

Table des matières

13 Calcul différentiel.	3
13.1 application différentiable	3
13.1.1 Dérivée suivant un vecteur	3
13.1.1.1 Définition, exemples	3
13.1.1.2 Dérivées partielles	4
13.1.1.3 Dérivée suivant un vecteur et composantes	4
13.1.2 Application différentiable	4
13.1.2.1 Différentiabilité en un point	4
13.1.2.2 Exemples d'applications différentiables	6
13.1.2.3 Dérivabilité et différentiabilité	6
13.1.3 Différentiabilité et dérivées partielles	7
13.1.4 Matrice jacobienne, jacobien	9
13.1.5 Opérations sur les applications différentiables	10
13.1.5.1 Combinaison linéaire	10
13.1.5.2 Composition	11
13.1.5.3 Produit via une application bilinéaire de deux fonctions différentiables	13
13.1.5.4 Composé avec une application linéaire	13
13.1.5.5 Différentiabilité et composantes	14
13.2 Fonction de classe $C^k, k \in \mathbb{N}^* \cup \{\infty\}$	15
13.2.1 Fonction de classe C^1	15
13.2.1.1 Définitions, caractérisation	15
13.2.1.2 Quelques propriétés	18
13.2.2 Fonctions de classe $C^k, k \in \mathbb{N} \cup \{\infty\}$	19
13.2.2.1 Dérivées partielles successives	19
13.2.2.2 Théorème de Schwarz	19
13.2.2.3 Opérations sur les fonctions de classe C^k	19
13.2.3 Formule de Taylor d'ordre 2	20
13.3 Fonction à valeur réelle	20
13.3.1 Gradient	20
13.3.1.1 Définition, exemple	21
13.3.2 Extrémums sans contrainte	21
13.3.2.1 Point critique, maximum, minimum, extremum	21
13.3.2.2 Condition nécessaire d'extremum	22
13.3.2.3 Conditions du second ordre d'extremum	22
13.3.2.4 Cas particulier de $n = 2$	24
13.3.3 extremums avec contrainte ou extrémums liés	25
13.3.4 Application à la géométrie différentielle	27
13.3.4.1 Vecteur tangent, variété tangente	27
13.3.4.2 Lignes de niveau, surfaces de niveau	28
13.3.4.3 Plan tangent à une surface de niveau $f(x, y, z) = 0$	28
13.3.4.4 Gradient et dérivée directionnelle	28

Chapitre 13

Calcul différentiel.

13.1 application différentiable

Dans tout ce qui suit les espaces vectoriels introduits sont des $\mathbb{R}$ –espaces vectoriels normés de dimensions finies. Quand on cite : $f : U \subset E \rightarrow F$ application, on entend par U un ouvert non vide de l'evn E et F un evn bien sûr tous de dimensions finies.

13.1.1 Dérivée suivant un vecteur

13.1.1.1 Définition, exemples

Proposition-Définition 13.1.1. Soit $f : U \subset E \rightarrow F$ une application, $a \in U$, e un vecteur de E et $\varphi_{a,e}$ l'application définie par $\varphi_{a,e}(t) = f(a + te)$. L'application $\varphi_{a,e}$ est bien définie sur un intervalle de la forme $] -\alpha, \alpha[$ avec $\alpha > 0$. Si $\varphi_{a,e}$ est dérivable au point 0 le vecteur $\varphi'_{a,e}(0)$ s'appelle dérivée de f au point a suivant le vecteur e . On le note $D_e f(a)$.

Ainsi, s'il existe, $D_e f(a) = \lim_{t \rightarrow 0} \frac{f(a+te) - f(a)}{t}$.

Preuve. Démontrons l'existence de $\alpha > 0$ tel que $\forall t \in]-\alpha, \alpha[, a + te \in U$. Supposons $e \neq 0$ (si $e = 0$ alors $a + te = a \in U$, pour tout $t \in \mathbb{R}$). Comme U est un ouvert et $a \in U$, il existe $R > 0$ tel que $B(a, R) \subset U$. Soit $\alpha = \frac{R}{\|e\|}$, alors pour tout réel t tel que $|t| < \alpha$, on a $\|(a + te) - a\| = |t|\|e\| < \alpha\|e\| = R$, donc $a + te \in B(a, R)$, et comme $B(a, R) \subset U$, on a $a + te \in U$. ■

Exemples. :

1. $f : \mathbb{R}^2 \rightarrow \mathbb{R}; (x, y) \mapsto f(x, y) = x^2y + 3x^3 + xy$; $a = (1, 2)$ et $e = (1, 1)$. Alors

$$\begin{aligned}\varphi_{a,e}(t) &= f(a + te) \\ &= f((1, 2) + (t, t)) \\ &= f(1 + t, 2 + t) \\ &= (1 + t)^2(2 + t) + 3(1 + t)^3 + (1 + t)(2 + t)\end{aligned}$$

Donc pour tout $t \in \mathbb{R}$, on a :

$$\varphi'_{a,e}(t) = 2(1 + t)(2 + t) + ((1 + t)^2 + 9(1 + t)^2 + 2t + 3)$$

de sorte que :

$$\varphi'_{a,e}(0) = 4 + 1 + 9 + 3 = 17$$

donc f est dérivable au point a suivant e et on a :

$$D_e f(a) = 17.$$

2. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3; (x, y) \mapsto (x, x+y, xy)$. Pour $a = (1, 1)$ et $e = (2, 3)$, on a $\varphi_{a,e}(t) = f(1+2t, 1+3t) = (1 + 2t, 2 + 5t, 6t^2 + 5t + 1)$, donc $\varphi'_{a,e}(t) = (2, 5, 12t + 5)$ et $D_e f(a) = (2, 5, 5)$.

3. $f : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}; M \mapsto \text{tr}(A)$; $A = E = I_n$; la dérivée de f en A suivant E est $\varphi'_{A,E}(0)$ avec : $\varphi_{A,E}(t) = f(A + tE) = \text{tr}((1+t)I_n) = n(t+1)$, donc : $\varphi'_{A,E}(t) = n$ et $D_E f(A) = n$.

13.1.1.2 Dérivées partielles

Soit E un $\mathbb{K}$ -espace vectoriel normé de dimension p , ($p \in \mathbb{N}^*$) et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E si $x \in E$, on note $x_1, \dots, x_p$ les coordonnées de x relativement à $\mathcal{B}$. Soit $f : U \subset E \rightarrow F$ une application et $a \in U$. Si pour $i \in \llbracket 1, p \rrbracket$, la dérivée $D_{e_i} f(a)$ de f au point a suivant e_i existe on la note $\frac{\partial f}{\partial x_i}(a)$ et on l'appelle i ème dérivée partielle de f au point a .

Remarque. Si $E = \mathbb{R}^p$ muni de sa base canonique et $f : U \subset \mathbb{R}^p \rightarrow F$ une application. La dérivée partielle $\frac{\partial f}{\partial x_i}(x)$ s'obtient en dérivant l'expression $f(x_1, \dots, x_p)$ par rapport à x_i . Par exemple $f(x, y) = (x^2 - y^2) + 2xyi$ considérée comme application de $\mathbb{R}^2$ vers $\mathbb{C}$ admet toutes les dérivées partielles en tout point. On a pour $X = (x, y) \in \mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(X) = 2(x + yi) \quad \text{et} \quad \frac{\partial f}{\partial y}(X) = 2i(x + yi)$$

13.1.1.3 Dérivée suivant un vecteur et composantes

On suppose que $\dim(F) = n$ et que $\mathcal{C} = (v_1, \dots, v_n)$ est une base de F . Soit $f : U \subset E \rightarrow F$ une applications de composantes dans $\mathcal{C}$, les applications $f_1, \dots, f_n$, donc $f = \sum_{i=1}^n f_i v_i$. En partant d'une proposition sur les fonctions vectorielles on a la :

Proposition 13.1.1. Soit $e \in E$ alors les assertions suivantes sont équivalentes :

- (1) f admet une dérivée au point a suivant e .
- (2) Pour tout $i \in \llbracket 1, n \rrbracket$, l'application f_i admet une dérivée au point a suivant e .

Si c'est le cas, on a :

$$D_e f(a) = \sum_{i=1}^n D_e f_i(a) v_i$$

Remarque. En particulier si $\dim(E) = p$ et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et on note $x_1, \dots, x_p$ les coordonnées de $x \in E$, avec les notations ci-dessus à savoir $\mathcal{C} = (v_1, \dots, v_n)$ est une base de F et $f_1, \dots, f_n$ les composantes de f alors les assertions suivantes sont équivalentes :

- (1) Les dérivées partielles $\frac{\partial f}{\partial x_j}(a)$ pour $j \in \llbracket 1, p \rrbracket$ existent.
- (2) Pour tout $i \in \llbracket 1, n \rrbracket$, Les dérivées partielles $\frac{\partial f_i}{\partial x_j}(a)$ pour $j \in \llbracket 1, p \rrbracket$ existent.

auquel cas on a

$$\frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_j}(a) v_i$$

13.1.2 Application différentiable

Pour le moment les études concernant la dérivabilité d'une application en un point sont faites sur des applications $f : I \rightarrow E$ où I est un intervalle de $\mathbb{R}$ et E un espace vectoriel normé de dimension finie. Nous allons généraliser cette notion et on parlera d'application différentiable; les applications concernées étant définies d'une partie ouverte U d'un espace vectoriel normé E vers un autre espace vectoriel normé F .

13.1.2.1 Différentiabilité en un point

Si $f : I \rightarrow \mathbb{R}$ est une application d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$ et a un point de I , on sait que si f est dérivable au point a alors pour h réel voisin de 0, on a :

$$f(a + h) = f(a) + L(h) + h\varepsilon(h)$$

où ε est une application de $V \setminus \{0\}$ vers $\mathbb{R}$ où V est un voisinage de 0 tel que

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \varepsilon(h) = 0$$

On s'inspire de cette idée pour donner la définition générale suivante :

Definition 13.1.1

Soit $f : U \subset E \rightarrow F$ une application et $a \in U$. On dit que f est différentiable au point a s'il existe une application linéaire $L \in \mathcal{L}(E, F)$, un voisinage V de 0 dans E et une application $\varepsilon : V \setminus \{0\} \rightarrow F$ tel que :

$$\begin{cases} \forall h \in V, a + h \in U \quad \text{et} \quad f(a + h) = f(a) + L(h) + \|h\|\varepsilon(h) \\ \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \varepsilon(h) = 0. \end{cases}$$

Remarque. Si U est un ouvert alors pour tout $a \in V$ l'ensemble

$$V_a = \{h \in E / a + h \in U\}$$

est un voisinage de a . On peut dire que f est différentiable au point a si et seulement si il existe $L \in \mathcal{L}(E, F)$ tel que l'application définie sur $V_a \setminus \{h\}$:

$$\forall h \in V_a, \quad \varepsilon(h) = \frac{1}{\|h\|} (f(a + h) - f(a) - L(h))$$

réalise :

$$\lim_{h \rightarrow 0} \varepsilon(h) = 0$$

Proposition 13.1.2. Si f est différentiable au point a , l'application L ci-dessus est unique.

Preuve. Si L, L' répondent à la définition ci-dessus, soit x un vecteur non nul de E alors il existe un voisinage $I_\alpha =]-\alpha, \alpha[$ de 0 dans $\mathbb{R}$ tel que $te \in V$ pour tout $t \in I_\alpha$. On a alors $f(a + te) = f(a) + tL(e) + \|te\|\varepsilon(te) = f(a) + tL'(e) + \|te\|\varepsilon'(te)$ de sorte que $|t|(L(e) - L'(e)) = \|te\|\varphi(t)$ avec $\lim_{t \rightarrow 0} \varphi(t) = 0$, donc $L(e) = L'(e)$. Par ailleurs $L(0) = L'(0)$ donc $L = L'$. ■

Definition 13.1.2

Si f est différentiable au point a , l'application linéaire L est notée $df(a)$ et appelée différentielle de f au point a .

Remarques. On fait les remarques suivantes :

1. Si f est différentiable au point a , alors pour h voisin de 0, on a :

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

2. f est différentiable au point a si et seulement si : il existe $L \in \mathcal{L}(E, F)$ tel que :

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a + h) - f(a) - L(h)}{\|h\|} = 0$$

auquel cas, $L = df(a)$.

Definition 13.1.3

Soit $f : U \subset E \rightarrow F$ une application. On dit que f est différentiable sur U si f est différentiable en tout point de U .

Proposition 13.1.3. Si f est différentiable au point a alors f est continue au point a

Preuve. Comme f est différentiable au point a , on a pour h voisin de 0 :

$$f(a + h) - f(a) = df(a)(h) + \|h\|\varepsilon(h)$$

avec $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \varepsilon(h) = 0$, donc : $\lim_{h \rightarrow 0} f(a+h) = f(a)$ puisque $\lim_{h \rightarrow 0} df(a)(h) = 0$ par continuité de l'application linéaire $df(a)$ (dimension finie). ■

13.1.2.2 Exemples d'applications différentiables

Proposition 13.1.4. E et F sont des espaces vectoriels normés de dimensions finies et $f \in \mathcal{L}(E, F)$ une application linéaire de E vers F . Pour tout ouvert U non vide de E , la restriction de f à U est différentiable sur U et on a (en notant f cette restriction) :

$$\forall a \in U, df(a) = f$$

Preuve. Soit $a \in U$ et h voisin de 0. Par linéarité, on a

$$f(a+h) = f(a) + f(h) + \|h\|\varepsilon(h)$$

avec $\varepsilon(h) = 0$ pour tout $h \in E$, donc en appliquant la définition, f est différentiable au point a et $df(a) = f$. ■

Proposition 13.1.5. E, F, G sont des espaces vectoriels normés de dimensions finies et $f : E \times F \rightarrow G$ une application bilinéaire. Pour tout ouvert non vide U de $E \times F$, la restriction de f à U est différentiable et pour tout $(a, b) \in U$ et tout $(h, k) \in E \times F$, on a :

$$df_{(a,b)}(h, k) = f(h, b) + f(a, k)$$

Preuve. Soit $(a, b) \in U$ et $(h, k) \in E \times F$, alors : en posant $A = (a, b)$ et $H = (h, k)$, on a :

$$\begin{aligned} f(A+H) &= f((a, b) + (h, k)) \\ &= f(a+h, b+k) \\ &= f(a, b) + f(a, k) + f(h, b) + f(h, k) \end{aligned}$$

On adopte par exemple la norme $\|H\| = \sup(\|h\|, \|k\|)$, alors par continuité de l'application bilinéaire f il existe une constante positive M tel que pour tout $(h, k) \in E \times F$, on a : $\|f(h, k)\| \leq M\|h\|\|k\| \leq M\|(h, k)\|^2$, de sorte que si on pose $L(H) = L(h, k) = f(a, k) + f(h, b)$ pour tout $H \in E \times F$, on a :

$$f(A+H) = f(A) + L(H) + \|H\|\varepsilon(H)$$

avec $\varepsilon(H) = \frac{1}{\|H\|} f(H) = \frac{1}{\|H\|} f(h, k)$ si $H \neq 0$ et $\varepsilon(0) = 0$ de sorte que d'après l'inégalité $\|f(h, k)\| \leq M\|(h, k)\|^2$, on a $\lim_{H \rightarrow 0} \varepsilon(H) = 0$, ce qui justifie la différentiabilité de f au point (a, b) et que $df(a, b)(h, k) = L(H) = f(a, k) + f(h, b)$.

Remarque. On a une généralisation : Si $F, E_1, \dots, E_m$ sont des espaces vectoriels normés de dimensions finies et $f : E = \prod_{k=1}^m E_k \rightarrow F$ une application m -linéaire alors la restriction de f à tout ouvert U de E est différentiable sur U et pour tout $A = (a_1, \dots, a_m) \in U$ et tout $H = (h_1, \dots, h_m) \in E$, on a :

$$df_A(H) = f(h_1, a_2, \dots, a_m) + \dots + f(a_1, \dots, a_{m-1}, h_m).$$

Proposition 13.1.6. Toute application constante $f : U \subset E \rightarrow F$ est différentiable sur U et $df(a) = \theta$ pour tout $a \in U$ où θ est l'application nulle de E vers F .

Preuve. Pour tout $(a, h) \in U \times E$, on a $f(a+h) = f(a) + \theta(h) + \|h\|\theta(h)$, ce qui prouve le résultat. ■

13.1.2.3 Dérivabilité et différentiabilité

Proposition 13.1.7. Soit I un intervalle ouvert non vide de $\mathbb{R}$ et $a \in I$. Soit $f : I \rightarrow F$ une application. Les assertions suivantes sont équivalentes :

- (1) f est dérivable au point a .
- (1) f est différentiable au point a .

Si c'est le cas, on a :

$$\forall h \in \mathbb{R}, df(a)(h) = hf'(a) \quad \text{et} \quad f'(a) = df(a)(1) = D_1 f(a)$$

Preuve. Supposons que f est dérivable au point a , alors pour h réel voisin de 0, on a : $f(a+h) = f(a) + hf'(a) + o(h)$, ce qui exprime que f est différentiable au point a et que pour tout h réel $df(a)(h) = hf'(a)$. Réciproquement, si f est différentiable au point a alors $df(a)$ est une application linéaire de $\mathbb{R}$ vers F et précisément, on a pour tout h réel : $df(a)(h) = hdf(a)(1)$. On a alors pour h voisin de 0 : $f(a+h) = f(a) + hdf(a)(1) + o(h)$, ce qui donne : f est dérivable au point a et $f'(a) = df(a)(1)$. Par définition de la dérivée suivant le vecteur 1, on a aussi $f'(a) = D_1 f(a)$ ■

13.1.3 Différentiabilité et dérivées partielles

Soit E un $\mathbb{K}$ – espace vectoriel normé de dimension p avec $p \in \mathbb{N}^*$ et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E . Soit F un $\mathbb{K}$ – espace vectoriel normé de dimension finie, U un ouvert non vide de E et $f : U \rightarrow F$ une application. On va voir que si f est différentiable en un point a de U alors les dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ existent et que $\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a) = df(a)(e_i)$ pour tout $i \in \llbracket 1, p \rrbracket$. On verra ensuite que la réciproque n'est pas vraie, mais on a un résultat si on ajoute des hypothèses supplémentaires.

Proposition 13.1.8. Si f est différentiable au point a alors pour tout vecteur e de E , la dérivée de f au point a suivant e existe et on a :

$$D_e f(a) = df(a)(e)$$

Preuve. En effet, supposons que f est différentiable au point a , alors $f(a+h) = f(a) + df(a)(h) + h\varepsilon(h)$ avec $\varepsilon(h) \rightarrow 0$ quand h tends vers 0. En particulier, pour t voisin de 0 dans $\mathbb{R}$, on a $f(a+te) = f(a) + tdf(a)(e) + |t||e|\varepsilon(te)$ donc :

$$\frac{1}{t}(f(a+te) - f(a)) = df(a)(e) + \|e\|\alpha(t)\varepsilon(te)$$

avec $\alpha(t) = \pm 1$, tends vers $df(a)(e)$ quand t tends vers 0. Il en découle que $D_e f(a) = df_a(e)$ ■

Proposition 13.1.9. Soit $\mathcal{B} = (e_1, \dots, e_p)$ une base de E , $f : U \subset E \rightarrow F$ une application et $a \in U$. Si f est différentiable au point a alors les dérivées partielles $\frac{\partial f}{\partial x_j}(a)$ pour $j \in \llbracket 1, p \rrbracket$ existent et, pour tout

$h = \sum_{j=1}^p h_j e_j \in E$, on a :

$$df(a)(h) = \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a)$$

Preuve. Conséquence immédiate de la proposition 13.1.8 appliquée aux vecteurs e_j de la base $\mathcal{B}$. Précisément, si $h = \sum_{j=1}^p h_j e_j$, alors : $df(a)(h) = \sum_{j=1}^p h_j df(a)(e_j) = \sum_{j=1}^p h_j D_{e_j} f(a) = \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a)$ ■

La réciproque de ce résultat n'est pas vraie. On peut même prouver qu'une application f peut admettre des dérivées suivant tout vecteur e au point a sans que f soit différentiable au point a . Cependant, on a le résultat suivant :

Théorème 13.1.1. Si les dérivées partielles de f au point a à savoir $\frac{\partial f}{\partial x_i}(a); 1 \leq i \leq p$ existent et si de plus,

pour tout $h = \sum_{i=1}^p h_i e_i$ voisin de 0, on a :

$$f(a+h) - f(a) - \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a) = o(\|h\|)$$

alors f est différentiable au point a et :

$$\forall h \in E, \quad df(a)(h) = \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a)$$

Preuve. Conséquence immédiate de la définition d'une application différentiable en un point a et vu que l'application :

$$E \rightarrow F; h = \sum_{i=1}^p h_i e_i \mapsto \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a)$$

est une application linéaire de E vers F . ■

Exemple. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}; (x, y) \mapsto \begin{cases} \frac{x^3 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$. Montrons que f est différentiable au

point $(0, 0)$ et que $df(0, 0) = \theta$ (θ est l'application nulle de $\mathbb{R}^2$ vers $\mathbb{R}$).

On a :

$$\lim_{\substack{x \rightarrow 0 \\ y \neq 0}} \frac{f(x, 0) - f(0, 0)}{x} = 0$$

, donc :

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

de même

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

On a f est différentiable au point $(0, 0)$ si et seulement si, pour (x, y) voisin de $(0, 0)$, on a :

$$f(x, y) - f(0, 0) - x \frac{\partial f}{\partial x}(0, 0) - y \frac{\partial f}{\partial y}(0, 0) = o(\sqrt{x^2 + y^2}),$$

c'est-à-dire :

$$f(x, y) = o(\sqrt{x^2 + y^2}).$$

En notant $X = (x, y)$ et $\|X\| = \sqrt{x^2 + y^2}$, et compte tenu du fait que : $\begin{cases} |x| \leq \|X\| \\ |y| \leq \|X\| \end{cases}$, on a :

$$\begin{aligned} \left| \frac{f(x, y)}{\sqrt{x^2 + y^2}} \right| &= \frac{|x^3 y|}{(x^2 + y^2)^{\frac{3}{2}}} \\ &= \frac{|x|^3 |y|}{\|X\|^3} \\ &\leq \frac{\|X\|^4}{\|X\|^3} \\ &= \|X\| \xrightarrow{X \rightarrow (0, 0)} 0 \end{aligned}$$

Il en découle que $f(X) = o(\|X\|)$ quand $X \rightarrow (0, 0)$, donc f est différentiable au point $(0, 0)$ et $df(0, 0) = \theta$.

Règle générale : Soit $f : U \subset \mathbb{R}^p \rightarrow F$ une application et $a \in U$. On note $\mathcal{B} = (e_1, \dots, e_p)$ la base canonique de $\mathbb{R}^p$.

- Si l'une des dérivées partielles de f au point a n'existe pas alors f n'est pas différentiable au point a .

- Si toutes les dérivées partielles $\frac{\partial f}{\partial x_i}(a); 1 \leq i \leq p$ existent, cela ne suffit pas encore pour dire que f est différentiable au point a , mais si on prouve que :

$$f(a+h) - f(a) - \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a) = o(\|h\|)$$

pour h voisin de 0 alors f est différentiable au point a et sa différentielle est définie par :

$$\forall h \in E, h = \sum_{i=1}^p h_i e_i \Rightarrow df_a(h) = \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(a)$$

13.1.4 Matrice jacobienne, jacobien

E et F sont deux espaces vectoriels normés de dimensions respectives p et n et $\mathcal{B} = (e_1, \dots, e_p)$ et $\mathcal{C} = (v_1, \dots, v_n)$ des bases respectives de E et F . Soit U un ouvert de E , $f : U \rightarrow F$ une application et $f_1, \dots, f_n$ les applications coordonnées de f relativement à $\mathcal{C}$, donc

$$\forall x \in E, f(x) = \sum_{i=1}^n f_i(x) v_i$$

qu'on abrège par :

$$f = \sum_{i=1}^n f_i v_i$$

Si $x \in E$ on note $x_1, \dots, x_p$ les coordonnées de x par rapport à $\mathcal{B}$ donc

$$x = \sum_{j=1}^p x_j e_j$$

On suppose que les dérivées partielles $\frac{\partial f}{\partial x_j}(a), 1 \leq j \leq p$ existent (ce qui revient à dire que les dérivées partielles des composantes de f existent). On dispose de la matrice :

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

Definition 13.1.4

Avec les notations et les conditions ci-dessus la matrice :

1. $J_f(a)$ s'appelle la matrice jacobienne de f au point a relativement aux bases $\mathcal{B}$ et $\mathcal{C}$.
2. Si $E = F$ et $\mathcal{B} = \mathcal{C}$ le déterminant : $\det(J_f(a))$ est appelé le jacobien au point a de f relativement à la base $\mathcal{B}$.

Exemple. Soit $U = \mathbb{R} \times \mathbb{R}^*$ (c'est un ouvert de $\mathbb{R}^2$ et

$$f : U \rightarrow \mathbb{R}^3; (x, y) \mapsto \left(xy, \frac{x}{y}, x + y^2 \right)$$

Alors pour tout $(x, y) \in U$, on a :

$$J_f((x, y)) = \begin{pmatrix} y & x \\ \frac{1}{y} & -\frac{x}{y^2} \\ 1 & 2y \end{pmatrix}$$

Lorsque la fonction $f : U \rightarrow F$ est différentiable en un point a , la matrice jacobienne $J_f(a)$ existe et elle représente la différentielle $df(a)$ dans le couple de bases $(\mathcal{B}, \mathcal{C})$ comme le précise le théorème suivant :

Théorème 13.1.2. Avec les notations ci-dessus, si $a \in U$ et f est différentiable au point a alors $J_f(a)$ est la matrice de l'application linéaire $df(a)$ relativement aux bases $\mathcal{B}$ et $\mathcal{C}$. C'est-à-dire :

$$J_f(a) = \text{mat}_{\mathcal{B}, \mathcal{C}}(df(a))$$

Preuve. Pour tout $j \in \llbracket 1, p \rrbracket$, on a :

$$df(a)(e_j) = D_{e_j} f(a) = \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_j}(a) v_i$$

ce qui prouve le théorème. ■

Exemple. Revenons à l'exemple ci-dessus de l'application $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par

$$f(x, y) = \left(xy, \frac{x}{y}, x + y^2 \right)$$

On verra plus tard que cette application est différentiable sur U . Si par exemple $a = (1, 1)$, on a la matrice jacobienne en a :

$$J_f(a) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix}$$

donc si $h = (h_1, h_2) \in \mathbb{R}^2$ alors :

$$df(a)(h) = h' = (h'_1, h'_2, h'_3)$$

tel que :

$$\begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} h_1 + h_2 \\ h_1 - h_2 \\ h_1 + 2h_2 \end{pmatrix}.$$

Donc, pour tout $h = (h_1, h_2) \in \mathbb{R}^2$, on a :

$$df((1, 1))(h_1, h_2) = (h_1 + h_2, h_1 - h_2, h_1 + 2h_2)$$

13.1.5 Opérations sur les applications différentiables

Les opérations sur les fonctions différentiables vont faciliter l'identification de celles-ci et le calcul de leur différentielles. Elles concernent, la combinaison linéaire, la composition et le produit via une application bilinéaire.

13.1.5.1 Combinaison linéaire

Proposition 13.1.10. Si $f, g : U \rightarrow F$ sont différentiables au point a alors $f + \lambda g$ est différentiable au point a et $d(f + \lambda g)(a) = df(a) + \lambda dg(a)$.

Si f et g sont différentiables sur U alors $f + \lambda g$ est différentiable sur U et pour tout $x \in U$, on a : $d(f + \lambda g)(x) = df(x) + \lambda dg(x)$

Remarques. La proposition ci-dessus nous laisse faire les remarques suivantes :

1. Si on note $\mathcal{D}_a(U, F)$ l'ensemble des applications de U vers F différentiables au point a alors $\mathcal{D}_a(U, F)$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(U, F)$ des applications de U vers F .
2. $\mathcal{D}(U, F)$ l'ensemble des applications de U vers F différentiables sur U est un sous-espace vectoriel de $\mathcal{F}(U, F)$. Remarquons que l'on a :

$$\mathcal{D}(U, F) = \bigcap_{a \in U} \mathcal{D}_a(U, F)$$

13.1.5.2 Composition

Théorème 13.1.3. E, F, G sont des espaces vectoriels normés, U, V des ouverts non vides respectifs de E et F et $f : U \rightarrow F$ et $g : V \rightarrow G$ des applications tel que $f(U) \subset V$. Soit $a \in U$. Si f est différentiable au point a et g est différentiable au point $f(a)$ alors $g \circ f$ est différentiable au point a et :

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

Preuve. f est différentiable au point a , donc il existe un voisinage V_1 de 0_E dans E , et une application $\varepsilon_1 : V_1 \rightarrow F$ tel que $\lim_{h \rightarrow 0} \varepsilon_1(h) = 0$ et pour tout $h \in V_1$:

$$(1) \quad f(a+h) = f(a) + \underbrace{df(a)(h) + \|h\|\varepsilon_1(h)}_{\varphi(h)}$$

De même, g est différentiable au point $f(a)$, donc il existe un voisinage V_2 de 0_F dans F , et une application $\varepsilon_2 : V_2 \rightarrow G$ tel que $\lim_{k \rightarrow 0} \varepsilon_2(k) = 0$ et pour tout $k \in V_2$:

$$(2) \quad g(f(a)+k) = g(f(a)) + dg(f(a))(k) + \|k\|\varepsilon_2(k)$$

Posons $\varphi(h) = df(a)(h) + \|h\|\varepsilon_1(h)$, pour tout $h \in V_1$. Comme $\lim_{h \rightarrow 0} \varphi(h) = 0$, il existe un voisinage V de 0_E dans E tel que $\varphi(V) \subset V_2$, donc d'après (2), on a pour tout $h \in V$:

$$\begin{aligned} (g \circ f)(a+h) &= g(f(a) + \varphi(h)) \\ &= g(f(a)) + dg(f(a))(\varphi(h)) \\ &\quad + \|\varphi(h)\|\varepsilon_2(\varphi(h)) \end{aligned}$$

Par linéarité de $dg(f(a))$, on a :

$$\begin{aligned} dg(f(a))(\varphi(h)) &= dg(f(a))(df(a)(h)) + \\ &\quad \|h\|dg(f(a))(\varepsilon_1(h)) \end{aligned}$$

Il en découle que :

$$(g \circ f)(a+h) = (g \circ f)(a) + [dg(f(a)) \circ df(a)](h) + \psi(h)$$

avec

$$\psi(h) = \underbrace{\|h\|dg(f(a))(\varepsilon_1(h))}_{\alpha_1(h)} + \underbrace{\|\varphi(h)\|\varepsilon_2(\varphi(h))}_{\alpha_2(h)}$$

Pour terminer la preuve du théorème, il suffit de démontrer que $\psi(h) = o(\|h\|)$ et comme $\psi(h) = \alpha_1(h) + \alpha_2(h)$ avec

$$\begin{cases} \alpha_1(h) = \|h\|dg(f(a))(\varepsilon_1(h)) = o(\|h\|) \\ \alpha_2(h) = \|\varphi(h)\|\varepsilon_2(\varphi(h)) \end{cases},$$

il suffit en fait de prouver que $\alpha_2(h) = o(\|h\|)$. Pour ce faire remarquons que la continuité de l'application linéaire $df(a)$ implique l'existence d'une constante $c > 0$ tel que $\|df(a)(h)(x)\| \leq c\|x\|$, pour tout $x \in E$. Cela dit, on a alors :

$$\begin{aligned} \|\alpha_2(h)\| &= \|\varphi(h)\|\|\varepsilon_2(\varphi(h))\| \\ &= \|df(a)(h) + \|h\|\varepsilon_1(h)\|\varepsilon_2(\varphi(h))\| \\ &\leq (c + \|\varepsilon_1(h)\|)\|\varepsilon_2(\varphi(h))\|\|h\| \end{aligned}$$

ce qui fournit aisément le résultat désiré, à savoir : $\alpha_2(h) = o(\|h\|)$ quand h tends vers 0 et achève en conséquence la preuve du théorème ci-dessus. ■

On dispose du cas particulier important suivant de le théorème 13.1.3 :

Proposition 13.1.11. Soit I un intervalle ouvert non vide de $\mathbb{R}$ et U un ouvert non vide de E . Soit $\varphi : I \rightarrow E$ et $f : U \rightarrow F$ des applications. Soit $t_0 \in I$. Si φ est dérivable au point t_0 et f est différentiable au point

$f(t_0)$ alors $g = f \circ \varphi$ est dérivable au point t_0 et on a :

$$g'(t_0) = (f \circ \varphi)'(t_0) = df(\varphi(t_0))(\varphi'(t_0)).$$

Remarques. Voici quelques remarques importantes sur la proposition ci-dessus :

1. Si $E = \mathbb{R}^p$, donc $\gamma : I \rightarrow \mathbb{R}^p$ avec $\gamma(t) = (\gamma_1(t), \dots, \gamma_p(t))$ alors la formule ci-dessus s'écrit :

$$g'(t_0) = \sum_{j=1}^p \gamma'_j(t_0) \frac{\partial f}{\partial x_j}(\gamma(t_0))$$

2. Une des conséquences de ce qui précède es la dérivation en chaîne :

Soit $f : U \subset E \rightarrow F$; $g : V \subset F \rightarrow G$ tel que $f(U) \subset V$. Soit $a \in U$ tel que f est différentiable au point a et g est différentiable au point $f(a)$. On considère $\mathcal{B} = (e_1, \dots, e_p)$, $\mathcal{C} = (u_1, \dots, u_r)$ et $\mathcal{D} = (v_1, \dots, v_n)$ bases respectives de E, F et G et on note $x_1, \dots, x_p$ (resp $y_1, \dots, y_r$ (resp $z_1, \dots, z_n$)) les coordonnées dans E (resp F (resp G)) , alors, si on note $f_1, \dots, f_r$ les composantes de f relativement à la base $\mathcal{C}$, on a :

$$\frac{\partial(g \circ f)}{\partial x_j}(a) = \sum_{k=1}^r \frac{\partial f_k}{\partial x_j}(a) \frac{\partial g}{\partial y_k}(f(a))$$

Preuve. On a :

$$\begin{aligned} \frac{\partial(g \circ f)}{\partial x_j}(a) &= D_{e_j}(g \circ f)(a) \\ &= d(g \circ f)(a)(e_j) \\ &= (dg(f(a)) \circ df(a))(e_j) \\ &= dg(f(a))(df(a)(e_j)) \\ &= dg(f(a))(D_{e_j}f(a)) \\ &= dg(f(a)) \left(\frac{\partial f}{\partial x_j}(a) \right) \end{aligned}$$

Rappelons que pour tout $h \in F$ tel que :

$$h = \sum_{k=1}^r h_k u_k,$$

on a :

$$dg(f(a))(h) = \sum_{k=1}^r h_k \frac{\partial g}{\partial y_k}(f(a)).$$

En appliquant pour

$$h = \frac{\partial f}{\partial x_j}(a) = \sum_{k=1}^r \frac{\partial f_k}{\partial x_j}(a) u_k$$

il vient :

$$\frac{\partial(g \circ f)}{\partial x_j}(a) = \sum_{k=1}^r \frac{\partial f_k}{\partial x_j}(a) \frac{\partial g}{\partial y_k}(f(a))$$

■

La proposition suivante donne la matrice jacobienne d'un composé :

E, F et G sont deux espaces vectoriels normés de dimensions respectives p, r et n et $\mathcal{E} = (e_1, \dots, e_p)$, $\mathcal{F} = (f_1, \dots, f_r)$ et $\mathcal{G} = (v_1, \dots, v_n)$ sont des bases respectives de E, F et G .

Proposition 13.1.12. Les notations ci-dessus étant prises en compte, soient U un ouvert de E , V un ouvert de F , $f : U \rightarrow F$ et $g : V \rightarrow G$ des applications tel que $f(U) \subset V$. Soit $a \in U$ tel que f est différentiable au point a et g est différentiable au point $f(a)$. Alors la matrice jacobienne de $g \circ f$ est :

$$J_{g \circ f}(a) = J_g(f(a)) \times J_f(a)$$

Les matrices jacobiniennes en question étant calculées par rapport aux bases correspondantes parmi celles fixées ci-dessus.

Preuve. C'est une conséquence immédiate du fait que :

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

et la propriété concernant la matrice du composé de ceux applications linéaires. ■

13.1.5.3 Produit via une application bilinéaire de deux fonctions différentiables

Soient E, F_1, F_2 et G des espaces vectoriels normés dimensions finies, U un ouverts non vide E . On considère une application bilinéaire

$$\Phi : F_1 \times F_2 \rightarrow G; (x, y) \mapsto \Phi(x, y)$$

de $F_1 \times F_2$ vers G . Si $f : U \rightarrow F_1$ et $g : U \rightarrow F_2$ sont des applications, on note $\Phi(f, g)$ l'application de U vers G définie par :

$$\forall x \in U, \quad \Phi(f, g)(x) = \Phi(f(x), g(x)).$$

Avec les notations ci-dessus, on a la :

Proposition 13.1.13. Si f et g sont différentiables au point a alors $\Phi(f, g)$ est différentiable au point a et :

$$d(\Phi(f, g))(a) = \Phi(f_a, dg(a)) + \Phi(df(a), g_a).$$

avec f_a et g_a les applications constantes de U vers F de valeurs respectives $f(a)$ et $g(a)$.

Exemples. Voici des exemples importants de l'application des propositions ci-dessus :

1. Si F est une algèbre normée de dimension finie et $f, g : U \subset E \rightarrow F$ alors si f et g sont différentiables en un point a de U , la produit fg est différentiable au point a et

$$d(fg)(a) = f(a)dg(a) + df(a)g(a).$$

2. Si $f : U \rightarrow \mathbb{R}$ est une application et $a \in U$ tel que $f(a) \neq 0$ et f différentiable au point a alors il existe un ouvert V contenu dans U tel que $f(x) \neq 0$ pour tout $x \in V$, donc $\frac{1}{f}$ est bien définie sur V . L'application $\frac{1}{f}$ est différentiable au point a et on a

$$d\left(\frac{1}{f}\right)(a) = -\frac{1}{(f(a))^2}df(a).$$

3. On en déduit que si $f, g : U \subset E \rightarrow \mathbb{R}$ différentiables au point a et $g(a) \neq 0$ alors $\frac{f}{g}$ est bien définie sur un ouvert V tel que $a \in V$ et elle est différentiable au point a et

$$d\left(\frac{f}{g}\right)(a) = \frac{1}{(g(a))^2}(df(a)g(a) - f(a)dg(a)).$$

13.1.5.4 Composé avec une application linéaire

Proposition 13.1.14. E, F, G sont trois espaces vectoriels normés, soit $f : U \subset E \rightarrow F$ une application et $\Phi : F \rightarrow G$ une application linéaire et $a \in U$. Si f est différentiable au point a alors $\Phi \circ f$ est différentiable au point a et :

$$d(\Phi \circ f)(a) = \Phi \circ df(a)$$

Preuve. Si f est différentiable au point a , comme Φ est linéaire elle est différentiable sur F donc au point a , donc, d'après le théorème 13.1.3, $\Phi \circ f$ est différentiable au point a et $d(\Phi \circ f)(a) = d\Phi(f(a)) \circ df(a)$ et comme $d\Phi(f(a)) = \Phi$, on a $d(\Phi \circ f)(a) = \Phi \circ df(a)$. ■

Exemple. Soit $n \in \mathbb{N}^*$ et

$$f : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathcal{M}_n(\mathbb{R}); X \mapsto X^2$$

et

$$g : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}; X \mapsto \text{tr}(X^2)$$

1. Montrer que f est différentiable sur $\mathcal{M}_n(\mathbb{R})$ et préciser sa différentielle en tout point A de $\mathcal{M}_n(\mathbb{R})$.
2. En déduire que g est différentiable sur $\mathcal{M}_n(\mathbb{R})$ et préciser sa différentielle en tout point.
3. Calculer $\frac{\partial g}{\partial x_{i,j}}(A)$, pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$.

Réponse :

1. Soit $A, H \in \mathcal{M}_n(\mathbb{R})$; alors :

$$f(A + H) = (A + H)^2 = A^2 + AH + HA + H^2 = f(A) + L(H) + \varphi(H)$$

avec $L(H) = AH + HA$, donc L est linéaire et $\varphi(H) = H^2$, donc, en choisissant une norme matricielle, on a $\|\varphi(H)\| \leq \|H\|^2$, par suite $\varphi(H) = o(\|H\|)$ quand H tends vers 0. Il en résulte que f est différentiable en tout point A , on a :

$$\forall H \in \mathcal{M}_n(\mathbb{R}), df(A)(H) = AH + HA$$

2. On remarque que $g = \text{tr} \circ f$ et comme tr est linéaire, l'application du théorème ci-dessus donne g est différentiable en tout point A de $\mathcal{M}_n(\mathbb{R})$ et $\forall H \in \mathcal{M}_n(\mathbb{R}), dg(A)(H) = \text{tr}(df(A)(H))$. Ainsi :

$$\forall (A, H) \in \mathcal{M}_n(\mathbb{R})^2, dg(A)(H) = 2 \text{tr}(AH)$$

3. On a

$$\frac{\partial g}{\partial x_{i,j}}(A) = D_{E_{i,j}}g(A) = dg(A)(E_{i,j}) = 2 \text{tr}(AE_{i,j}) = 2a_{ji}$$

13.1.5.5 Différentiabilité et composantes

Proposition 13.1.15. Soit $f : U \subset E \rightarrow F$ une application et $a \in E$. On suppose que $\dim(F) = n$ et que $\mathcal{V} = (V_1, \dots, V_n)$ est une base de F et que $f = \sum_{k=1}^n f_k V_k$, c'est-à-dire que $f_1, \dots, f_n$ sont les composantes de f relativement à la base $\mathcal{V}$. Alors f est différentiable au point a si et seulement si les applications f_k pour $k \in \llbracket 1, n \rrbracket$ sont différentiables au point a , auquel cas on a :

$$df(a)(h) = \sum_{i=1}^n df_i(a)(h) V_i$$

Preuve. Si f est différentiable au point a , remarquons que pour tout $i \in \llbracket 1, n \rrbracket$, on a :

$$f_i = \pi_i \circ f$$

où $\pi_i : F \rightarrow \mathbb{R}, x \mapsto \pi_i(x) = x_i$ où $x = \sum_{k=1}^n x_k V_k$, donc π_i étant linéaire, elle est différentiable et $d\pi_i = \pi_i$ de sorte que f_i est différentiable au point a et $df_i(a) = \pi_i \circ df(a)$.

Réciproquement, si pour tout $k \in \llbracket 1, n \rrbracket$, l'application f_k est différentiable au point a , il existe des applications $\varepsilon_k, k \in \llbracket 1, n \rrbracket$ définies sur $W \setminus \{0\}$ où W est un voisinage de 0 tel que :

$$\forall i \in \llbracket 1, n \rrbracket, \forall h \in W, f_i(a + h) = f_i(a) + df_i(a)(h) + \|h\| \varepsilon_i(h)$$

avec $\lim_{h \rightarrow 0} \varepsilon_i(h) = 0$, pour tout $i \in \llbracket 1, n \rrbracket$. En sommant, on a :

$$\sum_{i=1}^n f_i(a + h) V_i = \sum_{i=1}^n f_i(a) V_i + \sum_{i=1}^n df_i(a)(h) V_i + \|h\| \sum_{i=1}^n \varepsilon_i(h) V_i$$

ce qui donne :

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h)$$

avec

$$\forall h \in E, L(h) = \sum_{i=1}^n df_i(a)(h)V_i$$

donc L est linéaire de E vers F , et :

$$\forall h \in V, \varepsilon(h) = \sum_{i=1}^n \varepsilon_i(h)$$

de sorte que $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \varepsilon(h) = 0$ et par unicité de la différentielle, f est différentiable au point a et $df(a) = L$. ■

13.2 Fonction de classe $C^k, k \in \mathbb{N}^* \cup \{\infty\}$

13.2.1 Fonction de classe C^1

13.2.1.1 Définitions, caractérisation

Soient E et F deux espaces vectoriels normés de dimensions finies, U un ouvert de E et $f : U \rightarrow F$ une application. On dit que f est différentiable sur U si f est différentiable en tout point de U . Dans ce cas, on dispose d'une application notée df de U vers $\mathcal{L}(E, F)$ qui associe à tout $x \in U$ l'application linéaire $df(x)$.

Definition 13.2.1

Soit $f : U \subset E \rightarrow F$ une application. f est de classe C^1 sur U si f est différentiable sur U et l'application df est continue sur U .

Exemples. Voici des exemples importants :

1. Si $f \in \mathcal{L}(E, F)$ alors la restriction de f à tout ouvert U de E est de classe C^1 sur U
2. Toute application constante sur U est de classe C^1 sur U .

Proposition 13.2.1. 1. Si $f, g : U \rightarrow F$ sont deux applications de classe C^1 sur U et $\lambda \in \mathbb{R}$, alors $f + \lambda g$ est de classe C^1 sur U .
 2. Si E, F, G sont des espaces vectoriels normés de dimensions finies, U, V des ouverts respectifs de E et F et si $f : U \rightarrow F$ et $g : V \rightarrow G$ sont des applications de classe C^1 respectivement sur U et V et $f(U) \subset V$ alors $g \circ f$ est de classe C^1 sur U .

Preuve. 1. Si f et g sont de classe C^1 elles sont différentiables, donc $f + \lambda g$ est différentiable et

$$\forall x \in U, d(f + \lambda g)(x) = df(x) + \lambda dg(x)$$

et comme df et dg sont continues sur U , l'application $d(f + \lambda g) = df + \lambda dg$ est continue sur U . Donc $f + \lambda g$ est de classe C^1 sur U .

2. Comme f et g sont différentiables sur U et V respectivement et $f(U) \subset V$, on a $g \circ f$ est différentiable sur U et :

$$\forall x \in U, d(g \circ f)(x) = dg(f(x)) \circ df(x) = \Phi(dg(f(x)), df(x))$$

avec

$$\Phi : \mathcal{L}(F, G) \times \mathcal{L}(E, F); (u, v) \mapsto u \circ v$$

qui est une application bilinéaire, donc par un théorème qui concerne la continuité du produit de deux applications continues via une application bilinéaire, l'application

$$d(g \circ f) = \Phi(dg \circ f, df)$$

est continue sur U . Rappelons que si $\psi_1 : U \rightarrow F_1, \psi_2 : U \rightarrow F_2$ sont deux applications et $\Psi : F_1 \times F_2 \rightarrow G$ une application bilinéaire alors $\Psi(\psi_1, \psi_2)$ est l'application qui associé à tout vecteur x de U le vecteur de G , $y = \Psi(\psi_1(x), \psi_2(x))$. ■

Théorème 13.2.1. E et F sont deux espaces vectoriels normés de dimensions respectives p et n avec $p, n \in \mathbb{N}^*$ et $\mathcal{V} = (V_1, \dots, V_n)$ est une base de F . Soit U un ouvert de E et $f : U \rightarrow F$ une application de composantes $f_1, \dots, f_n$ relativement à la base $\mathcal{V}$ de F . Alors f est de classe C^1 sur U si et seulement si f_i est de classe C^1 sur U , pour tout $i \in \llbracket 1, n \rrbracket$.

Preuve.

• Si f est de classe C^1 sur U alors en particulier f est différentiable sur U et on a déjà vu que :

$$\forall x \in U, df(x) = \sum_{i=1}^n df(x)V_i$$

Pour tout $j \in \llbracket 1, n \rrbracket$, on note L_j l'application :

$$L_j : \mathcal{L}(E, F) \rightarrow \mathcal{L}(E, \mathbb{R}); g \mapsto \sum_{i=1}^n g_i V_i \mapsto L_j(g) = g_j$$

Alors g est linéaire donc continue (dimension finie). Par ailleurs, il est clair que :

$$\forall i \in \llbracket 1, n \rrbracket, \forall x \in U, df_i(x) = (L_i \circ df)(x)$$

Par continuité de df et L_i , on a donc df_i est continue sur U donc f_i est de classe C^1 sur U pour tout $i \in \llbracket 1, n \rrbracket$.

• Réciproquement supposons que pour tout $i \in \llbracket 1, n \rrbracket$, f_i est de classe C^1 . On a f est différentiable sur U et :

$$\forall x \in U, df(x) = \sum_{i=1}^n df(x)V_i$$

donc

$$\forall x \in U, df(x) = L(df_1(x), \dots, df_n(x))$$

où L est l'application :

$$L : (\mathcal{L}(E, \mathbb{R}))^n \rightarrow \mathcal{L}(E, F); (g_1, \dots, g_n) \mapsto \sum_{i=1}^n g_i V_i$$

qui est manifestement continue car linéaire. Par suite df est continue sur U donc f est de classe C^1 sur U . ■

Proposition 13.2.2. Soit E un espace vectoriel normé de dimension p et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E . Une application $f : U \rightarrow F$ est de classe C^1 si et seulement si les dérivées partielles $\frac{\partial f}{\partial x_j}; j \in \llbracket 1, p \rrbracket$ par rapport à la base $\mathcal{B}$ existent et sont continues sur U .

Preuve. Au cours de toute la démonstration E est rapporté à une base $\mathcal{E} = (e_1, \dots, e_p)$ et on adoptera la norme définie sur E par :

$$\forall x = \sum_{j=1}^p x_j e_j, \quad \|x\| = \sup_{1 \leq j \leq p} |x_j|$$

• Supposons que f est de classe C^1 sur U , donc f est différentiable et par suite ses dérivées partielles $\frac{\partial f}{\partial x_j}$ existent. Pour tout $j \in \llbracket 1, p \rrbracket$, on a :

$$\forall x \in U, \quad \frac{\partial f}{\partial x_j}(x) = df(x)(e_j) = (L_j \circ df)(x)$$

où

$$L_j : \mathcal{L}(E, F) \rightarrow F; u \mapsto L_j(u) = u(e_j)$$

de sorte que L_j est continue car linéaire en dimension finie. Comme

$$\frac{\partial f}{\partial x_j} = L_j \circ df$$

on a en vertu de la continuité de df supposée là haut que $\frac{\partial f}{\partial x_j}$ est continue sur U .

• Réciproquement, supposons que les dérivées partielles $\frac{\partial f}{\partial x_j}$ existent et sont continues, pour tout $j \in \llbracket 1, p \rrbracket$. Soit $a \in U$. Soit $\varepsilon > 0$. Par continuité des dérivées partielles au point a , il existe $\eta > 0$ tel que :

$$(\star) \quad \forall h \in E, \|h\| < \eta \Rightarrow \left\{ \begin{array}{l} a + h \in U \\ \text{et} \\ \left| \frac{\partial f}{\partial x_j}(a + h) - \frac{\partial f}{\partial x_j}(a) \right| < \frac{\varepsilon}{p} \end{array} \right. +$$

Soit alors $h \in E$ tel que $\|h\| < \eta$, alors on peut écrire :

$$\begin{aligned} f(a + h) - f(a) &= f(a_1 + h_1, \dots, a_p + h_p) - f(a_1, \dots, a_p) \\ &= f(a_1 + h_1, a_2 + h_2, \dots, a_p + h_p) - f(a_1, a_2 + h_2, \dots, a_p + h_p) \\ &\quad + f(a_1, a_2 + h_2, \dots, a_p + h_p) - f(a_1, a_2, \dots, a_p + h_p) \\ &\quad + \dots \\ &\quad + f(a_1, a_2, \dots, a_{p-1}, a_p + h_p) - f(a_1, a_2, \dots, a_p) \end{aligned}$$

Sans nuire à la généralité et compte tenu du théorème 13.2.1, on peut supposer que f est à valeurs dans $\mathbb{R}$. On peut donc appliquer l'égalité des accroissements finis aux fonctions partielles, donc il existe :

1. c_1 entre a_1 et $a_1 + h_1$.
2. c_2 entre $a_2 + h_2$ et a_2
3.
4. c_p entre a_p et $a_p + h_p$

tel que :

$$\begin{aligned} f(a + h) - f(a) &= h_1 \frac{\partial f}{\partial x_1}(c_1, a_2 + h_2, \dots, a_p + h_p) \\ &\quad + h_2 \frac{\partial f}{\partial x_2}(a_1, c_2, \dots, a_p + h_p) \\ &\quad + \dots \\ &\quad + h_p \frac{\partial f}{\partial x_p}(a_1, a_2, \dots, a_{p-1}, c_p) \end{aligned}$$

Si on note :

$$\Delta(h) = f(a + h) - f(a) - \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a)$$

ce qui précède permet de dire que pour tout $h \in E$ tel que $\|h\| < \eta$, on a :

$$\Delta(h) = \sum_{j=1}^p h_j \Phi_j(h)$$

Avec :

$$\Phi_j(h) = \frac{\partial f}{\partial x_j}(a + H_j) - \frac{\partial f}{\partial x_j}(a)$$

avec :

$$H_j = (0, \dots, 0, c_j - a_j, h_{j+1}, \dots, h_p)$$

de sorte que, pour tout $j \in \llbracket 1, p \rrbracket$, on a $\|H_j\| < \eta$ car la seule composante de H_j à examiner est $c_j - a_j$, et comme c_j est entre a_j et $a_j + h_j$, on a $|c_j - a_j| \leq |h_j| \leq \|h\| < \eta$. Il découle de $(\star)$ que :

$$|\Delta(h)| < \varepsilon \sum_{j=1}^p |h_j| = \varepsilon \|h\|$$

donc, quand h tends vers 0, on a

$$\Delta(h) = o(\|h\|)$$

Il en découle que f est différentiable au point a .

• Il reste à démontrer que df est continue sur U .

• Pour tout $x \in U$ on a :

$$df(x)(h) = \sum_{i=1}^n \pi_i(h) \frac{\partial f}{\partial x_i}(x)$$

où π_i désigne l'application linéaire : $\pi_i : E \rightarrow \mathbb{R}; x \mapsto x_i$, avec $x = \sum_{k=1}^n x_k V_k$. On a alors :

$$df = \sum_{i=1}^n \Phi(\varphi_i, g_i)$$

où :

• Pour tout $i \in \llbracket 1, p \rrbracket$, $\varphi_i : U \rightarrow \mathcal{L}(E, \mathbb{R}); x \mapsto \varphi_i(x) = \pi_i$, avec $\pi_i(h) = h_i$, pour tout $h = \sum_{j=1}^p h_j e_j$. Il en

découle que φ_i est continue car constante.

• Pour tout $i \in \llbracket 1, p \rrbracket$, $g_i : U \rightarrow F; x \mapsto \frac{\partial f}{\partial x_i}(x)$, qui est continue par hypothèse.

• $\Phi : \mathcal{L}(E, \mathbb{R}) \times F \rightarrow \mathcal{L}(E, F); (\varphi, x) \mapsto \Phi(\varphi, x)$ avec $\Phi(\varphi, x)(h) = \varphi(h)x$ pour tout $h \in E$. On a Φ est bilinéaire.

• En appliquant la proposition concernant le continuité du produit d'applications continues via une application bilinéaire, on obtient que df est continue, ce qui achève la preuve du théorème ci-dessus. ■

13.2.1.2 Quelques propriétés

Proposition 13.2.3. Si $f : U \subset E \rightarrow F$ est une application de classe C^1 de U dans F et $\gamma : I \rightarrow E$ une application de classe C^1 d'un intervalle non trivial I de $\mathbb{R}$ vers E tel que $\gamma(I) \subset U$ alors pour tout $\alpha, \beta \in I$, en posant $a = \gamma(\alpha)$ et $b = \gamma(\beta)$, on a :

$$f(b) - f(a) = \int_{\alpha}^{\beta} df(\gamma(t)) \cdot \gamma'(t) dt.$$

Preuve. On sait que $(f \circ \gamma)$ est de classe C^1 et pour tout $t \in I$, on a : $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$, donc $\int_{\alpha}^{\beta} df(\gamma(t)) \cdot \gamma'(t) dt = \int_{\alpha}^{\beta} (f \circ \gamma)'(t) dt = f(\gamma(\beta)) - f(\gamma(\alpha)) = f(b) - f(a)$. ■

Proposition 13.2.4. Soit U un ouvert non vide connexe par arcs de E et $f : U \rightarrow F$ une application de classe C^1 , alors f est constante sur U si et seulement si $df(a) = \theta$ pour tout $a \in U$. (θ désigne l'application nulle de E vers F).

Le sens direct est évident ; l'autre sens ne l'est pas. On donne la preuve de cette proposition uniquement dans le cas où U est un ouvert non vide convexe de E .

Preuve. Supposons U convexe et soit $(a, b) \in U$. Pour tout $t \in [0, 1]$, posons $\gamma(t) = (1-t)a + tb = t(b-a) + a$, donc γ est de classe C^1 sur $[0, 1]$ et pour tout $t \in [0, 1]$, on a $\gamma'(t) = b - a$. Remarquons que $\gamma(0) = a$ et $\gamma(1) = b$ donc en appliquant la proposition 13.2.4, il vient : $f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot (b-a) dt$, or $\gamma(t) \in U$ par convexité et par suite $df(\gamma(t)) = \theta$, donc $f(b) - f(a) = 0$ et $f(a) = f(b)$ pour tout $a, b \in U$, donc f est constante sur U . ■

13.2.2 Fonctions de classe $C^k, k \in \mathbb{N} \cup \{\infty\}$

13.2.2.1 Dérivées partielles successives

On considère un espace vectoriel normé E de dimension p muni d'une base $\mathcal{B} = (e_1, \dots, e_p)$, les coordonnées d'un vecteur x étant notées $x_1, \dots, x_p$. Pour $k \in \mathbb{N}^*$, on définit la fonction $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$, comme suit :

- Pour $k = 1$, on définit $\frac{\partial^1 f}{\partial x_{i_1}} = \frac{\partial f}{\partial x_{i_1}}$ si celles-ci existent.
- Pour $k \in \mathbb{N}^*$ tel que $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$ existe et admet la dérivée partielle par rapport à x_{k+1} , on pose :

$$\frac{\partial^{k+1} f}{\partial x_{i_{k+1}} \dots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_{k+1}}} \left(\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} \right)$$

Definition 13.2.2

Si celles-ci existent les $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$ sont appelées les dérivées partielles de f d'ordre k

Definition 13.2.3

On dit que f est de classe C^k sur U si f admet toutes les dérivées partielles d'ordre k sur U et celles-ci sont continues sur U .

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

13.2.2.2 Théorème de Schwarz

Théorème 13.2.2. E est un espace vectoriel normé de dimension p et $f : U \rightarrow F$ une application de classe C^k sur U . Alors pour toute permutation $\sigma \in \mathcal{S}(\llbracket 1, k \rrbracket)$, on a :

$$\frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \dots \partial x_{i_{\sigma(p)}}} = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_p}}$$

Dans le cas particulier de $k = 2$, on a le théorème suivant :

Théorème 13.2.3. E est un espace vectoriel normé de dimension p et $f : U \rightarrow F$ une application de classe C^2 sur U . Pour tout $i, j \in \llbracket 1, p \rrbracket$, on a :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Remarques. Les remarques suivantes peuvent être utiles :

- 1) Pour tout $j \in \llbracket 0, k \rrbracket$, et tout $\alpha = (\alpha_1, \dots, \alpha_p) \in \llbracket 0, k \rrbracket^p$ tel que $\sum_{i=1}^p \alpha_i = j$, on note ∂_α l'application :

$$\partial_\alpha : C^k(U, F) \rightarrow C^{k-j}(U, F); f \mapsto \partial_\alpha(f) = \frac{\partial^j f}{\partial x_1^{\alpha_1} \dots \partial x_p^{\alpha_p}}$$

et une application linéaire.

- 2) Dans le cas particulier où $p = \infty$, on peut, pour tout $\alpha = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^p$, considérer l'endomorphisme ∂_α de $C^\infty(U, F)$ défini par :

$$\forall f \in C^\infty(U, F), \quad \partial_\alpha f = \partial_\alpha(f) = \frac{\partial^j f}{\partial x_1^{\alpha_1} \dots \partial x_p^{\alpha_p}}$$

Le théorème de Schwarz permet de dire que si $\alpha, \beta \in \mathbb{N}^p$ tel que $\alpha = (\alpha_1, \dots, \alpha_p)$ et $\beta = (\beta_1, \dots, \beta_p)$ et $\gamma = \alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_p + \beta_p)$ alors

$$\partial_\alpha \circ \partial_\beta = \partial_\beta \circ \partial_\alpha = \partial_\gamma.$$

13.2.2.3 Opérations sur les fonctions de classe C^k

Proposition 13.2.5. Soit $f, g : U \subset E \rightarrow F$ des applications de classe C^k sur U et $\alpha \in \mathbb{R}$. Alors $f + \alpha g$ est de classe C^k sur U .

Proposition 13.2.6. E, F_1, F_2 et G sont des espaces vectoriels normés et U un ouvert de E . Soit $f : U \rightarrow F_1$ et $g : U \rightarrow F_2$ des applications de classe C^k sur U et soit $B : F_1 \times F_2 \rightarrow G$ une application bilinéaire ; Alors $B(f, g)$ est de classe C^k sur U .

On rappelle que $B(f, g)(x) = B(f(x), g(x))$, pour tout $x \in U$.

Proposition 13.2.7. Soient E et F deux espaces vectoriels normés de dimensions finies, $n = \dim(F)$, $\mathcal{C} = (v_1, \dots, v_n)$ une base de F et U un ouvert non vide de E . Soit $f : U \rightarrow F$ de composantes $f_1, \dots, f_n$. Pour tout $k \in \mathbb{N}^* \cup \{\infty\}$, f est de classe C^k si et seulement si les $f_i, i \in \llbracket 1, n \rrbracket$ sont de classe C^k .

Proposition 13.2.8. Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Si $f : U \subset E \rightarrow F$ et $g : V \subset F \rightarrow G$ sont deux applications de classe C^k sur U et V respectivement tel que $f(U) \subset V$ alors $g \circ f$ est de classe C^k sur U .

Definition 13.2.4

Soit U un ouvert de $\mathbb{R}^p$. On appelle fonction polynomiale de U vers $\mathbb{R}$, toute application $f : U \rightarrow \mathbb{R}$ définie par

$$f(x) = \sum_{m \in I} a_m x_1^{m_1} \cdots x_p^{m_p}$$

où $x = (x_1, \dots, x_p) \in U$, I une partie finie non vide de $\mathbb{N}^p$, $m \in I$ tel que $m = (m_1, \dots, m_p)$ et $a_m \in \mathbb{R}$ pour tout $m \in I$.

Proposition 13.2.9. Toute fonction polynomiale de U vers $\mathbb{R}$ est de classe C^∞ sur U . Si f et g sont deux fonctions polynomiales sur U tel que $\forall x \in U, g(x) \neq 0$ alors $\frac{f}{g}$ est de classe C^∞ sur U .

13.2.3 Formule de Taylor d'ordre 2

Proposition 13.2.10. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application de classe C^2 et $a \in U$. Alors pour h voisin de 0 dans $\mathbb{R}^p$, on a :

$$f(a+h) = f(a) + \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^p h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o(\|h\|^2)$$

13.3 Fonction à valeur réelle

Grâce aux propositions sur les composantes d'une application, l'étude d'une application $f : U \rightarrow F$ se ramène à celle d'une fonction $g : U \rightarrow \mathbb{R}$. C'est pour cela que nous réservons ce paragraphe à de telles fonctions. On se contentera du cas $E = \mathbb{R}^p$ puisque tout espace vectoriel normé réel de dimension p est isomorphe à $\mathbb{R}^p$.

Dans la suite $\mathbb{R}^p$ est muni de la norme de sa structure euclidienne canonique, c'est-à-dire celle qui provient du produit scalaire canonique : Pour $x = (x_1, \dots, x_p)$ et $y = (y_1, \dots, y_p)$, on a :

$$\langle x, y \rangle = \sum_{k=1}^p x_k y_k$$

13.3.1 Gradient

Remarquons que si $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ est une application différentiable en un point a de U , sa différentielle $df(a)$ est une forme linéaire sur $\mathbb{R}^p$. Par le théorème de représentation il existe un unique vecteur qu'on note $\nabla f(a)$ de $\mathbb{R}^p$ tel que

$$\forall h \in \mathbb{R}^p, \quad df(a).h = \langle \nabla f(a), h \rangle$$

13.3.1.1 Définition, exemple

Definition 13.3.1

Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ une application et $a \in U$. Si f est différentiable au point a , l'unique vecteur $\nabla f(a)$ tel que :

$$\forall h \in \mathbb{R}^p, \quad df(a).h = \langle \nabla f(a), h \rangle$$

est appelé gradient de f au point a .

Proposition 13.3.1. Si f est différentiable au point a et $\mathcal{B} = (e_1, \dots, e_p)$ est la base canonique de $\mathbb{R}^p$ alors :

$$\nabla f(a) = \sum_{k=1}^p \frac{\partial f}{\partial x_k}(a) e_k$$

Exemples. Quand l'espace d'arrivée n'est pas de la forme $\mathbb{R}^p$, on s'y ramène en considérant l'isomorphisme canonique.

1. $f : \mathbb{R}^2 \rightarrow \mathbb{R}; (x, y) \mapsto xy + x + y$ et $a = (3, 4)$ alors f est différentiable au point a car f est polynomiale des variables x et y . On a $\frac{\partial f}{\partial x}(x, y) = y + 1$ et $\frac{\partial f}{\partial y}(x, y) = x + 1$, donc $\nabla f(a) = (5, 4)$
2. Soit $f : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}, A \mapsto f(A) = \text{tr}(A)$. Alors pour tout $X \in \mathcal{M}_n(\mathbb{K})$ et tout $(i, j) \in \llbracket 1, n \rrbracket^2$, on a $\frac{\partial f}{\partial x_{ij}}(X) = \delta_{ij}$ le symbole de Kronecker. Il en découle que $\nabla f(X) = I_n$ pour tout $X \in \mathcal{M}_n(\mathbb{K})$.
3. Soit $g : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}; X \mapsto \det(X)$. Alors en développant suivant la colonne j , on a :

$$g(X) = \sum_{i=1}^n (-1)^{i+j} \Delta_{ij}$$

où Δ_{ij} est le mineur pour la ligne i et la colonne j . Comme Δ_{ij} ne dépend pas de X_{ij} , on a :

$$\frac{\partial g}{\partial x_{ij}}(X) = (-1)^{i+j} \Delta_{ij}$$

par suite :

$$\nabla g(X) = \text{Com}(X)$$

13.3.2 Extrêmes sans contrainte

13.3.2.1 Point critique, maximum, minimum, extremum

Definition 13.3.2

Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ une application différentiable sur U . Soit $a \in U$. On dit que a est un point critique de f si $df(a) = 0$, ce qui revient à $\nabla f(a) = 0$.

Remarque. Cela revient à dire que les dérivées partielles de f au point a sont nulles, donc :

$$\forall j \in \llbracket 1, p \rrbracket \quad \frac{\partial f}{\partial x_j}(a) = 0$$

Definition 13.3.3

Soit $f : U \rightarrow \mathbb{R}$ une application et $a \in U$. On dit que f admet un maximum (resp. minimum) local au point a s'il existe $r > 0$ tel que :

$$(1) \quad \forall x \in B(a, r), \quad f(x) \leq f(a) \text{ (resp. } f(a) \leq f(x) \text{)}.$$

On dit que f admet un extremum local au point a si f admet un minimum ou un maximum local au point a .

Si l'inégalité (1) ci-dessus est vérifiée pour tout $x \in U$, on parle de maximum (resp. minimum) global donc d'extremum global.

Si l'inégalité (1) ci-dessus est stricte sauf pour $x = a$, on parle de maximum (resp. minimum) local strict, donc d'extremum local strict.

13.3.2.2 Condition nécessaire d'extremum

Théorème 13.3.1. Soit $f : U \rightarrow \mathbb{R}$ une application différentiable sur U et $a \in U$. Si f admet un extremum au point a alors a est un point critique de f

Preuve. Supposons que f admet un extremum au point a . Soit e un vecteur quelconque de $\mathbb{R}^p$ et considérons la fonction réelle de variable réelle g définie par $g(t) = f(a + te)$. Alors g admet un extremum au point 0. Comme f est différentiable sur U , le fonction g est dérivable sur un intervalle ouvert contenant 0, donc $g'(0) = 0$ donc $\langle \nabla f(a), e \rangle = 0$. Ainsi on a :

$$\forall e \in \mathbb{R}^p, \quad \langle \nabla f(a), e \rangle = 0$$

donc $\nabla f(a) = 0$, donc a est un point critique de f au point a . ■

Remarques. 1. Cela se traduit dans la pratique par : Avant de chercher les extremums d'une applications $f : U \rightarrow \mathbb{R}$, il faut commencer par chercher les points critiques puis chercher parmi ceux-ci les extremums.
2. On donnera des conditions du second ordre dans le sous paragraphe ci-dessous.

13.3.2.3 Conditions du second ordre d'extremum

On commence par donner des rappels et compléments d'algèbre linéaire :

Théorème 13.3.2. Toute matrice réelle symétrique est diagonalisable dans une base orthonormée. Autrement dit si A est une matrice réelle symétrique d'ordre p , alors il existe des nombres réels $\lambda_1, \dots, \lambda_p$ et

une matrice orthogonale Ω tel $A = {}^t\Omega D \Omega$ où $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix}$ est la matrice diagonales de termes diagonaux $\lambda_1, \dots, \lambda_p$.

Dans tout ce qui suit U est un ouvert non vide de $\mathbb{R}^p$ et $f : U \rightarrow \mathbb{R}$ est une application de classe C^2 sur U . On rappelle la formule de Taylor-Young au voisinage de a à l'ordre 2 :

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^p h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \|h\|^2 \varepsilon(h)$$

On dispose de la matrice carrée :

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

appelée la hessienne de f au point a . Par le théorème de Shwarz, la hessienne au point a est une matrice réelle symétrique, donc par le théorème spectral elle est diagonalisable. Si on note X la colonne des coordonnées de h , la formule de Taylor-Young ci-dessus s'écrit :

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

Théorème 13.3.3. Soit $f : U \rightarrow \mathbb{R}$ une application de classe C^2 et $a \in U$ un point critique de f (c'est-à-dire $\nabla f(a) = 0$). On note $\lambda_1, \dots, \lambda_p$ les valeurs propres de

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

la hessienne de f au point a . Alors :

1. Si les $\lambda_j, j \in \llbracket 1, p \rrbracket$ sont strictement positives alors $f(a)$ est un minimum local strict de f .
2. Si les $\lambda_j, j \in \llbracket 1, p \rrbracket$ sont strictement négatives alors $f(a)$ est un maximum local strict de f .

Preuve. La formule de Taylor-Young s'écrit :

$$f(a+h) - f(a) = \frac{1}{2} \langle H_f(a)X, X \rangle + \|X\|^2 \varepsilon(X)$$

• Supposons que les valeurs propres $\lambda_1, \dots, \lambda_p$ de $H_f(a)$ sont strictement positives.

L'application $g : X \mapsto \langle H_f(a)X, X \rangle$ est continue sur $\mathbb{R}^p$ car polynomiale en les coordonnées de X dans n'importe quelle base orthonormée de $\mathbb{R}^p$. De plus, on a :

$$(\star) \quad \forall X \in \mathbb{R}^p, \quad X \neq 0 \Rightarrow g(X) > 0.$$

En effet si on écrit $X = \sum_{k=1}^p x_k V_k$ où $(V_k)_{1 \leq k \leq p}$ est une base orthonormée de vecteurs propres de $H_f(a)$ le vecteur V_k étant associé à la valeur propre λ_k , on a $H_f(a)X = \sum_{k=1}^p x_k \lambda_k V_k$, donc $g(X) = \sum_{k=1}^p \lambda_k x_k^2$ donc $g(X) \geq 0$, or $g(X) = 0 \Rightarrow \forall k \in \llbracket 1, p \rrbracket, x_k = 0$, or $X \neq 0$, donc $g(X) > 0$. Comme $S = \{X \in \mathbb{R}^p / \|X\| = 1\}$ est une partie compacte de $\mathbb{R}^p$, et g continue donc g est bornée sur S et atteint ses bornes, en particulier, il existe $X_0 \in S$ tel que $g(X_0) = \min_{X \in S} g(X)$. On a $X_0 \in S$, donc $\|X_0\| = 1$ et par suite $X_0 \neq 0$, donc, en vertu du $(\star)$ ci-dessus, on a $\alpha = g(X_0) > 0$. En particulier, pour tout $X \in \mathbb{R}^n \setminus \{0\}$, on a : $g\left(\frac{X}{\|X\|}\right) \geq \alpha$, donc $g(X) \geq \alpha \|X\|^2$. Comme $\lim_{X \rightarrow 0} \varepsilon(X) = 0$, il existe $r > 0$ tel que :

$$\|X\| < r \Rightarrow |\varepsilon(X)| < \frac{\alpha}{4}$$

En particulier :

$$\begin{aligned} \|X\| < r &\Rightarrow \varepsilon(X) > -\frac{\alpha}{4} \\ &\Rightarrow f(a+X) - f(a) = \frac{1}{2}g(X) + \|X\| \varepsilon(X) \geq \left(\frac{\alpha}{2} - \frac{\alpha}{4}\right) \|X\|^2 \\ &\Rightarrow f(a+X) - f(a) \geq \frac{\alpha}{4} \|X\|^2 \end{aligned}$$

Il en découle que pour tout $X \in B(0, r)$ on a $f(a+X) - f(a) \geq 0$ avec égalité si et seulement si $X = 0$, ce qui veut dire que $f(a)$ est un minimum local strict.

• Supposons maintenant qu'il existe $m, q \in \mathbb{N}^*$ tel que $m+q=p$ et les valeurs propres de $H_f(a)$ sont $\lambda_1, \dots, \lambda_m, -\lambda_{m+1}, \dots, -\lambda_{p+q} = -\lambda_n$ avec $\forall k \in \llbracket 1, p \rrbracket, \lambda_k > 0$. Si $X = \sum_{k=1}^p x_k V_k$ alors :

$$H_f(a)X = \sum_{k=1}^m \lambda_k x_k^2 - \sum_{k=1}^q \lambda_{m+k} x_{m+k}^2$$

de sorte que :

- si $V = \sum_{k=1}^m V_k$, alors :

$$\forall n \in \mathbb{N}^*, H_f(a) \left(\frac{1}{n} V \right) = \frac{1}{n^2} \lambda, \quad \text{avec} \quad \lambda = \sum_{k=1}^m \lambda_k > 0$$

- si $V' = \sum_{k=1}^q V_{m+k}$, alors :

$$\forall n \in \mathbb{N}^*, H_f(a) \left(\frac{1}{n} V' \right) = -\frac{1}{n^2} \lambda', \quad \text{avec} \quad \lambda' = \sum_{k=1}^q \lambda_{m+k} > 0$$

Il en découle et compte tenu de la formule de Taylor ci-dessus que pour $n \in \mathbb{N}^*$ suffisamment grand ($n \rightarrow +\infty$), on a :

$$\begin{cases} f\left(a + \frac{1}{n} V\right) = \frac{\lambda}{n^2} + o\left(\frac{1}{n^2}\right) \\ f\left(a + \frac{1}{n} V'\right) = -\frac{\lambda'}{n^2} + o\left(\frac{1}{n^2}\right) \end{cases}$$

ce qui prouve que $f(a+h) - f(a)$ ne garde pas un signe constant pour h sur un voisinage de 0 et explique pourquoi on a un point selle. ■

13.3.2.4 Cas particulier de $n = 2$

Proposition 13.3.2. Soit $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ une application de classe C^2 sur U et $a \in U$ un point critique de f . On note :

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

appelées notations de Monge. Alors, on a :

1. Si $rt - s^2 > 0$ alors :
 - Si $r > 0$ alors f admet un minimum local strict en a .
 - Si $r < 0$ alors f admet un maximum local strict en a .
2. Si $rt - s^2 < 0$ alors f admet un point selle au point a .

Preuve. Remarquons que $rt - s^2 = \det(H_f(a))$ et $r + t = \text{tr}(H_f(a))$. Si on note λ_1, λ_2 les valeurs propres de $H_f(a)$ alors on a :

$$\begin{aligned} \lambda_1 > 0 \quad \text{et} \quad \lambda_2 > 0 &\Leftrightarrow \lambda_1 \lambda_2 > 0 \quad \text{et} \quad \lambda_1 + \lambda_2 > 0 \\ &\Leftrightarrow \det(H_f(a)) > 0 \quad \text{et} \quad \text{tr}(H_f(a)) > 0 \\ &\Leftrightarrow rt - s^2 > 0 \quad \text{et} \quad r + t > 0 \\ &\Leftrightarrow rt - s^2 > 0 \quad \text{et} \quad r > 0 \end{aligned}$$

La dernière équivalence étant justifiée comme suit : Si $rt - s^2 > 0$ et $r + t > 0$ alors si $r \leq 0$, forcément $t > 0$ donc rt serait ≤ 0 et alors $rt - s^2 \leq 0$, ce qui n'est pas vrai, donc $r > 0$. réciproquement si $rt - s^2 > 0$ et $r > 0$ alors $r + t \leq 0 \Rightarrow t \leq 0$ donc $rt - s^2$ serait ≤ 0 , chose fausse, donc $r + t > 0$. ■

Exemple. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto f(x, y) = x^4 + y^4 - (x - y)^2$. La fonction f est de classe C^2 car elle est polynomiale en les variables x et y . On a :

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 - 2(x - y), \quad \frac{\partial f}{\partial y}(x, y) = 4y^3 + 2(x - y).$$

Donc :

$$\begin{aligned}
 (x, y) \text{ est un point critique} &\Leftrightarrow \begin{cases} 4x^3 = 2(x - y) \\ 4y^3 = -2(x - y) \end{cases} \\
 &\Leftrightarrow \begin{cases} x^3 = -y^3 \\ x - y = 2x^3 \end{cases} \\
 &\Leftrightarrow \begin{cases} y = -x \\ 2x = 2x^3 \end{cases} \\
 &\Leftrightarrow x = y = 0 \quad \text{ou} \quad \begin{cases} x = 1 \\ y = -1 \end{cases} \\
 &\quad \text{ou} \quad \begin{cases} x = -1 \\ y = 1 \end{cases}
 \end{aligned}$$

Il y'a donc trois points critiques : $(0, 0)$, $(1, -1)$ et $(-1, 1)$. En remarquant que $f(X) = f(-X)$ pour tout $X \in \mathbb{R}^2$, on fera l'étude uniquement en $(0, 0)$ et $(1, -1)$.

Les dérivées partielles d'ordre 2 de f sont :

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 12x^2 - 2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 2, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 12y^2 - 2,$$

Il en découle que :

- Au point $(1, -1)$, on a : $r = 10, s = 2, t = 10$, donc $rt - s^2 = 96 > 0$ et comme $r > 0$, $f(1, -1) = -2$ est un minimum local strict de f .
- Comme $f(1, -1) = f(-1, 1) = -2$ on a la même conclusion pour $(-1, 1)$.
- En $(0, 0)$: $r = -2, s = 2, t = -2$, donc $rt - s^2 = 0$. On ne peut pas conclure par le théorème ci-dessus, on fera une étude par une autre méthode : On remarque que si on pose $u_n = (\frac{1}{n}, \frac{1}{n})$, pour tout $n \in \mathbb{N}^*$ alors $\lim_{n \rightarrow +\infty} u_n = (0, 0)$ et $f(u_n) = \frac{2}{n^4} > 0$ pour tout $n \in \mathbb{N}^*$. Par ailleurs si on pose $v_n = (\frac{1}{n}, 0)$ pour tout $n \in \mathbb{N}^*$, on a $\lim_{n \rightarrow +\infty} v_n = (0, 0)$ et $f(v_n) = \frac{1-n^2}{n^4} < 0$, pour tout $n \in \mathbb{N}^*$ tel que $n \geq 2$. Il en résulte que f n'admet ni minimum ni maximum local en $(0, 0)$.

13.3.3 extremums avec contrainte ou extrêmes liés

On cherche dans cette partie à déterminer les extremas d'une fonction $f : U \subset E \rightarrow \mathbb{R}$ de classe C^1 sous la contrainte $g(x) = 0$, où g est elle-même une fonction numérique de classe C^1 sur l'ouvert U . Pour déterminer $\min_{g(x)=0} f(x)$, une approche naïve consiste à tirer de la contrainte $g(x_1, \dots, x_n) = 0$ une relation de la forme $x_n = \varphi(x_1, \dots, x_{n-1})$. On se ramène de la sorte aux techniques d'optimisation du paragraphe précédent :

$$\min_{g(x)=0} f(x) = \min_{(x_1, \dots, x_{n-1}) \in U'} f(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$$

où U' serait une partie de E à préciser.

Le théorème ci-dessous donne une réponse plus rigoureuse à la question en utilisant ce qu'on appelle multiplicateur de Lagrange :

Théorème 13.3.4. Soient f et g deux fonctions numériques de classe C^1 sur l'ouvert U de E et $X = \{x \in U \mid g(x) = 0\}$. Si la restriction de f à X admet un extremum local en $a \in X$ et $dg(a) \neq 0$, alors il existe $\lambda \in \mathbb{R}$ tel que $df(a) = \lambda dg(a)$.

Exemple 1 :

Déterminons $\min_{x^2+y^2=1} xy$. Pour tout $(x, y) \in \mathbb{R}^2$, posons $f(x, y) = xy$ et $g(x, y) = x^2 + y^2 - 1$.

- Les deux fonctions numériques f et g sont polynomiales donc de classe $\mathcal{C}^\infty$ sur l'ouvert $\mathbb{R}^2$.
- La partie Γ de $\mathbb{R}^2$ définie par $\Gamma = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$ est un compact car c'est la sphère unité de l'espace vectoriel normé $(\mathbb{R}, |\cdot|)$, il en découle que $f|_\Gamma$ admet un minimum atteint en un point $(x_0, y_0) \in \Gamma$.

- Pour tout $(x, y) \in \mathbb{R}^2$, on a les expressions des gradients de f et g respectivement au point (x, y) :

$$\nabla f(x, y) = \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{et} \quad \nabla g(x, y) = 2 \begin{pmatrix} x \\ y \end{pmatrix} \neq 0.$$

D'après le théorème ci-dessus, il existe $\lambda \in \mathbb{R}^*$ tel que :

$$\begin{cases} y_0 = 2\lambda x_0 \\ x_0 = 2\lambda y_0 \\ x_0^2 + y_0^2 = 1 \end{cases}.$$

Il en découle après calcul que $(x_0, y_0, \lambda) = \left(\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}, \frac{1}{2}\right)$ ou bien $(x_0, y_0, \lambda) = \left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, -\frac{1}{2}\right)$. D'où

$$\min_{x^2+y^2=1} xy = -\frac{1}{2}.$$

Exercice

Exercice 13.3.1. (CCP 2023)

Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f : (x, y) \mapsto 4x^2 + 12xy - y^2$. Soit $C = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 13\}$.

1. Justifier que f atteint un maximum et un minimum sur C .
2. Soit $(u, v) \in C$ un point où f atteint un de ses extremums.
 - (a) Justifier avec un théorème du programme qu'il existe un réel λ tel que le système (S) suivant soit vérifié :

$$(S) : \begin{cases} 4u + 6v = \lambda u \\ 6u - v = \lambda v \end{cases}$$

- (b) Montrer que $(\lambda - 4)(\lambda + 1) - 36 = 0$. En déduire les valeurs possibles de λ .
3. Déterminer les valeurs possibles de (u, v) , puis donner le maximum et le minimum de f sur C .

Solution:

1. C est la sphère de $\mathbb{R}^2$ de centre $(0, 0)$ et de rayon $\sqrt{13}$ pour la norme $\|\cdot\|_2$ usuelle. C est donc une partie fermée et bornée de $\mathbb{R}^2$. Comme $\mathbb{R}^2$ est de dimension finie, C est un compact de $\mathbb{R}^2$. On a aussi $C \neq \emptyset$. f est une application polynomiale donc f est continue sur le compact C . Donc f atteint un maximum et un minimum sur C .
2. (a) Soit $g : (x, y) \mapsto x^2 + y^2 - 13$. g est de classe C^1 sur $\mathbb{R}^2$ en tant que fonction polynomiale, tout comme f . De plus, le gradient de g est $\nabla g : (x, y) \mapsto (2x, 2y)$ de sorte que $\nabla g(x, y) = (0, 0)$ si et seulement si $(x, y) = (0, 0)$. On en déduit que ∇g ne s'annule pas sur C . Il découle alors, du théorème d'optimisation sous une contrainte, qu'il existe un scalaire λ (multiplicateur de Lagrange) tel que $\nabla f(u, v) = \lambda \nabla g(u, v)$. Et on a : $\nabla f(u, v) = \lambda \nabla g(u, v) \Leftrightarrow \begin{cases} 4u + 6v = \lambda u \\ 6u - v = \lambda v \end{cases}$
 - (b) (S) est équivalent à $\begin{cases} (4 - \lambda)u + 6v = 0 \\ 6u - (1 + \lambda)v = 0 \end{cases}$ que l'on peut voir comme un système linéaire en (u, v) avec un paramètre λ . Comme $(0, 0) \notin C$ et que f possède effectivement des extremums sur C , ce système linéaire a nécessairement au moins une solution non nulle. Ce qui implique que ce système n'est pas de Cramer. Son déterminant $(\lambda - 4)(\lambda + 1) - 36$ est donc nul. En développant, il vient $\lambda^2 - 3\lambda - 40 = 0$. Les solutions de cette équation sont 8 et -5, qui sont donc les deux seules valeurs possibles de λ .
3. On discute les cas :
 - Si $\lambda = 8$ on obtient, d'après (S), $v = \frac{2}{3}u$. Comme $(u, v) \in C$, il vient $u^2 + v^2 = u^2 + \frac{4}{9}u^2 = \frac{13}{9}u^2 = 13$. donc $u = \pm 3$ donc $(u, v) \in \{(3, 2), (-3, -2)\}$.
 - Si $\lambda = -5$ on obtient, d'après (S), $v = -\frac{3}{2}u$. Comme $(u, v) \in C$, il vient $\frac{13}{4}u^2 = 13$. donc $u = \pm 2$, donc $(u, v) \in \{(2, -3), (-2, 3)\}$. Or, $f(3, 2) = f(-3, -2) = 36 + 72 - 4 = 104$ et $f(2, -3) =$

$f(-2, 3) = 16 - 72 - 9 = -65$. Comme f atteint effectivement un maximum et un minimum sur C et qu'elle ne peut les atteindre qu'en ces points, on a donc $\max_C f = 104$ et $\min_C f = -65$.

Exercice 13.3.2. (inégalité arithmético-géométrique) :

Soit $n \geq 2$ et $f : \mathbb{R}^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto x_1 \cdots x_n$. On note :

$$\Gamma = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n; x_1 + \cdots + x_n = 1\}.$$

1. Démontrer que f admet un maximum global sur Γ et le déterminer.
2. En déduire l'inégalité arithmético-géométrique :

$$\forall (x_1, \dots, x_n) \in \mathbb{R}_+^n, \quad \sqrt[n]{\prod_{i=1}^n x_i} \leq \frac{1}{n} \sum_{i=1}^n x_i.$$

Solution:

1. On commence par remarquer que Γ est compact. En effet, Γ est clairement fermé. De plus, pour $(x_1, \dots, x_n) \in \mathbb{R}_+^n$, on a $\|(x_1, \dots, x_n)\|_1 = x_1 + \cdots + x_n \leq 1$, et donc Γ est borné. C'est une partie compacte de $\mathbb{R}^n$ et f , qui est continue sur Γ , admet un maximum global sur Γ , atteint en a . De plus, puisque $f(x_1, \dots, x_n)$ s'annule si un des x_i est nul, il est clair que $a \in]0, +\infty[^n$. Posons $g(x) = x_1 + \cdots + x_n$. Par le théorème des extrema liés (ou multiplicateurs de Lagrange), il existe $\lambda \in \mathbb{R}$ tel que, pour tout $i \in \{1, \dots, n\}$, on a

$$\frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial g}{\partial x_i}(a) = \lambda.$$

Or,

$$\frac{\partial f}{\partial x_i}(a) = \frac{f(a)}{a_i}$$

Puisque $f(a) \neq 0$, on a $\lambda \neq 0$ et tous les a_i sont égaux (à $f(a)/\lambda$). Mais, puisque $a_1 + \cdots + a_n = 1$, on en déduit que f atteint son maximum sur Γ en le point $(1/n, \dots, 1/n)$. Ce maximum vaut $1/n^n$.

2. La question précédente prouve que, si $(x_1, \dots, x_n) \in [0, +\infty[^n$ satisfait $x_1 + \cdots + x_n = 1$, alors

$$\sqrt[n]{\prod_{i=1}^n x_i} \leq \frac{1}{n} \sum_{i=1}^n x_i.$$

Dans le cas général, on se ramène à ceci en posant $x'_i = \frac{x_i}{\sum_{i=1}^n x_i}$,

qui vérifie $x'_1 + \cdots + x'_n = 1$. De $\sqrt[n]{\prod_{i=1}^n x_i} \leq \frac{1}{n}$, on déduit l'inégalité demandée.

13.3.4 Application à la géométrie différentielle

Dans tout ce sous-paragraphe p est un entier naturel non nul et $E = \mathbb{R}^p$ est muni de sa structure canonique d'espace euclidien.

13.3.4.1 Vecteur tangent, variété tangente

Soit Γ une partie non vide de $\mathbb{R}^p$. Un vecteur v de $\mathbb{R}^p$ est dit tangent à Γ en un point a de Γ s'il existe un arc paramétré $(] - \varepsilon, \varepsilon[, \gamma)$ de classe C^1 tel que $\gamma(I_\varepsilon) \subset \Gamma$ et $\gamma(0) = a$ et $\gamma'(0) = v$. (ici $I_\varepsilon =] - \varepsilon, \varepsilon[$). On note $T_a \Gamma$, l'ensemble des vecteurs de E tangents à Γ au point a . L'ensemble :

$$V_a = a + T_a \Gamma = t_a(T_a \Gamma)$$

image de $T_a \Gamma$ par la translation t_a de $\mathbb{R}^p$ de vecteur a , s'appelle la variété tangente à Γ au point a . Si $T_a \Gamma$ est un sous-espace vectoriel de $\mathbb{R}^p$ alors V_a s'appelle variété affine tangente à Γ au point a .

Remarque. Il existe une définition plus générale de vecteur tangent à une partie à savoir : Soit E un espace vectoriel normé réel et Γ une partie non vide de E et soit $a \in \Gamma$. Un vecteur v de E est dit vecteur tangent au point a à Γ s'il existe une suite $(x_n) \in (\Gamma \setminus \{a\})^{\mathbb{N}}$ et une suite $(\alpha_n) \in \mathbb{R}_+^{\mathbb{N}}$ tel que :

$$\left\{ \begin{array}{ll} (i) & \lim_{n \rightarrow +\infty} x_n = a \\ (ii) & \lim_{n \rightarrow +\infty} (\alpha_n(x_n - a)) = v \end{array} \right.$$

13.3.4.2 Lignes de niveau, surfaces de niveau

Si $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ est une application différentiable alors pour tout nombre réel c , le sous ensemble Γ_c de U tel que :

$$\Gamma_c = \{X \in U / f(X) = c\}$$

s'appelle ligne de niveau si $p = 2$ et surface de niveau si $p = 3$. Généralement on l'appelle une équipotentielle.

13.3.4.3 Plan tangent à une surface de niveau $f(x, y, z) = 0$

Soit $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ une application différentiable sur U et soit $\Gamma = \{X \in U / f(X) = 0\}$. Soit $a \in \Gamma$ tel que $\nabla f(a) \neq 0$. On peut démontrer et on admet que la plan $(\mathbb{R}\nabla f(a))^\perp$ est l'espace tangent à Γ au point a . On peut cependant démontrer une inclusion à savoir, tout vecteur tangent à Γ au point a est orthogonal au vecteur $\nabla f(a)$. En effet si (J, γ) est une arc de classe C^1 tracé sur Γ alors $\forall t \in J, f(\gamma(t)) = 0$, donc $f \circ \gamma$ est constante sur J , donc sa dérivée est nulle sur J , ce qui veut dire : $\forall t \in J, \langle \nabla f(t), \gamma'(t) \rangle = 0$ et traduit que $\gamma'(t) \perp \nabla f(t)$ pour tout $t \in J$. Si $J =]-\alpha, \alpha[$ et $\gamma(0) = a$ et $\gamma'(0) = V$ alors $\langle \nabla f(a), V \rangle = 0$, donc $V \in (\mathbb{R}\nabla f(a))^\perp$. Notons que le plan affine tangent à Γ au point a , est le plan affine $a + T_a\Gamma_c$ d'équation :

$$\langle X - a, \nabla f(a) \rangle = 0$$

c'est-à-dire :

$$(x_1 - a_1) \cdot \frac{\partial f}{\partial x_1}(a) + (x_2 - a_2) \cdot \frac{\partial f}{\partial x_2}(a) + (x_3 - a_3) \cdot \frac{\partial f}{\partial x_3}(a) = 0$$

13.3.4.4 Gradient et dérivée directionnelle

On aura besoin de la proposition suivante qui est une conséquence immédiate de l'inégalité de Cauchy-Schwarz et son cas d'égalité :

Proposition 13.3.3. Pour tout $x, y \in \mathbb{R}^p$, on a :

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

avec égalité si et seulement si x et y sont directement colinéaires.

Preuve. Rappelons que l'inégalité de Cauchy-Schwarz dans l'espace euclidien $\mathbb{R}^p$ dit que pour tout $x, y \in \mathbb{R}^p$ on a

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

avec égalité si et seulement si x et y sont colinéaires.

Il en découle en particulier que : $\langle x, y \rangle \leq \|x\| \|y\|$. Supposons que cette dernière inégalité est une égalité, alors on a le cas d'égalité de Cauchy-Schwarz à savoir $x = 0$ ou $x \neq 0$ et $y = \alpha x$ avec $\alpha \in \mathbb{R}$ et de plus $\langle x, y \rangle \geq 0$, ce qui donne $\lambda \geq 0$ donc x et y sont directement colinéaires. Réciproquement cette dernière condition donne l'égalité : $\langle x, y \rangle = \|x\| \|y\|$. ■

Proposition 13.3.4. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ différentiable sur U , et $a \in U$ tel que $\nabla f(a) \neq 0$. Alors la valeur maximale de $D_e f(a)$ où e est un vecteur unitaire correspond au cas où e et $\nabla f(a)$ sont directement colinéaires, c'est-à-dire où il existe $\alpha > 0$ tel que $\nabla f(a) = \alpha e$

Preuve. $g(t) = f(a + te)$. On sait qu'il existe un voisinage V de 0 tel que g est dérivable sur V et pour tout $t \in V$, on a :

$$g'(t) = \langle \nabla f(a + te), e \rangle$$

en particulier :

$$D_e f(a) = g'(0) = \langle \nabla f(a), e \rangle$$

donc par la conséquence ci-dessus de l'inégalité de Cauchy-Schwarz, on a :

$$D_e f(a) \leq \|\nabla f(a)\| \|e\|$$

avec égalité si $\nabla f(a)$ et e sont directement colinéaires. Or le cas d'égalité correspond au fait que $D_e f(a)$ est maximale, ce qui prouve la proposition. ■