

CALCUL DIFFÉRENTIEL

Dans tout ce chapitre,

- E, F, G désignent des **\mathbb{R} -espaces vectoriels de dimensions finies** ;
- \mathcal{U} et \mathcal{V} désignent des **ouverts** respectifs de E et F .

1 Différentiabilité

1.1 Dérivabilité selon un vecteur

Définition 1.1 Dérivée selon un vecteur

Soient $f : \mathcal{U} \rightarrow F, a \in \mathcal{U}$ et $v \in E$. On dit que f est **dérivable en a selon le vecteur v** si l'application $\varphi_{a,v} : t \mapsto f(a+tv)$ est dérivable en 0. Dans ce cas, on appelle **dérivée de f en a selon le vecteur v** le vecteur $\varphi'_{a,v}(0)$, que l'on note $D_v f(a)$.

REMARQUE. Si on note (f_1, \dots, f_n) les coordonnées de $f : \mathcal{U} \rightarrow F$ dans une base $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ de F (i.e. $f_i = \mathbf{f}_i^* \circ g$), alors f est dérivable en a selon le vecteur v si et seulement si les f_i le sont. De plus,

$$D_v f(a) = \sum_{i=1}^n D_v f_i(a) \mathbf{f}_i$$

Exemple 1.1

L'application $f : (x, y) \in \mathbb{R}^2 \mapsto (x^2 + y^2, 2xy)$ est dérivable en tout point $(a, b) \in \mathbb{R}^2$ selon tout vecteur $(u, v) \in \mathbb{R}^2$ et

$$D_{(u,v)} f(a, b) = 2(au + bv, av + bu)$$

Définition 1.2 Dérivées partielles dans une base

Soient $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ une base de $E, f : \mathcal{U} \rightarrow F$ et $a \in \mathcal{U}$. On appelle **dérivées partielles** de f dans la base $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ les applications $D_{\mathbf{e}_j} f$ si elles sont définies. On les note $\frac{\partial f}{\partial x_j}$ ou $\partial_j f$.

REMARQUE. Si $E = \mathbb{R}^p$ et qu'on ne précise pas la base dans laquelle on considère les dérivées partielles, c'est qu'on considère implicitement la **base canonique** de \mathbb{R}^p .

REMARQUE. Si on note $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ une base de E et (f_1, \dots, f_n) les coordonnées de $f : \mathcal{U} \rightarrow F$ dans une base $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ de F (i.e. $f_i = \mathbf{f}_i^* \circ g$), alors g admet des dérivées partielles en a dans la base $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ si et seulement si c'est également le cas pour les f_i . De plus,

$$\forall j \in \llbracket 1, p \rrbracket, \partial_j f(a) = \sum_{i=1}^n \partial_j f_i(a) \mathbf{f}_i$$

REMARQUE. Si $E = \mathbb{R}^2$, les variables d'une application $f : \mathbb{R}^2 \rightarrow F$ sont notées plus volontiers x et y que x_1 et x_2 . Les dérivées partielles dans la base canonique de \mathbb{R}^2 seront alors notées $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ plutôt que $\frac{\partial f}{\partial x_1}$ et $\frac{\partial f}{\partial x_2}$ ou $\partial_1 f$ et $\partial_2 f$.

De même, si $E = \mathbb{R}^3$, les dérivées partielles dans la base canonique seront plutôt notées $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ et $\frac{\partial f}{\partial z}$.

Méthode Calculer des dérivées partielles

Lorsque $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$, il est très aisé de calculer des dérivées partielles dans la base canonique. Il suffit de dériver chaque composante de la fonction par rapport à une variable les autres étant fixées.

Autrement dit, $\frac{\partial f}{\partial x_j}$ est la dérivée de l'application $x_j \mapsto f(x_1, \dots, x_n)$.

Exemple 1.2

L'application $f : (x, y) \in \mathbb{R}^2 \mapsto (x^2 + y^2, 2xy)$ admet des dérivées partielles dans la base canonique de \mathbb{R}^2 en tout point $(a, b) \in \mathbb{R}^2$ et

$$\frac{\partial f}{\partial x}(a, b) = 2(a, b) \quad \text{et} \quad \frac{\partial f}{\partial y}(a, b) = 2(b, a)$$

Exemple 1.3

On pose $\mathcal{U} = \{(x, y, z) \in \mathbb{R}^3, x + y^2 > 0\}$. L'application $f : (x, y, z) \in \mathcal{U} \mapsto (\ln(x + y^2), e^{xz})$ admet des dérivées partielles sur \mathcal{U} et

$$\frac{\partial f}{\partial x}(x, y, z) = \left(\frac{1}{x + y^2}, ze^{xz} \right) \quad \frac{\partial f}{\partial y}(x, y, z) = \left(\frac{2y}{x + y^2}, 0 \right) \quad \frac{\partial f}{\partial z}(x, y, z) = (0, xe^{xz})$$

Exemple 1.4

Les applications $\pi_i : (x_1, \dots, x_p) \in \mathbb{R}^p \mapsto x_i$ admettent des dérivées partielles en tout point de \mathbb{R}^p et

$$\frac{\partial \pi_i}{\partial x_j}(x_1, \dots, x_p) = \delta_{i,j}$$



ATTENTION ! Une fonction peut admettre des dérivées partielles sans être continue.

Exemple 1.5

Considérons la fonction

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

Alors f admet des dérivées partielles en tout point de \mathbb{R}^2 mais n'est pas continue en $(0, 0)$.

- Par opérations, f admet clairement des dérivées partielles en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$. De plus,

$$\forall x \in \mathbb{R}^*, \frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \quad \text{et} \quad \forall y \in \mathbb{R}^*, \frac{f(0, y) - f(0, 0)}{y - 0} = 0$$

donc $\frac{\partial f}{\partial x}(0, 0)$ et $\frac{\partial f}{\partial y}(0, 0)$ existent et $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$.

- Par contre, f n'est pas continue en $(0, 0)$ puisque, par exemple,

$$\forall t \in \mathbb{R}^*, f(t, t) = \frac{1}{2}$$

Ainsi $(t, t) \xrightarrow{t \rightarrow 0} (0, 0)$ mais $f(t, t) \xrightarrow{t \rightarrow 0} \frac{1}{2} \neq f(0, 0)$.



ATTENTION ! Une fonction peut même admettre des dérivées directionnelles selon tout vecteur sans être continue.

Exemple 1.6

Considérons la fonction

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Alors f admet des dérivées directionnelles selon tout vecteur en tout point de \mathbb{R}^2 mais n'est pas continue en $(0, 0)$.

- Par opérations, f admet clairement des dérivées directionnelles selon tout vecteur en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$. Soit $u = (h, k)$ un vecteur non nul de \mathbb{R}^2 . Alors

$$\forall t \in \mathbb{R}^*, \frac{f(tu) - f(0, 0)}{t} = \frac{hk^2}{h^2 + t^2k^4} \xrightarrow{t \rightarrow 0} \begin{cases} \frac{k^2}{h} & \text{si } k \neq 0 \\ 0 & \text{si } k = 0 \end{cases}$$

Donc f admet bien une dérivée directionnelle selon le vecteur u en $(0, 0)$.

- Par contre, f n'est pas continue en $(0, 0)$ puisque, par exemple,

$$\forall t \in \mathbb{R}^*, f(t^2, t) = \frac{1}{2}$$

Ainsi $(t^2, t) \xrightarrow{t \rightarrow 0} (0, 0)$ mais $f(t^2, t) \xrightarrow{t \rightarrow 0} \frac{1}{2} \neq f(0, 0)$.

1.2 Différentiabilité

Notation 1.1 Négligeabilité

Soit $f : \mathcal{U} \rightarrow F$. On suppose que $0_E \in \mathcal{U}$. Ecrire que $f(h) = o(h)$ signifie que $\lim_{h \rightarrow 0_E} \frac{f(h)}{\|h\|} = 0_F$.

REMARQUE. Les normes que l'on choisit sur E et F n'importent pas car toutes les normes sur un espace de dimension finie sont équivalentes.

Définition 1.3 Développement limité à l'ordre 1

Soit $f : \mathcal{U} \rightarrow F$. Une écriture du type

$$f(a + h) = c + L(h) + o(h)$$

avec $c \in F$ et $L \in \mathcal{L}(E, F)$ s'appelle un **développement limité** de f à l'ordre 1 en a . Si un tel développement limité existe, il est unique i.e. le vecteur c et l'application linéaire L sont uniques.

REMARQUE. Ceci signifie que

$$\lim_{h \rightarrow 0_E} \frac{f(a + h) - c - L(h)}{\|h\|} = 0_F$$

Définition 1.4 Différentiabilité en un point

Soit $f : \mathcal{U} \rightarrow F$. On dit que f est **différentiable** en $a \in \mathcal{U}$ si f admet un développement limité à l'ordre 1 en a . Dans ce cas, il existe une unique application linéaire $L \in \mathcal{L}(E, F)$ telle que

$$f(a+h) \underset{h \rightarrow 0_E}{=} f(a) + L(h) + o(h)$$

Cette application linéaire s'appelle la **différentielle** de f en a et se note $df(a)$.

REMARQUE. La différentielle de f en a est également appelée l'**application linéaire tangente** à f en a .

REMARQUE. Par souci de lisibilité, l'image d'un vecteur v par la différentielle de f en a se notera $df(a) \cdot v$ plutôt que $df(a)(v)$.

Proposition 1.1

Si on note (f_1, \dots, f_n) les coordonnées de $f : \mathcal{U} \rightarrow F$ dans une base $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ de F (i.e. $f_i = \mathbf{f}_i^* \circ g$), alors f différentiable en a si et seulement si les f_i le sont. De plus,

$$\forall v \in E, \quad df(a) \cdot v = \sum_{i=1}^n (df_i(a) \cdot v) \mathbf{f}_i$$

Exemple 1.7

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

$$\forall (h, k) \in \mathbb{R}^2, \quad f((a, b) + (h, k)) = f(a, b) + 2(ah + bk, bh + ak) + (h^2 + k^2, 2hk)$$

L'application

$$L : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R}^2 \\ (h, k) & \longmapsto 2(ah + bk, bh + ak) \end{cases}$$

est bien linéaire et

$$(h^2 + k^2, 2hk) \underset{(h,k) \rightarrow (0,0)}{=} o((h, k))$$

En effet, si l'on munit \mathbb{R}^2 de la norme définie par $\|(u, v)\| = |u| + |v|$

$$\|(h^2 + k^2, 2hk)\| = (|h| + |k|)^2 = \|(h, k)\|^2$$

de sorte que

$$\frac{\|(h^2 + k^2, 2hk)\|}{\|(h, k)\|} = \|(h, k)\| \underset{(h,k) \rightarrow (0,0)}{\longrightarrow} 0$$

On en déduit que f est différentiable en (a, b) et que $df(a, b)$ est l'endomorphisme $(h, k) \in \mathbb{R}^2 \mapsto 2(ah + bk, bh + ak)$.

Exemple 1.8 Différentielle de l'inversion matricielle

On considère l'application $f : M \in GL_n(\mathbb{R}) \mapsto M^{-1}$. On va montrer que f est différentiable sur $GL_n(\mathbb{R})$ et calculer sa différentielle.

Soit $M \in GL_n(\mathbb{R})$. Comme $GL_n(\mathbb{R})$ est ouvert, il existe un voisinage \mathcal{V} de 0 tel que $M + H$ est inversible pour tout $H \in \mathcal{V}$. Remarquons maintenant que

$$f(M + H) - f(M) = (M + H)^{-1} - M^{-1} = (M + H)^{-1}(I_n - (M + H)M^{-1}) = -(M + H)^{-1}HM^{-1}$$

puis

$$\begin{aligned} f(M + H) - f(M) + M^{-1}HM^{-1} &= M^{-1}HM^{-1} - (M + H)^{-1}HM^{-1} \\ &= (M + H)^{-1}((M + H)M^{-1} - I_n)HM^{-1} \\ &= f(M + H)(HM^{-1})^2 \end{aligned}$$

On munit $\mathcal{M}_n(\mathbb{R})$ d'une norme d'algèbre $\|\cdot\|$ de sorte que

$$\|f(M + H)(HM^{-1})^2\| \leq \|f(M + H)\| \|M^{-1}\|^2 \|H\|^2$$

puis pour $H \in \mathcal{V} \setminus \{0\}$,

$$\frac{\|f(M + H)(HM^{-1})^2\|}{\|H\|} \leq \|f(M + H)\| \|M^{-1}\|^2 \|H\|$$

f est continue sur $GL_n(\mathbb{R})$ ^a donc $\lim_{H \rightarrow 0} f(M + H) = f(M)$ puis $\lim_{H \rightarrow 0} \|f(M + H)\| = \|f(M)\|$. On en déduit que

$$\lim_{H \rightarrow 0} \frac{\|f(M + H)(HM^{-1})^2\|}{\|H\|} = 0$$

ou encore

$$f(M + H)(HM^{-1})^2 \underset{H \rightarrow 0}{=} o(H)$$

Finalement,

$$f(M + H) \underset{H \rightarrow 0}{=} f(M) - M^{-1}HM^{-1} + o(H)$$

L'application $H \mapsto -M^{-1}HM^{-1}$ est clairement linéaire : f est donc différentiable en M et $df(M)$ est l'application $H \mapsto -M^{-1}HM^{-1}$.

^aClassique : utiliser la formule de la comatrice.

Proposition 1.2

Si $f : \mathcal{U} \rightarrow F$ est **différentiable** en $a \in \mathcal{U}$, alors

- f est **continue** en a ;
- f admet des dérivées en a selon tout vecteur $v \in E$;
- $\forall v \in E, df(a) \cdot v = D_v f(a)$.

Définition 1.5 Différentiabilité sur un ouvert

Soit $f : \mathcal{U} \rightarrow F$. On dit que f est **différentiable** sur \mathcal{U} si f est différentiable en tout point de \mathcal{U} . L'application

$$df : \begin{cases} \mathcal{U} & \longrightarrow & \mathcal{L}(E, F) \\ a & \longmapsto & df(a) \end{cases}$$

s'appelle la **différentielle** de f sur \mathcal{U} .

Proposition 1.3 Cas particuliers

Soit $f : \mathcal{U} \rightarrow F$.

- Si f est **constante** sur \mathcal{U} , alors f est différentiable sur \mathcal{U} et df est **nulle** sur \mathcal{U} .
- Si f est la restriction à \mathcal{U} d'une **application linéaire** de E dans F , alors f est différentiable sur \mathcal{U} et $df = f$.
- Si \mathcal{U} est un intervalle ouvert de $E = \mathbb{R}$, alors f est différentiable en $a \in \mathcal{U}$ si et seulement si f est dérivable en a et, dans ce cas, $f'(a) = df(a) \cdot 1$.

Exemple 1.9

Les applications $\pi_i : (x_1, \dots, x_p) \in \mathbb{R}^p \mapsto x_i$ sont différentiables sur \mathbb{R}^p et $d\pi_i = \pi_i$.

1.3 Lien avec les dérivées partielles**Proposition 1.4 Lien entre différentielle et dérivées partielles**

Soient (e_1, \dots, e_p) une base de E et $f : \mathcal{U} \rightarrow F$. Si f est **différentiable** en a , alors f admet des **dérivées partielles** en a dans la base (e_1, \dots, e_p) et

$$\forall v \in E, df(a) \cdot v = \sum_{j=1}^p \partial_j f(a) e_j^*(v)$$

ou plus simplement

$$df(a) = \sum_{j=1}^p \partial_j f(a) e_j^*$$

REMARQUE. On en déduit un lien entre les dérivées directionnelles et les dérivées partielles si la fonction est **différentiable**. En effet

$$\forall v \in E, D_v f(a) = df(a) \cdot v = \sum_{j=1}^p \partial_j f(a) e_j^*(v)$$

REMARQUE. Dans le cas où $E = \mathbb{R}^n$, les formes linéaires $(x_1, \dots, x_n) \in \mathbb{R}^n \mapsto x_i$ sont souvent notées dx_i . Dans ce cas, on peut écrire pour $f : \mathcal{U} \mapsto F$ différentiable en $a \in \mathcal{U}$,

$$d(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

Exemple 1.10

Soit $f : (x, y) \in \mathbb{R}^2 \mapsto (x^2 + y^2, 2xy)$ Soit $(a, b) \in \mathbb{R}^2$. On a vu que

$$\frac{\partial f}{\partial x}(a, b) = 2(a, b) \quad \text{et} \quad \frac{\partial f}{\partial y}(a, b) = 2(b, a)$$

Comme f est différentiable sur \mathbb{R}^2 ,

$$df(a, b) \cdot (h, k) = 2h(a, b) + 2k(b, a) = 2(ah + bk, bh + ak)$$



ATTENTION ! Une fonction peut-être continue et admettre des dérivées selon tout vecteur sans pour autant être différentiable.

Exemple 1.11

Considérons la fonction

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Alors f admet des dérivées partielles en tout point de \mathbb{R}^2 mais n'est pas différentiable en $(0, 0)$.

- Par opérations, f est clairement continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Par ailleurs, on a classiquement $|xy| \leq \frac{1}{2}(x^2 + y^2)$ donc

$$\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}, |f(x, y)| \leq \frac{1}{2}|y|$$

On en déduit que

$$\lim_{(x, y) \rightarrow (0, 0)} f(x, y) = 0 = f(0, 0)$$

donc f est continue en $(0, 0)$.

- Par opérations, f admet clairement des dérivées directionnelles selon tout vecteur en tout point de $\mathbb{R}^2 \setminus \{(0, 0)\}$. Soit $u = (h, k)$ un vecteur non nul de \mathbb{R}^2 . Alors

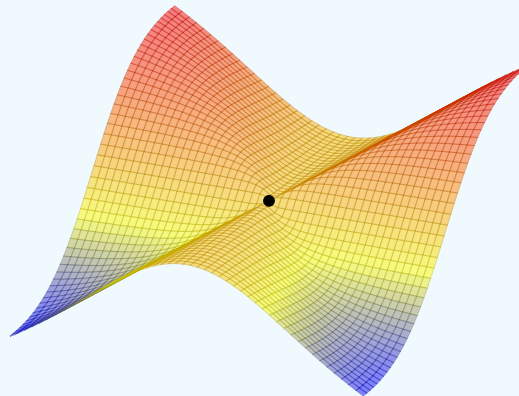
$$\forall t \in \mathbb{R}^*, \frac{f(tu) - f(0, 0)}{t} = f(h, k) = f(u)$$

Ainsi f admet une dérivée en $(0, 0)$ selon le vecteur u et $D_u f(0, 0) = f(u)$.

- Si f était différentiable en $(0, 0)$, alors on aurait

$$\forall u = (h, k) \in \mathbb{R}^2, D_u f(0, 0) = h \frac{\partial f}{\partial x}(0, 0) + k \frac{\partial f}{\partial y}(0, 0) = h D_{(1, 0)} f(0, 0) + k D_{(0, 1)} f(0, 0) = 0$$

Mais, par exemple, $D_{(1, 1)} f(0, 0) = \frac{1}{2} \neq (0, 0)$.

**Proposition 1.5 Matrice d'une différentielle dans un couple de bases**

Soient $\mathcal{E} = (\mathbf{e}_1, \dots, \mathbf{e}_p)$ une base de E , $\mathcal{F} = (\mathbf{f}_1, \dots, \mathbf{f}_n)$ une base de F et $f : \mathcal{U} \rightarrow F$. Notons (f_1, \dots, f_n) les coordonnées de f dans la base $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ (i.e. $f_i = \mathbf{f}_i^* \circ g$).

Si f est différentiable en $a \in \mathcal{U}$, alors la matrice de $df(a)$ dans les bases \mathcal{E} et \mathcal{F} est $(\partial_j f_i(a))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$.

Définition 1.6 Matrice jacobienne

Supposons $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$. Si $f : \mathcal{U} \rightarrow \mathbb{R}^n$ est différentiable en a , la matrice de $df(a)$ dans les bases canoniques de \mathbb{R}^p et \mathbb{R}^n s'appelle la **matrice jacobienne** de f en a .

Exemple 1.12

L'application $f : (x, y) \in \mathbb{R}^2 \mapsto (x^2 + y^2, xy)$ est différentiable sur \mathbb{R}^2 . Sa matrice jacobienne en un point $(a, b) \in \mathbb{R}^2$ est $\begin{pmatrix} 2a & 2b \\ 2b & 2a \end{pmatrix}$.

Exemple 1.13

On pose $\mathcal{U} = \{(x, y, z) \in \mathbb{R}^3, x + y^2 > 0\}$. L'application $f : (x, y, z) \in \mathcal{U} \mapsto (\ln(x + y^2), e^{xz})$ est différentiable sur \mathcal{U} . Sa matrice jacobienne en un point $(x, y, z) \in \mathcal{U}$ est $\begin{pmatrix} \frac{1}{x + y^2} & \frac{2y}{x + y^2} & 0 \\ ze^{xz} & 0 & xe^{xz} \end{pmatrix}$.

1.4 Gradient d'une fonction numérique

Dans ce paragraphe, $F = \mathbb{R}$.

Définition 1.7 Gradient

On suppose que E est un espace euclidien. Soit $f : \mathcal{U} \rightarrow \mathbb{R}$. Si f est différentiable en $a \in \mathcal{U}$, on appelle **gradient** de f en a , l'unique vecteur $\nabla f(a)$ de E tel que

$$\forall v \in E, df(a) \cdot v = \langle \nabla f(a), v \rangle$$

REMARQUE. On peut toujours munir un \mathbb{R} -espace vectoriel de dimension finie d'une structure d'espace euclidien. En effet, si (e_1, \dots, e_p) est une base de E , l'application

$$(x, y) \in E^2 \mapsto \sum_{k=1}^p e_k^*(x) e_k^*(y)$$

est un produit scalaire. De plus, (e_1, \dots, e_p) est alors une base orthonormale pour ce produit scalaire.

Exemple 1.14 Gradient du carré de la norme

Soit E un espace euclidien. Posons $f : x \in E \mapsto \|x\|^2$. Fixons $a \in E$.

$$\forall h \in E, f(a + h) = f(a) + 2\langle a, h \rangle + \|h\|^2$$

donc

$$f(a + h) \underset{h \rightarrow 0_E}{=} f(a) + 2\langle a, h \rangle + o(h)$$

Ainsi f est différentiable en a et $\nabla f(a) = 2a$.

Proposition 1.6 Coordonnées du gradient dans une base orthonormale

On suppose que E est un espace euclidien. Soit $f : \mathcal{U} \rightarrow \mathbb{R}$ différentiable en $a \in \mathcal{U}$. Les coordonnées de $\nabla f(a)$ dans une base orthonormale (e_1, \dots, e_p) de E sont les dérivées partielles de f dans cette base :

$$\nabla f(a) = \sum_{j=1}^p \partial_j f(a) e_j$$

REMARQUE. Si $E = \mathbb{R}^p$ est muni de son produit scalaire usuel, la base canonique est orthonormale. Par conséquent, si $f : \mathcal{U} \rightarrow \mathbb{R}$ est différentiable en $a \in \mathcal{U}$, alors

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_j}(a) \right)_{1 \leq j \leq n}$$

Interprétation géométrique du gradient

Si $\nabla f(a) \neq 0_E$, $\nabla f(a)$ est colinéaire et de même sens que le vecteur unitaire selon lequel la dérivée de f en a est maximale. Il suffit en effet de remarquer que pour tout vecteur v unitaire

$$D_v f(a) = \langle \nabla f(a), v \rangle \leq \|\nabla f(a)\| \|v\| = \|\nabla f(a)\|$$

avec égalité si et seulement si v et $\nabla f(a)$ sont colinéaires et de même sens (inégalité de Cauchy-Schwarz et cas d'égalité).

Exemple 1.15

Considérons l'application $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$. f est différentiable sur \mathbb{R}^2 car polynomiale. On munit \mathbb{R}^2 de son produit scalaire usuel de sorte que la base canonique de \mathbb{R}^2 est orthonormale. Alors

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) = (2x, 2y)$$

On peut retrouver la différentielle de f .

$$\forall (h, k) \in \mathbb{R}^2, df(x, y) \cdot (h, k) = \langle \nabla f(x, y), (h, k) \rangle = 2(xh + yk)$$

Exemple 1.16 Gradient et différentielle du déterminant

On munit $\mathcal{M}_n(\mathbb{R})$ de son produit scalaire usuel $(A, B) \mapsto \text{tr}(A^T B)$ et on considère l'application $\det : M \in \mathcal{M}_n(\mathbb{R}) \mapsto \det(M)$. Cette application est différentiable sur $\mathcal{M}_n(\mathbb{R})$ car polynomiale en les coefficients de la matrice.

Fixons $M \in \mathcal{M}_n(\mathbb{R})$. En développant le $\det(M)$ par rapport à sa $j^{\text{ème}}$ colonne, on obtient

$$\det(M) = \sum_{i=1}^n M_{i,j} \text{com}(M)_{i,j}$$

On en déduit notamment que

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \partial_{(i,j)} \det(M) = \text{com}(M)_{i,j}$$

où les $\partial_{(i,j)} \det$ désignent les dérivées partielles de \det dans la base canonique $(E_{i,j})_{1 \leq i, j \leq n}$ de $\mathcal{M}_n(\mathbb{R})$. Comme cette base est orthonormée,

$$\nabla \det(M) = \sum_{1 \leq i, j \leq n} \text{com}(M)_{i,j} E_{i,j} = \text{com}(M)$$

On peut alors retrouver la différentielle du déterminant :

$$\forall H \in \mathcal{M}_n(\mathbb{R}), d(\det)(M) \cdot H = \langle \nabla \det(M), H \rangle = \text{tr}(\text{com}(M)^T H)$$

2 Opérations sur les fonctions différentiables

Proposition 2.1 Combinaison linéaire

Soient $f : \mathcal{U} \rightarrow F$ et $g : \mathcal{U} \rightarrow F$. Si f et g sont différentiables en $a \in \mathcal{U}$, alors pour tout $(\lambda, \mu) \in \mathbb{R}^2$, $\lambda f + \mu g$ est différentiable en a et $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$.

Proposition 2.2

Soient F_1, \dots, F_p des \mathbb{R} -espaces vectoriels de dimensions finies, f_1, \dots, f_p des applications de \mathcal{U} à valeurs respectivement dans F_1, \dots, F_p et $M : \prod_{i=1}^p F_i \rightarrow G$ une application multilinéaire. Si f_1, \dots, f_p sont différentiables en $a \in \mathcal{U}$, alors $M(f_1, \dots, f_p)$ est différentiable en a et

$$d(M(f_1, \dots, f_p))(a) = M(df_1(a), f_2(a), \dots, f_p(a)) + M(f_1(a), df_2(a), \dots, f_p(a)) + \dots + M(f_1(a), \dots, f_{p-1}(a), df_p(a))$$

REMARQUE. De manière plus claire,

$$\forall h \in E, d(M(f_1, \dots, f_p))(a) \cdot h = M(df_1(a) \cdot h, f_2(a), \dots, f_p(a)) + \dots + M(f_1(a), \dots, f_{p-1}(a), df_p(a) \cdot h)$$

Exemple 2.1

Soit f un endomorphisme d'un espace euclidien E . Posons $\varphi(x) = \langle f(x), x \rangle$ pour tout $x \in E$. Comme f et Id_E sont différentiables sur E en tant qu'applications linéaires, φ est également différentiable sur E car le produit scalaire est bilinéaire. De plus, $df = f$ et $d\text{Id}_E = \text{Id}_E$ de sorte que

$$\forall (x, h) \in E^2, d\varphi(x) \cdot h = \langle f(h), x \rangle + \langle f(x), h \rangle$$

Proposition 2.3 Composition

Soient $f : \mathcal{U} \rightarrow F$ et $g : \mathcal{V} \rightarrow G$ telles que $f(\mathcal{U}) \subset \mathcal{V}$. Si f est différentiable en $a \in \mathcal{U}$ et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et $d(g \circ f)(a) = dg(f(a)) \circ df(a)$.

REMARQUE. Soient \mathcal{E} une base de E , \mathcal{F} une base de F et \mathcal{G} une base de G . Si A est la matrice de $df(a)$ dans les bases de \mathcal{E} et \mathcal{F} et B est la matrice de $dg(f(a))$ dans les bases \mathcal{F} et \mathcal{G} , alors BA est la matrice de $d(g \circ f)(a)$ dans les bases \mathcal{E} et \mathcal{G} .

Corollaire 2.1 Dérivée le long d'un arc

Soient I un intervalle de \mathbb{R} , $\gamma : I \rightarrow E$ et $f : \mathcal{U} \rightarrow F$ telles que $\gamma(I) \subset \mathcal{U}$. Si γ est dérivable en $t \in I$ et f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$.

Exemple 2.2 Dérivée le long d'une droite

Si $\gamma(t) = x + th$, alors $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot h = D_h f(\gamma(t))$.

Exemple 2.3

Si $E = \mathbb{R}^p$ et $\gamma = (x_1, \dots, x_p)$, alors

$$(f \circ \gamma)'(t) = \sum_{j=1}^p x_j'(t) \frac{\partial f}{\partial x_j}(\gamma(t))$$

Corollaire 2.2 Dérivées partielles d'une composée

Soient $f: \mathcal{U} \rightarrow F$ et $g: \mathcal{V} \rightarrow G$ telles que $f(\mathcal{U}) \subset \mathcal{V}$. On note $\frac{\partial f}{\partial x_j}$ les dérivées partielles de f dans une base (e_1, \dots, e_p) de E et $\frac{\partial g}{\partial y_i}$ les dérivées partielles de g dans une base (f_1, \dots, f_n) de F . Si f est différentiable en $a \in \mathcal{U}$ et si g est différentiable en $f(a)$, alors $g \circ f$ admet des dérivées partielles en a dans la base (e_1, \dots, e_p) et

$$\forall j \in \llbracket 1, p \rrbracket, \partial_j(g \circ f)(a) = \sum_{i=1}^n \partial_i g(f(a)) \partial_j f_i(a)$$

où (f_1, \dots, f_n) désignent les coordonnées de f dans la base (f_1, \dots, f_n) (i.e. $f_i = f_i^* \circ f$).

Méthode Règle de la chaîne

Soient x_1, \dots, x_n sont des fonctions différentiables sur un ouvert \mathcal{U} de \mathbb{R}^m à valeurs dans \mathbb{R} et f est une fonction différentiable sur un ouvert \mathcal{V} de \mathbb{R}^n . On suppose de plus que $(x_1, \dots, x_n)(\mathcal{U}) \subset \mathcal{V}$. Alors

$$\forall j \in \llbracket 1, m \rrbracket, \frac{\partial f}{\partial u_j}(u_1, \dots, u_m) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m)) \frac{\partial x_i}{\partial u_j}(u_1, \dots, u_m)$$

où on considère les dérivées partielles dans les bases canoniques de \mathbb{R}^m et \mathbb{R}^n .

Par abus de notation, on pourra tout simplement écrire

$$\forall j \in \llbracket 1, m \rrbracket, \frac{\partial f}{\partial u_j} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial u_j}$$

à condition de bien comprendre ce que l'on manipule.

Exemple 2.4 Coordonnées polaires

Soit f différentiable sur \mathbb{R}^2 . On pose

$$\forall (r, \theta) \in \mathbb{R}^2, g(r, \theta) = f(x(r, \theta), y(r, \theta))$$

avec

$$x(r, \theta) = r \cos \theta \quad \text{et} \quad y(r, \theta) = r \sin \theta$$

D'après la règle de la chaîne :

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) \frac{\partial x}{\partial r}(r, \theta) + \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \frac{\partial y}{\partial r}(r, \theta) \\ &= \cos \theta \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) + \sin \theta \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \\ \frac{\partial g}{\partial \theta}(r, \theta) &= \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) \frac{\partial x}{\partial \theta}(r, \theta) + \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \frac{\partial y}{\partial \theta}(r, \theta) \\ &= -r \sin \theta \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) + r \cos \theta \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \end{aligned}$$

Inversement, pour $r \neq 0$,

$$\begin{aligned} \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) &= \cos \theta \frac{\partial g}{\partial r}(r, \theta) - \frac{\sin \theta}{r} \frac{\partial g}{\partial \theta}(r, \theta) \\ \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) &= \sin \theta \frac{\partial g}{\partial r}(r, \theta) + \frac{\cos \theta}{r} \frac{\partial g}{\partial \theta}(r, \theta) \end{aligned}$$

Exemple 2.5 Gradient en coordonnées polaires

Soit f différentiable sur \mathbb{R}^2 . On pose

$$\forall (r, \theta) \in \mathbb{R}^2, g(r, \theta) = f(x(r, \theta), y(r, \theta))$$

avec

$$x(r, \theta) = r \cos \theta \quad \text{et} \quad y(r, \theta) = r \sin \theta$$

On a montré précédemment que, pour $r \neq 0$,

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \cos \theta \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) + \sin \theta \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \\ \frac{1}{r} \frac{\partial g}{\partial \theta}(r, \theta) &= -\sin \theta \frac{\partial f}{\partial x}(x(r, \theta), y(r, \theta)) + \cos \theta \frac{\partial f}{\partial y}(x(r, \theta), y(r, \theta)) \end{aligned}$$

En notant R la matrice de rotation d'angle θ , on a donc

$$\begin{pmatrix} \frac{\partial g}{\partial r}(r, \theta) \\ \frac{1}{r} \frac{\partial g}{\partial \theta}(r, \theta) \end{pmatrix} = R^{-1} \nabla f(x(r, \theta), y(r, \theta))$$

Ainsi en posant

$$\mathbf{u}_\theta = (\cos \theta, \sin \theta) \quad \text{et} \quad \mathbf{v}_\theta = (-\sin \theta, \cos \theta)$$

les coordonnées de $\nabla f(x(r, \theta), y(r, \theta))$ dans la base orthonormée $(\mathbf{u}_\theta, \mathbf{v}_\theta)$ sont

$$\left(\frac{\partial g}{\partial r}(r, \theta), \frac{1}{r} \frac{\partial g}{\partial \theta}(r, \theta) \right)$$

3 Applications de classe \mathcal{C}^k

3.1 Applications de classe \mathcal{C}^1

Définition 3.1 Application de classe \mathcal{C}^1

Soit $f : \mathcal{U} \rightarrow F$. On dit que f est de classe \mathcal{C}^1 sur \mathcal{U} si elle est différentiable sur \mathcal{U} et si df est continue sur \mathcal{U} .

Théorème 3.1 Classe \mathcal{C}^1 et dérivées partielles

Soit $f : \mathcal{U} \rightarrow F$. Alors f est de classe \mathcal{C}^1 sur \mathcal{U} si et seulement si

- f admet des dérivées partielles (dans une certaine base de E) en tout point de \mathcal{U} ;
- ces dérivées partielles sont continues sur \mathcal{U} .



ATTENTION ! Une fonction peut-être différentiable sans qu'elle soit de classe \mathcal{C}^1 . Notamment, les dérivées partielles d'une application différentiable ne sont pas nécessairement continues.

Exemple 3.1

Considérons la fonction

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} (x^2 + y^2) \sin\left(\frac{1}{x^2 + y^2}\right) & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Alors f est différentiable en $(0, 0)$ mais ses dérivées partielles n'y sont pas continues.

- Si l'on munit \mathbb{R}^2 de la norme euclidienne,

$$\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}, f(x, y) = \|(x, y)\|^2 \sin\left(\frac{1}{\|(x, y)\|^2}\right)$$

Comme \sin est bornée, il est clair que $f(x, y) \underset{(x, y) \rightarrow (0, 0)}{=} o(\|(x, y)\|)$. Ainsi f est bien différentiable en $(0, 0)$ et $df(0, 0)$ est nulle.

- Tout d'abord, pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 2x \sin\left(\frac{1}{x^2 + y^2}\right) - \frac{2x}{x^2 + y^2} \cos\left(\frac{1}{x^2 + y^2}\right) \\ \frac{\partial f}{\partial y}(x, y) &= 2y \sin\left(\frac{1}{x^2 + y^2}\right) - \frac{2y}{x^2 + y^2} \cos\left(\frac{1}{x^2 + y^2}\right) \end{aligned}$$

De plus,

$$\begin{aligned} \frac{f(x, 0) - f(0, 0)}{x - 0} &= x \sin\left(\frac{1}{x^2}\right) \xrightarrow{x \rightarrow 0} 0 \\ \frac{f(0, y) - f(0, 0)}{y - 0} &= y \sin\left(\frac{1}{y^2}\right) \xrightarrow{y \rightarrow 0} 0 \end{aligned}$$

Donc $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Pourtant,

$$\begin{aligned} \forall x \in \mathbb{R}^*, \frac{\partial f}{\partial x}(x, 0) &= 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x} \cos\left(\frac{1}{x^2}\right) \\ \forall y \in \mathbb{R}^*, \frac{\partial f}{\partial y}(0, y) &= 2y \sin\left(\frac{1}{y^2}\right) - \frac{2}{y} \cos\left(\frac{1}{y^2}\right) \end{aligned}$$

donc $x \mapsto \frac{\partial f}{\partial x}(x, 0)$ et $y \mapsto \frac{\partial f}{\partial y}(0, y)$ n'admettent pas de limite en 0 car la fonction $t \mapsto t \sin(1/t^2)$ admet une limite nulle en 0 mais la fonction $t \mapsto \frac{1}{t} \cos(1/t^2)$ n'admet pas de limite en 0. Ainsi $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ n'admettent pas de limite en $(0, 0)$. A fortiori, elles n'y sont pas continues.

Proposition 3.1 Intégrale curviligne

Soient $f : \mathcal{U} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 sur \mathcal{U} et $\gamma : [0, 1] \rightarrow \mathcal{U}$ de classe \mathcal{C}^1 sur $[0, 1]$. Alors

$$f(\gamma(1)) - f(\gamma(0)) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt$$

Corollaire 3.1 Applications constantes

On suppose \mathcal{U} connexe par arcs. Soit $f : \mathcal{U} \rightarrow \mathbb{F}$ de classe \mathcal{C}^1 sur \mathcal{U} . Alors f est constante sur \mathcal{U} si et seulement si df est nulle sur \mathcal{U} .

3.2 Applications de classe \mathcal{C}^k ($k \geq 1$)

On peut définir des dérivées partielles de dérivées partielles.

Définition 3.2 Dérivées partielles d'ordre k

Soit $f : \mathcal{U} \rightarrow \mathbb{F}$. On appelle **dérivée partielle d'ordre k** dans une base de E une dérivée partielle de la forme

$$\frac{\partial}{\partial x_{j_k}} \left(\frac{\partial}{\partial x_{j_{k-1}}} \left(\dots \left(\frac{\partial f}{\partial x_{j_1}} \right) \right) \right) = \partial_{j_k} (\partial_{j_{k-1}} (\dots (\partial_{j_1} f)))$$

que l'on notera plus simplement

$$\frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \dots \partial x_{j_1}} = \partial_{j_k} \partial_{j_{k-1}} \dots \partial_{j_1} f$$

REMARQUE. A priori, l'ordre des indices compte. Dans la définition, on dérive d'abord par rapport à la $j_1^{\text{ème}}$ coordonnée, puis par rapport à la $j_2^{\text{ème}}$ coordonnée, ..., et enfin par rapport à la $j_k^{\text{ème}}$ coordonnée.

REMARQUE. $\partial_i(\partial_i f)$ se note plus simplement $\partial_i^2 f$. De manière générale, $\partial_i^k f = \partial_i(\partial_i(\dots(\partial_i f)))$ (k dérivées partielles).

REMARQUE. $\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right)$ se note plus simplement $\frac{\partial^2 f}{\partial x_i^2}$. De manière générale, $\frac{\partial^k f}{\partial x_i^k} = \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_i} \left(\dots \left(\frac{\partial f}{\partial x_i} \right) \right) \right)$ (k dérivées partielles).

Exemple 3.2

Soit $f : (x, y) \mapsto xy^3 \ln(x^2 + y)$ définie sur l'ouvert $\mathcal{U} = \{(x, y) \in \mathbb{R}^2, x^2 + y > 0\}$. Alors f admet des dérivées partielles dans la base canonique en tout point de \mathcal{U} et

$$\begin{aligned} \forall (x, y) \in \mathcal{U}, \quad \frac{\partial f}{\partial x}(x, y) &= y^3 \ln(x^2 + y) + \frac{2x^2 y^3}{x^2 + y} \\ \forall (x, y) \in \mathcal{U}, \quad \frac{\partial f}{\partial y}(x, y) &= 3xy^2 \ln(x^2 + y) + \frac{xy^3}{x^2 + y} \end{aligned}$$

Ces dérivées partielles admettent elles-mêmes des dérivées partielles en tout point de \mathcal{U} et pour tout $(x, y) \in \mathcal{U}$,

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x, y) &= \frac{6xy^3}{x^2 + y} - \frac{4x^3 y^3}{(x^2 + y)^2} \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{6x^2 y^2}{x^2 + y} - \frac{2x^2 y^3}{(x^2 + y)^2} + 3y^2 \ln(x^2 + y) + \frac{y^3}{x^2 + y} \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) &= \frac{6x^2 y^2}{x^2 + y} - \frac{2x^2 y^3}{(x^2 + y)^2} + 3y^2 \ln(x^2 + y) + \frac{y^3}{x^2 + y} \\ \frac{\partial^2 f}{\partial y^2}(x, y) &= 6xy \ln(x^2 + y) + \frac{6xy^2}{x^2 + y} - \frac{xy^3}{(x^2 + y)^2} \end{aligned}$$

On constate notamment que $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$, ce qui n'est pas évident a priori même si ce n'est pas le fruit du hasard...

Définition 3.3 Applications de classe \mathcal{C}^k

Soit $f : \mathcal{U} \rightarrow F$. On dit que f est de classe \mathcal{C}^k ($k \in \mathbb{N}^*$) sur \mathcal{U} si toutes ses dérivées partielles d'ordre k existent et sont continues sur \mathcal{U} .

On dit que f est de classe \mathcal{C}^∞ sur \mathcal{U} si elle est de classe \mathcal{C}^k pour tout $k \in \mathbb{N}^*$.

Exemple 3.3

Toute application polynomiale sur \mathbb{R}^n est de classe \mathcal{C}^∞ sur \mathbb{R}^n .

Théorème 3.2 Schwarz

Soit $f : \mathcal{U} \rightarrow F$ de classe \mathcal{C}^2 sur \mathcal{U} . Alors

$$\forall (i, j) \in \llbracket 1, p \rrbracket^2, \partial_i \partial_j f = \partial_j \partial_i f$$



ATTENTION ! L'hypothèse que f est de classe \mathcal{C}^2 est primordiale.

Exemple 3.4

Soit en effet

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Tout d'abord, $\frac{\partial f}{\partial x}$ existe en $(x, y) \neq (0, 0)$ par opérations et en $(0, 0)$ (taux d'accroissement). De plus

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ \frac{\partial f}{\partial x}(0, 0) &= \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \end{aligned}$$

Comme $f(x, y) = -f(y, x)$, $\frac{\partial f}{\partial y}$ existe également en tout point $(x, y) \in \mathbb{R}^2$ et

$$\begin{aligned} \frac{\partial f}{\partial y}(x, y) &= -\frac{\partial f}{\partial x}(y, x) = -\frac{x(y^4 + 4x^2y^2 - x^4)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ \frac{\partial f}{\partial y}(0, 0) &= -\frac{\partial f}{\partial x}(0, 0) = 0 \end{aligned}$$

A l'aide de taux d'accroissement, on montre que $\frac{\partial^2 f}{\partial y \partial x}(0, 0)$ et $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$ existent et que

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x}(0, 0) &= \lim_{y \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0)}{y - 0} = -1 \\ \frac{\partial^2 f}{\partial x \partial y}(0, 0) &= \lim_{x \rightarrow 0} \frac{\frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0)}{x - 0} = 1 \end{aligned}$$

Ainsi

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) \neq \frac{\partial^2 f}{\partial x \partial y}(0, 0)$$

Le théorème de Schwarz permet en particulier d'affirmer que f n'est pas de classe \mathcal{C}^2 sur \mathbb{R}^2 .

Opérateurs différentiels

Hormis le gradient, on peut définir d'autres opérateurs différentiels.

- Si $f = (f_1, \dots, f_n)$ est un **champ de vecteurs** différentiable, autrement dit une application différentiable de \mathbb{R}^n dans \mathbb{R}^n , on peut définir sa **divergence** :

$$\mathbf{div} f = \sum_{i=1}^n \partial_i f_i$$

Par exemple, si $\vec{E} = (E_x, E_y, E_z)$ est un champ électrique,

$$\mathbf{div} \vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}$$

- Si $f = (f_x, f_y, f_z)$ est un champ de vecteurs différentiable de \mathbb{R}^3 dans \mathbb{R}^3 , on peut définir son **rotationnel** :

$$\mathbf{rot} f = \left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z}, \frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x}, \frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y} \right)$$

- Si f une application de \mathbb{R}^n dans \mathbb{R} de classe \mathcal{C}^2 , on peut définir son **laplacien** :

$$\Delta f = \sum_{i=1}^n \partial_i^2 f$$

Par exemple, si $V = (V_x, V_y, V_z)$ est un potentiel,

$$\Delta V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}$$

Exercice 3.1

1. Soit f une application de classe \mathcal{C}^2 de \mathbb{R}^n dans \mathbb{R} . Montrer que

$$\mathbf{div}(\nabla f) = \Delta f$$

2. Soit f une application de classe \mathcal{C}^2 de \mathbb{R}^3 dans \mathbb{R} . Montrer que

$$\mathbf{rot}(\nabla f) = 0$$

3. Soit $f = (f_x, f_y, f_z)$ une application de classe \mathcal{C}^2 de \mathbb{R}^3 dans \mathbb{R}^3 . Montrer que

$$\mathbf{rot}(\mathbf{rot} f) = \nabla(\mathbf{div} f) - \Delta f_x - \Delta f_y - \Delta f_z$$

3.3 Opérations

Proposition 3.2 Combinaison linéaire

Soient $f : \mathcal{U} \rightarrow F$ et $g : \mathcal{U} \rightarrow F$. Si f et g sont de classe \mathcal{C}^k sur \mathcal{U} , alors pour tout $(\lambda, \mu) \in \mathbb{R}^2$, $\lambda f + \mu g$ est de classe \mathcal{C}^k sur \mathcal{U} .

Proposition 3.3

Soient $f: \mathcal{U} \rightarrow F$, $g: \mathcal{U} \rightarrow G$ et $B: F \times G \rightarrow H$ une application **bilinéaire**. Si f et g sont de classe \mathcal{C}^k sur \mathcal{U} , alors $B(f, g)$ est de classe \mathcal{C}^k sur \mathcal{U} .

Proposition 3.4 Composition

Soient $f: \mathcal{U} \rightarrow F$ et $g: \mathcal{V} \rightarrow G$ telles que $f(\mathcal{U}) \subset \mathcal{V}$. Si f est de classe \mathcal{C}^k sur \mathcal{U} et g est différentiable sur \mathcal{V} , alors $g \circ f$ est de classe \mathcal{C}^k sur \mathcal{U} .

4 Tangence et orthogonalité

4.1 Vecteurs tangents

Définition 4.1 Vecteur tangent à une partie

Soient X une partie de E et $x \in X$. On dit que $v \in E$ est **tangent** à X en x s'il existe $\varepsilon > 0$ et un arc $\gamma:]-\varepsilon, \varepsilon[\rightarrow X$ dérivable en 0 tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

Notation 4.1

On note $T_x X$ l'ensemble des vecteurs tangents à X en x .

REMARQUE. Si $x \in \overset{\circ}{X}$, alors tout vecteur est tangent à X en x i.e. $T_x X = E$.

Soit $v \in E$. Comme $x \in X$, il existe $r > 0$ tel que $B(x, r) \subset X$. Posons $\varepsilon = \frac{r}{\|v\| + 1}$. Alors $\gamma:]-\varepsilon, \varepsilon[\mapsto x + tv$ est bien à valeurs dans $B(x, r) \subset X$, $\gamma(0) = x$ et $\gamma'(0) = v$.

Exemple 4.1 Vecteurs tangents à la sphère unité

On suppose que E est un espace euclidien. Soit $S = \{x \in E, \|x\| = 1\}$. Donnons-nous $x \in S$.

Soit v un vecteur tangent à S en x . Alors il existe $\gamma :]-\varepsilon, \varepsilon[\rightarrow S$ dérivable en 0 tel que $\gamma(0) = x$ et $\gamma'(0) = v$. L'application $\varphi = \|\gamma\|^2 = \langle \gamma, \gamma \rangle$ est donc constante égale à 1. Par bilinéarité du produit scalaire, φ est dérivable en 0 et

$$0 = \varphi'(0) = 2\langle \gamma'(0), \gamma(0) \rangle = \langle v, x \rangle$$

Ainsi $v \in \text{vect}(x)^\perp$.

Réciproquement soit $v \in \text{vect}(x)^\perp$.

- Posons $\varepsilon = \frac{\|x\|}{\|v\| + 1}$. Alors,

$$\forall t \in]-\varepsilon, \varepsilon[, \|x + tv\| \geq \|x\| - |t|\|v\| > \|x\| - \varepsilon\|v\| = \frac{\|x\|}{\|v\| + 1} \geq 0$$

donc $\gamma : t \mapsto \frac{x + tv}{\|x + tv\|}$ est bien défini sur $]-\varepsilon, \varepsilon[$ à valeurs dans S .

- Comme $\|x\| = 1$, $\gamma(0) = x$.
- Comme $\|x\| = 1$ et $\langle x, v \rangle = 0$, $\|x + tv\| = \sqrt{1 + t^2\|v\|^2}$. L'application $\varphi : t \mapsto x + tv$ est dérivable en 0 et $\varphi'(0) = v$ et l'application $\psi : t \mapsto \frac{1}{\sqrt{1 + t^2\|v\|^2}}$ est également dérivable en 0 et $\psi'(0) = 0$. Par opérations, γ est dérivable en 0 et

$$\gamma'(0) = \varphi(0)\psi'(0) + \varphi'(0)\psi(0) = v$$

Ainsi v est bien tangent à S en x .

Par double inclusion, $T_x S = \text{vect}(x)^\perp$.

REMARQUE. De manière générale, si $x \in S(a, r)$ (sphère de centre a et de rayon r), $T_x S(a, r) = \text{vect}(x - a)^\perp$.

Proposition 4.1 Cas du graphe d'une fonction de \mathbb{R}^2 dans \mathbb{R}

Soit Ω un ouvert de \mathbb{R}^2 et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω . Notons X le graphe de f , c'est-à-dire

$$X = \{(x, y, f(x, y)), (x, y) \in \Omega\} \subset \mathbb{R}^3$$

Alors pour tout $a \in \Omega$, les vecteurs tangents à X en a sont les vecteurs $(v, df(a) \cdot v) = (v, D_v f(a))$ où $v \in E$.

REMARQUE. Notamment, $\left(1, 0, \frac{\partial f}{\partial x}(a)\right)$ et $\left(0, 1, \frac{\partial f}{\partial y}(a)\right)$ sont des vecteurs tangents à X en a .

Plan tangent

Soit Ω un ouvert de \mathbb{R}^2 et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω . L'ensemble des vecteurs tangents en $(x_0, y_0, f(x_0, y_0))$ (avec $(x_0, y_0) \in \Omega$) au graphe de f est le plan vectoriel

$$P = \text{vect} \left(\left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0) \right), \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0) \right) \right)$$

On appelle **plan affine tangent** en $(x_0, y_0, f(x_0, y_0))$ au graphe de f le plan affine $\mathcal{P} = (x_0, y_0, z_0) + P$. On obtient une équation cartésienne de \mathcal{P} de la manière suivante :

$$\begin{aligned} (x, y, z) \in \mathcal{P} &\iff (x, y, z) - (x_0, y_0, z_0) \in \text{vect} \left(\left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0) \right), \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0) \right) \right) \\ &\iff \begin{vmatrix} x - x_0 & 1 & 0 \\ y - y_0 & 0 & 1 \\ z - f(x_0, y_0) & \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{vmatrix} = 0 \\ &\iff z = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) \end{aligned}$$

On peut également remarquer que \mathcal{P} est le plan passant par (x_0, y_0, z_0) et de vecteur normal $\left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right)$.

REMARQUE. On notera l'extrême similitude de cette équation avec l'équation de la tangente au graphe d'une fonction dérivable $f : \mathbb{R} \rightarrow \mathbb{R}$ en un point $(x_0, f(x_0))$, à savoir

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Exemple 4.2

Soit $f : (x, y) \in \mathbb{R}^2 \rightarrow x^2 + y^2$. Le plan affine tangent au graphe de f en $(x_0, y_0, f(x_0, y_0))$ admet pour équation

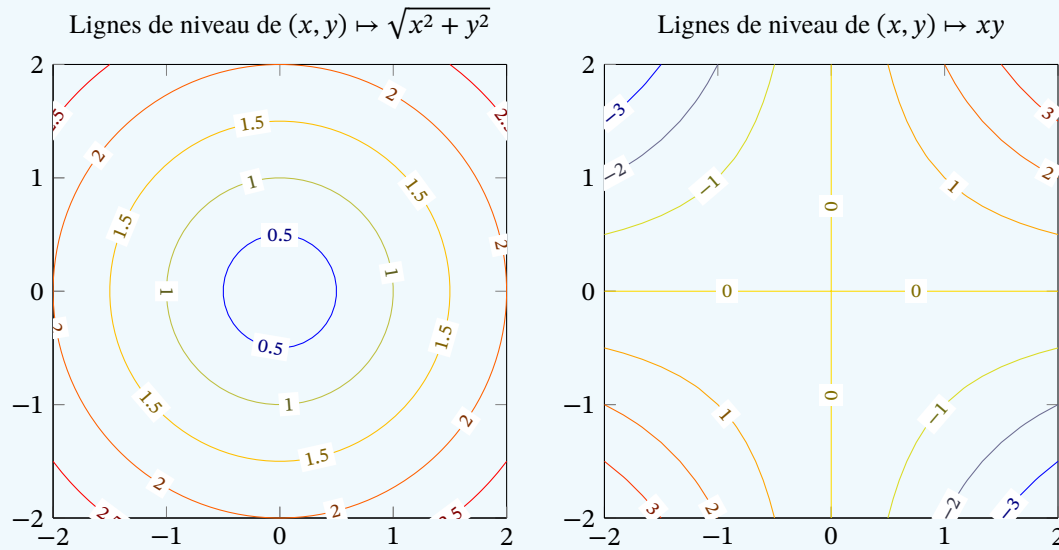
$$z = x_0^2 + y_0^2 + 2x_0(x - x_0) + 2y_0(y - y_0)$$

4.2 Lignes de niveau

Définition 4.2 Ligne de niveau

Soit $g : \mathcal{U} \rightarrow \mathbb{R}$. On appelle ligne de niveau $k \in \mathbb{R}$ de f l'ensemble

$$E_k = \{x \in \mathcal{U}, f(x) = k\}$$

Exemple 4.3**Proposition 4.2 Vecteurs tangents à une ligne de niveau**

Soient $g : \mathcal{U} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , X une ligne de niveau de g et $x \in X$ tel que $dg(x) \neq 0$. Alors $T_x X = \text{Ker } dg(x)$.

Proposition 4.3 Gradient et ligne de niveau

On suppose E euclidien. Soient $g : \mathcal{U} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , X une ligne de niveau de g et $x \in X$ tel que $\nabla g(x) \neq 0$. Alors $T_x X = \text{vect}(\nabla g(x))^\perp$.

Exemple 4.4 Plan tangent à une surface de \mathbb{R}^3

Si X est une surface de \mathbb{R}^3 d'équation cartésienne $g(x, y, z) = 0$ i.e.

$$X = \{(x, y, z) \in \mathcal{U}, g(x, y, z) = 0\}$$

alors le plan affine tangent à X en (x_0, y_0, z_0) est

$$(x_0, y_0, z_0) + \text{vect}(\nabla g(x_0, y_0, z_0))^\perp = (x_0, y_0, z_0) + \text{vect}\left(\frac{\partial g}{\partial x}(x_0, y_0, z_0), \frac{\partial g}{\partial y}(x_0, y_0, z_0), \frac{\partial g}{\partial z}(x_0, y_0, z_0)\right)^\perp$$

Il a donc pour équation

$$\frac{\partial g}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial g}{\partial z}(x_0, y_0, z_0)(z - z_0) = 0$$

REMARQUE. Notamment X est le graphe d'une application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, alors X est la ligne de niveau 0 de l'application $g : (x, y, z) \mapsto f(x, y) - z$. On en déduit que le plan affine tangent à X en un point $(x_0, y_0, f(x_0, y_0))$ est le plan passant par ce point et de vecteur normal

$$\nabla g(x_0, y_0, f(x_0, y_0)) = \left(\frac{\partial g}{\partial x}(x_0, y_0, f(x_0, y_0)), \frac{\partial g}{\partial y}(x_0, y_0, f(x_0, y_0)), \frac{\partial g}{\partial z}(x_0, y_0, f(x_0, y_0))\right) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1\right)$$

5 Optimisation des fonctions numériques

Dans tout ce paragraphe, $F = \mathbb{R}$.

5.1 Point critique

Définition 5.1 Point critique

Soit $f : \mathcal{U} \rightarrow \mathbb{R}$. Si f est différentiable en $a \in \mathcal{U}$, on dit que a est un **point critique** de f si $df(a) = 0_{E^*}$.

REMARQUE. a est un point critique de f si et seulement si toutes les dérivées partielles de f en a sont nulles.

REMARQUE. Si E est un espace euclidien (on peut toujours le supposer), a est un point critique de f si et seulement si $\nabla f(a) = 0_E$.

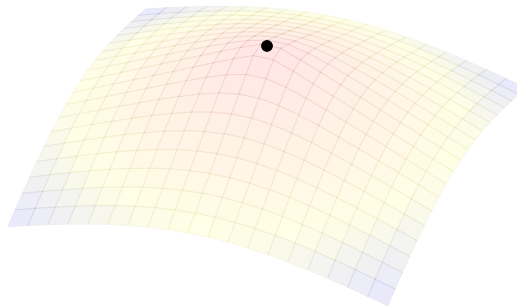
Exemple 5.1

Soit $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$. Puisque $\nabla f(x, y) = 2(x, y)$, l'unique point critique de f est $(0, 0)$.

Rappel Extremum local

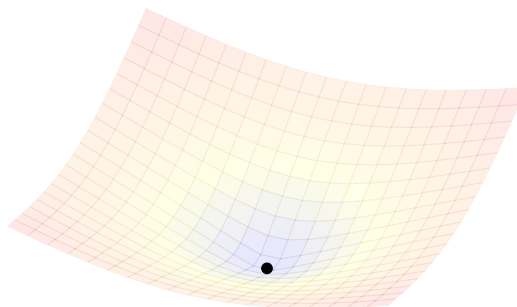
On dit que $f : D \rightarrow \mathbb{R}$ admet un **maximum local** en $a \in D$ s'il existe $\varepsilon > 0$ tel que

- $B(a, \varepsilon) \subset D$;
- $\forall x \in B(a, \varepsilon), f(x) \leq f(a)$.



On dit que $f : D \rightarrow \mathbb{R}$ admet un **minimum local** en $a \in D$ s'il existe $\varepsilon > 0$ tel que

- $B(a, \varepsilon) \subset D$;
- $\forall x \in B(a, \varepsilon), f(x) \geq f(a)$.



Si D est **ouvert**, il existe toujours $\varepsilon > 0$ tel que $B(a, \varepsilon) \subset D$. De manière générale, on ne parle d'extremum local en $a \in D$ que si a est un point **intérieur** à D .

Proposition 5.1 Condition nécessaire d'existence d'un extremum local

Soit $f: \mathcal{U} \rightarrow \mathbb{R}$ différentiable en $a \in \mathcal{U}$. Si f est différentiable en a et admet un **extremum local** en a , alors a est un **point critique** de f .



ATTENTION ! La réciproque est fausse. Considérons $f: (x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2$. Alors f est différentiable sur \mathbb{R}^2 et

$$\forall (x, y) \in \mathbb{R}^2, \nabla f(x, y) = 2(x, -y)$$

Ainsi $(0, 0)$ est bien l'unique point critique de f . Cependant

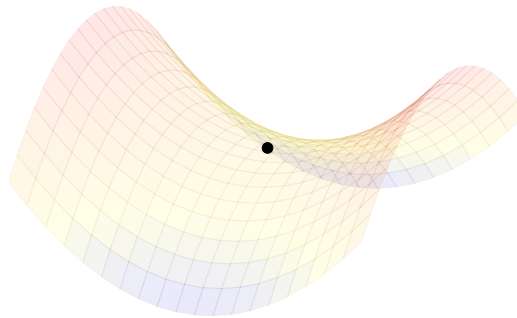
$$\forall \varepsilon > 0, f(\varepsilon, 0) = \varepsilon^2 > 0 = f(0, 0)$$

donc f n'admet pas de maximum local en $(0, 0)$ et

$$\forall \varepsilon > 0, f(0, \varepsilon) = -\varepsilon^2 < 0 = f(0, 0)$$

donc f n'admet pas non plus de minimum local en $(0, 0)$.

La fonction f n'admet donc pas d'extrema locaux sur \mathbb{R}^2 (et donc pas non plus d'extrema globaux) comme la représentation graphique suivante permet de s'en convaincre.

**Méthode Recherche d'extrema locaux**

1. On recherche les points critiques.
2. Pour chaque point critique a , on étudie le signe de $f(x) - f(a)$ pour x au voisinage de a . Pour simplifier, on pose généralement $u = x - a$ et on étudie le signe de $f(a + u) - f(a)$ pour u au voisinage de $0_{\mathbb{E}}$.

Exemple 5.2

Soit $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + xy + y^2 - 3x - 6y$. $(x, y) \in \mathbb{R}^2$ est un point critique de f si et seulement si

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$$

Ce système équivaut à

$$\begin{cases} 2x + y - 3 = 0 \\ x + 2y - 6 = 0 \end{cases}$$

L'unique solution de ce système i.e. l'unique point critique de f est $(0, 3)$. Pour $(u, v) \in \mathbb{R}^2$,

$$f(0 + u, 3 + v) - f(0, 3) = u^2 + uv + v^2 = \left(u + \frac{1}{2}v\right)^2 + \frac{3}{4}v^2 \geq 0$$

Ainsi f admet un minimum local (et même global) en $(0, 3)$.

Méthode Recherche d'extrema globaux

Soit D une partie de E (non nécessairement ouverte). Les extrema globaux d'une fonction f à valeurs réelles sur un domaine D sont

- soit atteints sur $D \setminus \mathring{D}$;
- soit atteints sur \mathring{D} , auquel cas ce sont des extrema locaux et donc **nécessairement** atteints en des points critiques de f .

Exemple 5.3

On recherche les extrema globaux de $f \mapsto xy(1 - x - y)$ sur

$$D = \{(x, y) \in \mathbb{R}^2, x \geq 0, y \geq 0, x + y \leq 1\}$$

- Tout d'abord, D est compact et f est continue donc f admet bien un minimum global et un maximum global sur D .
- On remarque d'abord que f est nulle sur la frontière de D (puisque alors $x = 0, y = 0$ ou $x + y = 1$). De plus, f est clairement positive sur D donc $\min_D f = 0$ et ce minimum est atteint en tout point de D .
- Recherchons les points critiques de f sur $\overset{\circ}{D}$. On résout le système

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$$

qui équivaut à

$$\begin{cases} y(1 - 2x - y) = 0 \\ x(1 - x - 2y) = 0 \end{cases}$$

Puisque l'on se situe sur la frontière de D , $x > 0$ et $y > 0$ donc le système équivaut à

$$\begin{cases} 1 - 2x - y = 0 \\ 1 - x - 2y = 0 \end{cases}$$

dont l'unique solution est $(1/3, 1/3)$ qui appartient bien à D . Comme $f(1/3, 1/3) > 0$, le maximum de f ne peut être atteint sur la frontière de D . Ce maximum global est donc un maximum local qui ne peut être atteint qu'en l'unique point critique $(1/3, 1/3)$ de f sur $\overset{\circ}{D}$. On en déduit que $\max_D f = f(1/3, 1/3) = 1/27$.

5.2 Optimisation sous contrainte**Proposition 5.2**

Soient $f : \mathcal{U} \rightarrow \mathbb{R}$ et $X \subset \mathcal{U}$. Si la restriction de f à X admet un extremum local en $x \in X$ et si f est différentiable en x , alors $df(x)$ est nulle sur $T_x X$.

Théorème 5.1 Extrema liés

Soient $f : \mathcal{U} \rightarrow \mathbb{R}$ et $g : \mathcal{U} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On pose $X = g^{-1}(\{0\})$. Si la restriction de f à X admet un extremum local en $x \in X$ et $dg(x) \neq 0$, alors $df(x)$ est colinéaire à $dg(x)$.

REMARQUE. Si E est un espace euclidien, alors la condition « $df(x)$ colinéaire à $dg(x)$ » équivaut à la condition « $\nabla f(x)$ colinéaire à $\nabla g(x)$ ».

Méthode Extrema locaux sous contrainte

Pour déterminer les extrema locaux d'une fonction f soumis à la contrainte $g(x) = 0$, on résout le système

$$\begin{cases} \nabla f(x) = \lambda \nabla g(x) \\ g(x) = 0 \end{cases}$$

sous réserve que $\nabla f(x) \neq 0$.

Exemple 5.4

Soit $f : (x, y, z) \in \mathbb{R}^3 \mapsto x^2 + y^2 + z^2$ et P le plan d'équation $x + 2y + 3z = 7$. Dans la suite, on munit \mathbb{R}^3 de sa structure euclidienne canonique. On souhaite déterminer le minimum éventuel de f sur P .

Justifions déjà l'existence de ce minimum. Remarquons déjà que $P = g^{-1}(\{0\})$ avec $g : (x, y, z) \in \mathbb{R}^3 \mapsto x + 2y + 3z - 1$. Notamment P est fermé car g est continue. Posons $K = [-7, 7]^3$. Il est clair que K est compact donc $K \cap P$ est fermé comme intersection d'un compact et d'un fermé. Ainsi f admet un minimum sur $K \cap P$. De plus, pour $(x, y, z) \in \mathbb{R}^3 \setminus K$, $f(x, y, z) > f(7, 0, 0)$ et $(7, 0, 0) \in P$. Ceci justifie que le minimum de f sur $K \cap P$ est bien le minimum de f sur P : la restriction de f à P admet donc un minimum en $(a, b, c) \in P$.

f est bien de classe \mathcal{C}^1 sur \mathbb{R}^3 et pour tout $(x, y, z) \in \mathbb{R}^3$, $\nabla f(x, y, z) = 2(x, y, z)$. g est également de classe \mathcal{C}^1 sur \mathbb{R}^3 et pour tout $(x, y, z) \in \mathbb{R}^3$, $\nabla g(x, y, z) = (1, 2, 3) \neq (0, 0, 0)$. D'après le théorème des extrema liés, $\nabla f(a, b, c)$ est colinéaire à $\nabla g(a, b, c)$ i.e. il existe $\lambda \in \mathbb{R}$ tel que $(a, b, c) = \lambda(1, 2, 3)$. Mais comme $(a, b, c) \in P$, $a + 2b + 3c = 7$ i.e. $\lambda = \frac{1}{2}$. Ainsi le minimum de la restriction de f à P est atteint en $\frac{1}{2}(1, 2, 3)$: il vaut $f\left(\frac{1}{2}, 1, \frac{3}{2}\right) = \frac{7}{2}$.

Exemple 5.5

Soit f un endomorphisme symétrique d'un espace euclidien E . On pose $\varphi : x \in E \mapsto \langle f(x), x \rangle$. On a déjà montré que φ était de classe \mathcal{C}^1 sur E est que $\nabla f(x) = 2f(x)$ pour tout $x \in E$.

La sphère unité S de E est compacte et f est continue donc la restriction de f à S admet un maximum en un point $u \in S$. Remarquons que $S = g^{-1}(\{0\})$ avec $g : x \in E \mapsto \|x\|^2 - 1$. De plus, g est de classe \mathcal{C}^1 sur E et pour tout $x \in E$, $dg(x) = 2x$. Notamment, g ne s'annule pas sur S .

D'après le théorème des extrema liés, $df(u)$ est colinéaire à $dg(u)$ i.e. $f(u)$ est colinéaire à u : u est donc un vecteur propre de f . On peut alors prouver le théorème spectral par récurrence sur la dimension de E .

5.3 Hessienne**Définition 5.2 Matrice hessienne**

Soient \mathcal{U} un ouvert de \mathbb{R}^n et $f : \mathcal{U} \rightarrow \mathbb{R}$ une application de classe \mathcal{C}^2 . On appelle **matrice hessienne** de f au point $x \in \mathcal{U}$ la matrice

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \leq i, j \leq n}$$

REMARQUE. En vertu du théorème de Schwarz, la matrice hessienne est une matrice **symétrique réelle**. Notamment, le théorème spectral permet d'affirmer qu'elle est orthodiagonalisable.

Proposition 5.3 Formule de Taylor-Young à l'ordre 2

Soient \mathcal{U} un ouvert de \mathbb{R}^n , $f : \mathcal{U} \rightarrow \mathbb{R}$ une application de classe \mathcal{C}^2 et $x \in \mathcal{U}$. Alors

$$f(x+h) \underset{h \rightarrow 0}{=} f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x) \cdot h, h \rangle + o(\|h\|^2)$$

REMARQUE. On munit implicitement \mathbb{R}^n de sa structure euclidienne canonique et on identifie $H_f(x)$ à l'endomorphisme de \mathbb{R}^n qui lui est canoniquement associé.

REMARQUE. En identifiant les vecteurs de \mathbb{R}^n aux matrices colonnes de $\mathcal{M}_{n,1}(\mathbb{R})$, le développement limité à l'ordre 2 peut également s'écrire :

$$f(x+h) \underset{h \rightarrow 0}{=} f(x) + \nabla f(x)^\top h + \frac{1}{2} h^\top H_f(x) h + o(\|h\|^2)$$

Proposition 5.4 Condition nécessaire pour un extremum local

Soient \mathcal{U} un ouvert de \mathbb{R}^n et $f : \mathcal{U} \rightarrow \mathbb{R}$ une application de classe \mathcal{C}^2 . Si f admet un **extremum local** en $x \in \mathcal{U}$, alors x est un **point critique** de f . De plus,

- si f admet un **minimum local** en x , alors $H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$;
- si f admet un **maximum local** en x , alors $-H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$.

REMARQUE. Dans le cas $n = 2$, si x est un point critique de f tel que $\det(H_f(x)) < 0$, f ne peut admettre un extremum local en x car $H_f(x)$ possède alors deux valeurs propres non nulles de signes opposés.

Proposition 5.5 Condition suffisante pour un extremum local

Soient \mathcal{U} un ouvert de \mathbb{R}^n , $f : \mathcal{U} \rightarrow \mathbb{R}$ une application de classe \mathcal{C}^2 et $x \in \mathcal{U}$.

- Si x est un point critique de f et si $H_f(x) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet un **minimum local strict** en x .
- Si x est un point critique de f et si $-H_f(x) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f admet un **maximum local strict** en x .

REMARQUE. Dire que f admet un **minimum local strict** en x , signifie qu'il existe $\varepsilon > 0$ tel que

$$\forall h \in B(0, \varepsilon) \setminus \{0\}, f(x+h) > f(x)$$

Dire que f admet un **maximum local strict** en x , signifie qu'il existe $\varepsilon > 0$ tel que

$$\forall h \in B(0, \varepsilon) \setminus \{0\}, f(x+h) < f(x)$$

Exercice 5.1

Soit $A \in \mathcal{S}_2(\mathbb{R})$. Montrer que $A \in \mathcal{S}_2^{++}(\mathbb{R})$ si et seulement si $\det(A) > 0$ et $\text{tr}(A) > 0$.

REMARQUE. Dans le cas $n = 2$,

$$\begin{aligned} H_f(x) \in \mathcal{S}_2^{++}(\mathbb{R}) &\iff (\det(H_f(x)) > 0 \text{ ET } \text{tr}(H_f(x)) > 0) \\ -H_f(x) \in \mathcal{S}_2^{++}(\mathbb{R}) &\iff (\det(H_f(x)) > 0 \text{ ET } \text{tr}(H_f(x)) < 0) \end{aligned}$$

Exemple 5.6

Soit $f : (x, y) \in \mathbb{R}^2 \mapsto 2x^3 - 6xy + 3y^2$. f est clairement de classe \mathcal{C}^2 sur \mathbb{R}^2 car polynomiale.

Les points critiques de f sont les solutions du système $\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$ i.e. $\begin{cases} 6x^2 - 6y = 0 \\ -6x + 6y = 0 \end{cases}$. On en déduit que les points

critiques de f sont $(0, 0)$ et $(1, 1)$.

De manière générale, $H_f(x, y) = \begin{pmatrix} 12x & -6 \\ -6 & 6 \end{pmatrix}$.

Notamment, $H_f(0, 0) = \begin{pmatrix} 0 & -6 \\ -6 & 6 \end{pmatrix}$. Alors $\det(H_f(0, 0)) = -36 < 0$ donc $H_f(0, 0)$ admet deux valeurs propres non nulles de signes opposés. Ainsi $H_f(0, 0) \notin \mathcal{S}_2^+(\mathbb{R})$ et $-H_f(0, 0) \notin \mathcal{S}_2^+(\mathbb{R})$. Par conséquent, f n'admet pas d'extremum local en $(0, 0)$.

Par ailleurs, $H_f(1, 1) = \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix}$ donc $\det(H_f(1, 1)) = 36 > 0$ et $\text{tr}(H_f(1, 1)) = 18 > 0$. On en déduit que $H_f(1, 1) \in \mathcal{S}_2^+(\mathbb{R})$. Ainsi f admet un minimum local strict en $(1, 1)$.

Le tracé des lignes de niveau de f permet sans doute de mieux comprendre la situation.

