Chapitre 13

Eléments de calcul différentiel

On poursuit l'étude des fonctions vectorielles avec ici les fonctions d'une variable vectorielle à valeurs vectorielles. On revient notamment sur la notion de différentiabilité qui nous permettra d'obtenir une approximation linéaire de la différence $f(a+h)-f(a)$... cela sera très utile dans l'étude des champs scalaires, car on pourra exhiber, à l'aide d'un développement limité vectoriel, des conditions pour l'existence d'un extremum.

Programmes 2022

Pour aller plus loin
Ce chapitre n'est pas facile, que ce soit en raison des notations ou du travail sur les vecteurs : il nécessite donc beaucoup de rigueur dans la compréhension des objets. Malgré tout, on définit ici une notion utile : la différentiabilité et on essaiera de retenir quelques exemples d'application très classiques pour les matrices, avec notamment une autre preuve du théorème spectral.

1 Notion de fonction différentiable

Dans tout ce chapitre, on travaille dans des espaces vectoriels réels, normés et de dimension finie, et les fonctions manipulées seront toutes définies sur un ouvert U inclus dans un espace vectoriel normé E de sorte que :

$$a \in U \Rightarrow \exists\, \epsilon > 0,\ B(a, \epsilon) \subset U$$

et ainsi, au voisinage de a, on pourra toujours considérer h tel que $a + h \in U$.

1.1 Existence et unicité de la différentielle

Définition Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $a \in U$.

- On dit que f est **différentiable en** a s'il existe une application linéaire $L_a \in \mathcal{L}(E, F)$ telle que :

$$f(a + h) = f(a) + L_a(h) + \underset{h \to 0}{o}(\|h\|)$$

- Plus généralement, on dit que f est **différentiable sur** U si f est différentiable en tout point de U.

Remarques

1. Vérifier que f est différentiable en un point revient à montrer qu'on peut donner une **approximation linéaire** de la différence $f(a + h) - f(a)$ au voisinage de a.
2. D'ailleurs, pour une fonction vectorielle d'une variable réelle, cette notion de différentiabilité équivaut à la dérivabilité. En effet, si $f : I \longrightarrow F$, on rappelle qu'elle est dérivable en a si et seulement si :
$$f(a + h) = f(a) + h.f'(a) + o(h)$$
et dans ce cas, on peut dire que f est différentiable en a avec $L_a(h) = h.f'(a)$. La réciproque est immédiate puisque la linéarité de L nous donne un développement limité de f en a, et donc sa dérivabilité :
$$f(a + h) = f(a) + L_a(h) + o(|h|) \Rightarrow f(a + h) = f(a) + h.L_a(1) + o(h)$$
3. Malheureusement, dans le cas plus général des fonctions d'une variable vectorielle, on ne pourra pas toujours parler de **dérivabilité** (car le taux d'accroissement peut ne pas avoir de sens) et on lui préférera la notion de **différentiabilité** : c'est là une des difficultés de ce chapitre.

Propriété 1 (unicité de la différentielle).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $a \in U$. On suppose de plus que f est différentiable en a. Alors, il existe une unique application linéaire $L_a \in \mathcal{L}(E, F)$ telle que :

$$f(a + h) = f(a) + L_a(h) + o(\|h\|)$$

Cette application désigne l'**application linéaire tangente** ou plus simplement, la **différentielle** de f en a qu'on notera $df_a(h)$.

▶ *L'existence est immédiate. Pour l'unicité, on raisonne par l'absurde de sorte que dans $\mathcal{L}_c(E, F)$, $\alpha = |||L_1 - L_2||| \neq 0$ et on se ramène à la différence $L_1(h) - L_2(h) = o(\|h\|)$ afin d'obtenir une contradiction sur la norme subordonnée.*

Propriété 2 (une fonction différentiable est continue).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $a \in U$. On suppose de plus que f est différentiable en a. Alors, f est continue en a.

▶ *On revient à la définition et on montre par continuité de la différentielle que $f(a + h) \longrightarrow f(a)$ quand $h \to 0$.*

Propriété 3 (cas particulier des applications linéaires et bilinéaires).

Soient E_1, E_2, F des espaces vectoriels normés de dimension finie.

1. Si L désigne une application linéaire de E_1 dans F, alors L est différentiable sur E_1 et on a en tout point $a \in E_1$:
$$dL_a : h \longmapsto L(h)$$
2. Si B désigne une application bilinéaire de $E_1 \times E_2$ dans F, alors B est différentiable sur $E_1 \times E_2$ et on a en tout point $a = (a_1, a_2) \in E_1 \times E_2$:
$$dB_a : h \longmapsto B(h_1, a_2) + B(a_1, h_2)$$

▶ *On revient à chaque fois la différence en $a+h$ et a, et on pensera à reconnaître une partie linéaire et une partie négligeable en $o(\|h\|)$, d'ailleurs en dimension finie on pourra invoquer la caractérisation des applications linéaires continues.*

Exemple 1 Dans cet exercice, $\|.\|$ désigne une norme d'algèbre sur $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire une norme vérifiant pour tout couple (A,B) de matrices de $\mathcal{M}_n(\mathbb{R})$, $\|AB\| \leqslant \|A\|\,\|B\|$.

1. Démontrer que pour toute matrice A de $\mathcal{M}_n(\mathbb{R})$, la série $\displaystyle\sum_{k\geq 0} \frac{1}{k!}A^k$ converge. On notera e^A sa somme.
2. Montrer que l'application $A \longmapsto e^A$ est continue sur $\mathcal{M}_n(\mathbb{R})$.
3. Si $H \in \mathcal{M}_n(\mathbb{R})$ est une matrice non nulle de la boule de centre 0 et de rayon $r>0$, déterminer la limite de $\displaystyle\frac{1}{\|H\|}\sum_{k=2}^{+\infty}\frac{1}{k!}H^k$ lorsque H tend vers 0.
4. En déduire que l'application $A \longmapsto e^A$ est différentiable en la matrice 0. On précisera sa différentielle en 0.

1.2 Opérations sur les fonctions différentiables

Propriété 4 (combinaison linéaire de fonctions différentiables).

Soient E,F deux espaces vectoriels normés de dimension finie et considérons $f,g : U\subset E \longrightarrow F$, $a\in U$. On suppose de plus que f,g sont différentiables en a. Alors, pour tout $\lambda\in\mathbb{R}$, $\lambda f+g$ est encore différentiable en a et :

$$d(\lambda f+g)_a : h \longmapsto \lambda df_a(h) + dg_a(h)$$

▶ *Comme pour les développements limités, on revient à la définition en $o(\|h\|)$.*

Propriété 5 (composée d'une application linéaire et d'une fonction différentiable).

Soit $f : U\subset E\longrightarrow F$, $a\in U$ et considérons $L : F\longrightarrow G$ une application linéaire. On suppose de plus que f est différentiable en a, alors $L\circ f$ est différentiable en a et on peut encore montrer que :

$$d(L\circ f)_a : h\longmapsto L\circ df_a(h)$$

▶ *Comme pour les développements limités, on revient à la définition en $o(\|h\|)$.*

Propriété 6 (composée d'une application bilinéaire et de deux fonctions différentiables).

Soient $f_1 : U\subset E\longrightarrow F_1$, $f_2 : U\subset E\longrightarrow F_2$, $a\in U$ et considérons $B : F_1\times F_2\longrightarrow G$ une application bilinéaire. On suppose de plus que f_1,f_2 sont différentiables en a, alors $B(f_1,f_2)$ est différentiable en a et on peut encore montrer que :

$$dB(f_1,f_2)_a : h\longmapsto B(d(f_1)_a(h), f_2(a)) + B(f_1(a), d(f_2)_a(h))$$

▶ *Comme pour les développements limités, on revient à la définition en $o(\|h\|)$.*

Remarque Ce dernier résultat se généralise et on pourra même décrire la différentielle d'une application multi-linéaire. Autrement dit, en notant M une application n-linéaire et si $f_1,\dots,f_n : U\subset E\longrightarrow F_i$ sont différentiables en a, alors $M(f_1,\dots,f_n)$ est différentiable en a et on a :

$$dM(f_1,\dots,f_n)_a = \sum_{k=1}^{n} M(f_1,\dots,d(f_k)_a,\dots,f_n)$$

Propriété 7 (composée de deux fonctions différentiables).

Soient E,F,G des espaces vectoriels normés de dimension finie, et considérons $f : U\subset E\longrightarrow F$, $g : V\subset F\longrightarrow G$, $a\in U$, $f(U)\subset V$. On suppose de plus que f est différentiable en a et g différentiale en $f(a)$, alors $g\circ f$ est encore différentiable en a et on a :

$$d(g\circ f)_a : h\longmapsto dg_{f(a)}\circ df_a(h)$$

▶ *On revient à la définition en $o(\|h\|)$, par contre on écrira d'abord la différentiabilité de f en a, de g en $f(a)$ avant d'aller chercher $g\circ f(a+h)$: on pourra alors utiliser la linéarité des applications différentielles.*

2 Dérivées partielles et fonctions de classe C^1

2.1 Définition et expression de la différentielle à l'aide des dérivées partielles

Définition Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $a \in U$.

- On dit que f admet une dérivée en a suivant le vecteur $h \in E$ si la fonction vectorielle $\phi : t \longmapsto f(a + th)$ est dérivable en 0. Dans ce cas, on appelle alors **dérivée de f en a suivant le vecteur h** le vecteur limite donné par :

$$D_h f(a) = \lim_{t \to 0} \frac{f(a + th) - f(a)}{t}$$

- Si de plus, $(e_1, \ldots, e_p)$ désigne une base de E, alors sous réserve d'existence, on appelle **j-ème dérivée partielle en a de f** la dérivée suivant le vecteur e_j et ainsi, on note :

$$D_j f(a) \text{ ou } \frac{\partial f}{\partial x_j}(a) = \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t}$$

Remarques

1. Concrètement, ce dernier taux d'accroissement peut, à l'aide de la base donnée, se réécrire :
$$\frac{f(a + te_j) - f(a)}{t} = \frac{f(a_1, \ldots, a_j + t, \ldots, a_n) - f(a_1, \ldots, a_n)}{t}$$
et ainsi, pour obtenir la dérivée partielle suivant e_j, il nous suffira de calculer la dérivée d'une fonction par rapport à sa j-ème variable, en fixant les autres variables. D'ailleurs, on peut aussi obtenir un développement limité suivant la direction e_j :
$$f(a + te_j) = f(a) + t.\frac{\partial f}{\partial x_j}(a) + o(t)$$

2. Si on reprend les notations du chapitre sur les fonctions vectorielles, alors la fonction vectorielle $\phi : t \longmapsto f(a + th)$ peut se décomposer de sorte que, dans F on a :
$$f(a + th) = \sum_{k=1}^{n} f_k(a + th)e'_k \text{ avec } f_k : E \longrightarrow \mathbb{R} \text{ les applications composantes, } (e'_k) \text{ une base de } F$$
En particulier, f admet des dérivées en a suivant le vecteur h si et seulement si les applications composantes f_k admettent des dérivées suivant le vecteur h et on a :
$$D_h f(a) = D_h f_1(a)e'_1 + \ldots + D_h f_n(a)e'_n$$
Par conséquent, déterminer la dérivée partielle d'une fonction reviendra ici à déterminer les dérivées partielles pour chacune des applications composantes.

Exemple 2 Les questions suivantes sont indépendantes.

1. On considère l'application $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = \arctan(xy)$, et on note (e_1, e_2) la base canonique de $\mathbb{R}^2$. Justifier que f possède des dérivées partielles suivant e_1 et e_2, puis donner l'expression de :
$$D_1 f(x, y) = \frac{\partial f}{\partial x}(x, y) \text{ et } D_2 f(x, y) = \frac{\partial f}{\partial y}(x, y)$$

2. On définit l'application $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ par :
$$F(x, y) = (\arctan(xy), e^{x+y})$$
et on note (e_1, e_2) la base canonique de $\mathbb{R}^2$. Justifier que F possède des dérivées partielles suivant e_1 et e_2, puis donner l'expression de :
$$D_1 F(x, y) \text{ et } D_2 F(x, y)$$

Théorème 8 (expression de l'application différentielle).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $a \in U$. On suppose de plus que f est différentiable en a, alors :

1. f admet au point a une dérivée suivant tout vecteur $h \in E$ et on a :
$$D_h f(a) = df_a(h)$$

2. en notant encore $(e_1, \ldots, e_p)$ une base de E, on peut exprimer la différentielle de f en a de sorte que :
$$df_a(h) = \sum_{i=1}^{p} h_i df_a(e_i) = \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a)$$

▶ *Pour le premier point, il suffit de revenir au taux d'accroissement et d'utiliser la différentiabilité. Pour le second point, on décompose h dans la base donnée et on conclut par linéarité.*

Remarque Pour une fonction différentiable en a, on peut donc toujours calculer ses dérivées partielles en a, mais attention la réciproque est fausse. En effet, il existe des fonctions pour lesquelles les dérivées partielles existent, sans pour autant que celles-ci soient différentiables : on essaiera donc de retenir le contre-exemple suivant.

Exemple 3 On considère la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{si } (x,y) \neq 0_{\mathbb{R}^2} \\ 0, & \text{sinon} \end{cases}$$

1. (a) On note $O(0,0)$. Montrer que f possède des dérivées partielles en O, puis préciser les valeurs de $\dfrac{\partial f}{\partial x}(0,0)$ et $\dfrac{\partial f}{\partial y}(0,0)$.
   (b) De la même façon, justifier que f admet des dérivées partielles en tout point $(x,y) \neq 0_{\mathbb{R}^2}$, puis donner l'expression de :
   $$\frac{\partial f}{\partial x}(x,y) \text{ et } \frac{\partial f}{\partial y}(x,y)$$
2. Etablir alors que f n'est pas continue en 0. Que peut-on en déduire ?

Théorème 9 (condition suffisante de différentiabilité).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$, $(e_1, \ldots, e_p)$ une base de E. On suppose de plus que :

$$\begin{cases} f \text{ possède des dérivées partielles en tout point de } U \\ \text{toutes les dérivées partielles sont continues sur } U \end{cases}$$

alors f est nécessairement différentiable sur U et on a encore :

$$df_a(h) = \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a)$$

▶ *On revient à l'étude de la différence $f(a+h) - f(a)$ et on utilise la dérivabilité de f suivant les vecteurs $e_1, \ldots, e_p$.*

Pour cela, fixons $a \in U$ et $h = \sum_{i=1}^{p} h_i e_i$, alors la dérivabilité suivant la direction e_p se réécrit quand $h \to 0$:

$$f(a+h) = f((a + \sum_{i=1}^{p-1} h_i e_i) + h_p e_p) = f(a + \sum_{i=1}^{p-1} h_i e_i) + h_p \frac{\partial f}{\partial x_p}(a + \sum_{i=1}^{p-1} h_i e_i) + \underbrace{o(h_p)}_{=o(\|h\|_\infty)}$$

De plus, les dérivées partielles étant continues en a, on a :

$$\frac{\partial f}{\partial x_p}(a + \sum_{i=1}^{p-1} h_i e_i) \underset{h \to 0}{\longrightarrow} \frac{\partial f}{\partial x_p}(a)$$

et ainsi,

$$f(a+h) = f((a + \sum_{i=1}^{p-1} h_i e_i) + h_p e_p) = f(a + \sum_{i=1}^{p-1} h_i e_i) + h_p \frac{\partial f}{\partial x_p}(a) + o(\|h\|_\infty)$$

En itérant le procédé, on obtient le développement limité :

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

Posons alors $df_a : h \longmapsto \sum_{i=1}^{p} h_i \dfrac{\partial f}{\partial x_i}(a)$, alors on vérifie que df_a est linéaire et ainsi f est différentiable en a et :

$$df_a(h) = \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a)$$

Définition Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$. On dit alors que f est de **classe** C^1 sur U si elle vérifie l'une de ces deux conditions équivalentes :

- f admet des dérivées continues sur U suivant tout vecteur $h \in E$,
- f admet des dérivées partielles continues sur U suivant une base de E.

Remarques

1. Pour vérifier qu'une fonction donnée est différentiable sur un ouvert, on préfèrera ainsi procéder en deux temps plutôt que de revenir à la différence $f(a+h) - f(a)$:

$$\begin{cases} \text{on vérifie que } f \text{ possède des dérivées partielles en chacune de ses variables} \\ \text{on établit alors que celles-ci sont bien continues sur } U \end{cases}$$

et ainsi, la fonction sera C^1 sur U et donc différentiable.

2. De plus, si f est de classe C^1 sur U, on en déduit que f est différentiable sur U et donc, on retrouve qu'elle est aussi continue sur U : on a naturellement $C^1 \Rightarrow C^0$.

Exemple 4 On se place dans $\mathcal{M}_p(\mathbb{R})$ muni de sa base canonique et on considère l'application $det : \mathcal{M}_p(\mathbb{R}) \longrightarrow \mathbb{R}$ défini par :

$$det : M \longmapsto det(M)$$

1. Soit $j \in [\![1,p]\!]$. Justifier que pour tout $M \in \mathcal{M}_p(\mathbb{R})$, $det(M) = \sum_{k=1}^{p} (-1)^{k+j} \Delta_{kj} m_{kj}$, où Δ_{kj} désigne le mineur d'indice (k,j).

2. Déterminer $D_{i,j}det(M)$ la dérivée partielle d'indice (i,j) du déterminant au point M, c'est à dire la dérivée en M suivant la matrice élémentaire E_{ij}.

3. Montrer alors que l'application det est différentiable sur $\mathcal{M}_p(\mathbb{R})$ et que pour tout $H \in \mathcal{M}_p(\mathbb{R})$,

$$d\,det_M(H) = \sum_{i=1}^{p} \sum_{j=1}^{p} H_{ij} (-1)^{i+j} \Delta_{ij} = tr(C(M)^T H)$$

En particulier, on pourra retenir cette **approximation du déterminant** au voisinage de M, puisque par définition de la différentiabilité, on a toujours :

$$det(M+H) = det(M) + tr(C(M)^T H) + o(\|H\|)$$

2.2 Matrice jacobienne et opérations sur les fonctions de classe C^1

Définition Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$. On note $(e_1, \dots, e_p)$ une base de E et $(e'_1, \dots, e'_n)$ une base de F et ainsi, on a pour tout $x = x_1 e_1 + \dots + x_p e_p \in E$,

$$f : \underbrace{(x_1, \dots, x_p)}_{=x} \longmapsto (f_1(x_1, \dots, x_p), \dots, f_n(x_1, \dots, x_p))$$

avec f_k les applications composantes associées. Si de plus, f est différentiable, alors on appelle **matrice jacobienne** de f en x la matrice de l'application différentielle df_x notée Jf_x et définie par :

$$Jf_x = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \dfrac{\partial f_1}{\partial x_2}(x) & \dots & \dfrac{\partial f_1}{\partial x_p}(x) \\ \dfrac{\partial f_2}{\partial x_1}(x) & \dfrac{\partial f_2}{\partial x_2}(x) & \dots & \dfrac{\partial f_2}{\partial x_p}(x) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(x) & \dfrac{\partial f_n}{\partial x_2}(x) & \dots & \dfrac{\partial f_n}{\partial x_p}(x) \end{pmatrix}$$

Et ainsi, dans chaque colonne, on retrouve les composantes du vecteur $D_j f(x) = \dfrac{\partial f}{\partial x_j}(x) = df_a(e_j)$.

Remarque La différentielle en un point étant une application linéaire de E dans F, on récupère alors l'image de tout vecteur h à l'aide de la matrice jacobienne :

$$df_a(h) = Jf_a.h$$

en identifiant ici h et la matrice colonne de ses composantes.

Corollaire 10 (immédiat).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$ qu'on suppose de classe C^1 sur U. On note $(e_1, \dots, e_p)$ une base de E et $(e'_1, \dots, e'_n)$ une base de F, alors f est différentiable en tout point $a \in U$ et on peut écrire plus généralement :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|) \Leftrightarrow f(a+h) = f(a) + Jf_a.h + o(\|h\|)$$

Remarque Autrement dit, toutes les opérations sur les fonctions différentiables se ramènent aussi à des opérations matricielles sur les matrices jacobiennes et on en déduit les résultats suivants.

Propriété 11 (combinaison linéaire de fonctions de classe C^1).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f, g : U \subset E \longrightarrow F$ qu'on suppose de classe C^1 sur U. Alors, pour tout $\lambda \in \mathbb{R}$, $\lambda f + g$ est encore de classe C^1 et pour tout point $a \in U$:

$$d(\lambda f + g)_a(h) = \lambda df_a(h) + dg_a(h) \Leftrightarrow J(\lambda f + g)_a = \lambda Jf_a + Jg_a$$

$$\Leftrightarrow \forall\, j \in [\![1, p]\!],\ \frac{\partial(\lambda f + g)}{\partial x_j}(a) = \lambda \frac{\partial f}{\partial x_j}(a) + \frac{\partial g}{\partial x_j}(a)$$

Propriété 12 (composée de deux fonctions de classe C^1).

Soient E, F, G des espaces vectoriels normés de dimension finie, et considérons $f : U \subset E \longrightarrow F$, $g : V \subset F \longrightarrow G$, $f(U) \subset V$. On suppose de plus que f et g sont de classe C^1, alors $g \circ f$ est encore de classe C^1 et pour tout point $a \in U$:

$$d(g \circ f)_a(h) = dg_{f(a)} \circ df_a(h) \Leftrightarrow J(g \circ f)_a = Jg_{f(a)} \times Jf_a$$

Et de cette dernière égalité, on en déduit un principe de dérivation très utile pour les dérivées partielles :

$$\Leftrightarrow \forall\, j \in [\![1, p]\!],\ \frac{\partial(g \circ f)}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)$$

Remarque Cette dernière propriété est aussi appelée "**règle de la chaîne**", autrement dit pour obtenir la j-ème dérivée partielle de $g \circ f$, on va chercher la dérivée partielle correspondante dans toutes les composantes de $g \circ f$.

Concrètement, si $g : (x, y) \longmapsto g(x, y)$ désigne une application de classe C^1 sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, alors on peut définir le changement de variable :

$$F : (r, \theta) \in \mathbb{R}_+ \times \mathbb{R} \longmapsto g(r\cos(\theta), r\sin(\theta))$$

et ainsi, F est encore de classe C^1 et on a :

- $\dfrac{\partial F}{\partial r}(r, \theta) = \dfrac{\partial g}{\partial x}(r\cos(\theta), r\sin(\theta)).\cos(\theta) + \dfrac{\partial g}{\partial y}(r\cos(\theta), r\sin(\theta)).\sin(\theta)$
- $\dfrac{\partial F}{\partial \theta}(r, \theta) = \dfrac{\partial g}{\partial x}(r\cos(\theta), r\sin(\theta)). - r\sin(\theta) + \dfrac{\partial g}{\partial y}(r\cos(\theta), r\sin(\theta)).r\cos(\theta)$

Exemple 5 On définit deux fonctions :

$$\begin{cases} \text{la fonction } f \text{ de } \mathbb{R}^2 \text{ dans } \mathbb{R} \text{ par } f(x, y) = \sin(x^2 - y^2) \\ \text{la fonction } g \text{ de } \mathbb{R}^2 \text{ dans } \mathbb{R}^2 \text{ par } g(x, y) = (x + y, x - y) \end{cases}$$

1. Justifier que les fonctions f et g sont différentiables en tout vecteur $(x, y) \in \mathbb{R}^2$ et écrire la matrice jacobienne de f puis de g en (x, y).
2. Pour tout $(x, y) \in \mathbb{R}^2$, déterminer l'image d'un vecteur $(u, v) \in \mathbb{R}^2$ par l'application linéaire $d(f \circ g)(x, y)$ en utilisant les deux méthodes suivantes :
   (a) en calculant $f \circ g$;
   (b) en utilisant le produit de deux matrices jacobiennes.

Théorème 13 (caractérisation des fonctions constantes).

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$ qu'on suppose de classe C^1 sur U. On note $(e_1, \ldots, e_p)$ une base de E et on suppose que U est un ouvert convexe de sorte que pour tout $(a, b) \in U^2$ et pour tout $t \in [0, 1]$,

$$a + t(b - a) \in U$$

Dans ce cas, f est constante sur U si et seulement si toutes ses dérivées partielles sont nulles.

▶ *Dans le sens direct, c'est immédiat puisque le taux d'accroissement est nul. Dans le sens réciproque, on montre que* $f(a) = f(b)$ *en utilisant la fonction d'une variable réelle* $\phi : t \longmapsto f(a + t(b - a))$ *: on pourra faire appel à la règle de la chaîne pour obtenir* $\phi'(t) = 0$.

Remarque On retrouve ici une généralisation d'un résultat très pratique sur les fonctions d'une variable réelle, et ainsi on retiendra qu'**une telle fonction est constante sur un ouvert convexe si et seulement si sa différentielle est nulle**. D'ailleurs, on admet que ce résultat peut être étendu à toute partie connexe par arcs.

2.3 Dérivées partielles d'ordre supérieur

Définition Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$. On dit alors que f est de **classe** C^2 sur U si elle vérifie l'une de ces deux conditions équivalentes :

- f admet des dérivées de classe C^1 sur U suivant tout vecteur $h \in E$,
- f admet des dérivées partielles de classe C^1 sur U suivant toute base de E.

De la même façon, on peut généraliser cette définition et ainsi, f est dite de **classe** C^k sur U si elle vérifie l'une de ces deux conditions équivalentes :

- f admet des dérivées de classe C^{k-1} sur U suivant tout vecteur $h \in E$,
- f admet des dérivées partielles de classe C^{k-1} sur U suivant toute base de E.

Remarque En notant $(e_1, \ldots, e_p)$ une base de E, on peut définir les dérivées partielles successives et par exemple, pour tout $(i, j) \in [\![1, p]\!]$:

$$D_i D_j f_a = D_i(D_j f_a) \text{ et } D_i^2 f_a = D_i(D_i f_a)$$

ce qui peut aussi s'écrire :

$$\frac{\partial f}{\partial x_i \partial x_j}(a) = \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(a) \text{ et } \frac{\partial^2 f}{\partial x_i^2}(a) = \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_i})(a)$$

Théorème 14 (de Schwarz)**.**

Soient E, F deux espaces vectoriels normés de dimension finie et considérons $f : U \subset E \longrightarrow F$ qu'on suppose de classe C^2 sur U. On note $(e_1, \ldots, e_p)$ une base de E et dans ce cas, on admet que pour tout $(i, j) \in [\![1, p]\!]$ et pour tout point $a \in U$,

$$D_i D_j f_a = D_j D_i f_a \text{ c'est à dire que les dérivées partielles vérifient : } \frac{\partial f}{\partial x_i \partial x_j}(a) = \frac{\partial f}{\partial x_j \partial x_i}(a)$$

Remarque Ce théorème est admis pour les fonctions de classe C^2 et il se généralise, mais ce dernier résultat n'est pas du tout au programme de MP : on pourra quand même retenir que pour une fonction de classe C^n, l'ordre dans lequel on calcule les dérivées partielles n'importe pas.

Exemple 6 On considère la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2}, \text{ si } (x, y) \neq 0_{\mathbb{R}^2} \\ 0, \text{ sinon} \end{cases}$$

1. (a) On note $O(0, 0)$. Montrer que f possède des dérivées partielles en O, puis préciser les valeurs de $\dfrac{\partial f}{\partial x}(0, 0)$ et $\dfrac{\partial f}{\partial y}(0, 0)$.

   (b) De la même façon, justifier que f admet des dérivées partielles en tout point $(x, y) \neq 0_{\mathbb{R}^2}$, puis donner l'expression de :
   $$\frac{\partial f}{\partial x}(x, y) \text{ et } \frac{\partial f}{\partial y}(x, y)$$

   (c) Etablir que f est de classe C^1 sur $\mathbb{R}^2$.

2. Montrer alors que f n'est pas de classe C^2 sur $\mathbb{R}^2$.

3 Cas particulier des champs scalaires

Dans cette dernière partie, on se place dans le cas particulier des champs scalaires, c'est à dire que les fonctions étudiées seront toujours définies sur un **ouvert** U d'une espace vectoriel normé E de dimension finie, mais elles seront toutes à valeurs dans $\mathbb{R}$. De plus, on munit E d'une structure euclidienne classique.

3.1 Définition du gradient et interprétation géométrique

Définition Soit E un espace vectoriel euclidien et considérons $f : U \subset E \longrightarrow \mathbb{R}$ qu'on suppose différentiable sur U.
Alors, pour tout point $a \in U$, l'application différentielle df_a est une forme linéaire su E, et d'après le théorème de représentation de Riesz, il existe donc un unique vecteur $\nabla f(a) \in E$ tel que :

$$\forall\ h \in E,\ df_a(h) = < \nabla f(a), h >$$

Cet unique vecteur est appelé le gradient de f en a et sera noté $\nabla f(a)$ ou $\vec{grad} f(a)$ de sorte que la différentiabilité de f nous donne :

$$f(a+h) = f(a) + < \nabla f(a), h > + o(\|h\|)$$

Propriété 15 (expression du gradient en base orthonormée)**.**

Soit E un espace vectoriel euclidien et considérons $f : U \subset E \longrightarrow \mathbb{R}$ qu'on suppose différentiable sur U. En notant $(e_1, \ldots, e_p)$ une base orthonormée de E, on a donc pour tout point $a \in U$:

$$\nabla f(a) = \sum_{i=1}^{p} < \nabla f(a), e_i > e_i = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) e_i$$

et ainsi, les composantes du gradient de f en a ne sont rien d'autres que les dérivées partielles de f en a.

► *C'est immédiat : on utilise la décomposition d'un vecteur en BON.*

Remarques

1. On rappelle que l'inégalité de **Cauchy-Schwarz** nous donne pour tout vecteur $(x, y) \in E^2$, $| < x, y > | \leq \|x\|_2 \, \|y\|_2$ et ainsi, on a l'encadrement suivant pour tout $(a, h) \in U^2$:

$$-\|h\|_2 \, \|\nabla f(a)\|_2 \leq \underbrace{< \nabla f(a), h >}_{=df_a(h)} \leq \|h\|_2 \, \|\nabla f(a)\|_2$$

   En fixant $a \in U$, et en considérant la différentielle comme une approximation linéaire de l'accroissement $f(a+h) - f(a)$, on en déduit que :

   - l'accroissement de f est maximale, là où la fonction croît le plus vite, lorsque h est colinéaire de même sens que le gradient en a,
   - l'accroissement de f est minimale, là où la fonction décroît le plus vite, lorsque h est colinéaire de sens contraire avec le gradient en a,
   - l'accroissement de f est nul lorsque h est orthogonal au gradient en a.

   En particulier, le gradient d'un champ scalaire rend compte géométriquement des variations de ce champ et **il est toujours tourné dans la direction des accroissements les plus grands, et orthogonal aux lignes de niveau**.

2. Le calcul du gradient en BON nous permet de récupérer des formules de dérivation pratiques, et on pourra utiliser par exemple les relations suivantes :

$$\nabla(\lambda f + g) = \lambda \nabla(f) + \nabla(g)$$
$$\nabla(fg) = \nabla(f) g + f \nabla(g)$$
$$\nabla(\phi \circ f) = (\phi' \circ f) \nabla(f)$$
$$\nabla(1/f) = -\nabla(f)/f^2$$

3.2 Formule de Taylor-Young et application à la recherche d'extremum

Définition Soit E un espace vectoriel euclidien et considérons $f : U \subset E \longrightarrow \mathbb{R}$,$a \in U$.

- On dit que f possède un **maximum local** en a s'il existe $r > 0$ tel que pour tout $x \in B(a, r)$, $f(x) \leq f(a)$.
- On dit que f possède un **minimum local** en a s'il existe $r > 0$ tel que pour tout $x \in B(a, r)$, $f(x) \geq f(a)$.

Plus généralement, on parle d'**extremum local** en a lorsque f possède un minimum ou un maximum au voisinage de a.

D'ailleurs,

- on dit aussi que f possède un **maximum absolu** ou **maximum global** en a si pour tout $x \in U$, $f(x) \leq f(a)$.
- on dit aussi que f possède un **minimum absolu** ou **minimum global** en a si pour tout $x \in U$, $f(x) \geq f(a)$.

Propriété 16 (condition nécessaire d'extremum).

Soit E un espace vectoriel euclidien et considérons $f : U \subset E \longrightarrow \mathbb{R}$ qu'on suppose différentiable sur U, $a \in U$.
Si de plus, f possède un extremum local en a, alors nécessairement :

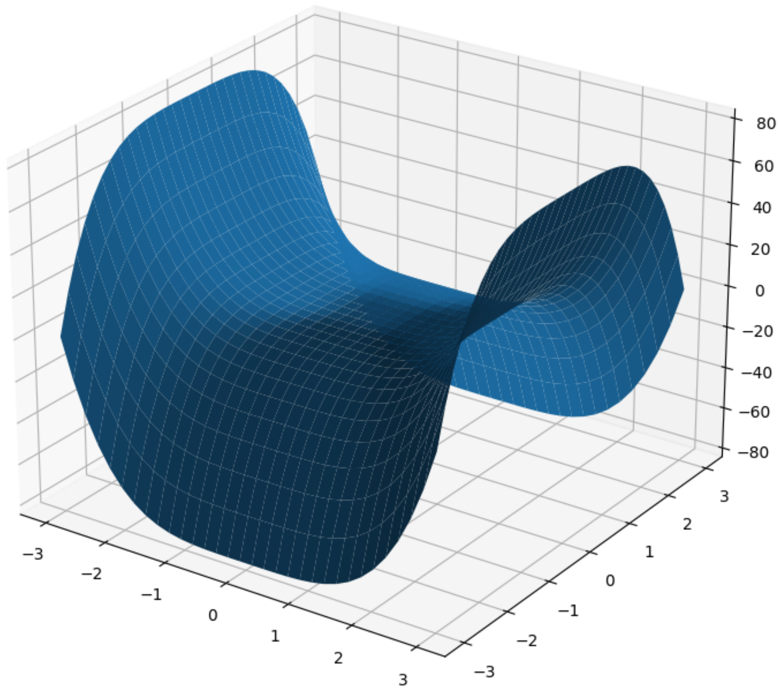
$$df_a = 0$$

On dit que a désigne un **point critique** et en ce point, le gradient associé en a est nul et dans une base donnée, toutes les dérivées partielles en a sont nulles.

▶ *On introduit encore la fonction $t \longmapsto f(a+th)$ et on invoque la CN d'extremum pour les fonctions d'une variable réelle en un point intérieur.*

Remarques

1. Attention, il ne s'agit là que d'une condition nécessaire. Par exemple, on pourra considérer la fonction $f : (x,y) \longmapsto x^4 - y^4$ pour laquelle on montre que $(0,0)$ est le seul point critique et pourtant, au voisinage de $O(0,0)$, la fonction change de signe ne permettant pas d'avoir un extremum local : on dit ici qu'il s'agit d'un **point col** ou d'un **point selle**.



2. Parfois, on pourra quand même garantir en amont l'**existence d'un tel extremum**. Pour cela, on rappelle que :

   - toute fonction continue sur une partie compacte et à valeurs réelles est bornée et atteint ses bornes. Dans ce cas, il faudra alors séparer le travail sur l'intérieur du compact (là où s'applique la condition nécessaire d'extremum) et l'étude sur le bord...
   - si f est définie sur E et si $f(x) \longrightarrow \pm\infty$ quand $\|x\| \to +\infty$, alors on montre que f possède un maximum ou minimum absolu : il suffit de traduire la limite avec $M = f(0_E)$ et d'appliquer le résultat précédent à $B_f(0_E, A)$.

On fera bien attention : les points critiques ne nous donnent pas a priori des extremas... une fois identifiés, il s'agira donc d'étudier le signe de la différence $f(x) - f(a)$.

Exemple 7

1. On considère la fonction f définie sur $\mathbb{R}^2$ par $f(x,y) = xy$. Déterminer les points critiques éventuels et montrer qu'elle n'admet pas d'extremum local.

2. On considère la fonction f définie sur $\mathbb{R}^2$ par $f(x,y) = x^2 + y^2 + 2x - 2y + 4$. Déterminer les points critiques éventuels et montrer qu'elle admet un extremum global.

Théorème 17 (formule de Taylor-Young).

Soit E un espace vectoriel euclidien et considérons $f : U \subset E \longrightarrow \mathbb{R}$, $a \in U$. On note $(e_1, \ldots, e_p)$ une base de E.

1. Si f est de classe C^1 sur U, alors on rappelle que f est différentiable et on a pour tout point $a \in U$:
$$f(a+h) = f(a) + \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|)$$
2. Si f est de classe C^2 sur U, alors on admet que pour tout point $a \in U$:
$$f(a+h) = f(a) + \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p} h_i h_j \frac{\partial^2 f}{\partial x_i x_j} + o(\|h\|^2)$$

Remarque Cette formule de Taylor-Young est admise, mais elle provient de la formule de Taylor avec reste intégral appliquée à la fonction vectorielle $t \longmapsto f(a+th)$ sur $[0,1]$... le calcul est laborieux d'autant qu'il n'est pas toujours simple de justifier rigoureusement que le reste intégral est négligeable devant $\|h\|^2$.

Définition Soit E un espace vectoriel normé de dimension finie et considérons $f : U \subset E \longrightarrow \mathbb{R}$. On note $(e_1, \ldots, e_p)$ une base de E et ainsi, on a pour tout $x = x_1 e_1 + \ldots + x_p e_p \in E$, $f : \underbrace{(x_1, \ldots, x_p)}_{=x} \longmapsto f(x_1, \ldots, x_p)$. Si de plus, f est de classe C^2, alors on définit encore :

- la **matrice jacobienne** appartenant à $\mathcal{M}_{1p}(\mathbb{R})$, et constituée des dérivées partielles premières :
$$Jf_x = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x) & \ldots & \dfrac{\partial f}{\partial x_p}(x) \end{pmatrix}$$
- la **matrice hessienne** appartenant à $\mathcal{M}_p(\mathbb{R})$, et constituée des dérivées partielles secondes par :
$$Hf_x = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \ldots & \dfrac{\partial^2 f}{\partial x_1 \partial x_p}(x) \\ \vdots & \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x) & \vdots \\ \dfrac{\partial^2 f}{\partial x_p \partial x_1}(x) & \ldots & \dfrac{\partial^2 f}{\partial x_p^2}(x) \end{pmatrix}$$

Remarques

1. Bien entendu, si f est de classe C^2 sur U, alors le théorème de Schwarz nous donne :
$$\forall (i,j) \in [\![1,p]\!]^2,\ \frac{\partial f}{\partial x_i \partial x_j}(a) = \frac{\partial f}{\partial x_j \partial x_i}(a)$$
et ainsi, la matrice hessienne est symétrique réelle et ses valeurs propres sont réelles.
2. On peut alors réécrire la **formule de Taylor-Young** à l'ordre 2 de sorte que :
$$f(a+h) = f(a) + Jf_a.h + \frac{1}{2}\underbrace{h^T.Hf_a.h}_{=Q(h)} + o(\|h\|^2)$$
En particulier, si a est un point critique, alors $df_a = 0$ et ainsi, au voisinage de a, le signe de $f(a+h) - f(a)$ dépend directement du signe de l'**expression quadratique** $Q(h)$.
Si on invoque le théorème spectral, alors il existe $P \in O_n(\mathbb{R})$ tel que $Q(h) = h^T.PDP^T.h = (P^T h)^T D\,(P^T h)$.
Et en notant $Y = P^T h \in \mathcal{M}_{n1}(\mathbb{R})$, on a finalement :
$$Q(h) = \sum_{i=1}^{n} \lambda_i y_i^2$$
C'est pour cela que le spectre de la matrice hessienne en un point est important : il peut, dans certains cas, nous donner des informations supplémentaires sur un point critique.

Corollaire 18 (étude locale en un point critique).

Soit E un espace vectoriel normé de dimension finie et considérons $f : U \subset E \longrightarrow \mathbb{R}$ qu'on suppose de classe C^2. On note $(e_1, \ldots, e_p)$ une base de E et a un point critique. Alors,

1. Si la matrice hessienne $Hf_a \in S_n^{++}(\mathbb{R})$, alors au voisinage de a, $f(a+h) - f(a) \geq 0$ et ainsi, f atteint un minimum local en a.

2. Si la matrice hessienne $Hf_a \in S_n^{--}(\mathbb{R})$, alors au voisinage de a, $f(a+h) - f(a) \leq 0$ et ainsi, f atteint un maximum local en a.

3. Si la matrice hessienne possède des valeurs propres de signes dictincts, alors au voisinage de a, $f(a+h) - f(a)$ peut changer de signe et il y a un point selle en a.

Remarques

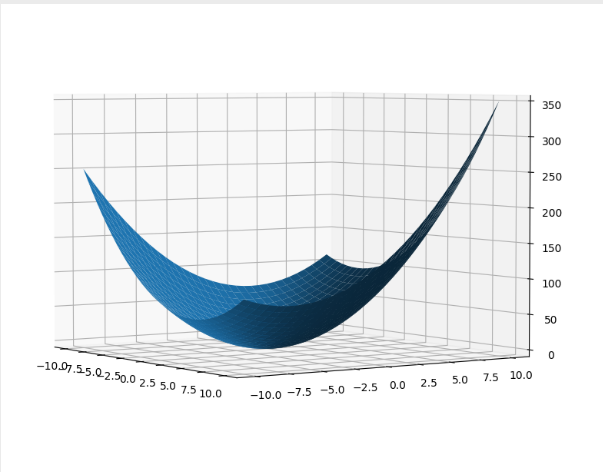
1. Dans tous les autres cas, on fera attention à ne pas aller trop vite : en effet, ces "dévelopements limités vectoriels" nous donnent des informations dans les différentes directions autour du point a, mais si une valeur propre est nulle, on perd l'information dans une direction et il faudrait alors augmenter l'ordre du développement limité pour conclure... On veillera donc à affiner l'étude autour du point considéré.

2. En dimension 2, on ne sera pas obligé de déterminer les valeurs propres pour obtenir des informations sur leurs signes et on pourra se contenter d'étudier la trace et le déterminant de la matrice hessienne.

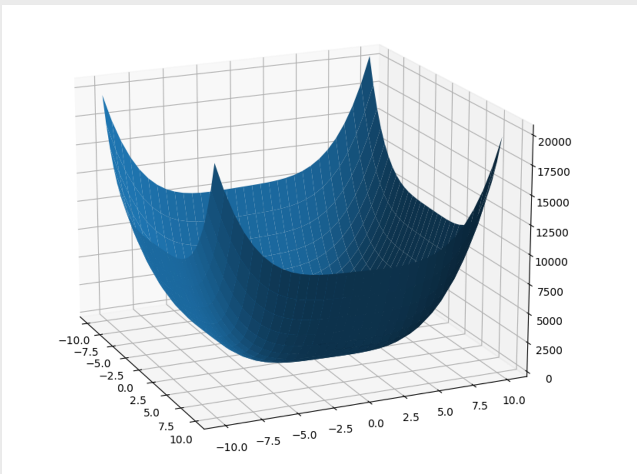
3. De la même façon, on n'a ici que des informations locales et on attendra d'autres arguments pour justifier qu'il s'agit d'un extremum global.

Exemple 8 Pour chacune de ces fonctions, déterminer les extremas locaux, puis étudier s'il s'agit d'extremas globaux :

1. $f : (x, y) \in \mathbb{R}^2 \longmapsto x^2 + xy + y^2 + 2x + 3y$



2. $f : (x, y) \in \mathbb{R}^2 \longmapsto x^4 + y^4 - 4xy$



3.3 Problème des extremas liés

Dans certains cas, on ne travaille pas sur une partie ouverte, mais sur des parties définies par une ou plusieurs équations reliant de fait les variables de notre fonction à plusieurs variables. Il existe alors une méthode très pratique pour identifier les points en lesquels se cachent un extremum : c'est le **théorème des extremas liés**, admis en MP.

Théorème 19 (des extremas liés)**.**

Soit E un espace vectoriel euclidien et considérons $f, g_1, \dots, g_p$ des fonctions qu'on suppose de classe C^1 sur un ouvert $U \subset E$ à valeurs réelles et on définit :

$$X = \{x \in U,\ g_1(x) = \dots = g_p(x) = 0\}$$

Si $f_{|X}$ admet un extremum en a et si les différentielles $dg_{1,a}, \dots, dg_{p,a}$ sont linéairement indépendantes, alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que :

$$df_{|X,a} = \sum_{i=1}^{p} \lambda_i dg_{i,a}$$

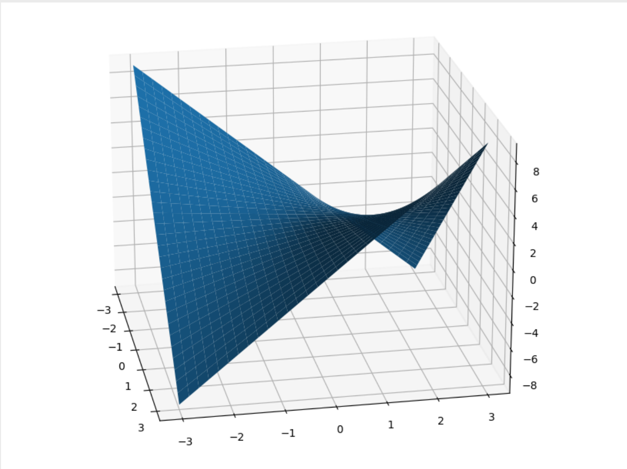
Remarques

1. La partie X n'étant pas ouverte, on ne peut pas invoquer la condition nécessaire d'extremum. Par contre, ce résultat est très efficace et les coefficients $\lambda_1, \dots, \lambda_p$ qui interviennent seront appelés **multiplicateurs de Lagrange**.

2. Cette méthode est une condition nécessaire : elle nous permettra encore une fois d'identifier des points en lesquels il peut y avoir un extremum, mais elle ne garantit pas la présence d'un maximum ou d'un minimum. Généralement, on essaiera de justifier leur existence autrement, avant de déterminer ces extremas.

Exemple 9 Les questions suivantes sont indépendantes.

1. Soit $f : (x, y) \in \mathbb{R}^2 \longmapsto xy$, on étudie ses extremas sur le cercle unité $S = \{(x, y) \in \mathbb{R}^2,\ x^2 + y^2 - 1 = 0\}$.



   (a) Justifier que f possède un maximum et un minimum sur S.
   (b) En utilisant le théorème des extremas liés, retrouver alors les extremas de f sur S.

2. Soit $f : (x, y, z) \in \mathbb{R}^3 \longmapsto x + y + z$, on étudie ses extremas sur $\mathcal{E} = \{(x, y, z) \in \mathbb{R}^3,\ \frac{x^2}{2} + \frac{y^2}{4} + \frac{z^2}{6} - 1 = 0\}$.

   (a) Justifier que $\mathcal{E}$ désigne une partie compacte de $\mathbb{R}^3$. En déduire que f possède un maximum et un minimum sur $\mathcal{E}$.
   (b) En utilisant le théorème des extremas liés, retrouver alors les extremas de f sur $\mathcal{E}$.

3. Soit $f : (x, y, z) \in \mathbb{R}^3 \longmapsto x + y + z$, on étudie ses extremas sur $\mathcal{E}' = \{(x, y, z) \in \mathbb{R}^3,\ -\frac{x^2}{2} + \frac{y^2}{4} + \frac{z^2}{6} - 1 = 0\}$.

   (a) En utilisant le théorème des extremas liés, déterminer les points en lesquels f possède un extremum éventuel sur $\mathcal{E}'$.
   (b) Considérons un point courant $M_t = (t, 2, t\sqrt{3}) \in \mathcal{E}'$, avec $t \in \mathbb{R}$. Montrer que f ne possède pas d'extremas sur $\mathcal{E}'$.