U de $\mathbb{R}^n$. On suppose que $U \cap C \neq \emptyset$.
Soit $a \in U \cap \mathcal{C}$.
On dit que f présente un **extremum** local (resp. global) **sous la contraire** $\mathcal{C}$ en a si la restriction de f à $U \cap \mathcal{C}$ présente un extremum local (resp. global) en a.

Remarque

Si $a \in \mathcal{C}$ alors $\mathcal{C} = \{a + h :\ h \in \mathcal{H}\}$. En particulier, si $a \in U \cap \mathcal{C}$,

$$a + h \in U \cap \mathcal{C} \iff \begin{cases} a + h \in U \\ h \in \mathcal{H} \end{cases}$$

Proposition

Si $g_i \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ avec

$$g_i : \ x = (x_1, ..., x_n) \longmapsto \sum_{j=1}^{n} \alpha_{i,j} x_j,$$

alors g_i est affine (et donc de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$), son gradient est constant, vérifie donc, pour tout $y \in \mathbb{R}^n$, $\nabla g_i(y) = \begin{pmatrix} \alpha_{i,1} & ... & \alpha_{i,n} \end{pmatrix} =: \nabla g_i$ et il suit que

$$\forall x \in \mathbb{R}^n, \qquad g_i(x) = 0 \Longleftrightarrow \langle \nabla g_i, x \rangle = 0.$$

En particulier,

$$\mathcal{H}^\perp = \text{Vect}(\nabla g_1, \nabla g_2, ..., \nabla g_n).$$

Théorème

Condition nécessaire de premier ordre sous contraire.

Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert U et $a \in U \cap \mathcal{C}$.
Si f présente un extremum en a sous la contrainte $\mathcal{C}$, alors le gradient de f en a est orthogonal à tout élément $h \in \mathcal{H}$. Notamment la dérivée directionnelle de f dans la direction de tout élément h de $\mathcal{H}$ est nulle. Autrement dit

$$\nabla f(a) \in \mathcal{H}^\perp = \text{Vect}(\nabla g_1, \nabla g_2, ..., \nabla g_n)$$

et il existe des réels $\lambda_1, ..., \lambda_p$ tels que

$$\nabla f(a) = \sum_{i=1}^{p} \lambda_i \nabla g_i$$

Un tel point a est appelé **point critique de f pour l'optimisation sous la contrainte $\mathcal{C}$.**

À retenir!

Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert U. Pour déterminer les *extrema* de f sous la contrainte $\mathcal{C}$, on va donc

- Chercher les points critiques a de f pour l'optimisation sous la contrainte $\mathcal{C}$;
- Déterminer si chaque point critique précédent forte effectivement ou non un extremum local (ou global) en estimant le signe de $f(a+h) - f(a)$ pour $h \in \mathcal{H}$ tel que $a + h \in U$.

Exercice 15. Quel est alors le rectangle de périmètre 2α avec une aire maximale ?

Exercice 16. Soit $f : \ (x, y, z) \in \mathbb{R}^3 \longmapsto x^2 - 2y^2 + z^2$. Déterminer les *extrema* de f :

(i) sous la contrainte $x - 2y + z = 1$.

(ii) sous la contrainte $\begin{cases} x + y + z &= 3 \\ x - y &= -1 \end{cases}$

Exercice 17. Soient $X_1, X_2, ..., X_n$ des variables aléatoires indépendantes de lois $\mathcal{B}(p_i)$.
On pose $S_n = X_1 + .. + X_n$ et on suppose que $E(S_n) = \alpha$. Déterminer $p_1, ..., p_n$ de sorte que $V(S_n)$ soit maximale.

6 Une annale récente : EDHEC 2020

Soit f la fonction définie par

$$\forall (x,y,z) \in \mathbb{R}^3, \qquad f(x,y,z) = x\mathrm{e}^{x(y^2+z^2+1)}.$$

(1) Montrer que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^3$.

(2) Déterminer le seul point critique A de f.

(3) (a) Calculer les valeurs des dérivées partielles d'ordre 2 de f en A.
(b) Former la hessienne de f au point A et vérifier qu'elle est diagonale. Montrer que f présente un minimum local en A. Préciser la valeur de ce minimum.

(4) (a) Montrer que, pour tout (x,y,z) de $\mathbb{R}^3$, $f(x,y,z) \geq x\mathrm{e}^x$.
(b) Que peut-on en déduire pour le minimum de f trouvé à la question 3.(b) ?

(5) On souhaite étudier les *extrema* de f sous la contrainte linéaire

$$(C) : \begin{cases} x = 1 \\ x + y = 0 \end{cases}.$$

Montrer que, sous la contrainte (C), f présente un minimum global au point $(1,0,0)$. Quelle est sa valeur?

(6) On souhaite maintenant étudier les *extrema* de f sous la contrainte $(C') : x(y^2+z^2+1) = 1$. Montrer que f possède un maximum global sous la contrainte (C'). En quel point est-il atteint? Quelle est sa valeur?