Khôlles : Semaine 24

- 01 - 05 Avril 2024 -

Sommaire

1 Questions de cours - Groupes A, B, C

1.1 Définitions de base : notion de fonction différentiable, de dérivée directionnelle et de dérivée partielle. Lien entre ces notions. (démo des liens)

Définition: Dérivée directionnelle

Soient E et F, deux Espaces vectoriels de dimension finie. Soit $U \subset E$ et $f : U \to F$.
Soit $a \in U$ et $u \in E$.

On dit que f admet une dérivée directionnelle en a suivant u si $\lim\limits_{t \to 0} \dfrac{f(a+tu)-f(a)}{t}$ existe et est finie.

Dans ce cas, on note $D_u(f)(a) = \lim\limits_{t \to 0} \dfrac{f(a+tu)-f(a)}{t}$

Définition: Fonction différentiable

Soient E et F, deux Espaces vectoriels de dimension finie. Soit $U \subset E$ et $f : U \to F$.
Soit $a \in U$.

On dit que f est différentiable en a si :

$$\exists \varphi \in \mathscr{L}(E,F),\ \forall h \in E,\ (a+h) \in U \Rightarrow\ f(a+h) = f(a) + \varphi(h) + \underbrace{o(h)}_{=o(\|h\|)}$$

On pose alors $\varphi = df(a)$. Donc, $df \in \mathscr{L}(E,F)$ et nous avons lorsque $h \to 0$: $f(a+h) = f(a) + df(a)(h) + o(h)$

Proposition fondamentale

Soit $E = \mathbb{R}^n$. Soit $U \subset E$ et soit F, $\mathbb{R}$-EVN de dimension finie. Soit $f : U \to F$.
Soit $B = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$.

Soit $a = (a_1, \ldots, a_n) \in U$. Alors, $\forall i \in [\![1, n]\!]$ et $\forall t \neq 0$:

$$\frac{f(a+te_i)-f(a)}{t} = \frac{f(a_1, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{t}$$

Si f admet une dérivée directionnelle selon e_i en a (i.e $\lim\limits_{t \to 0} \dfrac{f(a+te_i)-f(a)}{t}$ existe et est finie), alors
$D_{e_i}(f)(a) = \dfrac{\partial f}{\partial x_i}(a)$

Proposition

Soient E et F, deux Espaces vectoriels de dimension finie. Soit $U \subset E$ et $f : U \to F$.
Soit $a \in U$.

Si f est différentiable en a, alors $\forall v \in E,\ \ f$ admet une dérivée directionnelle selon v et $D_v f(a) = df(a)(v)$

Preuve :

$\forall t \neq 0$, on pose $h = tv \xrightarrow[t \to 0]{} 0$.

$$\begin{aligned} f(a+tv) &= f(a) + df(a)(tv) + o(tv) \\ &= f(a) + t df(a)(v) + o(t) \\ \frac{f(a+tv) - f(a)}{t} &= df(a)(v) + o(1) \xrightarrow[t \to 0]{} df(a)(v) \end{aligned}$$

Dès lors, la dérivée directionnelle existe et vaut $D_v f(a) = df(a)(v)$

1.2 Expression de la différentielle à l'aide des dérivées partielles (démo)

Proposition

Soit $E = \mathbb{R}^n$, $U \subset E$ et F : $\mathbb{R}$—EVN de dimension finie.
Soit $B = (e_1, \dots, e_n)$ base canonique de E. Soit $f : U \to F$ application.
Soit $a \in \mathring{U}$ tel que f soit différentiable en a.

1. $\forall i \in [\![1, n]\!],\ \ df(a)(e_i) = D_{e_i} f(a) = \dfrac{\partial f}{\partial x_i}(a)$
2. $\forall h \in \mathbb{R}^n,\ \ \exists (h_i)_i \in \mathbb{R}^n,\ \ h = \displaystyle\sum_{i=1}^{n} h_i e_i$.

$$df(a)(h) = df(a)\left(\sum_{i=1}^{n} h_i e_i\right) = \sum_{i=1}^{n} h_1 df(a)(e_i) = \sum_{i=1}^{n} h_1 \frac{\partial f}{\partial x_i}(a)$$

1.3 Théorème de représentation des formes linéaires en dimension finie (démo)

Théorème de Représentation des formes linéaires en dimension finie

Soit E, espace euclidien.

1. $\forall a \in E,\ \phi_a : \begin{cases} E \to \mathbb{R} \\ x \mapsto \langle a; x\rangle \end{cases}$ est une forme linéaire (i.e $\phi_a \in E^*$)
2. $\forall \psi \in \mathscr{L}(E,\mathbb{R}),\ \exists! a \in E,\ \psi = \phi_a$

Preuve :

1. Soit $a \in E$. Le produit scalaire est une forme Bilinéaire par définition. Ainsi, $x \mapsto \langle a; x\rangle$ est une forme linéaire.
2. Posons $\Phi : \begin{cases} E \to E^* \\ a \mapsto \phi_a \end{cases}$. Alors :
   - Φ est correctement définie d'après $1)$
   - Φ est linéaire : Soient $a, b \in E$ et $\lambda, \mu \in \mathbb{R}$. Soit $x \in E$:
   $$\phi_{\lambda a+\mu b}(x) = \langle \lambda a + \mu b; x\rangle = \lambda\langle a; x\rangle + \mu\langle b, x\rangle = (\lambda\phi_a + \mu\phi_b)(x)$$
   - Φ est injective : Soit $a \in \mathrm{Ker}(\Phi)$. $\phi_a = 0 \Rightarrow\ \forall x \in E,\ \phi_a(x) = \langle a; x\rangle = 0$.
   En particulier, $\phi_a(a) = \langle a, a\rangle = 0 \Rightarrow\ a = 0 \Rightarrow\ \mathrm{Ker}(\Phi) = \{0\}$
   - E est un espace Euclidien, donc $\dim(E) = n \in \mathbb{N}$. Or, $\dim(E^*) = \dim(E) \Rightarrow\ \Phi$ est bijective.
   - $\forall \varphi \in \mathscr{L}(E,\mathbb{R}),\ \exists! a \in E,\ \varphi = \Phi(a) = \phi_a$. i.e : $\forall \varphi \in \mathscr{L}(E,\mathbb{R});\ \exists! a \in E,\ \forall x \in E,\ \varphi(x) = \langle a; x\rangle$

1.4 Définition du gradient.

Définition: Gradient

Soit E euclidien, $U \subset E$ et $f : U \to \mathbb{R}$, application.
Soit $a \in U$, On suppose que f est différentiable en a.

Alors $df(a) \in E^*$, dès lors, $\exists!\ \nabla_f(a) \in E,\ \ df(a) = \phi_{\nabla_f(a)}$. i.e, $\forall h \in E,\ \ df(a)(h) = \langle \nabla_f(a), h\rangle$.

On appelle $\nabla_f(a)$ le Gradient de f en a.

Si E est muni d'une BON $(e_1, \dots e_p)$, dans cette base , on note $f : \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \mapsto f(x_1 \dots, x_p)$.

Alors, $\forall h = \displaystyle\sum_{i=1}^{p} h_i e_i \in E,\ \ df(a)(h) = \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a) = \langle \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_p}(a) \end{pmatrix}; \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix} \rangle$

Par unicité de $\nabla_f(a)$, $\ \nabla_f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_p}(a) \end{pmatrix}$

1.5 Définition et caractérisation d'une fonction de classe $\mathscr{C}^1$, de classe $\mathscr{C}^k$

Définition: Application de classe C^1

Soient E, F deux $\mathbb{R}-$espaces vectoriels de dimension finies. Soit $U \subset E$ ouvert. Soit $f : U \to F$.

On dit que f est de classe C^1 sur U si :

1. f est différentiable sur U
2. L'application $(a \mapsto df(a)) = df : \begin{cases} U \to \mathscr{L}(E, F) \\ a \mapsto df(a) \end{cases}$ est Continue

Théorème Caractérisation des applications de classe $\mathscr{C}^1$

Soient $n, p \in \mathbb{N}^*$. Soit $U \subset \mathbb{R}^n$ ouvert. Soit $f : U \to \mathbb{R}^p$. Alors :

$$\left[f \text{ est } \mathscr{C}^1 \text{ sur } U\right] \iff \begin{cases} 1. \quad \forall a \in U, \ \ f \text{ admet des dérivées partielles par rapport à } x_1, \dots, x_n \text{ en } a \\ \\ 2. \quad \dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_n} \text{ sont } \mathscr{C}^0 \text{ sur } U \text{ (Comme fonctions de plusieurs variables)} \end{cases}$$

Définition: Application de classe $\mathscr{C}^k$

Notre programme n'évoquant pas la notion de différentielle d'ordre supérieur, la caractérisation des applications de classe $\mathscr{C}^k$ devient notre définition d'application de classe $\mathscr{C}^k$:

Soit $U \subset \mathbb{R}^n$. Soit $f : U \to F$ avec U ouvert. Soit $k \in \mathbb{N}$ $(k \geqslant 2)$.

On dit que f est de classe $\mathscr{C}^k$ sur U si :

1. f admet des dérivées partielles par rapport à tout $k-$uplet de variables en tout point de U
2. Toutes ces dérivées partielles sont $\mathscr{C}^0$ sur U

1.6 Théorème de Schwarz

Théorème de Schwarz

Soit $U \subset \mathbb{R}^n$ ouvert. Soit F : $\mathbb{R}$-EVN de dimension finie. Soit $f : U \to F$.
Soit $a \in U$. On suppose f de classe $\mathscr{C}^2$ au voisinage de a.

Alors, $\forall i, j \in [\![1, n]\!]$, $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$

1.7 Définition d'un vecteur tangent à une partie.

Définition: Vecteur tangent à une partie

Soit $E : \mathbb{R}-$EVN de dimension finie. Soit $X \subset E$ une partie.
Soit $a \in X$ et $u \in E$.

On dit que u est un vecteur tangent à X en a si :

$$\exists \varepsilon > 0,\ \exists \gamma : \begin{cases}]-\varepsilon, \varepsilon[\to X \\ t \mapsto \gamma(t) \end{cases} \quad \text{Arc } \mathscr{C}^1 \text{ tel que } \gamma(0) = a \text{ et } \gamma'(0) = u$$

On note $T_a(X)$ voir $T_a X$ l'ensemble des vecteurs tangents à X en a.

1.8 Définition de la matrice Hessienne et formule de Taylor à l'ordre 2.

Définition: Matrice Hessienne

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

On appelle Matrice Hessienne de f en a la matrice :

$$H_f(a) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j}$$

Proposition Formule de Taylor-Young : Ordre 2

Soit $E = \mathbb{R}^n$, muni de sa structure Euclidienne canonique. Soit $U \subset \mathbb{R}^n$ Ouvert. Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur U.
Soit $a \in U$.

$$\begin{aligned} \forall h \in E,\ a + h \in U \Rightarrow\ f(a+h) &= f(a) + df(a)(h) + \frac{1}{2}\langle H_f(a)h; h\rangle + o(h^2) \\ &= f(a) + \langle \nabla_f(a); h\rangle + \frac{1}{2}\langle H_f(a)h; h\rangle + o(h^2) \\ &= f(a) + h^\top \left(\nabla_f(a) + \frac{1}{2} H_f(a)h \right) + o(h^2) \end{aligned}$$

1.9 Lien entre extrema et caractère positif/défini positif de la matrice Hessienne.

Proposition

Si f admet un Minimum Local ou Global en a, Alors :

1. $df(a) = 0$, i.e : $\dfrac{\partial f}{\partial x_1}(a) = \cdots = \dfrac{\partial f}{\partial x_n}(a) = 0$
2. $H_f(a) \in S_n^+(\mathbb{R})$

Proposition

Si :

1. $df(a) = 0$
2. $H_f(a) \in S_n^{++}(\mathbb{R})$

Alors a est un Minimum Local de f.

Idem, si $df(a) = 0$ et $-H_f(a) \in S_n^{++}(\mathbb{R})$, alors a est un Maximum Local de f.

2 Questions de cours, groupes B et C

2.1 Différentiabilité d'une composée (démo)

Proposition

Soient E, F, G trois $\mathbb{R}$-EVN de dimension finies. Soit $U \subset E$ et $V \subset F$. Soit $f : U \to F$ et $g : V \to G$, deux applications telles que $f(U) \subset V$.
Soit $a \in U$. On pose $b = f(a)$. On suppose f différentiable en a et g différentiable en b.

Alors $g \circ f$ est différentiable en a et $d(g \circ f)(a) = dg(b) \circ df(a)$

Preuve :

$\forall h \in E$ tel que $a + h \in U$. Alors $f(a+h) = f(a) + df(a)(h) + o(h)$. Dès lors :

$$\begin{aligned} g(f(a+h)) &= g(b + \underbrace{df(a)(h) + o(h)}_{=h'}) \\ &= g(b) + dg(b)(h') + o(h') \\ &= g(b) + dg(b)(df(a)(h)) + dg(b)(o(h)) + o(h') \\ &= g(b) + (dg(b) \circ df(a))(h) + o(h) \end{aligned}$$

$dg(b)$ Linéaire

Or, $dg(b) \circ df(a)$ est Linéaire : $g \circ f$ est alors différentiable en a et $d(g \circ f)(a) = dg(f(a)) \circ df(a)$.

Montrons que $dg(b)(o(h)) = o(h)$: $dg(b) \in \mathscr{L}(F, G)$, en dimension finie donc est $\mathscr{C}^0$.
Donc, $\exists K \in \mathbb{R}_+,\ \|dg(b)(o(h))\|_G \leqslant K \times o(\|h\|) \Rightarrow\ dg(b)(o(h)) = o(h)$

2.2 Règle de la chaîne. (démo)

Proposition Règle de la Chaîne

Soient $n, p, q \in \mathbb{N}^*$. Soit $g : U \subset \mathbb{R}^n \to \mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \to \mathbb{R}^q$.
On notera $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$. On suppose U et V ouverts.
On suppose $f(U) \subset V$, $g \circ f$ existe. Soit $a \in U$, on pose $b = f(a)$.
On suppose f différentiable en a et g différentiable en b. Donc $g \circ f$ est différentiable en a et $d(g \circ f)(a) = dg(b) \circ df(a)$.
On munit $\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^q$ de leur base canonique.

Dans ces bases : $\mathrm{Jac}(g \circ f)(a) = \mathrm{Jac}(g)(b) \times \mathrm{Jac}(f)(a)$. Or :

$$\mathrm{Jac}(f)(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} \quad \text{et} \quad \mathrm{Jac}(g)(b) = \left(\frac{\partial g_i}{\partial y_j}(b)\right)_{\substack{1 \leqslant i \leqslant q \\ 1 \leqslant j \leqslant p}}$$

$$\mathrm{Jac}(g \circ f)(a) = \left(\sum_{k=1}^{p} \mathrm{Jac}(g)(b)_{i,k} \times \mathrm{Jac}(f)(a)_{k,j}\right)_{\substack{1 \leqslant i \leqslant q \\ 1 \leqslant j \leqslant n}} = \left(\sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(b) \times \frac{\partial f_k}{\partial x_j}(a)\right)_{\substack{1 \leqslant i \leqslant q \\ 1 \leqslant j \leqslant n}}$$

2.3 Caractérisation des fonctions de classe $\mathscr{C}^1$.

Théorème Caractérisation des applications de classe $\mathscr{C}^1$

Soient $n, p \in \mathbb{N}^*$. Soit $U \subset \mathbb{R}^n$ ouvert. Soit $f : U \to \mathbb{R}^p$. Alors :

$$\left[f \text{ est } \mathscr{C}^1 \text{ sur } U\right] \iff \begin{cases} 1. \quad \forall a \in U,\ \ f \text{ admet des dérivées partielles par rapport à } x_1, \ldots, x_n \text{ en } a \\ 2. \quad \dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n} \text{ sont } \mathscr{C}^0 \text{ sur } U \text{ (Comme fonctions de plusieurs variables)} \end{cases}$$

2.4 Lien entre espace tangent et noyau de la différentielle. (démo d'une inclusion)

Proposition

Soit $E : \mathbb{R}-$EVN et $g : E \to \mathbb{R}$ de classe $\mathscr{C}^1$.
On note $X = \{x \in E \mid g(x) = 0\}$.

Alors $\forall x_0 \in X,\ \ [dg(x_0) \neq 0 \Rightarrow T_{x_0}(X) = \ker(dg(x_0))]$

***Preuve* Inclusion directe Uniquement :**

Soit $a \in X$, $u_0 \in T_a(X)$. Alors, $\exists \varepsilon > 0$ et $\exists \gamma : \begin{cases}]-\varepsilon;\varepsilon[\to X \\ t \mapsto \gamma(t) \end{cases}$, Arc $\mathscr{C}^1$ tel que $\gamma(0) = a$ et $\gamma'(0) = u_0$.

Alors, $\forall t \in]-\varepsilon, \varepsilon[,\ \ \gamma(t) \in X \Rightarrow\ g(\gamma(t)) = 0$. Dès lors :
$\forall t \in]-\varepsilon, \varepsilon[,\ \ d(g \circ \gamma)(t) = dg(\gamma(t))(\gamma'(t)) = 0$

$\Rightarrow$ pour $t = 0$: $dg(a)(u_0) = 0 \Rightarrow\ u_0 \in \ker(dg(a))$

2.5 Le gradient donne la direction de variation maximale d'une fonction scalaire (démo)

***Preuve* :**

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$. Soit $a \in \mathring{U}$ tel que f soit différentiable en a.
On cherche u, unitaire de direction correspondant à "l'augmentation maximale de f" : Tel que $D_u f(a)$ soit maximale.

$\forall h \in E$ tel que $a + h \in U$, nous avons :

$$\begin{aligned} f(a+h) &= f(a) + df(a)(h) + o(h) \\ &= f(a) + D_h f(a) + o(h) \\ &= f(a) + \langle \nabla_f(a); h\rangle + o(h) \end{aligned}$$

Afin de maximiser la quantité $\langle \nabla_f(a); h\rangle$, il faut que h soit colinéaire à $\nabla_f(a)$ et de même sens par le cas d'égalité de Cauchy-Schwartz :

$|\langle \nabla_f(a); h\rangle| \leqslant \|\nabla_f(a)\| \|h\|$ avec égalité si et seulement si h est colinéaire à $\nabla_f(a)$.

Si h est anticolinéaire à $\nabla_f(a)$, alors le produit scalaire devient négatif. Ainsi, il faut h de même direction et sens que $\nabla_f(a)$ pour maximiser $\langle \nabla_f(a); h \rangle$. La contrainte de h unitaire donne l'unicité du vecteur recherché.

2.6 Espace tangent à une partie donnée par l'équation scalaire $g(x) = 0$. (démo d'une inclusion)

C.F la démo "Lien entre espace tangent et noyau de la différentielle"

2.7 Théorème d'optimisation sous une contrainte. (démo)

> **Théorème** d'optimisation sous Contrainte
> Soit $U \subset \mathbb{R}^n$ un ouvert. Soient f et $g : U \to \mathbb{R}$ de classe $\mathscr{C}^1$.
> On note $X = \{x \in E \mid g(x) = 0\}$. Soit $a \in X$, $dg(a) \neq 0$.
> On note $\tilde{f} = f_{|X}$.
>
> Si $\tilde{f}$ admet un extrémum local en a, alors $df(a)$ est colinéaire à $dg(a)$ et $\nabla_f(a)$ est colinéaire à $\nabla_g(a)$

Preuve :

Nous savons que si $\tilde{f}$ admet un extrémum en a, alors $\forall u \in T_a(X),\ df(a)(u) = 0$. Donc $T_a(X) \subset \text{Ker}(df(a))$.

Or, $T_a(X) = \text{Ker}(dg(a)) \Rightarrow \text{Ker}(dg(a)) \subset \text{Ker}(df(a))$.

Or, $dg(a) \neq 0$, i.e : $dg(a) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R})$ et est non-nulle : Donc $\text{Ker}(dg(a))$ est un Hyperplan de $\mathbb{R}^n$:
$\exists v_0 \in \mathbb{R}^n,\ \mathbb{R}^n = \text{Ker}(dg(a)) \oplus \text{Vect}(\{v_0\})$ et $g(v_0) \neq 0$.

$\forall x \in \mathbb{R}^n, \exists!(x_0, \lambda) \in \text{Ker}(dg(a)) \times \mathbb{R},\ x = x_0 + \lambda v_0$. Alors $dg(a)(x) = \lambda dg(a)(v_0)$.

$df(a)(x) = \lambda df(a)(v_0)$ car $x_0 \in \text{Ker}(dg(a)) \subset \text{Ker}(df(a))$.

Or, $\lambda df(a)(v_0) = df(a)(v_0) \times \dfrac{dg(a)(x)}{dg(a)(v_0)} = \alpha dg(a)(x)$ avec $\alpha = \dfrac{df(a)(v_0)}{dg(v_0)}$.
Ainsi, $\exists \alpha \in \mathbb{R},\ \forall x \in E,\ df(a)(x) = \alpha dg(a)(x)$, donc $df(a) = \alpha dg(a)$.

Si on munit $\mathbb{R}^n$ de sa structure Euclidienne canonique :

$$\begin{aligned} \forall x \in \mathbb{R}^n,\ dg(a)(x) &= \langle \nabla_g(a); x \rangle \\ df(a)(x) &= \langle \nabla_f(a); x \rangle \\ &= \alpha \langle \nabla_g(a); x \rangle \end{aligned}$$

Donc, $\forall x \in \mathbb{R}^n,\ \langle \nabla_f(a) - \alpha \nabla_g(a); x \rangle = 0 \Rightarrow \nabla_f(a) - \alpha \nabla_g(a) = 0 \Rightarrow \nabla_f(a) = \alpha \nabla_g(a)$

2.8 Sur un ouvert, les extrema d'une fonction scalaire différentiable sont des points critiques (démo)

Définition

Soient E, F deux EVN de dimension finie. Soit $U \subset E$ ouvert. Soit $f : U \to F$ différentiable.
Soit $a \in U$.

On dit que a est un point critique de f si $df(a) = 0$.

En particulier, si $E = \mathbb{R}^n$: $\left[df(a) = 0 \iff \frac{\partial f}{\partial x_1}(a) = \cdots = \frac{\partial f}{\partial x_n}(a) = 0\right]$

Théorème

Soit $E : \mathbb{R}-$EVN de dimension finie. Soit $U \subset E$ ouvert. Soit $f : U \to \mathbb{R}$ application différentiable.
Soit $a \in U$.

1. On dit que f admet un Minimum local en a si :

$$\exists r>0,\ \forall x \in B_f(a, r),\ f(a) \leqslant f(x)$$

Idem, on dit que f admet un Maximum local en a si :

$$\exists r>0,\ \forall x \in B_f(a, r),\ f(a) \geqslant f(x)$$

2. On dit que f admet un Minimum Global en a si :

$$\forall x \in U,\ f(x) \geqslant f(a)$$

Idem, On dit que f admet un Maximum Global en a si :

$$\forall x \in U,\ f(x) \leqslant f(a)$$

Si f admet un Extremum Local ou Global en a, alors $df(a) = 0$. i.e a est un point critique de f

Preuve :

Soit $a \in U$. On suppose que f admet un Minimum local en a.

Alors $\exists r>0,\ B_f(a, r) \subset U$ et tel que $\forall x \in B_f(a, r),\ f(a) \leqslant f(x)$.

Soit $u_0 \in E$. Alors pour t assez petit : $\|tu_0\| \leqslant r$. Donc $a + tu_0 \in B_f(a, r)$.

Or, $0 \leqslant f(a + tu_0) - f(a) = df(a)(tu_0) + o(t) = tdf(a) + o(t)$ (car a est un Min local).

Si $df(a)(u_0) \neq 0$, alors $f(a+tu_0) - f(a) \sim tdf(a)(u_0)$. Donc $f(a+tu_0) - f(a)$ change de signe au voisinage de 0.
Or, $0 \leqslant f(a + tu_0) - f(a) \Rightarrow$ Absurde. Donc $df(a)(u_0) = 0$

Ainsi, a est un point critique de f : $df(a) = 0$

2.9 Exemple d'équation aux dérivées partielles sur un convexe.

Exemple

Soit $U = \mathbb{R}^2$ (convexe). Résolvons l'équation différentielle $2\dfrac{\partial f}{\partial x} - 7\dfrac{\partial f}{\partial y} = 0$.

On pose u et v, autre système de coordonées tel que $x(u,v) = 2u + v$ et $y(u,v) = -7u$.

Ce changement de variables est bien bijectif. On pose alors $g(u,v) = f(x(u,v), y(u,v))$.

Alors $\dfrac{\partial g}{\partial u} = \dfrac{\partial x}{\partial u}\dfrac{\partial f}{\partial x} + \dfrac{\partial y}{\partial u}\dfrac{\partial f}{\partial y} = 2\dfrac{\partial f}{\partial x} - 7\dfrac{\partial f}{\partial y}$.

Ainsi, notre équation différentielle se ramène à $\dfrac{\partial g}{\partial u} = 0$, Puisque $\mathbb{R}^2$ est Convexe : $\exists h : \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^1$ telle que $g(u,v) = h(v)$ pour tout $(u,v) \in \mathbb{R}^2$.

Dès lors, $\forall (x,y) \in \mathbb{R}^2,\ \ f(x,y) = h\left(x + \dfrac{2}{7}y\right)$

3 Questions de cours du groupe C

3.1 Différentielle de $B(f,g)$ où B est bilinéaire en dimension finie. (démo)

Proposition

Soient E, F, G, trois $\mathbb{R}-$EVN de dimension finie. Soit $U \subset E$. Soient $f : U \to F$ et $g : U \to G$, deux applications.
Soit $B : F \times G \to H$, une application Bilinéaire (Continue car en dimension finie).

On considère $B(f,g) : \begin{cases} U \to H \\ x \mapsto B(f(x), g(x)) \end{cases}$. Soit $a \in U$.

Si f et g sont différentiables en a, alors $B(f,g)$ est différentiable en a et $dB(f,g)(a) = B(df(a), g(a)) + B(f(a), dg(a))$

Preuve :

$$\begin{aligned} B(f(a+h), g(a+h)) &= B(f(a) + df(a)(h) + o(h), g(a) + dg(a)(h) + o(h)) \\ &= B(f(a), g(a)) + B(f(a), dg(a)(h)) + B(df(a)(h), g(a)) + o(h) \end{aligned}$$

Or, $B(f(a), dg(a)(h)) + B(df(a)(h), g(a))$ est linéaire en h.

Justifions les $o(h)$: B est Continue et Bilinéaire, donc $\exists K \in \mathbb{R},\ \forall x, y \in F \times G,\ \ \|B(x,y)\|_H \leqslant K \times \|x\|_F \times \|y\|_G$.

Ainsi, $\|B(df(a)(h), dg(a)(h))\|_H \leqslant K \times \|df(a)(h)\|_F dg(a)(h)_G$.

Or, $df(a)$ et $dg(a)$ sont linéaires continues : $\exists K_1, K_2 \in \mathbb{R}^2$ tels que $\|df(a)(h)\|_F \leqslant K_1 \times \|h\|_E$ et Idem pour $dg(a)(h)$.

Ainsi, $\|B(df(a)(h), dg(a)(h))\|_H \leqslant K \times K_1 \times \|h\| \times K_2 \times \|h\| = \mathcal{O}(\|h^2\|) = o(h)$.

Donc, $B(df(a)(h), dg(a)(h)) = o(h)$. Idem, $B(X, o(h)) = o(h)$ pour les mêmes raisons.

3.2 Théorème de Schwarz. (démo)

***Preuve* La démo n'est pas dans le cours (Admise) :**

Ramenons nous au cas où f est une fonction de deux variables. La démonstration est analogue pour plus de variables. Ainsi, soit $f : U \subset \mathbb{R}^2 \to F$.
Soit $a \in U$ et f de classe $\mathscr{C}^2$ au voisinage de $a = (x, y)$.

Considérons $\Delta : t \mapsto [f(x+t, y+t) - f(x+t, y)] - [f(x, y+t) - f(x, y)]$. Cette application correspond à une variation selon l'axe des y (Remarquez le carré de sommets $f(a,b), f(a+t,b), f(a,b+t), f(a+t,b+t)$).

Alors, pour t assez petit, $f(x+t, y+t), f(x+t, y)$ et $f(x, y+t) \in U$. Ainsi, Δ est une fonction d'une variable dérivable par hypothèse sur f.

Posons donc t assez petit, et considérons donc $\delta_t : s \mapsto [f(a+t, b+s) - f(a+t, b)] - [f(a, b+s) - f(a, b)]$
Ainsi, nous pouvons appliquer l'égalité des accroissements finis à δ_t. (Cette fonction est continue, et sa dérivée correspond à une dérivée partielle de f, ici selon y). Ainsi, $\exists \eta \in]0, t[$ tel que :

$$\begin{aligned}
\Delta(t) &= \frac{\partial f}{\partial y}(a+t, b+\eta) \times t - \frac{\partial f}{\partial y}(a, b+\eta) \times t \\
&= t \times \left[\frac{\partial f}{\partial y}(a,b) + \frac{\partial^2 f}{\partial x \partial y}(a,b)t + \frac{\partial^2 f}{\partial y^2}(a,b)\eta + o(t) - \frac{\partial f}{\partial y}(a,b) - \frac{\partial^2 f}{\partial y^2}(a,b)\eta + o(t)\right] \\
&= \frac{\partial^2 f}{\partial x \partial y}(a,b) \times t^2 + o(t^2)
\end{aligned}$$

Formule de Taylor (Hyp

Ainsi, $\dfrac{\Delta(t)}{t^2} = \dfrac{\partial^2 f}{\partial x \partial y}(a,b) + o(1) \to \dfrac{\partial^2 f}{\partial x \partial y}(a,b)$

Nous pouvons effectuer le même calcul en considérant cette fois une variation selon l'axe des x : On considèrera $\tilde{\Delta} : t \mapsto [f(a+t, b+t) - f(a, b+t)] - [f(a+t, b) - f(a, b)]$.

Le calcul est identique à l'exception de l'ordre des dérivées partielles : Nous obtenons $\dfrac{\tilde{\Delta}(t)}{t^2} \to \dfrac{\partial^2 f}{\partial y \partial x}(a,b)$.

Nous pouvons remarquer (en développant) que $\Delta = \tilde{\Delta}$, d'où finalement :

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

3.3 Caractérisation des fonctions de classe $\mathscr{C}^1$ sur un ouvert à l'aide des dérivées partielles. (démo)

Théorème Caractérisation des applications de classe $\mathscr{C}^1$

Soient $n, p \in \mathbb{N}^*$. Soit $U \subset \mathbb{R}^n$ ouvert. Soit $f : U \to \mathbb{R}^p$. Alors :

$$\left[f \text{ est } \mathscr{C}^1 \text{ sur } U\right] \iff \begin{cases} 1. \quad \forall a \in U,\ f \text{ admet des dérivées partielles par rapport à } x_1, \dots, x_n \text{ en } a \\ 2. \quad \dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_n} \text{ sont } \mathscr{C}^0 \text{ sur } U \text{ (Comme fonctions de plusieurs variables)} \end{cases}$$

Preuve :

$\underline{\Rightarrow}$: Si f est différentiable sur U et que $a \mapsto df(a)$ est $\mathscr{C}^0$:

1. $\forall a \in U$, (f est différentiable en $a \Rightarrow$ f admet des dérivées partielles en a) et $\dfrac{\partial f}{\partial x_i} = df(a)(e_i)$ où e_i est le i−ème vecteur de la Base canonique.
2. $a \mapsto df(a)$ est $\mathscr{C}^0$ par hypothèse :
   $\forall i \in [\![1, n]\!]$, $\begin{cases} \mathscr{L}(E, F) \to F \\ \varphi \mapsto \varphi(e_i) \end{cases}$ est Continue car linéaire en dimension finie. Par composition : $a \mapsto df(a)(e_i)$ est Continue sur U. Donc $a \mapsto \dfrac{\partial f}{\partial x_i}(a)$ est Continue sur U.

$\underline{\Leftarrow}$: Prenons le cas particulier $n = 2$. La démo est analogue pour $n \geqslant 2$ quelconque.

Soit donc $U \subset \mathbb{R}^2$, soit $a = (x_a, y_a) \in U$. Montrons que f est différentiable en a :

Soit $h = (\alpha, \beta)$. Posons $\|h\| = \|h\|_\infty = \max(|\alpha|, |\beta|)$. Alors, pour $\|h\|$ assez petit, $a + h \in U$ car U est un ouvert.

On pose $B = a + \alpha e_1$ et $C = a + \alpha e_1 + \beta e_2 = a + h$.

Alors $f(B) - f(a) = f(a + \alpha e_1) - f(a) = \displaystyle\int_{x_a}^{x_a+\alpha} \frac{\partial f}{\partial x}(t, y_a)dt$.

Donc : $f(B) - f(a) - \alpha\dfrac{\partial f}{\partial x}(a) = \displaystyle\int_{x_a}^{x_a+\alpha} \left(\frac{\partial f}{\partial x}(t, y_a) - \frac{\partial f}{\partial x}(x_a, y_a)\right) dt$

Or, $\dfrac{\partial f}{\partial x}$ est Continue en a. Ainsi, $\displaystyle\lim_{t \to x_a} \frac{\partial f}{\partial x}(t, y_a) = \frac{\partial f}{\partial x}(x_a, y_a)$.

Ainsi, pour $\|h\|$ assez petit : $\|\dfrac{\partial f}{\partial x}(t, y_a) - \dfrac{\partial f}{\partial x}(x_a, y_a)\| \leqslant \varepsilon$ pour $t \in [x_a, x_a + \alpha]$.

Dès lors, $\|\displaystyle\int_{x_a}^{x_a+\alpha} \frac{\partial f}{\partial x}(t, y_a)dt - \int_{x_a}^{x_a+\alpha} \frac{\partial f}{\partial x}(x_a, y_a)dt\| = \|\int_{x_a}^{x_a+\alpha} \frac{\partial f}{\partial x}(t, y_a) - \frac{\partial f}{\partial x}(x_a, y_a)dt\| \leqslant \varepsilon|\alpha| \leqslant \varepsilon\|h\|$

Or, ceci est vrai pour tout $\varepsilon>0$. Donc $\displaystyle\int_{x_a}^{x_a+\alpha} \left(\frac{\partial f}{\partial x}(t, y_a) - \frac{\partial f}{\partial x}(x_a, y_a)\right) dt = o(h)$.

Donc, $f(B) - f(a) = \alpha\dfrac{\partial f}{\partial x}(a) + o(h)$. De même, $f(C) - f(B) = \beta\dfrac{\partial f}{\partial y} + o(h)$.

Par somme : $f(C) - f(B) + f(B) - f(a) = f(C) - f(a) = f(a + h) - f(a) = \alpha\dfrac{\partial f}{\partial x}(a) + \beta\dfrac{\partial f}{\partial y}(a) + o(h)$.

Ainsi, f est différentiable en a et $df(a) = (\alpha, \beta) \mapsto \alpha\dfrac{\partial f}{\partial x}(a) + \beta\dfrac{\partial f}{\partial y}(a)$. Or, $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues, donc $a \mapsto df(a)$ est Continue par continuité des coefficients.

3.4 Lien entre extrema et caractère positif/défini positif de la matrice Hessienne. (démo)

Proposition
Si f admet un Minimum Local ou Global en a, Alors :

1. $df(a) = 0$, i.e : $\frac{\partial f}{\partial x_1}(a) = \cdots = \frac{\partial f}{\partial x_n}(a) = 0$
2. $H_f(a) \in S_n^+(\mathbb{R})$

Preuve :

1. Déjà fait
2. $\forall h \in E$ tel que $a + h \in U$ (avec a extrémal) :

$$f(a+h) = f(a) + 0 + \frac{1}{2}\langle H_f(a)h; h\rangle + o(h^2)$$

Donc, pour h assez petit : $f(a+h) - f(a)$ est du signe de $\langle H_f(a)h; h\rangle$ (si $\neq 0$).

S'il existe h assez petit tel que $\langle H_f(a)h; h\rangle < 0$, alors a n'est pas un minimum, ce qui est absurde.

Si $\exists \lambda \in \mathrm{Sp}(H_f(a))$ tel que $\lambda < 0$, alors $\exists h_0 \in \mathbb{R}^n$ tel que $H_f(a)h_0 = \lambda h_0$.

Posons alors $h = th_0$. Pour t assez petit : $a + h \in U$ car U est un ouvert et $f(a+h) - f(a)$ est du signe de $\langle H_f(a)h; h\rangle = t^2\lambda\|h_0\|^2 < 0$.

Ainsi, pour t assez petit, $f(a+h) - f(a) < 0$: Ce qui est absurde car a est un Min.

Ainsi, $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}_+$. Or, $H_f(a) \in S_n(\mathbb{R})$, donc $H_f(a) \in S_n^+(\mathbb{R})$

Proposition
Si :

1. $df(a) = 0$
2. $H_f(a) \in S_n^{++}(\mathbb{R})$

Alors a est un Minimum Local de f.

Idem, si $df(a) = 0$ et $-H_f(a) \in S_n^{++}(\mathbb{R})$, alors a est un Maximum Local de f.

Preuve :

Pour $h \neq 0$ assez petit :
$f(a+h) - f(a)$ est du signe de $\langle H_f(a)h; h\rangle (\neq 0)$ car $H_f(a) \in S_n^{++}(\mathbb{R}) \Rightarrow$ a est un Minimum local de f.

3.5 Formule de Taylor à l'ordre 2. (démo HP)

Proposition Formule de Taylor-Young : Ordre 2

Soit $E = \mathbb{R}^n$, muni de sa structure Euclidienne canonique. Soit $U \subset \mathbb{R}^n$ Ouvert. Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur U. Soit $a \in U$.

$$\begin{aligned}\forall h \in E,\ \ a+h \in U \Rightarrow\ f(a+h) &= f(a) + df(a)(h) + \frac{1}{2}\langle H_f(a)h; h\rangle + o(h^2)\\ &= f(a) + \langle \nabla_f(a); h\rangle + \frac{1}{2}\langle H_f(a)h; h\rangle + o(h^2)\\ &= f(a) + h^\top \left(\nabla_f(a) + \frac{1}{2}H_f(a)h\right) + o(h^2)\end{aligned}$$

Preuve :

Soit $a \in U$ et h tel que $a+h \in U$.

Posons $g : x \mapsto f(a+x) - f(a) - df_a(x) - \frac{1}{2}d^2f_a(x,x)$. Alors cette application est différentiable.

De plus, $d_x g(h) = df_{a+x}(h) - df_a(h) - d^2f_a(x,h)$ (le terme en $f(a)$ saute et le terme en $d^2f_a(x,x)$ se différentie en $d^2f_a(x,h)$ et $d_a^2f(h,x)$ et le théorème de Schwarz permet d'intervertir).

Or, nous pouvons appliquer la formule de Taylor Young (Ordre 1) à df, car f est en particulier $\mathscr{C}^1$, donc différentiable au point a. Ainsi, $df_{a+h} - df_a = d^2f_a(h,\bullet) + o(h)$. En norme : $\|dg_h\| \leqslant \varepsilon\|h\|$ pour $\varepsilon \xrightarrow[h\to 0]{} 0$

Nous pouvons donc appliquer l'inégalité des accroissements finis :

$$\begin{aligned}\|f(a+h) - f(a) - df_a(h) - \frac{d^2f_a(h,h)}{2}\| &= \|g(h) - g(0)\|\\ &\leqslant \sup_{x\in B_f(0,h)} \|dg_x\| \times \|h\|\\ &\leqslant \varepsilon\|h\|^2\end{aligned}$$

D'où l'inégalité souhaitée, et la formule de Taylor-Young à l'ordre deux. (Voir ci-dessous pour généraliser cette démonstration à tout ordre).

3.6 BONUS : Formule de Taylor Young - ordre n

Théorème

Soit $E = \mathbb{R}^n$, muni de sa structure Euclidienne canonique. Soit $p \in \mathbb{N}$. Soit $U \subset \mathbb{R}^n$ Ouvert. Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^p$ sur U. Soit $a \in U$

$$\begin{aligned}\forall x \text{ tq } a+x \in U,\ \ f(a+x) &= \sum_{k=0}^{p} \sum_{\substack{\alpha\in\mathbb{N}^k \\ |\alpha|=k}} \frac{1}{\alpha!} D^p_\alpha f_a(\underbrace{x, \cdots, x}_{\text{k fois}}) + o(\|h\|^p) \\ &= \sum_{k=0}^{p} \frac{1}{k!} D^k f_a(x, \cdots, x) + o(\|h\|^p)\end{aligned}$$

Preuve **(BONUS) Formule Générale, à tout ordre *:***

Procédons par récurrence sur l'ordre de la formule de taylor : Nous avons déjà démontré la formule à l'ordre 1.

Posons premièrement la notion de multi-indice pour commodité de calcul : On appelle multi-indice d'ordre p tout n-uplet $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ tel que $\sum_i \alpha_i = p$. On note $p = |\alpha|$.

Ceci nous permet de noter $D^p_\alpha f$, les dérivées partielles successives $D^p_{x_1^{\alpha_1} \cdots x_n^{\alpha_n}} f$.

Ceci nous permet également de noter plus aisément les polynômes à n variables de degré p :
On note $X^\alpha := {X_1}^{\alpha_1} \times \cdots \times {X_n}^{\alpha_n}$

Par hypothèse, nous supposons f de classe $\mathscr{C}^p$, donc df est de classe $\mathscr{C}^{p-1}$, nous pouvons donc appliquer l'hypothèse de récurrence afin de déduire un DL de df en a :

$$df_{a+h}(\bullet) = \sum_{k=0}^{p-1} c_k(\underbrace{h, \ldots, h}_{\text{k fois}}, \underbrace{\bullet, \ldots, \bullet}_{\text{p}-\text{k fois}}) + o(\|h\|^{p-1})$$

où c_k désigne l'application multilinéaire obtenue en sommant toutes les applications dérivées partielles d'ordre k. D'après le théorème de schwarz, nous pouvons intervertir l'ordre de dérivation partielle, et les applications c_k sont alors symétriques (nous pouvons échanger deux paramètres sans modifier le résultat, d'où l'écriture avec h en premier).

Alors, en posant $g : h \mapsto f(a+h) - \displaystyle\sum_{k=0}^{p-1} \frac{1}{k+1} c_k(\underbrace{h, \ldots, h}_{\text{k+1 fois}}, \underbrace{\bullet, \ldots, \bullet}_{\text{p}-\text{k}-\text{1 fois}})$, g est différentiable et :

$$dg_h = df_{a+h} - \sum_{k=0}^{p-1} c_k(\underbrace{h, \ldots, h}_{\text{k fois}}, \underbrace{\bullet, \ldots, \bullet}_{\text{p}-\text{k fois}}) = o(\|h\|^{p-1})$$

Il nous suffit alors d'appliquer le corollaire de l'IAF : Du fait que dg est linéaire en dimension finie, nous avons $dg_h \leqslant \|h\|^{p-1} \times \varepsilon(h) \Rightarrow\ g(h) \leqslant \|h\|^p \times \varepsilon'(h)$.

Finalement, $f(a+h) = \displaystyle\sum_{k=0}^{p} \frac{1}{k!} D^k f_a(h, \ldots, h) + o(\|h\|^p)$