Notes de Calcul Différentiel
2022–2023

Vincent Perrollaz

27 février 2023

Table des matières

0 Rappels

On rappelle dans cette section les résultats et définitions de topologie et algèbre linéaire dont on aura besoin dans la suite du cours.

On développera le cours dans le cadre des espaces vectoriels de dimension finie. Une théorie très similaire peut, de fait, être construite lorsqu'on utilise des espaces de Banach.

Définition 1. Soit E un espace vectoriel réel. Une norme sur E est une application définie sur E à valeurs dans $\mathbb{R}$ (traditionnellement notée $x \mapsto \|x\|$) qui vérifie :

$$\begin{aligned} \forall x \in E, \qquad & x \neq 0 \implies \|x\| > 0, \\ \forall x \in E, \quad \forall \lambda \in \mathbb{R}, \qquad & \|\lambda\, x\| = |\lambda|\, \|x\|, \\ \forall (x,y) \in E^2, \qquad & \|x+y\| \leq \|x\| + \|y\|. \end{aligned} \tag{1}$$

Définition 2. Soient $(E, \|\cdot\|)$ un espace vectoriel normé, x_0 un point de E et V un sous ensemble de E. On dit que V est un voisinage de x_0 lorsque

$$\exists \epsilon > 0, \quad \forall x \in E, \qquad \|x - x_0\| < \epsilon \Rightarrow x \in V.$$

Définition 3. Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

- On dit qu'un sous ensemble U de E est un ouvert lorsqu'il est voisinage de chacun de ses points. Formellement cela revient à exiger

$$\forall x_0 \in U, \quad \exists \epsilon > 0, \quad \forall x \in E, \qquad \|x - x_0\| < \epsilon \Rightarrow x \in V.$$

- Un sous ensemble C de E est dit fermé lorsque $E \setminus C$ est un ouvert.

Proposition 1. Dans un espace vectoriel de dimension finie

1. toutes les normes sont équivalentes,
2. les ensembles fermés et bornés sont compacts,
3. toute application linéaire $L : E \to F$ est continue et même lipschitzienne de constante :

$$\|L\|_F^E := \sup_{\|x\|_E \leq 1} \|L(x)\|_F.$$

Proposition 2. Soit E un espace vectoriel de dimension finie. Soit $\mathfrak{e} = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)$ une base de E.

Alors on a

$$\forall (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d, \quad \exists ! l \in E^*, \quad \forall i \in \{1, \ldots, d\}, \quad l(\mathfrak{e}_i) = \alpha_i. \tag{2}$$

De plus, l'application $(\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d \mapsto l \in E^*$ est un isomorphisme.

Définition 4. Soit E un espace de dimension finie. Soit $\mathfrak{e} = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)$ une base de E.

Il existe une unique famille $\mathfrak{e}^* = (\mathfrak{e}_1^*, \ldots, \mathfrak{e}_d^*)$ de E^* vérifiant

$$\forall\, (i,j) \in \{1, \ldots, d\}^2, \quad \mathfrak{e}_i^*(\mathfrak{e}_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases} \tag{3}$$

De plus, $\mathfrak{e}^*$ est une base de E^* qu'on appelle la duale de $\mathfrak{e}$.

Attention. *Il faut bien réaliser que, contrairement à ce que la notation pourrait le laisser supposer, la forme linéaire $\mathfrak{e}_i^*$ dépend de* toute *la base $\mathfrak{e}$ et* pas seulement *du vecteur $\mathfrak{e}_i$.*

1 Notion de dérivée directionnelle

Définition 5. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- v un vecteur de E,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans F.

On dit que f admet une dérivée directionnelle au point x_0 en la direction v lorsque la limite

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + t\,v) - f(x_0)}{t}, \tag{4}$$

existe. On notera $\mathrm{D}_v f(x_0)$ cette limite.

Remarque 1. *De manière plus précise, $\mathrm{D}_v f(x_0)$ est l'unique vecteur $w \in F$ satisfaisant*

$$\forall \epsilon > 0, \quad \exists r > 0, \quad \forall t \in \mathbb{R}, \qquad 0 < |t| \leq r \Rightarrow \begin{cases} x_0 + t\,v \in U, \\ \left\| \frac{f(x_0 + t\,v) - f(x_0)}{t} - w \right\|_F \leq \epsilon. \end{cases} \tag{5}$$

On voit alors que les espaces étant de dimensions finies, le fait d'être dérivable et la dérivée elle même ne dépendent pas du choix de la norme.

Remarque 2. *On voit que f admet une une dérivée directionnelle en x_0 suivant la direction v si et seulement si l'application γ définie près de $t = 0$ par $\gamma(t) := f(x_0 + tv)$, admet une dérivée classique en 0 et alors $\mathrm{D}_v f(x_0) = \gamma'(0)$.*

Exemple 1. *Si on définit f de $\mathbb{R}^2$ dans $\mathbb{R}$ par*

$$\forall (x, y) \in \mathbb{R}^2, \quad f(x, y) := x^3 + y^4 + 3xy,$$

on vérifie que si $v := (a, b)$ est un vecteur de $\mathbb{R}^2$, f admet une dérivée directionnelle suivant v en tout point (x, y) et

$$\mathrm{D}_v f(x, y) = 3x^2 a + 4y^3 b + 3xb + 3ay.$$

Proposition 3 (Homogénéité par rapport au vecteur)**.** Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- v un vecteur de E,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans F admettant une dérivée directionnelle en x_0 suivant v.

Alors pour tout réel s, f admet une dérivée directionnelle en x_0 suivant sv et

$$\mathrm{D}_{sv} f(x) = s\mathrm{D}_v f(x).$$

Proposition 4. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- v un vecteur de E,
- x_0 un point de E,

- U un ouvert de E contenant x_0,
- f et g des applications définies sur U à valeurs dans F admettant chacune une dérivée directionnelle en x_0 suivant v,
- α et β deux nombres réels.

Alors l'application c définie sur U par

$$\forall x \in U, \quad c(x) := \alpha f(x) + \beta g(x),$$

admet une dérivée directionnelle en x_0 suivant v et

$$\mathrm{D}_v c(x_0) = \alpha \mathrm{D}_v f(x_0) + \beta \mathrm{D}_v g(x_0).$$

Proposition 5. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U et à valeurs dans F, ayant une dérivée directionnelle en x_0 suivant tout vecteur.

Si f admet un extremum local en x_0 alors

$$\forall v \in E, \quad \mathrm{D}_v f(x_0) = 0_F. \tag{6}$$

Exemple 2. *Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par*

$$\forall (x,y) \in \mathbb{R}^2, \qquad f(x,y) := \begin{cases} 0 & si\ x = y = 0, \\ 1 & si\ y = x^2, \\ 0 & sinon. \end{cases}$$

La fonction f a des dérivées directionnelles (nulles) en 0 mais n'y est même pas continue.

De plus, si on définit une fonction $\gamma : \mathbb{R} \to \mathbb{R}^2$ par

$$\forall t \in \mathbb{R}, \quad \gamma(t) := \begin{cases} (0,0) & si\ t = 0, \\ \left(t, t^2\left(1 + \sin\left(\frac{1}{t}\right)\right)\right) & sinon, \end{cases}$$

On peut montrer que γ est dérivable en 0 mais pas $f \circ \gamma$.

Remarque 3. *Considérons une fonction $f : \mathbb{R}^2 \to F$.*

Notons d'abord que la notation usuelle $f(x,y)$ est abusive. En effet, la fonction a un seul argument qui est un élément de $\mathbb{R} \times \mathbb{R}$. Il faudrait donc plutôt écrire $f((x,y))$!

Considérons maintenant $\mathfrak{e}_1 := (1,0)$ et $\mathfrak{e}_2 := (0,1)$ les vecteurs de la base canonique de $\mathbb{R}^2$. On a de fait "$f(x,y)''= f((x,y)) = f(x\mathfrak{e}_1 + y\mathfrak{e}_2)$ pour tout couple (x,y).

On peut alors constater que les dérivées directionnelles $\mathrm{D}_{\mathfrak{e}_1} f$ et $\mathrm{D}_{\mathfrak{e}_2} f$ ne sont autres que les objets $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ vues en première année.

La nouvelle notation est cependant moins sujette à précautions. Par exemple, la valeur de $\frac{\partial f}{\partial x}(x,x)$ est, à priori, ambigüe. S'agit il de

$$\lim_{h \to 0} \frac{f(x+h,x) - f(x,x)}{h}$$

ou de

$$\lim_{h\to 0} \frac{f(x+h,x+h)-f(x,x)}{h}.$$

On verra parfois D_i *au lieu de* D_{e_i} *lorsqu'on est dans* $\mathbb{R}^n$ *et qu'on considère la base canonique. Cela reste un peu dangereux si plusieurs espaces cartésiens de dimensions différentes apparaissent dans le même problème.*

2 Notion de différentielle

Définition 6. Soient
- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- D un sous ensemble de E contenant x_0,
- f une application définie sur D à valeurs dans F.

On dit que f est différentiable en x_0 lorsqu'il existe une application linéaire L de E dans F vérifiant

$$\forall \epsilon > 0, \quad \exists r > 0, \quad \forall x \in D, \qquad \|x - x_0\|_E \leq r \Rightarrow \|f(x) - f(x_0) - L(x - x_0)\|_F \leq \epsilon \|x - x_0\|_E. \tag{7}$$

Remarque 4. *De manière informelle, on peut interpréter intuitivement* (7) *comme signifiant*

$$\frac{\|f(x) - f(x_0) - L(x - x_0)\|_F}{\|x - x_0\|_E} \to 0,$$

ou de manière encore plus abrégée

$$f(x) = f(x_0) + L(x - x_0) + o(\|x - x_0\|).$$

L'analogie avec le cas des fonctions allant de $\mathbb{R}$ dans $\mathbb{R}$ est alors plus visible.

Remarque 5. *Si l'on compare* (7) *et* (5)*, on peut constater que la différence entre les deux notions vient du fait que pour la différentielle la différence entre f et son approximation linéaire est petite uniformément par rapport au choix de la direction.*

Proposition 6. Soient
- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- D un sous ensemble de E contenant x_0,
- f une application définie sur D, à valeurs dans F, différentiable en x_0,
- L_1 et L_2 deux applications linéaires vérifiant (7).

Si D est un voisinage de x_0 on a

$$\forall v \in E, \qquad L_1(v) = L_2(v).$$

Notation. *Soient*
- *$(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,*
- *x_0 un point de E,*
- *U un ouvert de E contenant x_0,*
- *f une application définie sur U, à valeurs dans F et différentiable en x_0.*

L'unique application linéaire satisfaisant (7) *sera notée* $\mathrm{d}f(x_0)$*. On l'appellera la différentielle de f en x_0.*

Remarque 6. *On voit là encore qu'en dimension finie le fait d'être différentiable en un point et la valeur de la différentielle ne dépendent pas des normes choisies.*

Exemple 3. *Si on définit f de $\mathbb{R}^2$ dans $\mathbb{R}$ par*

$$\forall (x,y) \in \mathbb{R}^2, \quad f(x,y) := x^3 + y^4 + 3xy,$$

on vérifie que f est différentiable en tout point et que

$$\forall (x,y) \in \mathbb{R}^2, \quad \forall (a,b) \in \mathbb{R}^2, \qquad \mathrm{d}f((x,y))((a,b)) = 3x^2 a + 4y^3 b + 3xb + 3ay.$$

Proposition 7. On commence par deux exemples d'applications différentiables, élémentaires mais très utiles.

- Soit $f : E \to F$ une application constante. Alors f est différentiable en tout point de E et sa différentielle en chaque point est l'application linéaire nulle. C'est à dire qu'on a
$$\forall x \in E, \quad \forall v \in E, \qquad \mathrm{d}f(x)(v) = 0.$$
- Soit $g : E \to F$ une application linéaire. Elle est différentiable en tout point de E. Et en chaque point sa différentielle est g elle même. C'est à dire qu'on a
$$\forall x \in E, \quad \forall v \in E, \qquad \mathrm{d}g(x)(v) = g(v).$$

Proposition 8. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- D un sous ensemble de E contenant x_0,
- f une application définie sur D à valeurs dans F.

Si f est différentiable en x_0 alors f est continue en x_0.

Proposition 9. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U, à valeurs dans F et différentiable en x_0.

Quelque soit le vecteur v de E, f admet une dérivée directionnelle suivant v en x_0 et on a

$$\mathrm{D}_v f(x_0) = \mathrm{d}f(x_0)(v). \tag{8}$$

Remarque 7. *La réciproque n'est **absolument pas vraie**, comme le montre le contre-exemple suivant. Considérons la fonction $g : \mathbb{R}^2 \to \mathbb{R}$ définie par*

$$\forall (x,y) \in R^2, \qquad g((x,y)) := \begin{cases} \frac{x^2 y}{x^2+y^2} & si\ (x,y) \neq (0,0) \\ 0 & sinon \end{cases}$$

On constate que g admet des dérivées directionnelles en $(0,0)$ suivant toute direction et

$$\forall (a,b) \in \mathbb{R}^2, \quad \mathrm{D}_{(a,b)} g((0,0)) = \begin{cases} \frac{a^2 b}{a^2+b^2} & si\ b \neq 0, \\ 0 & sinon \end{cases}$$

L'application $v \in \mathbb{R}^2 \mapsto \mathrm{D}_v g((0,0))$ n'est pas additive pour autant.

3 Premiers résultats sur la différentielle

Théorème 1. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f et g deux applications définies sur U, à valeurs dans F, différentiables en x_0,
- α et β deux nombres réels.

L'application c définie sur U par

$$\forall x \in U, \quad c(u) := \alpha f(u) + \beta g(u),$$

est différentiable en x_0 et on a

$$\forall v \in E, \qquad \mathrm{d}c(x_0)(v) = \alpha \, \mathrm{d}f(x_0)(v) + \beta \, \mathrm{d}g(x_0)(v). \tag{9}$$

Proposition 10. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- D un sous ensemble de E contenant x_0,
- f une application définie sur D, à valeurs dans F et différentiable en x_0.

L'application f est continue en x_0, c'est à dire que l'on a

$$\forall \epsilon > 0, \quad \exists r > 0, \quad \forall x \in D, \qquad \|x - x_0\|_E \leq r \Rightarrow \|f(x) - f(x_0)\|_F \leq \epsilon.$$

Théorème 2 (Règle de la chaîne)**.** Soient

- $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans F qui soit différentiable en x_0.
- V un ouvert de F contenant $f(x_0)$
- g une application définie sur V à valeurs dans G qui soit différentiable en $f(x_0)$

Alors l'ensemble $W := U \cap f^{-1}(V)$ est un voisinage de x_0 et l'application h définie sur W par

$$\forall x \in W, \quad h(x) = g(f(x)),$$

est différentiable en x_0. Elle vérifie

$$\forall v \in E, \qquad \mathrm{d}h(x_0)(v) = \mathrm{d}g\big(f(x_0)\big)\big(\mathrm{d}f(x_0)(v)\big). \tag{10}$$

Cela revient à dire que l'on a

$$\mathrm{d}\,(g \circ f)\,(x) = \mathrm{d}g(f(x)) \circ \mathrm{d}f(x),$$

ou, de manière informelle, *la différentielle de la composée est la composée des différentielles.*

Remarque 8. *Si $\gamma : \mathbb{R} \to E$ est dérivable et $f : E \to \mathbb{R}$ est différentiable, alors $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ est dérivable et*

$$(f \circ \gamma)'\,(t) = \mathrm{d}f(\gamma(t))(\gamma'(t)).$$

Exemple 4. *On considère* $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, $(G, \|\cdot\|_G)$ *et* $(H, \|\cdot\|_H)$ *des espaces vectoriels de dimension finie.*

- *On se donne* x_0 *un point de* E *et* U *un ouvert de* E *contenant* x_0*. Si* $f : U \to F$ *et* $g : U \to G$ *sont différentiables en* x_0 *alors l'application* $h : U \to F \times G$ *définie par*

$$\forall x \in U, \qquad h(x) := (f(x), g(x)),$$

est différentiable en x_0*. On a de plus*

$$\forall v \in E, \qquad \mathrm{d}h(x_0)(v) = (\mathrm{d}f(x_0)(v), \mathrm{d}g(x_0)(v)).$$

- *Si* $b : F \times G \to H$ *est une application bilinéaire alors elle différentiable en tout point. On a de plus*

$$\forall (z, y) \in F \times G, \quad \forall (u, w) \in F \times G, \qquad \mathrm{d}b((z, y))((u, w)) = b((z, w)) + b((u, y)).$$

- *On peut alors déduire en utilisant le théorème précédent que* $p : x \mapsto b((f(x), g(x)))$ *est différentiable en* x_0*. Et de plus*

$$\forall v \in E, \qquad \mathrm{d}p(x_0)(v) = b((f(x_0), \mathrm{d}g(x_0)(v))) + b((\mathrm{d}f(x_0)(v), g(x_0))).$$

Proposition 11. Soient

- E et F deux espaces vectoriels de dimension finie,
- $\mathfrak{f} = (\mathfrak{f}_1, \ldots, \mathfrak{f}_n)$ une base de F,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans F.

Alors f est différentiable en x_0 si et seulement si pour tout indice $j \in \{1, \ldots, n\}$ l'application $\mathfrak{f}_j^* \circ f$ (définie sur U à valeurs dans $\mathbb{R}$) l'est. Dans ce cas on a alors

$$\forall u \in E, \qquad \mathrm{d}f(x_0)(u) = \sum_{j=1}^{n} \mathrm{d}\left(\mathfrak{f}_j^* \circ f\right)(x_0)(u).$$

Remarque 9. *Le résultat précédent montre que l'on peut toujours se ramener à étudier la différentiabilité d'applications à valeurs dans* $\mathbb{R}$*.*

En particulier si $F = \mathbb{R}^n$ *et* $\mathfrak{f}$ *est la base canonique on a littéralement*

$$\forall x \in U, \quad f(x) = (\mathfrak{f}_1^* \circ f(x), \ldots, \mathfrak{f}_n^* \circ f(x)),$$

c'est à dire que $\mathfrak{f}_j^* \circ f$ *est la* j*-ième composante de* f*.*

4 Accroissements finis

Théorème 3 (Inégalité des accroissements finis 1d)**.** Soient

- $(E, \|\cdot\|_E)$ un espace vectoriel normé,
- a et b des réels tels que $a < b$,
- f une fonction définie sur $[a, b]$ à valeurs dans E dérivable en tout point, c'est à dire qu'on a une fonction que l'on notera $f' : [a, b] \to E$ vérifiant

$$\forall t_0 \in [a, b], \quad \forall \epsilon > 0, \quad \exists r > 0, \quad \forall t_1 \in [a, b],$$
$$|t_1 - t_0| \leq r \implies \|f(t_1) - f(t_0) - f'(t_0)(t_1 - t_0)\|_E \leq \epsilon |t_1 - t_0|.$$

Alors on a

$$\|f(b) - f(a)\|_E \leq |b - a| \sup_{\theta \in [0,1]} \|f'(a + \theta(b - a))\|_E. \tag{11}$$

Corollaire 1. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés de dimension finie,
- U un ouvert de E,
- v un vecteur de E,
- f une application définie sur U, à valeurs dans F, admettant des dérivées directionnelles en tout point suivant la direction v,
- x un point de U,
- θ un nombre réel tel que
$$\forall \sigma \in [0, 1], \quad x + \sigma\theta v \in U.$$

Alors on a

$$\|f(x + \theta v) - f(x)\|_F \leq |\theta| \sup_{\sigma \in [0,1]} \|\mathrm{D}_v f(x + \sigma\theta v)\|_F.$$

Théorème 4 (Inégalités des accroissement finis)**.** Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés de dimension finie,
- U un ouvert de E,
- f une application définie sur U, à valeurs dans F différentiable en tout point de U,
- C un sous ensemble convexe de U, c'est à dire vérifiant

$$\forall (x, y) \in C, \quad \forall \theta \in [0, 1], \qquad \theta y + (1 - \theta)x \in C.$$

Alors on a

$$\forall (x, y) \in C^2, \quad \|f(y) - f(x)\|_F \leq \|y - x\|_E \sup_{\theta \in [0,1]} \|\mathrm{d}f(x + \theta(y - x))\|_F^E.$$

5 Des dérivées directionnelles à la différentielle

On a vu à la fin de la Section 2 que la différentiabilité en un point impliquait l'existence des dérivées directionnelles en ce point mais pas la réciproque. On a par contre le résultat suivant constamment utilisé dans la pratique.

Théorème 5. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- U un ouvert de E,
- $(\mathfrak{e}_1, \ldots, \mathfrak{e}_n)$ une base de E.
- f une application définie sur U à valeurs dans F.

On suppose de plus que

- l'application f admet des dérivées directionnelles suivant tout vecteur de la base et en tout point de U,
- quelque soit l'indice i dans $\{1, \ldots, n\}$, l'application $x \in U \mapsto \mathrm{D}_{\mathfrak{e}_i} f(x)$ est continue.

Alors on en conclut que

1. l'application f est différentiable en tout point de U,
2. l'application $x \in U \mapsto \mathrm{d}f(x) \in \mathcal{L}(E, F)$ est continue,
3. on a l'égalité

$$\forall x \in U, \quad \forall v \in E, \qquad \mathrm{d}f(x)(v) = \sum_{i=1}^{n} \mathrm{D}_{\mathfrak{e}_i} f(x) \mathfrak{e}_i^*(v). \tag{12}$$

avec $\{\mathfrak{e}_1^*, \ldots, \mathfrak{e}_n^*\}$ la base duale associée à $\{\mathfrak{e}_1, \ldots, \mathfrak{e}_n\}$.

Remarque 10. *La réciproque est clairement vraie grâce à la Proposition 9. Dans cette situation, on dira que f est de classe $\mathcal{C}^1$. On notera $f \in \mathcal{C}^1(U, F)$.*

Lemme 1. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels de dimension finie,
- U un ouvert de E,
- v un vecteur de E,
- f une application définie sur U, à valeurs dans F et admettant des dérivées dans la direction v en tout point de U,
- x un point de U,
- θ un nombre réel tel que pour tout réel $\sigma \in [0, 1]$ on a $x + \sigma\theta v \in U$.

Alors on a

$$\|f(x + \theta v) - f(x) - \theta \mathrm{D}_v f(x)\|_F \leq |\theta| \sup_{\sigma \in [0,1]} \|\mathrm{D}_v f(x + \sigma\theta v) - \mathrm{D}_v f(x)\|_F \tag{13}$$

Proposition 12 (Règle de la chaîne)**.** Soient

- E, F et G des espaces vectoriels de dimension finie,
- $\mathfrak{e} = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)$ une base de E,
- $\mathfrak{f} = (\mathfrak{f}_1, \ldots, \mathfrak{f}_n)$ une base de F,
- U un ouvert de E,
- V un ouvert de F,
- f une application définie sur U à valeurs dans V de classe $\mathcal{C}^1$,

- g une application définie sur V à valeurs dans G de classe $\mathcal{C}^1$.

Alors on a la formule de composition suivante

$$\forall x \in U, \quad \forall i \in \{1, \ldots, d\}, \qquad \mathrm{D}_{\mathfrak{e}_i}\left(g \circ f\right)(x) = \sum_{j=1}^{n} \mathfrak{f}_j^* \left(\mathrm{D}_{\mathfrak{e}_i} f(x)\right) \mathrm{D}_{\mathfrak{f}_j} g(f(x)) \tag{14}$$

Exemple 5. *Si $E = \mathbb{R}^3$ et $F = \mathbb{R}^2$ munis de leurs bases canoniques :*

$$\mathfrak{e}_1 = (1,0,0)\,, \quad \mathfrak{e}_2 = (0,1,0)\,, \quad \mathfrak{e}_3 = (0,0,1)\,, \quad \mathfrak{f}_1 = (1,0)\,, \quad \mathfrak{f}_2 = (0,1)\,.$$

Une application $f : E \to F$ peut donc s'écrire $f((x,y,z)) = (f_1((x,y,z)), f_2((x,y,z)))$.
Pour simplifier, on considère $g : F \to \mathbb{R}$.
On considère que l'on notera génériquement (x,y,z) un élément de E et (a,b) un élément de F.
La formule (14) *devient en notation* traditionnelle

$$\begin{aligned}
\frac{\partial (g \circ f)}{\partial x}((x,y,z)) &= \frac{\partial g}{\partial a}(f((x,y,z)))\frac{\partial f_1}{\partial x}((x,y,z)) + \frac{\partial g}{\partial b}(f((x,y,z)))\frac{\partial f_2}{\partial x}((x,y,z)),\\
\frac{\partial (g \circ f)}{\partial y}((x,y,z)) &= \frac{\partial g}{\partial a}(f((x,y,z)))\frac{\partial f_1}{\partial y}((x,y,z)) + \frac{\partial g}{\partial b}(f((x,y,z)))\frac{\partial f_2}{\partial y}((x,y,z)),\\
\frac{\partial (g \circ f)}{\partial z}((x,y,z)) &= \frac{\partial g}{\partial a}(f((x,y,z)))\frac{\partial f_1}{\partial z}((x,y,z)) + \frac{\partial g}{\partial b}(f((x,y,z)))\frac{\partial f_2}{\partial z}((x,y,z)).
\end{aligned}$$

6 Passage en coordonnées vectorielles

Définition 7 (« Vecteurs » lignes et colonnes)**.** Soit E un espace vectoriel de dimension finie. Soit $\mathfrak{e} = (\mathfrak{e}_1, \dots, \mathfrak{e}_d)$ une base de E.

On considère alors les applications $\mathfrak{C}_\mathfrak{e} : E \to \mathcal{M}_{d,1}(\mathbb{R})$ et $\mathfrak{C}_\mathfrak{e}^* : E^* \to \mathcal{M}_{1,d}(\mathbb{R})$ vérifiant :

$$\begin{aligned} \forall x \in E, \qquad & \mathfrak{C}_\mathfrak{e}(x) = (e_i^*(x))_{\substack{1 \le i \le d \\ 1 \le j \le 1}} \\ \forall l \in E^*, \qquad & \mathfrak{C}_\mathfrak{e}^*(l) = (l(e_j))_{\substack{1 \le i \le 1 \\ 1 \le j \le d}} \end{aligned} \tag{15}$$

On constate que

- l'application $\mathfrak{C}_\mathfrak{e}$ envoie $\mathfrak{e}$ sur la base canonique de $\mathcal{M}_{d,1}(\mathbb{R})$,
- l'application $\mathfrak{C}_\mathfrak{e}^*$ envoie $\mathfrak{e}^*$ sur la base canonique de $\mathcal{M}_{1,d}(\mathbb{R})$,
- ces applications sont des isomorphismes d'espaces vectoriels.

De manière informelle, on associe les éléments de E aux *« vecteurs colonnes »* et les formes linéaires sur E aux *« vecteurs lignes »* .

Attention. *On notera que ces applications dépendent de la base. On ne peut donc pas faire les identifications $E \leftrightarrow \mathcal{M}_{d,1}(\mathbb{R})$, $\quad E^* \leftrightarrow \mathcal{M}_{1,d}(\mathbb{R})$ pour deux bases distinctes en même temps.*

Exemple 6. *Soient E l'espace $\mathbb{R}^3$ et $\mathfrak{e}$ la base canonique*

$$\mathfrak{e}_1 = (1,0,0), \quad \mathfrak{e}_2 = (0,1,0), \quad \mathfrak{e}_3 = (0,0,1).$$

On constate alors

$$\forall (x,y,z) \in \mathbb{R}^3, \qquad \mathfrak{C}_\mathfrak{e}((x,y,z)) = \mathfrak{C}_\mathfrak{e}(x\mathfrak{e}_1 + y\mathfrak{e}_2 + z\mathfrak{e}_3) = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$
$$\forall (a,b,c) \in \mathbb{R}^3, \qquad \mathfrak{C}_\mathfrak{e}^*(a\mathfrak{e}_1^* + b\mathfrak{e}_2^* + c\mathfrak{e}_3^*) = \begin{pmatrix} a & b & c \end{pmatrix}.$$

Proposition 13. Soit E un espace vectoriel de dimension finie. Soit $\mathfrak{e}$ une base de E. Alors

$$\forall (x,l) \in E \times E^*, \qquad l(x) = \mathfrak{C}_\mathfrak{e}^*(l) \cdot \mathfrak{C}_\mathfrak{e}(x). \tag{16}$$

Au second membre, $\cdot$ représente la multiplication matricielle et on a fait l'identification *(canonique)* entre $\mathcal{M}_{1,1}(\mathbb{R})$ et $\mathbb{R}$.

Définition 8 (Gradient)**.** Soient E un espace de dimension finie, $\mathfrak{e} = (\mathfrak{e}_1, \dots, \mathfrak{e}_d)$ une base de E, U un ouvert de E et $f : U \to \mathbb{R}$ une application de classe $\mathcal{C}^1$.

Pour tout point x de U, on a $\mathrm{d}f(x) \in E^*$. On appelle gradient de f en x le *« vecteur colonne »* associé que l'on note

$$\nabla f(x) := \mathfrak{C}_\mathfrak{e}^*(\mathrm{d}f(x))^T.$$

En appliquant la Proposition 13, on constate

$$\forall x \in U, \quad \forall u \in E, \qquad \mathrm{d}f(x)(u) = \nabla f(x)^T \cdot \mathfrak{C}_\mathfrak{e}(u) = \langle \nabla f(x) | u \rangle.$$

Le terme de droite est le produit scalaire canonique entre les deux vecteurs colonnes : $\nabla f(x)$ et $\mathfrak{C}_\mathfrak{e}(u)$.

Exemple 7. *En poursuivant l'Exemple 6, on considère U un ouvert de $\mathbb{R}^3$ et $f \in \mathcal{C}^1(U,\mathbb{R})$. On constate alors que*

$$\forall x_0 \in U,\, \forall (a,b,c) \in \mathbb{R}^3, \quad \begin{aligned} \mathrm{d}f(x_0)((a,b,c)) &= a\mathrm{D}_{\mathfrak{e}_1} f(x_0) + b\mathrm{D}_{\mathfrak{e}_2} f(x_0) + c\mathrm{D}_{\mathfrak{e}_3} f(x_0) \\ &= a\frac{\partial f}{\partial x}(x_0) + b\frac{\partial f}{\partial y}(x_0) + c\ \frac{\partial f}{\partial z}(x_0) \\ &= \nabla f(x)^T \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} \end{aligned}.$$

Remarque 11. *Le fait d'avoir considère le gradient de f en x comme le vecteur colonne $\nabla f(x)^T$ peut avoir des conséquences subtiles.*

Cela revient en effet à introduire un isomorphisme $\psi_{\mathfrak{e}} : E \to E^*$ défini par

$$\forall u \in E, \qquad \psi_{\mathfrak{e}}(u) = \sum_{i=1}^{n} e_i^*(u) e_i^*.$$

Cela permet une interprétation géométrique plus simple de $\nabla f(x)$ mais l'isomorphisme $\psi_{\mathfrak{e}}$ est là encore totalement dépendant de la base $\mathfrak{e}$ choisie (i.e. n'est pas canonique).

Pour bien faire comprendre le sens du terme canonique *donnons l'exemple de l'isomorphisme (qui lui est* canonique*) ϕ entre E et E^{**} défini par*

$$\forall u \in E, \quad \forall l \in E^*, \qquad \phi(u)(l) = l(u).$$

On voit qu'aucune base n'intervient dans la définition ce qui se traduit, entre autre, par le fait qu'on a quelque soit la base $\bar{\mathfrak{e}}$ de E la propriété

$$\forall i \in \{1,\dots,d\}, \quad \phi(\bar{\mathfrak{e}}_i) = \bar{\mathfrak{e}}_i,$$

Dans le cas de $\psi_{\mathfrak{e}}$ on a par contre juste la relation

$$\forall i \in \{1,\dots,d\}, \quad \psi_{\mathfrak{e}}(\mathfrak{e}_i) = \mathfrak{e}_i^*,$$

mais à priori pas

$$\forall i \in \{1,\dots,d\}, \quad \psi_{\mathfrak{e}}(\bar{\mathfrak{e}}_i) = \bar{\mathfrak{e}}_i^*.$$

Définition 9. Soient E et F deux espaces vectoriels de dimension finie, $\mathfrak{e} = (\mathfrak{e}_1,\dots,\mathfrak{e}_d)$ une base de E, $\mathfrak{f} = (\mathfrak{f}_1,\dots,\mathfrak{f}_n)$ une base de F.

On introduit alors l'application $\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}$ de $\mathcal{L}(E,F)$ vers $\mathcal{M}_{n,d}(\mathbb{R})$ par

$$\forall L \in \mathcal{L}(E,F), \qquad \mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}(L) := \left(\mathfrak{f}_i^*\left(L\left(\mathfrak{e}_j\right)\right)\right)_{\substack{1\le i\le n \\ 1\le j\le d}}. \tag{17}$$

Proposition 14. Soient E et F deux espaces vectoriels de dimension finie, $\mathfrak{e} = (\mathfrak{e}_1,\dots,\mathfrak{e}_d)$ une base de E, $\mathfrak{f} = (\mathfrak{f}_1,\dots,\mathfrak{f}_n)$ une base de F. L'application $\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}$ est un isomorphisme d'espace vectoriel et de plus

$$\forall L \in \mathcal{L}(E,F), \quad \forall x \in E, \qquad \mathfrak{C}_{\mathfrak{f}}(L(x)) = \mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}(L) \cdot \mathfrak{C}_{\mathfrak{e}}(x). \tag{18}$$

Exemple 8. *On considère $E=\mathbb{R}^3$ et $F=\mathbb{R}^2$ munis de leurs bases canoniques :*

$$\mathfrak{e}_1=(1,0,0)\,,\quad \mathfrak{e}_2=(0,1,0)\,,\quad \mathfrak{e}_3=(0,0,1)\,,\quad \mathfrak{f}_1=(1,0)\,,\quad \mathfrak{f}_2=(0,1)\,.$$

Si L est une application linéaire de E vers F on a

$$\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}(L)=\begin{pmatrix}\mathfrak{f}_1^*\left(L\left(\mathfrak{e}_1\right)\right) & \mathfrak{f}_1^*\left(L\left(\mathfrak{e}_2\right)\right) & \mathfrak{f}_1^*\left(L\left(\mathfrak{e}_3\right)\right)\\ \mathfrak{f}_2^*\left(L\left(\mathfrak{e}_1\right)\right) & \mathfrak{f}_2^*\left(L\left(\mathfrak{e}_2\right)\right) & \mathfrak{f}_2^*\left(L\left(\mathfrak{e}_3\right)\right)\end{pmatrix}\in\mathcal{M}_{2,3}\left(\mathbb{R}\right).$$

Dit autrement si on introduit les réels a_1,a_2,a_3,b_1,b_2,b_3 tels que

$$\forall\,(x,y,z)\in E,\quad L((x,y,z))=(a_1x+a_2y+a_3z,b_1x+b_2y+b_3z),$$

alors

$$\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}(L)=\begin{pmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{pmatrix}\in\mathcal{M}_{2,3}\left(\mathbb{R}\right).$$

Définition 10 (Jacobienne)**.** Soient

- E et F des espaces de dimension finie,
- $\mathfrak{e}=(\mathfrak{e}_1,\ldots,\mathfrak{e}_d)$ une base de E,
- $\mathfrak{f}=(\mathfrak{f}_1,\ldots,\mathfrak{f}_n)$ une base de F,
- U un ouvert de E,
- $f:U\to F$ une application de classe $\mathcal{C}^1$.

Pour tout point x de U, l'application $\mathrm{d}f(x)$ est un élément de $\mathcal{L}(E,F)$. On appelle alors jacobienne de f en x la matrice associée par $\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}$

$$\mathrm{J}f(x):=\mathfrak{C}_{\mathfrak{f}}^{\mathfrak{e}}(\mathrm{d}f(x)).$$

On notera qu'en fait on a

$$\mathrm{J}f(x)=\begin{pmatrix}\mathfrak{f}_1^*\left(\mathrm{D}_{\mathfrak{e}_1}f(x)\right) & \cdots & \mathfrak{f}_1^*\left(\mathrm{D}_{\mathfrak{e}_d}f(x)\right)\\ \vdots & & \vdots\\ \mathfrak{f}_n^*\left(\mathrm{D}_{\mathfrak{e}_1}f(x)\right) & \cdots & \mathfrak{f}_n^*\left(\mathrm{D}_{\mathfrak{e}_d}f(x)\right)\end{pmatrix}\in\mathcal{M}_{n,d}(\mathbb{R}),$$

Finalement, en appliquant la Proposition 13 on constate

$$\forall x\in U,\quad\forall u\in E,\qquad \mathfrak{C}_{\mathfrak{f}}\left(\mathrm{d}f(x)(u)\right)=\mathrm{J}f(x)\cdot\mathfrak{C}_{\mathfrak{e}}(u).$$

Remarque 12. *En relation avec la matrice jacobienne on pourra noter que par linéarité on a*

$$\mathfrak{f}_j^*\left(\mathrm{D}_{\mathfrak{e}_i}f(x)\right)=\mathrm{D}_{\mathfrak{e}_i}\left(\mathfrak{f}_j^*\circ f\right)(x).$$

Proposition 15. Soient

- E, F et G des espaces vectoriels de dimension finie,
- $\mathfrak{e}=(\mathfrak{e}_1,\ldots,\mathfrak{e}_d)$ une base de E,
- $\mathfrak{f}=(\mathfrak{f}_1,\ldots,\mathfrak{f}_n)$ une base de F,
- $\mathfrak{g}=(\mathfrak{g}_1,\ldots,\mathfrak{g}_m)$ une base de G,
- U un ouvert de E,
- V un ouvert de F,
- f une application définie sur U à valeurs dans V de classe $\mathcal{C}^1$,
- g une application définie sur V à valeurs dans G de classe $\mathcal{C}^1$.

La règle de composition des différentielles énoncée dans le Théorème 2 devient alors :

$$\mathrm{J}\left(g\circ f\right)(x)=\mathrm{J}g(f(x))\cdot\mathrm{J}f(x),\tag{19}$$

où de manière informelle « *la jacobienne de la composée est le produit matriciel des jacobiennes* » .

7 Problème intrinsèque : de l'intérêt de choisir ses coordonnées

L'objectif de cette section est d'expliquer rapidement pourquoi on a construit la première partie du cours avec des espaces vectoriels abstraits et des bases quelconque plutôt que de travailler directement avec l'espace cartésien $\mathbb{R}^d$ muni de sa base canonique.

Aspect « philosophique ». L'espace cartésien est d'une certaine façon un peu trop spécial. Il s'agit simultanément

- d'un espace affine,
- d'un espace vectoriel,
- d'un espace euclidien.

Sa base canonique est de la même façon

- un repère affine,
- une base vectorielle,
- une base orthonormée.

De plus, ces points de vue se mélangent. Ainsi lorsque $\mathfrak{e}_1, \dots, \mathfrak{e}_d$ est la base canonique de $\mathbb{R}^d$, on a

$$\forall i \in \{1, \dots, d\}, \quad \forall u \in \mathbb{R}^d, \quad \mathfrak{e}_i^*(u) = \langle \mathfrak{e}_i | u \rangle \,.$$

Il devient alors compliqué de savoir ce que l'on utilise réellement lorsqu'on fait des calculs dans un espace cartésien avec sa base canonique.

Intérêt de choisir ses coordonnées. Lorsque le problème qu'on cherche à résoudre a en fait un sens indépendamment de tout choix de base, on dit qu'il est intrinsèque. On a alors tout intérêt à choisir une base rendant les calculs les plus simples possibles.

Donnons un exemple : soient A, B et C trois points du plan, non alignés. Montrer que les médianes du triangle ABC sont concourantes et se coupent dans une proportion 1/3, 2/3. On va esquisser trois façons d'aborder le problème.

1. Commençons par une approche *« Euclidienne »*. On se référera à la figure 1 page 17 pour suivre la construction.
   On notera $\mathcal{A}(T)$ l'aire d'un triangle T. Considérons B' et C' les milieux des segments AC et AB. Soient I le point d'intersection des droites BB' et CC' et P le point d'intersection de AI avec BC.

   Fait : P est le milieu de BC (et donc que AP est la troisième médiane).
   Les triangles APB et APC partagent la même hauteur issue de A. On en déduit alors que

   $$\frac{BP}{CP} = \frac{\mathcal{A}(ABP)}{\mathcal{A}(ACP)}.$$

   La même remarque permet aussi de voir

   $$\frac{BP}{CP} = \frac{\mathcal{A}(IBP)}{\mathcal{A}(ICP)}.$$

   Une manipulation algébrique élémentaire permet alors d'obtenir

   $$\frac{BP}{CP} = \frac{\mathcal{A}(ABP) - \mathcal{A}(IBP)}{\mathcal{A}(ACP) - \mathcal{A}(ICP)} = \frac{\mathcal{A}(ABI)}{\mathcal{A}(ACI)}.$$

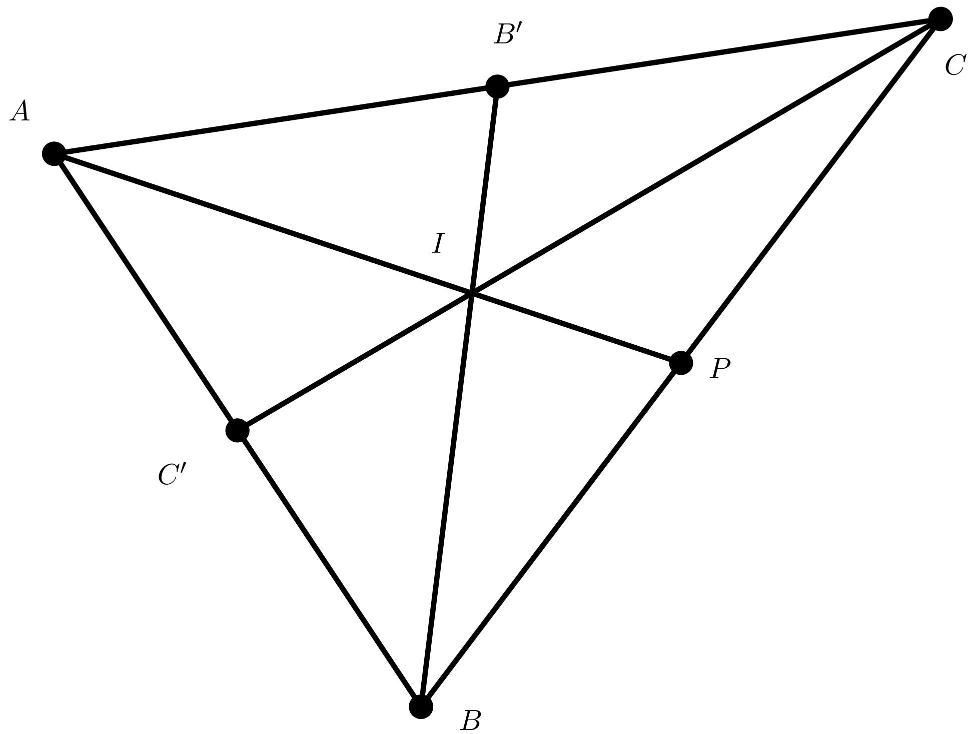

Figure 1 – Construction

Le même raisonnement permet d'obtenir

$$\frac{CB'}{AB'} = \frac{\mathcal{A}(BCI)}{\mathcal{A}(BAI)}, \qquad \frac{AC'}{BC'} = \frac{\mathcal{A}(CAI)}{\mathcal{A}(CBI)}.$$

On en conclut donc facilement que

$$\frac{BP}{CP}\frac{CB'}{AB'}\frac{AC'}{BC'} = 1.$$

Mais par définition des points B' et C' on a $CB' = AB'$ et $AC' = BC'$. On en conclut alors bien $BP = CP$.

Fait : Les médianes se coupent dans un rapport 1/3 2/3.
On voit avec les mêmes idées que

$$\begin{gathered}
\mathcal{A}\left(AIB'\right) = \mathcal{A}\left(CIB'\right), \\
\mathcal{A}\left(AIC'\right) = \mathcal{A}\left(BIC'\right), \\
\mathcal{A}(BIP) = \mathcal{A}(CIP), \\
\mathcal{A}\left(ACC'\right) = \mathcal{A}\left(BCC'\right), \\
\mathcal{A}\left(ABB'\right) = \mathcal{A}\left(CBB'\right), \\
\mathcal{A}(BAP) = \mathcal{A}(CAP).
\end{gathered}$$

De là on voit que

$$\begin{aligned}
\mathcal{A}(BIA) = 2\mathcal{A}(BIP) \quad &\Longrightarrow \quad IA = 2IP,\\
\mathcal{A}(CIB) = 2\mathcal{A}\left(CIB'\right) \quad &\Longrightarrow \quad IB = 2IB',\\
\mathcal{A}(AIC) = 2\mathcal{A}\left(AIC'\right) \quad &\Longrightarrow \quad IC = 2IC'.
\end{aligned} \tag{20}$$

Conclusion. La preuve ne relève pas vraiment d'une méthodologie générale. Elle s'appuie beaucoup sur la figure. En particulier, il faudrait sans doute prouver l'existence des points d'intersection I puis P, le fait que P est entre B et C, etc.

2. On va maintenant *esquisser* une approche « *Cartésienne* » mais naïve.
   On associe au plan l'espace cartésien $\mathbb{R}^2$. On introduit les coordonnées $A = (a, \alpha)$, $B = (b, \beta)$, $C = (c, \gamma)$. Le fait que les points ne soient pas alignés se traduit alors par le fait que
   $$a\beta + b\gamma + c\alpha - a\gamma - b\alpha - c\beta \neq 0.$$
   On peut alors calculer que $A' = \left(\frac{b+c}{2}, \frac{\beta+\gamma}{2}\right)$, $B' = \left(\frac{a+c}{2}, \frac{\alpha+\gamma}{2}\right)$, $C' = \left(\frac{a+b}{2}, \frac{\alpha+\beta}{2}\right)$.
   On devra maintenant déterminer des équations cartésiennes pour les droites AA', BB' et CC'. On calculera ensuite les coordonnées des points d'intersection des couples de droites. On pourra enfin prouver le résultat en montrant que les points coïncident et en calculant les distances.

   Conclusion. Pour un résultat relativement élémentaire, les calculs sont relativement lourds.

3. Finissons par mentionner la « vraie » version « *Cartésienne* ».
   Dire que A' est le milieu de BC revient en fait à avoir la relation vectorielle $\vec{BC} = 2\vec{BA'}$.
   Dire que le rapport est en $2/3$, $1/3$ revient à demander $3\vec{IA'} = \vec{AA'}$. Le problème est donc clairement de nature vectorielle.
   Comme les points A, B, C ne sont pas alignés, on peut choisir comme origine A, puis $\left(\vec{AB}, \vec{AC}\right)$ comme base.
   En passant en coordonnées, on a donc $A = (0, 0)$, $B = (1, 0)$, $C = (0, 1)$.
   On en déduit alors $A' = \left(\frac{1}{2}, \frac{1}{2}\right)$, $B' = \left(0, \frac{1}{2}\right)$, $C' = \left(\frac{1}{2}, 0\right)$.
   Puis on détermine, à vue, des équations de droite
   $$AA' : \; x - y = 0, \qquad BB' : \; x + 2y = 1, \qquad CC' : \; 2x + y = 1.$$
   On calcule alors instantanément
   $$BB' \cap AA' = \left(\frac{1}{3}, \frac{1}{3}\right), \quad CC' \cap AA' = \left(\frac{1}{3}, \frac{1}{3}\right), \quad BB' \cap CC' = \left(\frac{1}{3}, \frac{1}{3}\right).$$
   On appelle $I = \left(\frac{1}{3}, \frac{1}{3}\right)$ l'intersection des trois médianes. Il est alors clair que $2\vec{AA'} = (1, 1) = 3\vec{AI}$ etc.

   Conclusion. On voit qu'en choisissant une base adaptée au problème, on a réduit les calculs au strict minimum.

Du point de vue plus sophistiqué des actions de groupes, on a utilisé le fait que tout triangle non dégénéré de $\mathbb{R}^2$ est dans l'orbite du triangle de sommets $(0, 0)$, $(1, 0)$, $(0, 1)$ sous l'action du groupe des transformations affines.

8 Optimisation : exemples et premières définitions

Exemple 9. *Commençons par des exemples « concrets » de problèmes d'optimisation issus de divers domaines.*

- *Parmi les courbes fermées de longueur fixée L, quelles sont celles qui « englobent » un domaine d'aire maximale A ? On de fait l'inégalité dite* isopérimétrique

$$A \leq \frac{L^2}{4\pi},$$

  l'égalité n'étant satisfaite que pour les cercles.
  Historiquement ce problème était connu comme le « problème de la reine Didon ».
- *Étant donné deux points A et B quelle est la courbe reliant A à B telle que le temps de trajet d'une particule parcourant la courbe sous l'action de la seule gravité soit minimal. La réponse est que cette courbe qu'on qualifie de « brachistochrone » est un arc de cycloïde.*
- *Quel est l'angle que doit choisir un archer pour tirer sa flèche de telle façon que celle-ci retombe le plus loin possible ?*
- *Étant donné un nuage de points $(x_1, y_1), \dots, (x_n, y_n)$ quelle est la droite « représentant le mieux » ce nuage ? On se ramène à minimiser la fonction $J : \mathbb{R}^2 \to \mathbb{R}$ définie par*

$$\forall (a,b) \in \mathbb{R}^2, \quad J((a,b)) := \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

  On parle de « problème des moindre carrés ».
- *Une entreprise possède n entrepôts contenant des quantités $S_1, \dots, S_n$ d'un certains produits. Elle a m clients à qui elle doit fournir des quantités $Q_1, \dots, Q_m$ de ce produit. Le cout d'acheminement entre l'entrepôt j et le client i est de $C_{i,j}$. On cherche le meilleur plan de transport à savoir la matrice $(T_{i,j})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ vérifiant*

$$\begin{cases} \forall (i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}, \quad T_{i,j} \geq 0, \\ \forall i \in \{1,\dots,m\}, \quad \sum_{j=1}^{n} T_{i,j} = Q_i, \\ \forall j \in \{1,\dots,n\}, \quad \sum_{i=1}^{m} T_{i,j} \leq S_j. \end{cases} \tag{21}$$

  telle que le cout total $\sum_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} C_{i,j} T_{i,j}$ soit minimal. Ici $T_{i,j}$ représente évidemment la quantité de produit transportée de l'entrepôt j vers le client i.
- *Un consommateur doit répartir son revenu $R > 0$ entre deux produits. Le premier sera acheté en quantité q_1 pour un prix unitaire p_1, le deuxième en quantité q_2 pour un prix unitaire p_2. On considère que la satisfaction du consommateur pour une répartition donnée sera donnée par la fonction dite d'utilité*

$$\forall (q_1, q_2) \in \mathbb{R}^2, \qquad U((q_1, q_2)) := \alpha q_1 + \beta q_2 - \frac{a q_1^2 + b q_2^2 - 2c q_1 q_2}{2}.$$

  On se ramène donc à trouver la plus grande valeur possible de la fonction U sur l'ensemble $\{(q_1, q_2) \in \mathbb{R}^2 \,:\, q_1 \geq 0,\, q_2 \geq 0,\, p_1 q_1 + p_2 q_2 \leq R\}$.

Définition 11. Soient

- $(E, \|\cdot\|_E)$ un espace vectoriel normé de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans $\mathbb{R}$,
- D un sous ensemble de U contenant x_0.

On dit alors que x_0 est

1. un point de minimum global pour f relativement à D lorsque
$$\forall x \in D, \quad f(x) \geq f(x_0),$$
2. un point de maximum global pour f relativement à D lorsque
$$\forall x \in D, \quad f(x) \leq f(x_0),$$
3. un point de minimum local pour f relativement à D lorsque
$$\exists r > 0, \quad \forall x \in D, \quad \|x - x_0\|_E \leq r \implies f(x) \geq f(x_0),$$
4. un point de maximum local pour f relativement à D lorsque
$$\exists r > 0, \quad \forall x \in D, \quad \|x - x_0\|_E \leq r \implies f(x) \leq f(x_0).$$

Remarque 13. *Complétons ces définitions par les faits suivant.*

- *Lorsque les inégalités sont strictes on parle de minimum/maximum stricts.*
- *Les valeurs correspondant $f(x_0)$ sont les minimum/maximum* **pas** *les points où ceux-ci sont atteints.*
- *On qualifie d'extremum un maximum ou un minimum.*
- *Un extremum global est* **aussi** *local.*
- *Lorsque on ne précise par l'ensemble D cela sous entend qu'il s'agit de l'ensemble de définition de f.*

Notation. *On appelle* $\arg\max_D f$ *l'ensemble des points où un maximum global de f par rapports à D est atteint. On appelle* $\arg\min_D f$ *l'ensemble des points où un minimum global de f par rapports à D est atteint.*

Remarque 14. *Si* $\arg\max_D f \neq \emptyset$ *alors* $\forall x_0, x_1 \in \arg\max_D f$ *on a* $f(x_0) = f(x_1)$*, on note cette unique valeurs de f,* $\max_D f$*. De la même façon si* $\arg\min_D f \neq \emptyset$ *on note la valeur de f sur cet ensemble* $\min_D f$*.*

Exemple 10. *On considère l'application $f : \mathbb{R} \to \mathbb{R}$ définie par*
$$\forall x \in \mathbb{R}, \quad f(x) := x^2.$$
On considère $D_1 := [-1, 1]$ et $D_2 := [1, 2[$, on peut alors prouver que

- $\arg\min_{D_1} f = \{0\}$, $\min_{D_1} f = 0$, $\arg\max_{D_1} f = \{-1, 1\}$ *et* $\max_{D_1} f = 1$.
- $\arg\min_{D_2} f = \{1\}$, $\min_{D_2} f = 1$, *alors que f n'admet pas de maximum global par rapport à D_2 c'est à dire* $\arg\max_{D_2} f = \emptyset$*. Par contre on a* $\sup_{D_2} f = 4$.

On va dans la suite des moyens de déterminer les points de maximum/minimum.

9 Optimisation : information d'ordre 1

Dans cette partie on va fournir des conditions nécessaires sur les extrema. Cela permettra souvent en pratique de restreindre la recherche de ces extrema à un nombre fini de valeurs que l'on peut ensuite évaluer manuellement.

Définition 12. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U différentiable en x_0,

on dit que x_0 est un point critique de f lorsque $\mathrm{d}f(x_0)$ n'est pas surjective de E vers F.

Lorsque $F = \mathbb{R}$ cela revient à demander à ce que $\mathrm{d}f(x_0) = 0$.

Théorème 6. Soient

- $(E, \|\cdot\|_E)$ un espace vectoriel normé de dimension finie,
- x_0 un point de E,
- D un sous ensemble de E contenant x_0 et voisinage de celui-ci,
- f une application définie sur D à valeurs dans $\mathbb{R}$ différentiable en x_0.

Si x_0 est un extremum local pour f alors on a $\mathrm{d}f(x_0) = 0$, c'est à dire que x_0 est un point critique de f.

Remarque 15. *La réciproque est clairement fausse comme le montre le cas de $f : \mathbb{R} \to \mathbb{R}$ définie par*

$$\forall x \in \mathbb{R}, \quad f(x) := x^3,$$

et $x_0 = 0$.

Exemple 11. *On considère la fonction f définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ par*

$$\forall (x,y) \in \mathbb{R}^2, \quad f(x,y) = 3x + 4y - \frac{2x^2 + 3y^2 + 2xy}{2}.$$

Le seul point critique est alors $(1,1)$.

Exemple 12. *On considère d'abord l'ensemble*

$$D := \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) \, : \, \mathrm{Tr}(A) = 0 \ \textit{ou} \ \det(A) = 0\}.$$

On considère ensuite la fonction g définie sur $\mathcal{M}_{2\times 2}(\mathbb{R})$ par

$$\forall M \in \mathcal{M}_{2\times 2}(\mathbb{R}), \quad g(M) := M^2.$$

Alors D est exactement l'ensemble des points critiques de g.

10 Optimisation : Existence

On a, jusqu'à présent, vu des moyens de déterminer des « candidats à être extremum ». Cependant, il faudra commencer par montrer que ces points existent, sous peine d'aboutir à des aberrations comme le montre le paradoxe suivant.

Attention. *On considère la fonction $f \in \mathcal{C}^1(\mathbb{R};\mathbb{R})$ définie par*

$$\forall x \in \mathbb{R}, \quad f(x) := x^2.$$

On considère maintenant $M := \max_{\mathbb{R}} f$ et un point $x_0 \in \mathbb{R}$ tel que $f(x_0) = M$. Comme $\mathbb{R}$ est un ouvert on peut appliquer le Théorème 6 et en déduire que $0 = f'(x_0) = 2x_0$. Mais donc $x_0 = 0$ dont on déduit $M = 0$. Au final, on peut en conclure

$$\forall x \in \mathbb{R}, \quad x^2 = f(x) \leq M = 0.$$

Le problème est évidemment que le maximum considéré n'existe pas !

Théorème 7. Soient

- $(E, \|\cdot\|_E)$ un espace vectoriel normé de dimension finie,
- D un sous ensemble **non vide** de E, à la fois **fermé** et **borné**,
- f une application continue définie sur D à valeurs dans $\mathbb{R}$.

Alors il existe deux points $\underline{x}$ et $\bar{x}$ de D tels que

$$\forall x \in D, \qquad f(\underline{x}) \leq f(x) \leq f(\bar{x}). \tag{22}$$

C'est à dire que

$$\underline{x} \in \arg\min_D f, \qquad \bar{x} \in \arg\max_D f. \tag{23}$$

De manière plus informelle : « f est bornée sur D et atteint ses bornes ».

Remarque 16. *Il convient de noter que le résultat est de nature topologique et non différentielle.*

Remarque 17. *Si E était de dimension infinie, il faudrait demander à ce que D soit compact, pas seulement fermé borné.*

Exemple 13. *Soit $\lambda > 0$ on considère l'ensemble*

$$D_\lambda := \{(x,y) \in \mathbb{R}^2 \,:\, x^2 + y^2 \leq \lambda\}.$$

On se donne également la fonction définie sur $\mathbb{R}^2$ par

$$\forall (x,y) \in \mathbb{R}^2, \quad f(x,y) := x(1-x) + y(1-y) + xy.$$

Alors quelque soit λ, la fonction f admet toujours un maximum global sur D. Mais pour $\lambda < 2$, il n'y a pas de point critique. On pourra identifier les extrema en exercice.

Exemple 14. *Parfois on n'applique pas directement ce théorème mais on s'y ramène ou on adapte la démonstration.*

On peut ainsi chercher à montrer que si f une fonction continue de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\lim_{x\to\pm\infty} f(x) = 0$ alors $x \mapsto |f(x)|$ admet un maximum global sur $D = \mathbb{R}$.

11 Différentiation d'ordre 2

Définition 13. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- U un ouvert de E,
- f une application définie sur U à valeurs dans F de classe $\mathcal{C}^1$.

On dit que f est de classe $\mathcal{C}^2$ lorsque l'application $\mathrm{d}f : U \to \mathcal{L}(E, F)$ est de classe $\mathcal{C}^1$.

Notation. *Si f est de classe $\mathcal{C}^2$ on notera*

$$\forall x \in U, \quad \forall (u, v) \in E^2, \quad \mathrm{d}^2 f(x)((u, v)) := [\mathrm{d}(\mathrm{d}f)(x)(u)]\,(v). \tag{24}$$

On constatera facilement que pour tout $x \in U$, l'application $(u, v) \mapsto \mathrm{d}^2 f(x)((u, v))$ est bilinéaire.

Théorème 8. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- $\mathfrak{e}_1, \ldots \mathfrak{e}_d$ une base de E,
- U un ouvert de E,
- f une application définie sur U à valeurs dans F de classe $\mathcal{C}^1$.

Alors f est de classe $\mathcal{C}^2$ si et seulement si pour tout i entre 1 et d la fonction $x \mapsto \mathrm{D}_{e_i} f(x)$ définie sur U est de classe $\mathcal{C}^1$.

On a alors

$$\forall x \in U, \quad \forall (i, j) \in \{1, \ldots, d\}^2, \quad \mathrm{D}_{e_i}\mathrm{D}_{e_j} f(x) = \mathrm{d}^2 f(x)((e_i, e_j)). \tag{25}$$

Notation. *Soient*

- *$(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,*
- *$\mathfrak{e}_1, \ldots \mathfrak{e}_d$ une base de E,*
- *U un ouvert de E,*
- *f une application définie sur U à valeurs dans F de classe $\mathcal{C}^2$.*

On appelle la ***Hessienne*** *de f relativement à la base $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ l'application définie sur U et à valeurs dans $\mathcal{M}_{d,d}(\mathbb{R})$ définie par*

$$\forall x \in U, \quad \mathrm{H}f(x) := \left(\mathrm{d}^2 f(x)((e_i, e_j))\right)_{1 \le i, j \le d}. \tag{26}$$

Exemple 15. *Si on pose*

$$\forall (x, y, z) \in \mathbb{R}^3, \quad f(x, y, z) = x^2 + xy + y^3 + e^{z-x},$$

les théorèmes usuels garantissent que f est de classe $\mathcal{C}^1$ et que

$$\forall (x, y, z) \in \mathbb{R}^3, \quad \forall (a, b, c) \in \mathbb{R}^3, \quad \mathrm{d}f(x, y, z)(a, b, c) = a(2x + y - e^{z-x}) + b(x + 3y^2) + ce^{z-x}.$$

La jacobienne dans la base canonique est donc donnée par

$$\forall (x, y, z) \in \mathbb{R}^3, \quad \mathrm{J}f(x, y, z) = \begin{pmatrix} 2x + y - e^{z-x} & x + 3y^2 & e^{z-x} \end{pmatrix}.$$

Les théorèmes usuels assurent alors que f est de classe $\mathcal{C}^2$ et que

$$\forall (x, y, z) \in \mathbb{R}^3, \quad \mathrm{H}f(x, y, z) = \begin{pmatrix} 2 + e^{z-x} & 1 & -e^{z-x} \\ 1 & 6y & 0 \\ -e^{z-x} & 0 & e^{z-x} \end{pmatrix}.$$

Théorème 9 (Lemme de Schwarz)**.** Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- $\mathfrak{e}_1, \dots \mathfrak{e}_d$ une base de E,
- U un ouvert de E,
- f une application définie sur U à valeurs dans F de classe $\mathcal{C}^2$.

Quelques soient les indices i et j entre 1 et d on a

$$\forall x \in U, \qquad \mathrm{D}_{e_i}\mathrm{D}_{e_j}f(x) = \mathrm{D}_{e_j}\mathrm{D}_{e_i}f(x). \tag{27}$$

Ce qui est équivalent à dire que pour tout point x de U, l'application bilinéaire $\mathrm{d}^2 f(x)$ est symétrique.

Remarque 18. *Le lemme de Schwarz revient à dire que la Hessienne prend des valeurs symétriques.*

Théorème 10. Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- U un ouvert de E,
- f une application définie sur U à valeurs dans F de classe $\mathcal{C}^2$.

Alors on a

$$\begin{aligned}
&\forall x \in U, \quad \forall \epsilon > 0, \quad \exists r > 0, \quad \forall y \in U,\\
&\|y - x\|_E \leq r \implies \|f(y) - f(x) - \mathrm{d}f(x)(y-x) - \frac{\mathrm{d}^2 f(x)((y-x, y-x))}{2}\|_F \leq \epsilon \|y - x\|_E^2.
\end{aligned} \tag{28}$$

ou de manière informelle

$$f(x+v) = f(x) + \mathrm{d}f(x)(v) + \frac{\mathrm{d}^2 f(x)((v,v))}{2} + o(\|v\|_E^2). \tag{29}$$

Remarque 19. *Dans le cas d'une application f de classe $\mathcal{C}^2$ définie sur un ouvert U d'un espace vectoriel de dimension finie, si on considère une base $\mathfrak{e}_1, \dots, \mathfrak{e}_d$ on a alors*

$$f(x+v) = f(x) + \mathrm{J}f(x) \cdot \mathfrak{C}_{\mathfrak{e}}(v) + \frac{1}{2}\mathfrak{C}_{\mathfrak{e}}(v)^T \cdot \mathrm{H}f(x) \cdot \mathfrak{C}_{\mathfrak{e}}(v) + o(\|v\|^2).$$

12 Optimisation : information d'ordre 2

Théorème 11. Soient

- $(E, \|\cdot\|_E)$ un espace vectoriel normé de dimension finie,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans $\mathbb{R}$ de classe $\mathcal{C}^2$.

alors

1. Si x_0 est un point de maximum local pour f on a
$$\forall v \in E, \qquad \begin{cases} \mathrm{d}f(x)(v) = 0, \\ \mathrm{d}^2 f(x)((v,v)) \leq 0. \end{cases} \tag{30}$$
2. Si x_0 est un point de minimum local pour f on a
$$\forall v \in E, \qquad \begin{cases} \mathrm{d}f(x)(v) = 0, \\ \mathrm{d}^2 f(x)((v,v)) \geq 0. \end{cases} \tag{31}$$
3. Si x_0 satisfait
$$\forall v \in E \setminus \{0\}, \qquad \begin{cases} \mathrm{d}f(x_0)(v) = 0, \\ \mathrm{d}^2 f(x_0)((v,v)) > 0. \end{cases} \tag{32}$$
c'est un minimum local.
4. Si x_0 satisfait
$$\forall v \in E \setminus \{0\}, \qquad \begin{cases} \mathrm{d}f(x_0)(v) = 0, \\ \mathrm{d}^2 f(x_0)((v,v)) < 0. \end{cases} \tag{33}$$
c'est un maximum local.

Lemme 2. Soit E un espace vectoriel normé de dimension finie. On considère $b : E^2 \to \mathbb{R}$ une application bilinéaire symétrique.

Alors b est définie positive c'est à dire que
$$\forall x \in E, \quad x \neq 0 \implies b(x,x) > 0, \tag{34}$$
si et seulement si il existe un réel positif C tel que
$$\forall x \in E, \quad \|x\|^2 \leq C\, b(x,x). \tag{35}$$

Remarque 20. *Si $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ est de classe $\mathcal{C}^2$, on a en un point $(x,y) \in U$ pour (u,v) suffisamment proche de $(0,0)$*
$$f((x+u, y+v)) = f((x,y)) + \frac{\partial f}{\partial x}((x,y))u + \frac{\partial f}{\partial y}((x,y))v + \frac{\partial^2 f}{\partial x^2}((x,y))\frac{u^2}{2} + \frac{\partial^2 f}{\partial y^2}((x,y))\frac{v^2}{2} + \frac{\partial^2 f}{\partial x \partial y}((x,y))uv + o(u^2+v^2). \tag{36}$$

Et on a minimum local en (x,y) lorsque

- *à l'ordre 1 :*
$$\mathrm{Jac}f((x,y)) = \begin{pmatrix} \frac{\partial f}{\partial x}((x,y)) & \frac{\partial f}{\partial y}((x,y)) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix},$$
- *à l'ordre 2 la hessienne*
$$\mathrm{H}f((x,y)) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}((x,y)) & \frac{\partial^2 f}{\partial x \partial y}((x,y)) \\ \frac{\partial^2 f}{\partial x \partial y}((x,y)) & \frac{\partial^2 f}{\partial y^2}((x,y)) \end{pmatrix}$$
est définie positive.
Cela revient à demander que ses valeurs propres soit strictement positives.
De manière équivalente, cela revient également à demander que sa trace et son déterminant sont positifs (mais ***ATTENTION*** *ce dernier résultat est spécifique à la dimension 2).*

Exemple 16. *On considère la fonction f de classe $\mathcal{C}^2$ définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ par*
$$\forall (x,y) \in \mathbb{R}^2, \quad f(x,y) := x^2 + 2y^2 - xy - x - 3y.$$
On constate tout d'abord que
$$\forall (x,y) \in \mathbb{R}^2, \quad \mathrm{J}f(x,y) = \begin{pmatrix} 2x - y - 1 & -x + 4y - 3 \end{pmatrix}.$$
On en déduit donc que les points critiques sont les solutions du système
$$\begin{cases} 2x - y = 1 \\ -x + 4y = 3 \end{cases},$$
et donc que $(1,1)$ est le seul point critique. Mais la Hessienne est donnée par
$$\forall (x,y) \in \mathbb{R}^2, \quad \mathrm{H}f(x,y) = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}.$$
La Hessienne est constante et comme sa trace et son déterminant sont positifs, f admet un minimum local en $(1,1)$.

On pourrait procéder ***alternativement*** *via le développement limité en $(1,1)$ et constater*
$$\begin{aligned} \forall (a,b) \in \mathbb{R}^2, \quad f(1+a, 1+b) &= -2 + a^2 + b^2 - ab + o(a^2 + b^2) \\ &= f(1,1) + \left(a + \frac{b}{2}\right)^2 + \frac{7b^2}{8} + o(a^2 + b^2). \end{aligned}$$
On voit alors que pour (a,b) petit le terme $\left(a + \frac{b}{2}\right)^2 + \frac{7b^2}{8} + o(a^2 + b^2)$ est forcément positif et donc qu'on a un minimum local en $(1,1)$. En fait, le développement limité est en fait une égalité sans o, on en déduit donc que le minimum est global.

13 Inversion locale

Théorème 12 (dit d'inversion locale)**.** Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- x_0 un point de E,
- Ω un ouvert de E contenant x_0,
- f une application définie sur Ω à valeurs dans F de classe $\mathcal{C}^1$.

Alors si $\mathrm{d}f(x_0)$ est un isomorphisme de E vers F, il existe

1. un ouvert U de E tel que $x_0 \in U \subset \Omega$,
2. un ouvert V de F tel que $f(x_0) \in V$,
3. une application p définie sur V à valeurs dans U de classe $\mathcal{C}^1$,

satisfaisant

$$\forall x \in U, \quad p(f(x)) = x, \tag{37}$$
$$\forall y \in V, \quad f(p(y)) = y. \tag{38}$$

Remarque 21. *De manière informelle : si la différentielle en un point est inversible alors l'application non linéaire est localement inversible près de ce point.*

Notons que la régularité de l'application réciproque montre facilement l'implication réciproque.

Remarque 22. *En dimension finie, notons que l'hypothèse implique en particulier que E et F ont la même dimension.*

Remarque 23. *Le théorème reste vrai en dimension infinie à condition de supposer que E et F sont des espaces de Banach et qu'un isomorphisme signifie que l'inverse aussi est une application linéaire continue.*

Démonstration. On va fournir un squelette de la preuve. Un **excellent** exercice consistera à combler les trous.

On commence par définir $y_0 := f(x_0)$, puis $L := \mathrm{d}f(x_0)^{-1} \in \mathcal{L}(F; E)$. On constate tout de suite que $L \neq 0$ et donc que $\|L\|_E^F \neq 0$.

On considère ensuite la fonction g définie sur Ω à valeurs dans E par

$$\forall x \in \Omega, \quad g(x) := x - L(f(x)). \tag{39}$$

Fait 1. La fonction g est de classe $\mathcal{C}^1$ et $\mathrm{d}g(x_0) = 0$.

Grâce au Fait 1 et à la continuité de $x \mapsto \|\mathrm{d}g(x)\|_E^E$, on considère $R > 0$ tel que

$$\forall x \in \bar{\mathrm{B}}(x_0, R), \quad \|\mathrm{d}g(x)\|_E^E \leq \frac{1}{2}. \tag{40}$$

Fait 2. Quelque soit $y \in \bar{\mathrm{B}}(y_0, \frac{R}{2\|L\|_E^F})$ il existe un unique $x \in \bar{\mathrm{B}}(x_0, R)$ tel que $y = f(x)$. On notera dans la suite $p(y)$ ce point x.

On pose maintenant $V := \mathrm{B}(y_0, \frac{R}{2\|L\|_E^F})$ et $U := \mathrm{B}(x_0, R) \cap f^{-1}(V)$.

Fait 3. L'ensemble V est un ouvert de F, U est un ouvert de E et p est une bijection de V sur U telle que

$$\begin{cases} f \circ p = \mathrm{Id}_{|V}, \\ p \circ f_{|U} = \mathrm{Id}_{|U}. \end{cases} \tag{41}$$

Fait 4. L'application p est $2\|L\|_E^F$ lipschitzienne sur V.

Fait 5. Pour tout $x \in U$, l'application linéaire $\mathrm{Id} - \mathrm{d}(g)(x)$ est inversible d'inverse donné par la série (convergeant normalement sur $\bar{B}(x_0, R)$)

$$S(x) := \sum_{k=0}^{\infty} \mathrm{d}g(x)^k \tag{42}$$

et l'application $x \in U \mapsto S(x) \in L(E; E)$ est continue.

Fait 6. L'application p est de classe $\mathcal{C}^1$ sur V. $\square$

Remarque 24. *Les* Faits *1 et 2 établissent l'existence d'une réciproque locale p construite « point par point ».*

Les Faits *suivants permettent d'établir que cette réciproque a de bonnes propriétés, et ce indépendamment de la façon dont elle a été construite.*

Définition 14. Soient
- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés,
- x_0 un point de E,
- U un ouvert de E contenant x_0,
- f une application définie sur U à valeurs dans F de classe $\mathcal{C}^1$.

On dit que f un difféomorphisme sur U lorsque $V := f(U)$ est un ouvert de F, f est une bijection de U sur V et sa réciproque est une application de classe $\mathcal{C}^1$ sur V.

On dit que f est un difféomorphisme local près de x_0 s'il existe un ouvert U' de E contenu dans U et contenant x_0 telle que $f_{|U'}$ est un difféomorphisme.

Remarque 25. *Le théorème d'inversion local signifie qu'il faut et qu'il suffit que la différentielle d'une fonction en un point soit inversible pour que la fonction soit un difféomorphisme local près de ce point.*

Exemple 17. *La fonction C définie sur $\mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 \,:\, x = y\}$ par*

$$C((x,y)) := (-x-y, x\,y)$$

est un difféomorphisme local près de tout point mais n'est pas un difféomorphisme sur son ouvert de définition. C'est par contre un difféomorphisme sur $\{(x,y) \in \mathbb{R}^2 \,:\, x < y\}$. On l'a appelé C à cause des relations coefficients car on a $(X-x)(X-y) = X^2 + (-x-y)X + xy$.

Exemple 18. *La fonction P (pour polaire) définie sur $(0, +\infty) \times \mathbb{R}$ et à valeurs dans $\mathbb{R}^2$ par $P((r,\theta)) := (r\cos(\theta), r\sin(\theta))$ est un difféomorphisme local près de tout point mais n'est pas un difféomorphisme sur son ouvert de définition.*

14 Applications du théorème d'inversion locale

Théorème 13 (Théorème d'inversion globale)**.** Soient

- $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés de dimension finie,
- Ω un ouvert non vide de E,
- f une application définie sur Ω à valeurs dans F de classe $\mathcal{C}^1$,
- U un ouvert non vide inclu dans Ω.

On suppose que

(i) $f_{|U}$ est injective,

(ii) pour tout point x de U, la différentielle $\mathrm{d}f(x)$ est inversible.

Alors $V := f(U)$ est un ouvert et $f_{|U}^{-1} : V \to U$ est de classe $\mathcal{C}^1$.

Exemple 19 (Fonctions implicites en dimensions $2 \to 1$)**.** *Soit f est une application de classe $\mathcal{C}^1$ définie sur U, ouvert non vide de $\mathbb{R}^2$, et à valeurs dans $\mathbb{R}$. On se donne $(\bar{x}, \bar{y})$ dans U et on pose $\bar{z} := f(\bar{x}, \bar{y})$. Si $\frac{\partial f}{\partial x}(\bar{x}, \bar{y}) \neq 0$ alors il existe $\epsilon > 0$ et une fonction $g :]\bar{y} - \epsilon, \bar{y} + \epsilon[\times]\bar{z} - \epsilon, \bar{z} + \epsilon[\to \mathbb{R}$ de classe $\mathcal{C}^1$ tels que*

$$\forall (x, y, z) \in]\bar{x} - \epsilon, \bar{x} + \epsilon[\times]\bar{y} - \epsilon, \bar{y} + \epsilon[\times]\bar{z} - \epsilon, \bar{z} + \epsilon[, \qquad f(x, y) = z \iff x = g(y, z). \tag{43}$$

Exemple 20 (Fonctions implicites en dimensions $3 \to 1$)**.** *Soit f est une application de classe $\mathcal{C}^1$ définie sur U, ouvert non vide de $\mathbb{R}^3$, et à valeurs dans $\mathbb{R}$. On se donne $(\bar{x}, \bar{y}, \bar{z})$ dans U et on pose $\bar{w} := f(\bar{x}, \bar{y}, \bar{z})$. Si $\frac{\partial f}{\partial x}(\bar{x}, \bar{y}, \bar{z}) \neq 0$ alors il existe $\epsilon > 0$ et une fonction $g :]\bar{y} - \epsilon, \bar{y} + \epsilon[\times]\bar{z} - \epsilon, \bar{z} + \epsilon[\times]\bar{w} - \epsilon, \bar{w} + \epsilon[\to \mathbb{R}$ de classe $\mathcal{C}^1$ tels que*

$$\begin{aligned} \forall (x, y, z) \in]\bar{x} - \epsilon, \bar{x} + \epsilon[\times]\bar{y} - \epsilon, \bar{y} + \epsilon[\times]\bar{z} - \epsilon, \bar{z} + \epsilon[\times]\bar{w} - \epsilon, \bar{w} + \epsilon[,& \\ f(x, y, z) = w \iff x = g(y, z, w).& \end{aligned} \tag{44}$$

Exemple 21 (Fonctions implicites en dimensions $3 \to 2$)**.** *Soient f et g deux applications de classe $\mathcal{C}^1$ définie sur U, ouvert non vide de $\mathbb{R}^3$, et à valeurs dans $\mathbb{R}$. On se donne $(\bar{x}, \bar{y}, \bar{z})$ dans U et on pose $\bar{a} := f(\bar{x}, \bar{y}, \bar{z})$ et $\bar{b} := g(\bar{x}, \bar{y}, \bar{z})$. Si la matrice*

$$\begin{pmatrix} \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) & \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \\ \frac{\partial g}{\partial x}(\bar{x}, \bar{y}) & \frac{\partial g}{\partial y}(\bar{x}, \bar{y}) \end{pmatrix}$$

est inversible, alors il existe $\epsilon > 0$ et deux fonctions

$$k, l :]\bar{z} - \epsilon, \bar{z} + \epsilon[\times]\bar{a} - \epsilon, \bar{a} + \epsilon[\times]\bar{b} - \epsilon, \bar{b} + \epsilon[\to \mathbb{R}$$

de classe $\mathcal{C}^1$ tels que

$$\begin{aligned} \forall (x, y, z, a, b) \in]\bar{x} - \epsilon, \bar{x} + \epsilon[\times]\bar{y} - \epsilon, \bar{y} + \epsilon[\times]\bar{z} - \epsilon, \bar{z} + \epsilon[\times]\bar{a} - \epsilon, \bar{a} + \epsilon[\times]\bar{b} - \epsilon, \bar{b} + \epsilon[,& \\ \begin{cases} f(x, y, z) = a \\ g(x, y, z) = b \end{cases} \iff \begin{cases} x = k(z, a, b) \\ y = l(z, a, b) \end{cases}.& \end{aligned} \tag{45}$$