CHAPITRE

XI

CALCUL DIFFÉRENTIEL

Sommaire

A - Éléments de topologie de $\mathbb{R}^n$

Définition XI-1

Soit $\mathbf{a} \in \mathbb{R}^n$ et $r \in]0, +\infty[$.

1. On appelle ***boule ouverte de centre* a *et de rayon*** r l'ensemble :

$$\mathrm{B}(a, r) = \{\mathbf{x} \in \mathbb{R}^n \ ; \ \|\mathbf{x} - \mathbf{a}\| < r\}.$$

C'est l'ensemble des vecteurs de $\mathbb{R}^n$ situés à une distance strictement inférieure à r de **a**.

2. On appelle ***boule fermée de centre* a *et de rayon*** r l'ensemble :

$$\mathrm{B}_f(a, r) = \{\mathbf{x} \in \mathbb{R}^n \ ; \ \|\mathbf{x} - \mathbf{a}\| \leqslant r\}.$$

C'est l'ensemble des vecteurs de $\mathbb{R}^n$ situés à une distance inférieure ou égale à r de **a**.

3. On appelle ***sphère de centre* a *et de rayon*** r l'ensemble :

$$\{\mathbf{x} \in \mathbb{R}^n \ ; \ \|\mathbf{x} - \mathbf{a}\| = r\}.$$

C'est l'ensemble des vecteurs de $\mathbb{R}^n$ situés à une distance égale à r de **a**.

Exemples

1 ▹ Dans $\mathbb{R}$:

2 ▹ Dans $\mathbb{R}^2$:

3 ▹ Dans $\mathbb{R}^3$:

Définition XI-2

On dit qu'une partie $\mathscr{U}$ de $\mathbb{R}^n$ est ***ouverte*** si :

$$\forall \mathbf{a} \in \mathscr{U},\ \exists r \in]0,+\infty[\ /\ \mathrm{B}(\mathbf{a},r) \subset \mathscr{U}.$$

Exemples

1 ▹ Dans $\mathbb{R}$, un intervalle $]a,b[$ et $\mathbb{R}^*$ sont des ouverts.

2 ▹ Dans $\mathbb{R}^n$, une boule ouverte est un ouvert mais une boule fermée n'est pas ouverte.

3 ▹ Dans $\mathbb{R}^n$, $\mathbb{R}^n$ est un ouvert.

4 ▹ Une union d'ouverts de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.

5 ▹ Une intersection **finie** d'ouverts de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.

Définition XI-3

On dit qu'une partie $\mathscr{F}$ de $\mathbb{R}^n$ est ***fermée*** si son complémentaire $\mathbb{R}^n \setminus \mathscr{F}$ est un ouvert de $\mathbb{R}^n$.

Exemples

1 ▹ Dans $\mathbb{R}$, un intervalle $[a, b]$, $\mathbb{R}_+$, $\mathbb{R}_-$ sont des fermés.

2 ▹ Dans $\mathbb{R}^n$, une boule fermée est un fermé.

3 ▹ Dans $\mathbb{R}^n$, $\mathbb{R}^n$ est un fermé.

4 ▹ Une intersection de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.

5 ▹ Une réunion **finie** de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.

Remarques

1 ▹ Un produit cartésien d'ouverts de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.

Un produit cartésien de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.

Par exemple, $(\mathbb{R}^*)^n$ est un ouvert de $\mathbb{R}^n$, $(\mathbb{R}_+)^2$ est un fermé de $\mathbb{R}^2$, $]a,b[\times]c,d[$ est un ouvert de $\mathbb{R}^2$.

2 ▹ Il existe des parties ni ouvertes, ni fermées (par exemple $\mathbb{R}_+^* \times \mathbb{R}_+$ dans $\mathbb{R}^2$).

3 ▹ Soit $\varphi : \mathbb{R}^n \to \mathbb{R}$ une fonction continue et $r > 0$:

- ▹ la partie $\{x \in \mathbb{R}^n \ ; \ \varphi(x) < r\}$ est un ouvert de $\mathbb{R}^n$;
- ▹ la partie $\{x \in \mathbb{R}^n \ ; \ \varphi(x) \leqslant r\}$ est un fermé de $\mathbb{R}^n$;
- ▹ la partie $\{x \in \mathbb{R}^n \ ; \ \varphi(x) = r\}$ est un fermé de $\mathbb{R}^n$.

Définition XI-4

On dit qu'une partie A de $\mathbb{R}^n$ est ***bornée*** s'il existe un réel $K \geqslant 0$ tel que, pour tout $\mathbf{x} \in A$, $\|\mathbf{x}\| \leqslant K$.

Cela revient à dire que A est contenue dans une boule centrée en l'origine.

Exercice C-131

Préciser si les ensembles suivants sont des parties ouvertes, fermées, bornées :

1. $A_1 = \left\{(x,y) \in \mathbb{R}^2 \ ; \ x + 2y = 5\right\}$;
2. $A_2 = \left\{(x,y) \in \mathbb{R}^2 \ ; \ x > 0, y > 0 \text{ et } x + y < 1\right\}$;
3. $A_3 = \{\mathbf{x} \in \mathbb{R}^n \ ; \ \langle \mathbf{u}, \mathbf{x} \rangle \geqslant 0\}$;
4. $A_4 = \{\mathbf{x} \in \mathbb{R}^n \ ; \ \langle \mathbf{u}, \mathbf{x} \rangle = 0\}$ où $\mathbf{u} \in \mathbb{R}^n \backslash \{0\}$.

Exercice C-132

Montrer que tout sous-espace vectoriel de $\mathbb{R}^n$ est un fermé.

Remarque

Ce qui a été étudié dans le premier chapitre sur les fonctions de plusieurs variables (les définitions et résultats concernant la continuité ou le caractère $\mathscr{C}^1$, ainsi que la définition des dérivées partielles) se généralise ou est encore vrai pour des fonctions définies sur un ouvert (non vide) de $\mathbb{R}^n$.

Par exemple, si $\mathscr{U}$ est un ouvert de $\mathbb{R}^n$ alors une fonction $f : \mathscr{U} \to \mathbb{R}$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ si les dérivées partielles existent et sont continues en chaque point de $\mathscr{U}$.

Pour tout $\mathbf{a} \in \mathscr{U}$, la fonction f admet alors en $\mathbf{a}$ un unique développement limité à l'ordre 1 : il existe $\varepsilon : \mathscr{U} \to \mathbb{R}$ telle que, pour $\mathbf{h} \in \mathbb{R}^n$ vérifiant $\mathbf{a} + \mathbf{h} \in \mathscr{U}$,

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \langle \nabla(f)(\mathbf{a}), \mathbf{h} \rangle + \|\mathbf{h}\|\,\varepsilon(\mathbf{h}) \quad \text{avec} \quad \varepsilon(\mathbf{h}) \underset{\mathbf{h}\to 0_{\mathbb{R}^n}}{\longrightarrow} 0.$$

B - Fonctions de classe $\mathscr{C}^2$ sur un ouvert de $\mathbb{R}^n$

B.1 - Classe $\mathscr{C}^2$ et matrice hessienne

Définition XI-5

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

On dit qu'une fonction $f : \mathscr{U} \to \mathbb{R}$ est ***de classe $\mathscr{C}^2$ sur $\mathscr{U}$*** si, pour tout $(i, j) \in [\![1, n]\!]^2$, la dérivée partielle d'ordre 2 d'indice (i, j) existe et est continue sur $\mathscr{U}$.

Remarque

Comme pour le cas $\mathscr{C}^1$, les combinaisons linéaires, les produits et quotients de fonctions de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ ou encore la composition par une fonction d'une variable de classe $\mathscr{C}^2$ restent de classe $\mathscr{C}^2$ sur $\mathscr{U}$.

Exemples

1 ▹ Montrons que $f :]0, +\infty[^2 \to \mathbb{R}, x \mapsto \ln(x) + \ln(y) - xy^2$ est de classe $\mathscr{C}^2$.

2 ▹ Les fonctions polynomiales sont de classe $\mathscr{C}^2$ sur tout ouvert de $\mathbb{R}^n$.

3 ▹ Montrons que $f : \mathbb{R}^n \to \mathbb{R}, (x_1, \ldots, x_n) \mapsto \mathrm{e}^{-(x_1^2+\cdots+x_n^2)}$ est de classe $\mathscr{C}^2$.

Définition XI-6

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$ et $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles d'ordre 2 pour tout couple $(i,j) \in [\![1,n]\!]^2$.

La ***matrice hessienne*** de f en $\mathbf{a} \in \mathscr{U}$, notée $\nabla^2 f(\mathbf{a})$, est la matrice de $\mathscr{M}_n(\mathbb{R})$, définie par :

$$\nabla^2 f(\mathbf{a}) = \left(\partial^2_{i,j} f(\mathbf{a})\right)_{(i,j)\in[1,n]]^2} = \begin{pmatrix} \partial^2_{1,1} f(\mathbf{a}) & \partial^2_{1,2} f(\mathbf{a}) & \cdots & \partial^2_{1,n} f(\mathbf{a}) \\ \partial^2_{2,1} f(\mathbf{a}) & \partial^2_{2,2} f(\mathbf{a}) & \cdots & \partial^2_{2,n} f(\mathbf{a}) \\ \vdots & \vdots & \vdots & \vdots \\ \partial^2_{n,1} f(\mathbf{a}) & \partial^2_{2,2} f(\mathbf{a}) & \cdots & \partial^2_{n,n} f(\mathbf{a}) \end{pmatrix}.$$

Exemples

1 ▹ Considérons la fonction polynomiale (de classe $\mathscr{C}^2$) définie sur $\mathbb{R}^2$ par :

$$f(x,y) = ax^2 + by^2 + cxy + dx + ey + f.$$

2 ▹ Considérons la fonction $g : \mathbb{R}^n \to \mathbb{R}, \mathbf{x} \mapsto \|\mathbf{x}\|^2$ c'est-à-dire :

$$g(x_1, \ldots, x_n) = x_1^2 + x_2^2 + \ldots + x_n^2 = \sum_{i=1}^{n} x_i^2.$$

3 ▹ Considérons la fonction $h : \mathbb{R}^n \to \mathbb{R}, (x_1, \dots, x_n) \mapsto \sum_{i=1}^{n} x_i^2 + \left(\sum_{i=1}^{n} x_i \right)^2 - \sum_{i=1}^{n} x_i$.

Théorème XI-7 (de Schwarz)

Soit f définie sur un ouvert $\mathscr{U}$ de $\mathbb{R}^n$.

Si f est de classe $\mathscr{C}^2$ sur $\mathscr{U}$ alors, pour tout $(i, j) \in [\![1, n]\!]^2$ et tout $\mathbf{a} \in \mathscr{U}$: $\partial^2_{j,i} f(\mathbf{a}) = \partial^2_{i,j} f(\mathbf{a})$.

Corollaire XI-8

La matrice hessienne en un point d'une fonction de classe $\mathscr{C}^2$ est symétrique.

Corollaire XI-9

La matrice hessienne en un point d'une fonction de classe $\mathscr{C}^2$ est diagonalisable dans une base orthonormée.

Exercice C-133

Soit f la fonction définie sur $\mathbb{R}^3$ par : $f(x, y, z) = x\mathrm{e}^{x(y^2+z^2+1)}$.

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

2. Déterminer le seul point critique $\mathbf{a}$ de f.

3. Écrire la hessienne de f au point $\mathbf{a}$ et vérifier qu'elle est diagonale.

B.2 - Forme quadratique et développement limité d'ordre 2

On rappelle que la forme quadratique associée à une matrice symétrique $A \in \mathscr{M}_n(\mathbb{R})$ est l'application définie sur $\mathbb{R}^n$ par : $q(\mathbf{h}) = {}^t\mathrm{HAH}$ où H est la matrice des coordonnées de $\mathbf{h}$ dans la base canonique de $\mathbb{R}^n$.

On a vu que :

- $(\forall \mathbf{u} \in \mathbb{R}^n,\ q(\mathbf{u}) \geqslant 0) \iff \mathrm{Sp}(A) \subset \mathbb{R}_+$;
- $(\forall \mathbf{u} \in \mathbb{R}^n,\ q(\mathbf{u}) \leqslant 0) \iff \mathrm{Sp}(A) \subset \mathbb{R}_-$.

Formes quadratiques et matrices hessiennes

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$ et $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^2$.

On peut considérer la forme quadratique associée à la matrice hessienne de f en $\mathbf{a}$, notée $q_{\mathbf{a}}$ et définie par :

$$\forall \mathbf{h} = (h_1, \dots, h_n) \in \mathbb{R}^n,\ q_{\mathbf{a}}(\mathbf{h}) = {}^t\mathrm{H}\left(\nabla^2 f(\mathbf{a})\right)\mathrm{H} \quad \text{où } \mathrm{H} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

c'est-à-dire :

$$q_{\mathbf{a}}(\mathbf{h}) = \sum_{(i,j)\in[\![1,n]\!]^2} \partial^2_{i,j} f(\mathbf{a}) h_i h_j = \sum_{i=1}^{n} \partial^2_{ii} f(\mathbf{a}) h_i^2 + 2 \sum_{i<j} \partial^2_{i,j} f(\mathbf{a}) h_i h_j.$$

Dérivées directionnelles

Soit $\mathbf{h} \in \mathbb{R}^n$ tel que le segment $[\mathbf{a}; \mathbf{a}+\mathbf{h}] = \{\mathbf{a} + t\mathbf{h}\ ;\ t \in [0,1]\}$ soit inclus dans $\mathscr{U}$.

On considère la fonction $g_{\mathbf{a},\mathbf{h}} : [0,1] \to \mathscr{U}, t \mapsto f(\mathbf{a} + t\mathbf{h})$.

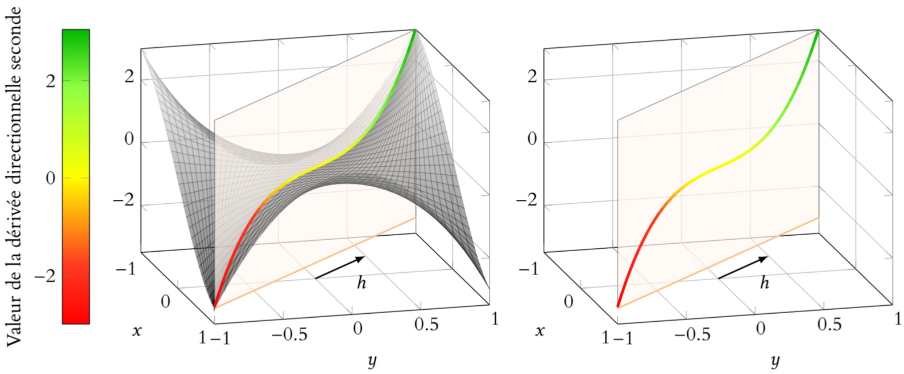
Alors la fonction $g_{\mathbf{a},\mathbf{h}}$ est de classe $\mathscr{C}^2$ sur $[0,1]$ et : $g''_{\mathbf{a},\mathbf{h}}(t) = q_{\mathbf{a}+t\mathbf{h}}(\mathbf{h})$.

En particulier : $g''_{\mathbf{a},\mathbf{h}}(0) = q_{\mathbf{a}}(\mathbf{h})$.

Remarque

Le nombre $q_{\mathbf{a}}(\mathbf{h}) = g''_{\mathbf{a},\mathbf{h}}(0)$ est appelé la *dérivée directionnelle seconde* de f en $\mathbf{a}$ dans la direction $\mathbf{h}$.

Son signe indique la convexité de la fonction $t \mapsto f(\mathbf{a} + t\mathbf{h})$.



Démonstration

On sait que $g_{\mathbf{a},\mathbf{h}}$ est une fonction de classe $\mathscr{C}^1$ sur $[0,1]$ et, pour tout $t \in [0,1]$, on a :

$$g'_{\mathbf{a},\mathbf{h}}(t) = \langle \nabla f(\mathbf{a} + t\mathbf{h}), \mathbf{h} \rangle = \sum_{i=1}^{n} \partial_i f(\mathbf{a} + t\mathbf{h}) h_i \quad \text{où } \mathbf{h} = (h_1, \dots, h_n).$$

Comme f est de classe $\mathscr{C}^2$, les $\partial_i f$ sont de classe $\mathscr{C}^1$ donc la fonction $t \mapsto \partial_i f(\mathbf{a} + t\mathbf{h})$ est dérivable et l'expression de sa dérivée est :

$$\sum_{j=1}^{n} \partial^2_{j,i} f(\mathbf{a} + t\mathbf{h}) h_j.$$

On en déduit que $g'_{\mathbf{a},\mathbf{h}}$ est dérivable et l'expression de sa dérivée est :

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\partial^2_{j,i}f(\mathbf{a}+t\mathbf{h})h_j\right)h_i = q_{\mathbf{a}+t\mathbf{h}}(\mathbf{h}).$$

Proposition XI-10 (développement limité à l'ordre 2)

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$ et $f:\mathscr{U}\to\mathbb{R}$ de classe $\mathscr{C}^2$.

Pour tout $\mathbf{a}\in\mathscr{U}$, il existe un voisinage $\mathscr{V}$ de $0_{\mathbb{R}^n}$ dans $\mathbb{R}^n$, une fonction $\varepsilon:\mathscr{V}\to\mathbb{R}$ telle que pour tout $\mathbf{h}\in\mathscr{V}$, on ait :

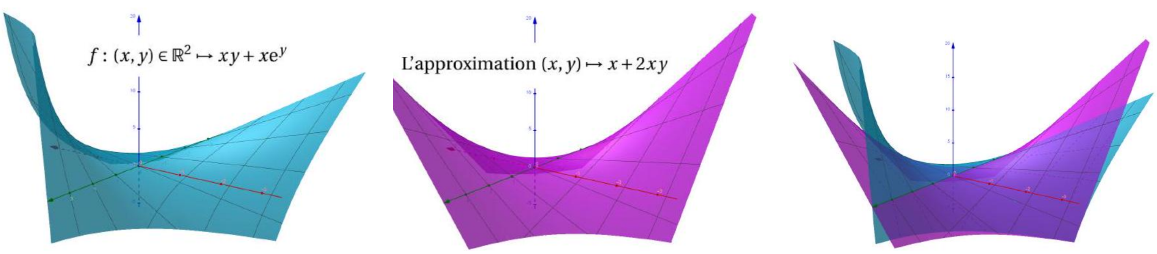
$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \langle\nabla f(\mathbf{a}),\mathbf{h}\rangle + \frac{1}{2}q_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|^2\varepsilon(\mathbf{h}) \quad \text{et} \quad \varepsilon(\mathbf{h}) \underset{\mathbf{h}\to 0_{\mathbb{R}^n}}{\longrightarrow} 0$$

où $q_{\mathbf{a}}$ est la forme quadratique associée à $\nabla^2 f(\mathbf{a})$.

Exemple

Soit $f:\mathbb{R}^2\to\mathbb{R}, (x,y)\mapsto xy+xe^y$.

Il s'agit d'une fonction de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$. Écrivons son développement limité en $(0,0)$.



C - Applications à l'optimisation

C.1 - Notion d'extremum

Définitions

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ et $\mathbf{a} \in \mathscr{U}$.

1. On dit que f a un ***maximum local*** en $\mathbf{a}$ s'il existe un réel $r > 0$ tel que :

$$\forall \mathbf{x} \in \mathrm{B}(\mathbf{a}, r),\ f(\mathbf{a}) \geqslant f(\mathbf{x}).$$

2. On dit que f a un ***minimum local*** en $\mathbf{a}$ s'il existe un réel $r > 0$ tel que :

$$\forall \mathbf{x} \in \mathrm{B}(\mathbf{a}, r),\ f(\mathbf{a}) \leqslant f(\mathbf{x}).$$

3. On dit que f a un ***extremum local*** en $\mathbf{a}$ si f a un maximum local ou un minimum local en $\mathbf{a}$.

Si de plus f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ alors on dit que $\mathbf{a}$ est un ***point critique*** de f lorsque $\nabla f(\mathbf{a}) = 0_{\mathbb{R}^n}$, c'est-à-dire :

$$\forall k \in [\![1, n]\!],\ \partial_k f(\mathbf{a}) = 0.$$

Proposition XI-11

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathscr{U}$ et $\mathbf{a} \in \mathscr{U}$.

Si f a un extremum en $\mathbf{a}$ alors $\mathbf{a}$ est un point critique de f.

Proposition XI-12 (condition d'existence d'un extremum global)

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ et A une partie de $\mathscr{U}$.

Si f est continue et si A est fermé borné alors la restriction de f à A admet un maximum global et un minimum global.

C'est-à-dire qu'il existe $\mathbf{a}$ et $\mathbf{b}$ dans A tels que :

$$\forall \mathbf{x} \in \mathrm{A},\ f(\mathbf{a}) \leqslant f(\mathbf{x}) \leqslant f(\mathbf{b}).$$

C.2 - Condition suffisante d'ordre 2

Proposition XI-13

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur $\mathscr{U}$ et $\mathbf{a} \in \mathscr{U}$ un point critique de f.

1. Si $\mathrm{Sp}\left(\nabla^2 f(\mathbf{a})\right) \subset \mathbb{R}_+^*$, alors f admet un minimum local en $\mathbf{a}$.

2. Si $\mathrm{Sp}\left(\nabla^2 f(\mathbf{a})\right) \subset \mathbb{R}_-^*$, alors f admet un maximum local en $\mathbf{a}$.

3. Si $\mathrm{Sp}\left(\nabla^2 f(\mathbf{a})\right)$ contient deux réels non nuls de signes distincts, alors f n'admet pas d'extremum en $\mathbf{a}$.

Démonstration

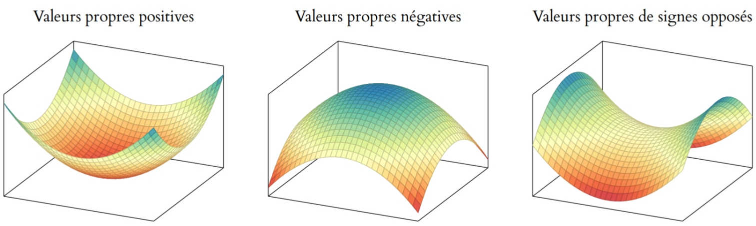
Remarques

1 ▹ Les réciproques sont fausses. Par exemple, la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = x^4 + y^4$ admet un minimum en $(0,0)$ mais la matrice hessienne est nulle en ce point et le spectre est donc réduit à $\{0\}$.

2 ▹ Considérons dans $\mathbb{R}^2$, la fonction $f_{\alpha,\beta}$ définie par : $f_{\alpha,\beta}(x,y) = \alpha x^2 + \beta y^2$.

Sa matrice hessienne est $\mathrm{H}_{\alpha,\beta} = \begin{pmatrix} 2\alpha & 0 \\ 0 & 2\beta \end{pmatrix}$.

Il y a donc trois cas selon que α et β soient tous deux positifs, tous deux négatifs ou de signes opposés.



Exercice C-134

Étudier les éventuels extremums de la fonction :

$$f :]0,+\infty[^2 \to \mathbb{R},\ (x,y) \mapsto f(x,y) = \frac{x}{y^2} + y^2 + \frac{1}{x}.$$

Exercice C-135

Étudier les éventuels extremums de la fonction :

$$f : \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto f(x,y) = x(x+1)^2 - y^2.$$

Exercice C-136

Étudier les éventuels extremums de la fonction :

$$f : \mathbb{R}^n \to \mathbb{R},\ (x_1,\dots,x_n) \mapsto f(x_1,\dots,x_n) = \sum_{k=1}^{n} x_k^2 + \left(\sum_{k=1}^{n} x_k\right)^2 - \sum_{k=1}^{n} x_k.$$

Exercice C-137

Soit f de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$, $(a,b) \in \mathscr{U}$ un point critique de f et $\mathrm{H} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$ la matrice hessienne de f en (a,b).

Comment peut-on caractériser les différents cas de la proposition à l'aide des réels r, s et t ?

C.3 - Exploitation de la convexité

Définition XI-14

Une partie C de $\mathbb{R}^n$ est dite ***convexe*** si, pour tout couple $(\mathbf{a},\mathbf{b})$ de C, le segment $[\mathbf{a},\mathbf{b}]$ est contenu dans C, autrement dit :

$$\forall(\mathbf{a},\mathbf{b})\in C^2,\ \forall\lambda\in[0,1],\ \lambda\mathbf{a}+(1-\lambda)\mathbf{b}\in C.$$

Proposition XI-15

Soit $\mathscr{U}$ un ouvert **convexe** de $\mathbb{R}^n$, $f:\mathscr{U}\to\mathbb{R}$ de classe $\mathscr{C}^2$ sur $\mathscr{U}$ et $\mathbf{a}\in\mathscr{U}$ un point critique de f.

1. Si pour tout $\mathbf{x}\in\mathscr{U}$, $\mathrm{Sp}\left(\nabla^2 f(\mathbf{x})\right)\subset\mathbb{R}_+$, alors f admet un minimum global en $\mathbf{a}$.

2. Si pour tout $\mathbf{x}\in\mathscr{U}$, $\mathrm{Sp}\left(\nabla^2 f(\mathbf{x})\right)\subset\mathbb{R}_-$, alors f admet un maximum global en $\mathbf{a}$.

Exemple

Étudions les extremums éventuels de $f:\mathbb{R}^3\to\mathbb{R},(x,y,z)\mapsto 11x^2+3y^2+3z^2-10xy+10xz-6yz+1$.

C.4 - Cas des extremums avec contrainte d'égalités linéaires

On considère le système linéaire suivant :

$$\begin{cases} g_1(\mathbf{x}) = b_1 \\ \vdots \\ g_p(\mathbf{x}) = b_p \end{cases}$$

où les g_i sont des formes linéaires sur $\mathbb{R}^n$ donc sont de la forme :

$$g_i : \mathbf{x} = (x_1, \dots, x_n) \mapsto \lambda_{i,1} x_1 + \dots + \lambda_{i,n} x_n = \langle (\lambda_{i,1}, \dots, \lambda_{i,n}), \mathbf{x} \rangle.$$

L'ensemble $\mathscr{C}$ des solutions de ce système correspond à $\displaystyle\bigcap_{i=1}^{p} \left\{ \mathbf{x} \in \mathbb{R}^n \ ; \ g_i(\mathbf{x}) = b_i \right\}$ donc est fermé.

Notons que la fonction g_i est de gradient constant : $\forall \mathbf{x} \in \mathbb{R}^n, \ \nabla g_i(\mathbf{x}) = (\lambda_{i,1}, \dots, \lambda_{i,n})$.

On note $\mathscr{H}$ l'ensemble des solutions du système linéaire homogène associé :

$$\begin{cases} g_1(\mathbf{x}) = 0 \\ \vdots \\ g_p(\mathbf{x}) = 0 \end{cases},$$

c'est un sous-espace vectoriel de $\mathbb{R}^n$ et on a (en notant ∇g_i au lieu de $\nabla g_i(\mathbf{x})$) :

$$\mathscr{H} = \bigcap_{i=1}^{p} \ker(g_i) = \bigcap_{i=1}^{p} \mathrm{Vect}(\nabla g_i)^{\perp} \ \ \textit{i.e.} \ \ \mathscr{H} = \mathrm{Vect}(\nabla g_1, \dots, \nabla g_p)^{\perp}.$$

Définition XI-16

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$ et $f : \mathscr{U} \to \mathbb{R}$.

On dit que f admet un ***extremum sous la contrainte*** $\mathscr{C}$ si la restriction de f à $\mathscr{C}$ admet un extremum.

Par exemple, f admet un maximum local en $\mathbf{a} \in \mathscr{C}$ sous la contrainte $\mathscr{C}$ s'il existe un réel $r > 0$ tel que :

$$\forall \mathbf{x} \in \mathrm{B}(\mathbf{a}, r) \cap \mathscr{C}, \ f(\mathbf{x}) \leqslant f(\mathbf{a}).$$

De même pour un minimum local, un maximum global ou un minimum global.

Proposition et définition XI-17

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ et $\mathbf{a} \in \mathscr{C}$.

Si f admet un extremum en $\mathbf{a}$ sous la contrainte $\mathscr{C}$ alors $\nabla f(\mathbf{a}) \in \mathscr{H}^{\perp}$.

Un vecteur $\mathbf{a} \in \mathscr{C}$ tel que $\nabla f(\mathbf{a}) \in \mathscr{H}^{\perp}$ est appelé un ***point critique de f sous la contrainte*** $\mathscr{C}$.

Corollaire XI-18

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ et $\mathbf{a} \in \mathscr{C}$.

Si f admet un extremum en $\mathbf{a}$ sous la contrainte $\mathscr{C}$ alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que :

$$\nabla f(\mathbf{a}) = \lambda_1 \nabla g_1(\mathbf{a}) + \cdots + \lambda_p \nabla g_p(\mathbf{a}).$$

On peut parfois écrire par abus de notation : $\nabla f(\mathbf{a}) = \lambda_1 \nabla g_1 + \cdots + \lambda_p \nabla g_p$.

Démonstration

Justifions la proposition dans le cas, par exemple, d'un maximum local. Il existe donc un réel $r > 0$ tel que :

$$\forall \mathbf{x} \in \mathrm{B}(\mathbf{a}, r) \cap \mathscr{C},\ f(\mathbf{x}) \leqslant f(\mathbf{a}).$$

Soit $\mathbf{h} \in \mathscr{H}$ et $g : t \mapsto f(\mathbf{a} + t\mathbf{h})$ (qui est définie au voisinage de 0 puisque $\mathbf{a} \in \mathscr{U}$ et $\mathscr{U}$ est un ouvert).

Pour t proche de 0, on a $\mathbf{a} + t\mathbf{h} \in \mathrm{B}(a, r)$ et d'autre part $\mathbf{a} + t\mathbf{h} \in \mathscr{C}$ (puisque $\mathbf{a} \in \mathscr{C}$ et $t\mathbf{h} \in \mathscr{H}$) donc :

$$g(t) = f(\mathbf{a} + t\mathbf{h}) \leqslant f(\mathbf{a}) = g(0)$$

donc g admet un maximum local en 0.

On a donc $g'(0) = 0$ or $g'(t) = \langle \nabla f(\mathbf{a} + t\mathbf{h}), \mathbf{h} \rangle$ donc :

$$\langle \nabla f(\mathbf{a}), \mathbf{h} \rangle = 0 \quad i.e. \quad \nabla f(\mathbf{a}) \perp \mathbf{h}.$$

On a donc $\nabla f(\mathbf{a}) \in \mathscr{H}^{\perp}$.

Exemples

1 ▹ Déterminons les extremums de la fonction $f : \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto x^2 + y^2 + 2z^2$ sous la contrainte $\mathscr{C} :\ 2x - y + z = 3$.

On note $\mathscr{H}$ l'ensemble des solutions de l'équation $2x - y + z = 0$ alors $\mathscr{H}^{\perp} = \mathrm{Vect}\big((2, -1, 1)\big)$.

Un vecteur $\mathbf{a} = (a_0, a_1, a_2)$ est un point critique de f sous la contrainte $\mathscr{C}$ si seulement si :

$$\begin{cases} 2a_0 - a_1 + a_2 = 3 \\ \exists \lambda \in \mathbb{R} \ /\ (2a_0, 2a_1, 4a_2) = \lambda(2, -1, 1) \end{cases} \qquad i.e. \qquad \begin{cases} 2a_0 - a_1 + a_2 = 3 \\ \exists \lambda \in \mathbb{R} \ /\ \begin{cases} a_0 = \lambda \\ 2a_1 = -\lambda \\ 4a_2 = \lambda \end{cases} \end{cases}$$

c'est-à-dire qu'il existe $\lambda \in \mathbb{R}$ tel que :

$$2\lambda - \Big(-\frac{1}{2}\lambda\Big) + \frac{1}{4}\lambda = 3,\ a_0 = \lambda,\ a_1 = -\frac{1}{2}\lambda,\ a_2 = \frac{1}{4}\lambda$$

soit :

$$\frac{11}{4}\lambda = 3,\ a_0 = \lambda,\ a_1 = -\frac{1}{2}\lambda,\ a_2 = \frac{1}{4}\lambda$$

donc si et seulement si :

$$a_0 = \frac{12}{11}, a_1 = -\frac{6}{11}, a_2 = \frac{3}{11}$$

donc f possède un unique point critique sous la contrainte $\mathscr{C}$, c'est $\mathbf{a} = \left(\frac{12}{11}, -\frac{6}{11}, \frac{3}{11}\right)$.

Soit $\mathbf{h}$ tel que $\mathbf{a}+\mathbf{h}\in\mathscr{C}$, on a :

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{h}\rangle + \frac{1}{2} q_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|^2\,\varepsilon(\mathbf{h})$$

avec $\varepsilon(\mathbf{h}) \xrightarrow[\mathbf{h}\to 0_{\mathbb{R}^n}]{} 0$.

On a $\mathbf{a}+\mathbf{h}\in\mathscr{C}$ or $\mathbf{a}\in\mathscr{C}$ donc $\mathbf{h}\in\mathscr{H}$ or $\nabla f(\mathbf{a})\in\mathscr{H}^{\perp}$ donc $\langle \nabla f(\mathbf{a}), \mathbf{h}\rangle = 0$.

On a $\nabla^2 f(\mathbf{a}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ donc $q_{\mathbf{a}}(\mathbf{h}) > 0$ pour tout $\mathbf{h}\in\mathscr{H}$ non nul.

Donc f admet un minimum en $\mathbf{a}$ sous la contrainte $\mathscr{C}$.

2 ▹ Déterminons les extremums de la fonction $f : \mathbb{R}^3 \to \mathbb{R}, (x,y,z) \mapsto 3x^2 + 2y^2 + z^2$ sous la contrainte $\mathscr{C} :\ 3x - 2y + z = 6$.

⋆ On note $\mathscr{H}$ l'ensemble des solutions de l'équation $3x - 2y + z = 0$ alors $\mathscr{H}^{\perp} = \mathrm{Vect}\big((3,-2,1)\big)$.

Un vecteur $\mathbf{a} = (a_0, a_1, a_2)$ est un point critique de f sous la contrainte $\mathscr{C}$ si seulement si :

$$\begin{cases} 3a_0 - 2a_1 + a_2 = 6 \\ \exists \lambda \in \mathbb{R} \ / \ (6a_0, 4a_1, 2a_2) = \lambda(3,-2,1) \end{cases} \qquad \textit{i.e.} \quad \begin{cases} 3a_0 - 2a_1 + a_2 = 6 \\ \exists \lambda\in\mathbb{R} \ / \begin{cases} 6a_0 = 3\lambda \\ 4a_1 = -2\lambda \\ 2a_2 = \lambda \end{cases} \end{cases}$$

c'est-à-dire qu'il existe $\lambda\in\mathbb{R}$ tel que :

$$\frac{3}{2}\lambda - 2\Big(-\frac{1}{2}\lambda\Big) + \frac{1}{2}\lambda = 6,\ a_0 = \frac{1}{2}\lambda,\ a_1 = -\frac{1}{2}\lambda,\ a_2 = \frac{1}{2}\lambda$$

soit :

$$3\lambda = 6,\ a_0 = \frac{1}{2}\lambda,\ a_1 = -\frac{1}{2}\lambda,\ a_2 = \frac{1}{2}\lambda$$

donc si et seulement si :

$$a_0 = 1,\ a_1 = -1,\ a_2 = 1$$

donc f possède un unique point critique sous la contrainte $\mathscr{C}$, c'est $\mathbf{a} = (1,-1,1)$.

Soit $\mathbf{h}$ tel que $\mathbf{a}+\mathbf{h}\in\mathscr{C}$, on a comme dans l'exemple précédent :

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + 0 + \frac{1}{2} q_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|^2\,\varepsilon(\mathbf{h})$$

avec $\varepsilon(\mathbf{h}) \xrightarrow[\mathbf{h}\to 0_{\mathbb{R}^n}]{} 0$.

De plus $\nabla^2 f(\mathbf{a}) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ donc $q_{\mathbf{a}}(\mathbf{h}) > 0$ pour tout $\mathbf{h}\in\mathscr{H}$ non nul.

Donc f admet un minimum en $\mathbf{a}$ sous la contrainte $\mathscr{C}$.

⋆ On peut aussi déterminer une base de $\mathscr{H}$: $\mathscr{H} = \mathrm{Vect}\big((1,0,-3),(0,1,2)\big)$ et l'on est amené à résoudre le système :

$$\begin{cases} 3x - 2y + z = 6 \\ \nabla f(x,y,z) \perp (1,0,-3) \\ \nabla f(x,y,z) \perp (0,1,2) \end{cases},$$

ce qui conduit au système :

$$\begin{cases} 3x - 2y + z = 6 \\ 6x - 6z = 0 \\ 4y + 4z = 0 \end{cases} \quad i.e. \quad \begin{cases} x & = 1 \\ y & = -1 \\ z & = 1 \end{cases}.$$

Puis on continue comme dans la première méthode.

⋆ On peut aussi se ramener à la détermination des extremums de la fonction $g : \mathbb{R}^2 \to \mathbb{R}$ définie par $g(x,y) = f(x,y,-3x+2y+6)$.

Exercice C-138

Étudier les extremums de $f : (x,y,z) \mapsto xy + yz + zx$ sous la contrainte $x + y + z = 1$.

Exercice C-139

On considère la fonction f définie sur $\mathbb{R}^4$ par $f : (x,y,z,t) \mapsto x^2 + y^2 + z^2 + t^2$.

Déterminer l'unique point critique de f sous les contraintes $x + y + z - t = 3$ et $2x - y + z + t = -6$, puis étudier si f admet un extremum sous ces contraintes.