

Calcul différentiel et optimisation

Binyze Mohamed

MP 2025-2026

Sommaire

1	Différentiabilité	1
2	Classe d'une fonction	4
3	Vecteurs tangents à une partie	6
4	Optimisation	7

Dans ce chapitre est sauf indication contraire :

- E, F, G et H désignent des $\mathbb{R}$ -espaces vectoriels normés de dimensions finies non nulles.
- I désigne un intervalle ouvert de $\mathbb{R}$.
- U désigne un ouvert de E et V désigne un ouvert de F .

1 Différentiabilité

Soit $f : U \subset E \longrightarrow F$ et $a \in U$.

Dérivée selon un vecteur, dérivées partielles

Définition 1.1.

dérivée selon un vecteur, dérivées partielles

1. Soit $v \in E$ non nul. On dit que f est **dérivable en a selon v** si l'application $\varphi_v : t \longmapsto f(a + t.v)$ est dérivable en 0. Dans ce cas, on note $D_v f(a) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{1}{t} (f(a + t.v) - f(a))$ appelé **la dérivée de f en a selon le vecteur v** .
2. Soit $\mathcal{E} = (e_1, \dots, e_n)$ une base de E . Lorsqu'elle existe, on appelle **j -ème dérivée partielle de f en a dans la base $\mathcal{E}$** la dérivée de f en a selon le vecteur e_j . Celle-ci est notée $D_j f(a)$ ou $\frac{\partial f}{\partial x_j}(a)$.

Remarque 1.1. ■ Pour $x = \sum_{i=1}^n x_i e_i \in E$ on peut identifier $x \longmapsto f(x)$ et la fonction de n variables réelles $(x_1, \dots, x_n) \longmapsto$

$f(x_1, \dots, x_n)$. Ainsi
$$D_j f(x) = \frac{\partial f}{\partial x_j}(x) = \frac{d}{dt} (f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_n)) \Big|_{t=x_j}.$$

Exemple 1.1. ■ Soit $a = (-1, 1)$, $v = (1, 0)$ et $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $f(x, y) = (x + y, x - y)$. Pour $t \neq 0$:

$$\frac{1}{t} (f(a + t.v) - f(a)) = \frac{1}{t} (f(t - 1, 1) - f(-1, 1)) = \frac{1}{t} ((t, t - 2) - (0, -2)) \xrightarrow{t \rightarrow 0} (1, 1).$$

Donc $D_v f(a) = (1, 1)$.

■ Dérivées partielles de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^2 y - e^{x+y}$ dans la base canonique.

$$D_1 f(x, y) = \frac{\partial f}{\partial x}(x, y) = \lim_{t \rightarrow 0} \frac{1}{t} (f((x, y) + t.e_1) - f(x, y)) = \lim_{t \rightarrow 0} \frac{1}{t} ((x + t)^2 y - e^{x+y+t} - x^2 y + e^{x+y}) = 2xy - e^{x+y}.$$

De même $D_2 f(x, y) = \frac{\partial f}{\partial y}(x, y) = x^2 - e^{x+y}$.

Applications différentiables

Définition 1.2.

application différentiable

1. On dit que f est **différentiable en** a s'il existe $\ell \in \mathcal{L}(E, F)$ tel que $\frac{1}{\|h\|} (f(a + h) - f(a) - \ell(h)) \xrightarrow{h \rightarrow 0_E} 0_F$.

Dans ce cas ℓ est **unique** et appelé ¹ la **différentielle** de f au point a . On note $\ell = df(a) \in \mathcal{L}(E, F)$.

2. On dit que f est **différentiable sur** U si f est différentiable en tout point de U .

1. ou l'application linéaire **tangente**.

Remarque 1.2 (Matrice jacobienne).

■ Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable en a , on appelle **matrice jacobienne de f en a** la matrice de l'application linéaire $df(a)$ dans les bases canoniques. On la note $J_f(a)$.

Proposition 1.1.

différentiabilité vs DL1

f est différentiable en a si, et seulement si, f admet ¹ un **développement limité à l'ordre 1 en a** :

$$f(a + h) = f(a) + df(a).h + o(\|h\|).$$

1. $o(\|h\|) = \|h\|.\varepsilon(h)$ avec $\varepsilon(h) \rightarrow 0_F$ quand $h \rightarrow 0_E$.

Exemple 1.2. ■ Si f est constante, alors $df(a) = 0$.

■ Différentielle de $f : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathcal{M}_n(\mathbb{R})$ définie par $f(M) = M^2$. Pour $A \in \mathcal{M}_n(\mathbb{R})$:

$$f(A + H) = (A + H)^2 = A^2 + AH + HA + H^2 = f(A) + \underbrace{AH + HA}_{\ell(H)} + H^2.$$

D'une part, l'application ℓ est linéaire. D'autre part, en introduisant une norme sous-multiplicative ¹ sur $\mathcal{M}_n(\mathbb{R})$:

$$\|H^2\| \leq \|H\|^2 = \|H\| \times \|H\| \underset{H \rightarrow 0_n}{=} o(\|H\|).$$

Donc f est différentiable et $df(A).H = AH + HA$ pour tout $H \in \mathcal{M}_n(\mathbb{R})$.

Proposition 1.2.

différentielle vs dérivée selon un vecteur

Soit $\mathcal{E} = (e_1, \dots, e_n)$ une base de E et f **différentiable** en a .

1. f est continue en a .

2. f admet une dérivée en a selon tout vecteur v non nul de E et $D_v f(a) = df(a).v$. En particulier, f admet des dérivées partielles en a . De plus

$$\forall i \in \llbracket 1, n \rrbracket, \quad \frac{\partial f}{\partial x_i}(a) = df(a).e_i \quad \text{et} \quad \forall h = \sum_{i=1}^n h_i e_i \in E, \quad df(a).h = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a).$$

1. Par exemple, la norme $\|\cdot\|_2$ sur $\mathcal{M}_n(\mathbb{R})$ vérifie $\|AB\|_2 \leq \|A\|_2 \|B\|_2$. (sous-multiplicativité)

Proposition 1.3.**cas linéaire et bilinéaire**

1. Toute application linéaire $f : E \longrightarrow F$ est différentiable sur E et $\boxed{\forall a \in E, \, df(a) = f}$.
2. Toute application bilinéaire $f : E \times F \longrightarrow G$ est différentiable sur $E \times F$ et

$$\boxed{\forall (a, h) \in E, \, \forall (b, k) \in F, \, df(a, b).(h, k) = f(a, k) + f(h, b)}.$$

Proposition 1.4.**dérivée vs différentielle**

Soit $f : I \subset \mathbb{R} \longrightarrow F$ et $a \in I$. f est différentiable en a si, et seulement si, f est dérivable en a . Dans ce cas :

$$\boxed{\forall h \in \mathbb{R}, \, df(a).h = h.f'(a)}. \text{ En particulier } \boxed{f'(a) = df(a).1}.$$

Opérations**Proposition 1.5.****linéarité**

Si $f, g : U \longrightarrow F$ sont différentiables en a , alors pour tout $\lambda \in \mathbb{R}$, $f + \lambda g$ est différentiable en a et

$$\boxed{\forall h \in E, \, d(f + \lambda g)(a).h = df(a).h + \lambda dg(a).h}.$$

Théorème 1.1.**composée**

Soit $f : U \longrightarrow F$ et $g : V \longrightarrow G$ tel que $f(U) \subset V$. Si f est différentiable en a et g différentiable en $f(a)$ alors $g \circ f$ est différentiable en a et $\boxed{\forall h \in E, \, d(g \circ f)(a).h = (dg(f(a)) \circ df(a)).h}$.

Exemple 1.3. ■ $d(f^\alpha) = \alpha f^{\alpha-1} df$, $d(e^f) = e^f df$, $d(\ln f) = \frac{df}{f}$, $d\left(\frac{1}{f}\right) = -\frac{df}{f^2}$, etc.

Corollaire 1.1.**dérivée le long d'un arc**

Si $\gamma : I \longrightarrow E$ est dérivable en t et $f : E \longrightarrow F$ est différentiable en $\gamma(t)$, alors l'application $f \circ \gamma : I \longrightarrow F$ est dérivable en t et $\boxed{(f \circ \gamma)'(t) = df(\gamma(t)).\gamma'(t)}$. En particulier $\boxed{(f(x + th))' = df(x + th).h}$

Corollaire 1.2.**dérivation de $t \mapsto f(x_1(t), \dots, x_n(t))$**

Soient $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable sur l'ouvert U et $x_1, \dots, x_n : I \longrightarrow \mathbb{R}$ des fonctions dérivables sur I , telles que $\forall t \in I, \, (x_1(t), \dots, x_n(t)) \in U$. Alors la fonction $g : t \in I \mapsto f(x_1(t), \dots, x_n(t))$ est dérivable sur I et

$$\boxed{g'(t) = \sum_{i=1}^n x'_i(t) \cdot \frac{\partial f}{\partial x_i}(x_1(t), \dots, x_n(t))}.$$

Corollaire 1.3.**règle de la chaîne**

Soient $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable sur l'ouvert U et $x_1, \dots, x_n : V \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ des fonctions différentiables sur l'ouvert V , telles que $\forall (t_1, \dots, t_p) \in V, \, (x_1(t_1, \dots, t_p), \dots, x_n(t_1, \dots, t_p)) \in U$.

Alors l'application $g : (t_1, \dots, t_p) \in V \mapsto f(x_1(t_1, \dots, t_p), \dots, x_n(t_1, \dots, t_p))$ est différentiable sur V et

$$\boxed{\forall i \in \llbracket 1, p \rrbracket, \, \frac{\partial g}{\partial t_i}(t_1, \dots, t_p) = \sum_{k=1}^n \frac{\partial x_k}{\partial t_i}(t_1, \dots, t_p) \cdot \frac{\partial f}{\partial x_k}(x_1(t_1, \dots, t_p), \dots, x_n(t_1, \dots, t_p))}.$$

Exemple 1.4. ■ Si $f : (x, y) \mapsto f(x, y)$ est différentiable sur $\mathbb{R}^2$ alors, $g : t \mapsto f(\cos(t), \sin(t))$ est dérivable sur $\mathbb{R}$ et on a

$$g'(t) = -\sin(t) \frac{\partial f}{\partial x}(\cos(t), \sin(t)) + \cos(t) \frac{\partial f}{\partial y}(\cos(t), \sin(t)).$$

■ Si $f : (x, y) \mapsto f(x, y)$ est différentiable sur $\mathbb{R}^2$ alors, $g : (r, \theta) \mapsto f(r \cos(\theta), r \sin(\theta))$ est différentiable sur $\mathbb{R}^2$ et on a

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

Proposition 1.6.

fonctions coordonnées

Soit $(\varepsilon_1, \dots, \varepsilon_p)$ une base de F . On écrit pour tout $x \in U$, $f(x) = \sum_{i=1}^p f_i(x) \varepsilon_i$. Alors f est différentiable en a si, et seulement si, pour tout $i \in \llbracket 1, p \rrbracket$, f_i est différentiable en a . Dans ce cas

$$\forall h \in E, \quad df(a).h = \sum_{i=1}^p (df_i(a).h) \cdot \varepsilon_i.$$

Proposition 1.7.

composée par application bilinéaire

Soit $B : F \times G \rightarrow H$ une application bilinéaire. Si $f : U \rightarrow F$ et $g : U \rightarrow G$ sont différentiables en a alors, l'application $B(f, g)$ est différentiable en a et

$$\forall h \in E, \quad dB(f, g)(a).h = B(df(a).h, g(a)) + B(f(a), dg(a).h).$$

Proposition 1.8.

composée par application multilinéaire

Soit $m : F_1 \times \dots \times F_k \rightarrow H$ une application multilinéaire. Si, pour tout $i \in \llbracket 1, k \rrbracket$, $f_i : U \rightarrow F_i$ est différentiable en a alors, l'application $m(f_1, \dots, f_k)$ est différentiable en a et

$$\forall h \in E, \quad dm(f_1, \dots, f_k)(a).h = \sum_{i=1}^k m(f_1(a), \dots, f_{i-1}(a), df_i(a).h, f_{i+1}(a), \dots, f_k(a)).$$

Exemple 1.5. ■ Toute application polynomiale est différentiable.

■ $\det : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ est différentiable car polynomiale en les coefficients de la matrice.

Gradient

$(E, (\cdot | \cdot))$ désigne un espace euclidien de dimension $n \in \mathbb{N}^*$. Soit $f : U \subset E \rightarrow \mathbb{R}$ et $a \in U$.

Théorème 1.2.

vecteur gradient

Si f est différentiable en a alors il existe un **unique** vecteur de E noté $\nabla f(a)$, appelé **gradient** de f en a vérifiant

$$\forall h \in E, \quad df(a).h = (\nabla f(a) | h).$$

De plus, si $(e_1, \dots, e_n)$ est une b.o.n de E , alors

$$\nabla f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot e_i.$$

Remarque 1.3. ■ Si f est une fonction de p variables réelles alors

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a) \right).$$

2 Classe d'une fonction

Soit $f : U \subset E \rightarrow F$ et $a \in U$.

Fonctions de classe $\mathcal{C}^1$ **Définition 2.1.****fonction de classe $\mathcal{C}^1$**

On dit que f est de **classe $\mathcal{C}^1$** sur U si elle est différentiable sur U et ¹ si l'application $df : a \mapsto df(a)$ est continue sur U .

1. Les fonctions de classe $\mathcal{C}^1$ sont continues car différentiables.

Théorème 2.1.**théorème fondamental**

On a équivalence entre :

- (i) f est de classe $\mathcal{C}^1$ sur U ;
- (ii) les dérivées partielles de f dans une base ¹ de E existent en tout point de U et sont continues sur U .

1. La notion ne dépend pas du choix de la base utilisée.

Proposition 2.1.**formule d'intégration**

1. Soit $\gamma \in \mathcal{C}^1(I, U)$ et f de classe $\mathcal{C}^1$ sur U . $\forall (\alpha, \beta) \in I^2, \int_{\alpha}^{\beta} df(\gamma(t)) \cdot \gamma'(t) dt = f(\gamma(\beta)) - f(\gamma(\alpha))$.
2. Si U est convexe alors $\forall (a, b) \in U^2, \int_0^1 df(a + t(b-a)) \cdot (b-a) dt = f(b) - f(a)$.

Proposition 2.2.**caractérisation des applications constantes**

Soit U un ouvert connexe par arcs et f de classe $\mathcal{C}^1$ sur U .

f est constante sur U si, et seulement si, pour $x \in U$, $df(x) = 0$.

Dérivées partielles successives

Soit $k \in \mathbb{N}$ et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E .

Définition 2.2.**dérivées partielles d'ordre supérieur**

1. La fonction f est appelée **dérivée partielle d'ordre 0 de f** .
2. Sous réserve d'existence, on appelle **dérivées partielles d'ordre $k+1$ de f** les dérivées partielles des dérivées partielles d'ordre k de f .

Notations :

- Si l'on note $x_1, \dots, x_n$ les coordonnées de x dans $\mathcal{B}$, on note $\frac{\partial^k f}{\partial x_i^k} = \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_i} \left(\dots \left(\frac{\partial f}{\partial x_i} \right) \dots \right) \right)$.
- Pour tout $(\alpha_1, \dots, \alpha_p) \in (\mathbb{N}^*)^p$ avec $\alpha = \alpha_1 + \dots + \alpha_p$ et $(i_1, \dots, i_p) \in \llbracket 1, n \rrbracket^p$, on note

$$\frac{\partial^{\alpha} f}{\partial x_{i_1}^{\alpha_1} \partial x_{i_2}^{\alpha_2} \dots \partial x_{i_p}^{\alpha_p}} = \frac{\partial^{\alpha_1}}{\partial x_{i_1}^{\alpha_1}} \left(\frac{\partial^{\alpha_2}}{\partial x_{i_2}^{\alpha_2}} \left(\dots \left(\frac{\partial^{\alpha_p} f}{\partial x_{i_p}^{\alpha_p}} \right) \dots \right) \right)$$

et lorsqu'elle existe, l'application $\frac{\partial^{\alpha} f}{\partial x_{i_1}^{\alpha_1} \partial x_{i_2}^{\alpha_2} \dots \partial x_{i_p}^{\alpha_p}} : U \longrightarrow F$ est appelée **dérivée partielle d'ordre α de f** .

Définition 2.3.**fonction de classe $\mathcal{C}^k$**

1. On dit que f est de **classe $\mathcal{C}^k$** sur U si¹ ses dérivées partielles d'ordre k existent et sont continues sur U . On note $\mathcal{C}^k(U, F)$ l'ensemble des applications de classe $\mathcal{C}^k$ de U dans F .
2. On dit que f est de **classe $\mathcal{C}^\infty$** sur U si f est de classe $\mathcal{C}^k$ sur U pour tout $k \in \mathbb{N}$. On note $\mathcal{C}^\infty(U, F)$ l'ensemble des applications de classe $\mathcal{C}^\infty$ de U dans F .

1. Les fonctions de classe $\mathcal{C}^0$ correspondent aux fonctions continues.

Proposition 2.3.**propriétés**

1. L'ensemble $\mathcal{C}^k(U, F)$ est un sous-espace vectoriel de $\mathcal{C}(U, F)$.
2. Si F est une $\mathbb{R}$ -algèbre alors $\mathcal{C}^k(U, F)$ est de plus une $\mathbb{R}$ -algèbre.
3. La composition d'applications de classe $\mathcal{C}^k$ est de classe $\mathcal{C}^k$.
4. Une application est de classe $\mathcal{C}^k$ si, et seulement si, ses fonctions coordonnées le sont.

Théorème 2.2.**théorème de Schwartz**

Si f est de classe $\mathcal{C}^2$ sur U alors¹ $\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$

1. Lorsque f est de classe $\mathcal{C}^k$, $k \geq 2$, l'ordre de dérivation n'importe pas dans le calcul d'une dérivée partielle d'ordre k .

3 Vecteurs tangents à une partie

$\mathcal{A}$ désigne une partie de E .

Définition 3.1.**vecteur tangent à une partie**

Un vecteur $v \in E$ est dit **tangent** à $\mathcal{A}$ en $a \in \mathcal{A}$ s'il existe $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\rightarrow \mathcal{A}$ dérivable en 0 tel que $\gamma(0) = a$ et $\gamma'(0) = v$. On note $T_a \mathcal{A}$ l'ensemble des **vecteurs tangents** à $\mathcal{A}$ en a .

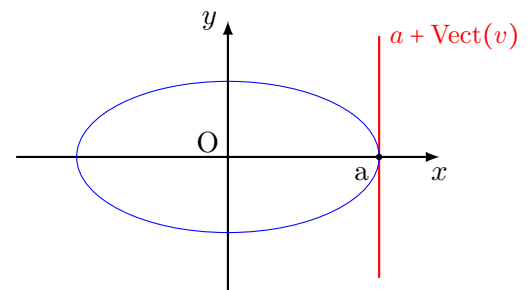
Exemple 3.1. ■ Soit I un intervalle ouvert de $\mathbb{R}$ et $f : I \rightarrow \mathbb{R}$ dérivable en $t_0 \in I$. Notons $\mathcal{A} = \{(t, f(t)), t \in I\}$ le graphe de f . Le vecteur $v = (1, f'(t_0))$ est tangent à $\mathcal{A}$ en $a = (t_0, f(t_0))$. En effet :

- $\exists \varepsilon > 0, \forall t \in]-\varepsilon, \varepsilon[, t + t_0 \in I$. (car I est un intervalle ouvert)
- Soit $\gamma : t \in]-\varepsilon, \varepsilon[\mapsto (t + t_0, f(t + t_0)) \in \mathcal{A}$. On a γ est dérivable en 0 et vérifie : $\gamma(0) = a$ et $\gamma'(0) = v$.

Exemple 3.2.

■ Soit $\mathcal{A} = \{(x, y) \in \mathbb{R}^2, x^2 + 4y^2 = 4\}$, $a = (2, 0)$ et $v = (0, 1)$.

- Soit $\gamma :]-1, 1[\rightarrow \mathcal{A}$ définie par $\gamma(t) = (2\sqrt{1-t^2}, t)$.
- γ est dérivable en 0 et $\gamma(0) = a$, $\gamma'(0) = v$, donc v est un vecteur tangent à $\mathcal{A}$ en a .

**Définition 3.2.****surface**

On appelle **surface représentative** de $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ l'ensemble $\mathcal{S}_f \stackrel{\text{déf}}{=} \{(x, y, z) \in \mathbb{R}^3, z = f(x, y)\}$.

Théorème 3.1.**plan tangent à une surface**

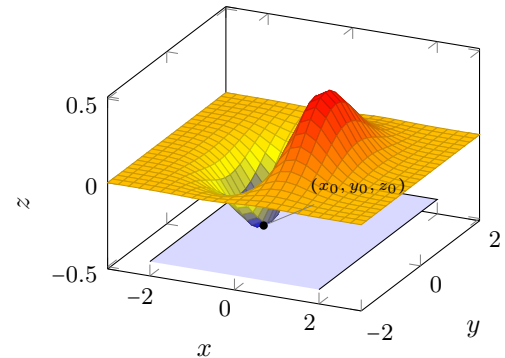
Soit $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ différentiable en $(x_0, y_0) \in U$. La surface $\mathcal{S}_f$ admet au point (x_0, y_0, z_0) avec $z_0 = f(x_0, y_0)$

un plan affine tangent d'équation :
$$z = (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) + f(x_0, y_0).$$

Exemple 3.3.

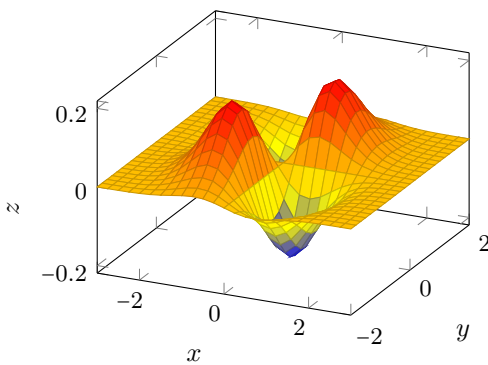
■ Soit la surface de $\mathbb{R}^3$ d'équation $z = x e^{-x^2 - y^2}$.

- On pose $f(x, y) = x e^{-x^2 - y^2}$ et soit $(x_0, y_0) = \left(-\frac{\sqrt{2}}{2}, 0\right)$.
- f est différentiable en (x_0, y_0) et $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$.
- L'équation du plan tangent à $\mathcal{S}_f$ en (x_0, y_0, z_0) est $z = -\frac{1}{\sqrt{2}e}$.

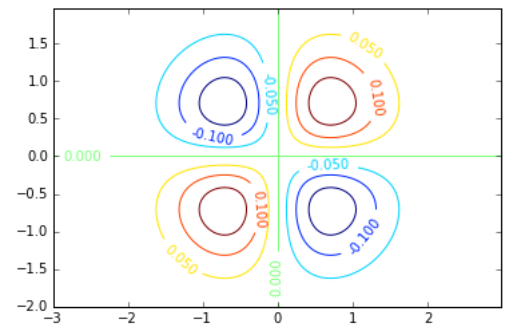
**Définition 3.3.****ligne de niveau**

Soit $f : U \subset E \rightarrow \mathbb{R}$ et $c \in \mathbb{R}$.

1. L'ensemble $\mathcal{A} = \{x \in U, f(x) = c\}$ est appelé la **ligne de niveau** c de f .
2. Si $\dim E = 3$ (resp. $\dim E = 2$), $\mathcal{A}$ est appelé la **surface de niveau** (resp. la **courbe de niveau**) c de f .



(a) Tracé de $f(x, y) = xy e^{-x^2 - y^2}$



(b) courbes de niveau de f

Théorème 3.2.

Soit $f : U \subset E \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ et $\mathcal{A}$ une ligne de niveau de f . Si $a \in \mathcal{A}$ tel que $df(a) \neq 0$, alors

$$T_a \mathcal{A} = \ker(df(a)).$$

Proposition 3.1.**gradient vs vecteurs tangents**

Soit $(E, (\cdot | \cdot))$ un espace euclidien et $\mathcal{A}$ une ligne de niveau de f . Si f est différentiable en $a \in \mathcal{A}$, alors

$$\forall v \in T_a \mathcal{A}, (v | \nabla f(a)) = 0.$$

4 Optimisation

Dans ce paragraphe, les applications considérées sont définies sur l'**ouvert** U de E à valeurs réelles (numériques).

Étude au premier ordre

Définition 4.1.

minimum, maximum

Soit $\mathcal{A}$ une partie de E .

1. On dit que $f : \mathcal{A} \rightarrow \mathbb{R}$ admet un **minimum global** (ou **absolu**) (resp. **maximum global**) en $a \in \mathcal{A}$ si

$$\forall x \in \mathcal{A}, f(x) \geq f(a) \quad (\text{resp. } f(x) \leq f(a))$$

2. On dit que $f : \mathcal{A} \rightarrow \mathbb{R}$ admet un **minimum local** (resp. **maximum local**) en $a \in \mathcal{A}$ si

$$\exists r > 0, \forall x \in \mathcal{A} \cap B(a, r), f(x) \geq f(a) \quad (\text{resp. } f(x) \leq f(a))$$

3. On dit que $f : \mathcal{A} \rightarrow \mathbb{R}$ admet un **extremum local** (resp. **extremum global**) en $a \in \mathcal{A}$ si f admet un maximum ou un minimum local (resp. un maximum ou un minimum global) en a .

Définition 4.2.

point critique

On dit que $f : U \rightarrow \mathbb{R}$ admet un **point critique**¹ en $a \in U$ si f est différentiable en a et $df(a) = 0$.

1. On rappelle ici que U est un **ouvert**.

Proposition 4.1.

caractérisation des points critiques

Si $f : U \rightarrow \mathbb{R}$ est différentiable en $a \in U$ alors a point critique de $f \iff \forall i \in \llbracket 1, n \rrbracket, \frac{\partial f}{\partial x_i}(a) = 0$.

Proposition 4.2.

points critiques vs extremums

Si $f : U \rightarrow \mathbb{R}$ est différentiable en $a \in U$ et admet un extremum local en a alors¹ f admet un point critique en a .

1. La réciproque est fausse.

Exemple 4.1.

■ Extremums de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 - xy + 1$.

- f est différentiable sur l'ouvert $\mathbb{R}^2$ car polynomiale.
- Points critiques de f :

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0 \iff (x, y) = (0, 0)$$

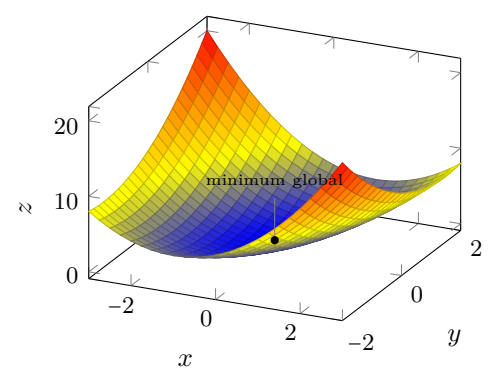
donc $(0, 0)$ est le seul point critique de f .

- Étude du signe de la différence $f(x, y) - f(0, 0)$: on a

$$f(x, y) - f(0, 0) = x^2 + y^2 - xy = \left(x - \frac{y}{2}\right)^2 + \frac{3}{4}y^2 \geq 0.$$

donc f admet un minimum global en $(0, 0)$.

Tracé de $z = f(x, y)$



Proposition 4.3.

Soit $\mathcal{A}$ une partie de U et $f : U \rightarrow \mathbb{R}$ différentiable en $a \in \mathcal{A}$.

Si la restriction de f à $\mathcal{A}$ admet un extremum local en a alors $\forall v \in T_a \mathcal{A}, df(a).v = 0$.

Théorème 4.1.

optimisation sous une contrainte

Soit $f, g : U \rightarrow \mathbb{R}$ deux applications de classe $\mathcal{C}^1$ sur l'ouvert U . Si la restriction de f à $\mathcal{A} = \{x \in U, g(x) = 0\}$ admet un extremum local en $a \in \mathcal{A}$ tel que $dg(a) \neq 0$ alors $df(a)$ est colinéaire à $dg(a)$.

Remarque 4.1 (Cas où E est euclidien).

■ Lorsque E est un espace euclidien alors le résultat du théorème précédent s'écrit $\boxed{\exists \lambda \in \mathbb{R}, \nabla f(a) = \lambda \cdot \nabla g(a)}$.

Exemple 4.2. ■ Extréma de $f(x, y) = xy$ sur $\mathcal{A} = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$.

- $\mathcal{A}$ est compact car fermé borné de $\mathbb{R}^2$ donc les extréma de f existent.
- Cherchons d'abord les points critiques de f . Si $a = (x, y)$ est un point critique de f alors

$$\begin{cases} y &= 2\lambda x \\ x &= 2\lambda y \\ x^2 + y^2 &= 1 \end{cases} \implies a = \left(\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}} \right).$$

Il y a donc 4 points critiques. Or $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right) = \frac{1}{2}$ et $f\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right) = \frac{-1}{2}$.

Le maximum de f est $\frac{1}{2}$ et son minimum est $\frac{-1}{2}$.

Étude au second ordre

Définition 4.3.

matrice hessienne

Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ sur U . On appelle **matrice hessienne** de f en a , la matrice notée $H_f(a)$ définie par¹ :
$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}.$$

1. D'après le théorème de Schwartz, $H_f(a) \in \mathcal{S}_n(\mathbb{R})$.

Théorème 4.2.

développement de Taylor à l'ordre 2

Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in U$.

$$\begin{aligned} f(a+h) &= f(a) + (\nabla f(a) \mid h) + \frac{1}{2} (H_f(a).h \mid h) + o_{h \rightarrow 0}(\|h\|^2) \\ &= f(a) + \nabla f(a)^\top . h + \frac{1}{2} h^\top . H_f(a).h + o_{h \rightarrow 0}(\|h\|^2) \end{aligned}$$

Théorème 4.3.

matrice hessienne vs extremums

Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in U$.

1. Si f admet un minimum (resp. maximum) local en a alors $\text{Sp}(H_f(a)) \subset \mathbb{R}_+$ (resp. $\text{Sp}(H_f(a)) \subset \mathbb{R}_-$).
2. Supposons f admet un point critique en $a \in U$.
 - a. Si $\text{Sp}(H_f(a)) \subset \mathbb{R}_+^*$ (resp. $\text{Sp}(H_f(a)) \subset \mathbb{R}_-^*$) alors f admet un minimum (resp. maximum) local strict¹ en a .
 - b. Si $\text{Sp}(H_f(a)) \subset \mathbb{R}_+^* \cap \mathbb{R}_-^*$ alors f n'admet pas d'extremum local en a .

1. minimum (resp. maximum) local strict en $a : \exists r > 0, \forall x \in U \cap B(a, r) \setminus \{a\}, f(x) > f(a)$ (resp. $f(x) < f(a)$).

Corollaire 4.1.

cas de la dimension 2

Soit $f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et a un point critique de f .

1. Si $\det(H_f(a)) > 0$ et $\text{Tr}(H_f(a)) > 0$ alors f admet un minimum local strict en a .
2. Si $\det(H_f(a)) > 0$ et $\text{Tr}(H_f(a)) < 0$ alors f admet un maximum local strict en a .
3. Si $\det(H_f(a)) < 0$ alors f n'admet pas d'extremum en a .

Théorème 4.4.**extremums en dimension 2**

Soit $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et a un point critique de f . On note

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a), \quad t = \frac{\partial^2 f}{\partial y^2}(a). \quad (\text{notation de Hesse en dimension 2})$$

1. Si $rt - s^2 > 0$ et $r > 0$ alors f admet un minimum local en a .
2. Si $rt - s^2 > 0$ et $r < 0$ alors f admet un maximum local en a .
3. Si $rt - s^2 < 0$ alors f n'a pas d'extremum en a . (on dit que a est un **point-col** ou **point-selle**)
4. Si $rt - s^2 = 0$, on ne peut pas conclure.

Exemple 4.3.

■ Extremums de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^3 - 2xy - y$.

- f est différentiable sur l'ouvert $\mathbb{R}^2$ car polynomiale.

- Points critiques de f :

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0 \iff (x, y) = (1, 1) \text{ ou } (x, y) = \left(-\frac{1}{3}, -\frac{1}{3}\right)$$

donc $(1, 1)$ et $\left(-\frac{1}{3}, -\frac{1}{3}\right)$ sont les seuls points critiques de f .

- Au point $(1, 1)$, $r = 2$, $s = -2$ et $t = 6$ donc $rt - s^2 > 0$ et $r > 0$ donc f admet un minimum local en $(1, 1)$.
- Au point $\left(-\frac{1}{3}, -\frac{1}{3}\right)$, $r = 2$, $s = -2$ et $t = -2$ donc $rt - s^2 < 0$ donc $\left(-\frac{1}{3}, -\frac{1}{3}\right)$ est un point selle de f .

Tracé de $z = f(x, y)$ 