Support de cours de calcul différentiel

Table des matières

1 Fonctions de classe C^1 sur un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ 1
1.1 Graphe, lignes de niveau et applications partielles . . . 1
1.2 Dérivées partielles et gradient . . . 2
1.2.1 Dérivées partielles d'ordre 1 en un point . . . 2
1.2.2 Fonctions dérivées partielles et gradient . . . 2
1.3 Fonctions de classe C^1 . . . 2
1.4 Développement limité à l'ordre 1 d'une fonction de classe C^1 et différentielle en un point . . . 2
1.5 Différentielle . . . 4
1.6 Opérations sur les fonctions de classe C^1 . . . 5
1.7 Dérivées selon un chemin, dérivées directionnelles . . . 5

2 Calcul différentiel à l'ordre 2 6
2.1 Définition des dérivées partielles secondes . . . 6
2.2 Fonctions de classe C^2 . . . 6
2.3 Matrice Hessienne . . . 6
2.4 Développement limité du second ordre . . . 7

3 Extrema 7
3.1 Généralités . . . 7
3.2 Condition nécessaire du premier ordre . . . 8
3.3 Conditions du second ordre . . . 8
3.4 Compléments sur les extrema . . . 9
3.4.1 Passer des extrema locaux aux extrema globaux . . . 9
3.4.2 Une condition suffisante de minimum global . . . 9

4 Le cas des fonctions d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$. 10
4.1 Fonction de classes C^1 d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$, matrice jacobienne. . . . 10
4.2 Le point de vue de la différentielle . . . 10

5 Difféomorphismes et inversion locale 11
5.1 Difféomorphismes . . . 11
5.2 Le théorème d'inversion locale . . . 11
5.3 Applications importantes . . . 12

Cadre et notations

Dans toute la suite.
- n est un entier supérieur (ou égal) à 1 ;
- $\mathbb{R}^n$ est muni de sa structure euclidienne canonique, $\langle\ ,\ \rangle$ désignant le produit scalaire canonique ;
- $\mathcal{O}$ désigne un ouvert non vide de $\mathbb{R}^n$.

1 Fonctions de classe C^1 sur un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$

1.1 Graphe, lignes de niveau et applications partielles

Définition 1 *Soient $A \subset \mathbb{R}^n$ et $f : A \to \mathbb{R}$.*

1. *Le graphe de f, noté G_f, est l'ensemble des points $(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ tels que $f(x_1, \dots, x_n) = x_{n+1}$.*
2. *Si λ est un réel, la ligne de niveau λ de f est :*

$$L_\lambda(f) = \{m \in \mathbb{R}^n \mid f(m) = \lambda\}$$

Remarque 1.1

1. Les lignes de niveaux permettent de mieux comprendre le graphe d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$. Elles peuvent se visualiser pour $n = 2$ sur le graphe

de f comme l'intersection de G_f et de l'hyperplan plan $z = \lambda$ de $\mathbb{R}^3$ mais ce sont des parties de $\mathbb{R}^2$ et non de $\mathbb{R}^3$.

2. Lorsque f est continue, $L_\lambda(f)$ est un fermé de $\mathbb{R}^n$ comme image réciproque d'un fermé de $\mathbb{R}$.

• Applications partielles en un point.
Soient $f : \mathcal{O} \to \mathbb{R}$ et $m_0 = (a_1, \ldots, a_n)$ un point de $\mathcal{O}$. Les applications partielles de f en m_0 sont les n applications à valeurs réelles :

$$\begin{array}{rcl} f_1 : t & \mapsto & f_1(t) = f(t, a_2, \ldots, a_n) \\ f_2 : t & \mapsto & f_2(t) = f(a_1, t, a_3, \ldots, a_n) \\ & \vdots & \\ f_n : t & \mapsto & f_n(t) = f(a_1, \ldots, a_{n-1}, t) \end{array}$$

Chaque f_i est définie au moins sur un intervalle ouvert de $\mathbb{R}$ contenant a_i puisque $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$.

1.2 Dérivées partielles et gradient

1.2.1 Dérivées partielles d'ordre 1 en un point

Définition 2 *Soient $f : \mathcal{O} \to \mathbb{R}$ et $m_0 = (a_1, \ldots, a_n)$ un point de $\mathcal{O}$. Soit $i \in \{1, \ldots, n\}$. Si la fonction partielle f_i en m_0 est dérivable en a_i, la dérivée partielle de f par rapport à la i-ème variable en m_0 est :*

$$\partial_i f(m_0) = f_i'(a_i).$$

On note aussi $\dfrac{\partial f}{\partial x_i}(m_0)$ à la place de $\partial_i f(m_0)$.

1.2.2 Fonctions dérivées partielles et gradient

Lorsque f admet en tout point m de $\mathcal{O}$ une dérivée partielle par rapport à la i-ème variable pour tout $i \in \{1, \ldots, n\}$, on dispose de n fonctions de $\mathcal{O}$ dans $\mathbb{R}$:

$$\partial_i f = \begin{pmatrix} \mathcal{O} & \longrightarrow & \mathbb{R} \\ m & \longmapsto & \partial_i f(m) \end{pmatrix}.$$

Définition 3 *Sous les mêmes conditions on a encore une fonction $\nabla f : \mathcal{O} \to \mathbb{R}^n$ définie, pour $m \in \mathcal{O}$, par $\nabla f(m) = (\partial_1 f(m), \ldots, \partial_n f(m))$. Cette fonction ∇f est appelée le gradient de f.*

Remarque 1.2
L'existence des dérivées partielles en m n'implique pas la continuité en m.

1.3 Fonctions de classe C^1

Définition 4 *Soient $f : \mathcal{O} \to \mathbb{R}$. On dit que f est de classe C^1 sur $\mathcal{O}$ lorsque :*

1. *f admet en tout point m de $\mathcal{O}$ des dérivées partielles $\partial_i f(m)$ pour $i \in \{1, \ldots, n\}$.*
2. *Chacune des fonctions $\partial_i f$ est continue sur $\mathcal{O}$.*

Remarque 1.3
Les applications coordonnées sont de classe C^1.

1.4 Développement limité à l'ordre 1 d'une fonction de classe C^1 et différentielle en un point

Définition 5 *Soient $m_0 \in \mathcal{O}$ et $f : \mathcal{O} \to \mathbb{R}$. On dit que f admet un développement limité à l'ordre 1 en m_0 lorsque qu'il existe une forme linéaire $L : \mathbb{R}^n \to \mathbb{R}$ et une fonction $\varepsilon : \mathcal{O} \to \mathbb{R}$ continue et nulle en m_0 telle que pour tout $m \in \mathcal{O}$ on ait :*

$$f(m) = f(m_0) + L(m - m_0) + \|m - m_0\| \varepsilon(m).$$

Théorème 1
Avec les notations de la définition 5, lorsque f admet en m_0 un développement limité à l'ordre 1, la fonction L est unique et s'appelle la différentielle de f en m_0. On la note df_{m_0} ou $Df(m_0)$ ou $df(m_0)$ ou ...

Démonstration. Soient $m_0 \in \mathcal{O}$ et $f : \mathcal{O} \to \mathbb{R}$. On suppose qu'il existe deux formes linéaires L_1 et L_1 sur $\mathbb{R}^n$ ainsi que deux applications ε_1 et ε_1 de $\mathcal{O}$ dans $\mathbb{R}$, nulles et continues en m_0 telles que pour tout $m \in \mathcal{O}$ on ait :

$$\begin{cases} f(m) = f(m_0) + L_1(m - m_0) + \|m - m_0\| \varepsilon_1(m) & (1) \\ f(m) = f(m_0) + L_2(m - m_0) + \|m - m_0\| \varepsilon_2(m) & (2). \end{cases}$$

Soit $v \in \mathbb{R}^n$. Comme $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$, il existe $\alpha > 0$ tel que pour tout $t \in [-\alpha, \alpha]$ le vecteur $m_0 + tv$ appartienne à $\mathcal{O}$. Pour tout $t \in [-\alpha, \alpha]$, et en prenant $m = m_0 + tv$, il vient en égalant (1) et (2) :

$$t \underbrace{(L_1(v) - L_2(v))}_{=L(v)} = \|tv\| \underbrace{(\varepsilon_2(m_0 + tv) - \varepsilon_1(m_0 + tv))}_{=\varepsilon(t)}.$$

□

Théorème 2
Soit $f : \mathcal{O} \to \mathbb{R}$ une fonction de classe C^1. Alors f admet en tout point m_0 de $\mathcal{O}$ un développement limité à l'ordre 1 et la différentielle de f en m_0 est la forme linéaire :

$$df_{m_0} = \left(\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ v & \longmapsto & \langle \nabla f(m_0), v \rangle \end{array} \right).$$

Démonstration. Par hypothèse f est de classe C^1. Cela signifie que pour tout $m_0 = (a_1, \ldots, a_n) \in \mathbb{R}^n$, et tout $i \in \{1, \ldots, n\}$, l'application :

$$f_i : t \mapsto f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n)$$

est dérivable en a_i, de dérivée $f_i'(a_i) = \partial_i f(m_0)$ et que $m_0 \mapsto \partial_i f(m_0)$ est continue de $\mathbb{R}^n$ dans $\mathbb{R}$.

Soit $m_0 \in \mathbb{R}^n$. On pose pour v dans $\mathbb{R}^n$:

$$\varepsilon(v) = \begin{cases} \dfrac{1}{\|v\|}\big(f(m_0+v) - f(m_0) - \langle \nabla f(m_0), v\rangle\big) & \text{si } \|v\| \neq 0 \\ 0 \quad \text{sinon} \end{cases}$$

Cela définit bien une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}$ et on va démontrer que cette fonction est continue en 0.

Fixons $\alpha > 0$. Par continuité de $\partial_i f$, il existe $\eta > 0$ tel que les relations $m \in \mathbb{R}^n$ et $\|m - m_0\| \leqslant \eta$ impliquent :

$$|\partial_i f(m) - \partial_i(m_0)| \leqslant \alpha \qquad (*)$$

Soit $v \in \mathbb{R}^n$ tel que $\|v\| \leqslant \eta$. On va aller de m_0 à $m_0 + v$ en suivant les axes de coordonnées (voir figure 1.4).

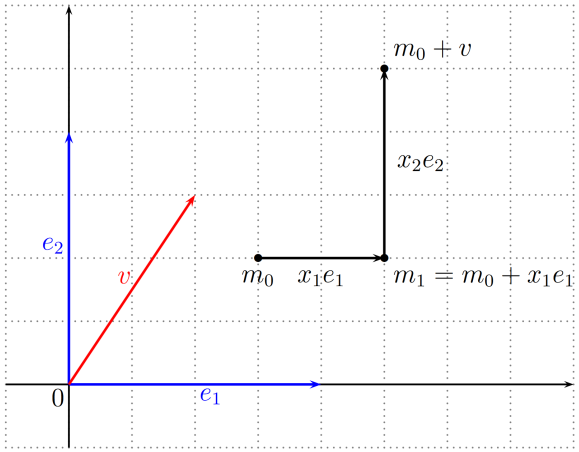


FIGURE 1 – Figure pour n=2

On écrit $v = (x_1, \ldots, x_n)$ et on considère les points :

$$\begin{cases} m_1 &= m_0 + x_1 e_1 \\ m_2 &= m_0 + x_1 e_1 + x_2 e_2 \\ &\vdots \\ m_n &= m_0 + \sum_{i=1}^{n} x_i e_i = m_0 + v \end{cases}$$

Soit $i \in \{1, \ldots, n\}$. On considère la fonction :

$$g_i : \left(\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & f(m_{i-1} + te_i) - f(m_{i-1}) - t\partial_i f(m_0) \end{array} \right).$$

Montrons que la fonction g_i est dérivable sur $\mathbb{R}$.
Il s'agit en fait de montrer que $h_i = t \mapsto f(m_{i-1} + te_i)$ est dérivable sur $\mathbb{R}$. Fixons $t \in \mathbb{R}$. Pour tout h non nul dans $\mathbb{R}$ on a :

$$\frac{h_i(t+h) - h_i(t)}{h} = \frac{1}{h}\big(f(m_{i-1} + te_i + he_i) - f(m_{i-1} + te_i)\big).$$

Ainsi, par hypothèse sur les dérivées partielles qui existent en tout point de $\mathbb{R}^n$, en particulier en $m_i + te_i$, on a :

$$\exists \lim_{h \to 0} \frac{h_i(t+h) - h_i(t)}{h} = \partial_i f(m_i + te_i).$$

On peut donc affirmer que h_i est dérivable en t, donc g_i aussi. De plus, pour tout t réel :

$$g_i'(t) = \partial_i f(m_{i-1} + te_i) - \partial_i f(m_0).$$

Supposons maintenant que $t \in [0, x_i]$. On a alors :

$$\|m_{i-1} + te_i - m_0\|^2 = \|x_1 e_1 + \cdots + x_{i-1}e_{i-1} + te_i\|^2 = x_1^2 + \cdots + x_{i-1}^2 + t^2 \leqslant \|v\|^2 .$$

Il vient ainsi $\|m_{i-1} + te_i - m_0\| \leqslant \eta$ ce qui permet d'obtenir, via $(*)$:

$$|g_i'(t)| \leqslant |\partial_i f(m_{i-1} + te_i) - \partial_i f(m_0)| \leqslant \alpha.$$

Puis les accroissement finis ponctués fournissent t_i entre 0 et x_i tel que :

$$|g_i(x_i)| = |g_i(x_i) - g_i(0)| = |(x_i - 0)g_i'(t_i)| \leqslant |x_i|\,\alpha.$$

Or on a :

$$\begin{aligned} f(m_0 + v) - f(m_0) - \langle \nabla f(m_0), v\rangle &= \sum_{i=1}^{n} \big(f(m_i) - f(m_{i-1}) - x_i \partial_i f(m_0)\big) \\ &= \sum_{i=1}^{n} g_i(x_i). \end{aligned}$$

Ainsi, par inégalité triangulaire :

$$|f(m_0 + v) - f(m_0) - \langle \nabla f(m_0), v\rangle| \leqslant \sum_{i=1}^{n} |g_i(x_i)| \leqslant \alpha \sum_{i=1}^{n} |x_i|$$

Enfin on applique l'inégalité de Cauchy-Schwarz-Bouniakovsky aux vecteurs $w_1 = (|x_1|, \ldots, |x_n|)$ et $w_2 = (1, \ldots, 1)$ pour obtenir :

$$\sum_{i=1}^{n} |x_i| \leqslant \sqrt{n}\,\|v\| .$$

Il vient alors $|\varepsilon(v)| \leqslant \alpha\sqrt{n}$.
En changeant dans $(*)$ α par $\alpha/\sqrt{n}$, on obtient alors la continuité de la fonction ε en 0. □

Remarque 1.4

1. Avec les notation du théorème 2, on a, pour tout $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$:

$$df_{m_0}(v) = \sum_{i=1}^{n} \partial_i f(m_0) v_i.$$

2. **Le gradient, une notion riemannienne**.
Lorsque φ est un produit scalaire sur $\mathbb{R}^n$, le théorème de représentation de Riesz assure l'existence d'un unique vecteur $\nabla_\varphi f(m_0)$ tel que pour tout $v \in \mathbb{R}^n$ on ait :

$$df_{m_0}(v) = \varphi(\nabla_\varphi f(m_0), v).$$

Ce vecteur est le gradient de f en m_0 pour le produit scalaire φ. Lorsque f est de classe C^1, le gradient vu précédemment est donc le gradient de f pour le produit scalaire canonique de $\mathbb{R}^n$.

Corollaire 2.1
Si $f : \mathcal{O} \to \mathbb{R}$ une fonction de classe C^1 alors f est continue sur $\mathcal{O}$.

Démonstration. C'est une conséquence immédiate du théorème 2. □

1.5 Différentielle

Définition 6 *On dit que $f : \mathcal{O} \to \mathbb{R}$ est différentiable lorsqu'elle admet en tout point de $\mathcal{O}$ un développement limité à l'ordre 1. La différentielle de f est alors l'application*

$$df = \begin{pmatrix} \mathcal{O} & \longrightarrow & \mathcal{L}(E, \mathbb{R}) \\ m & \longmapsto & df_m \end{pmatrix}.$$

Remarque 1.5

1. En notant dx_i les fonctions coordonnées de $\mathbb{R}^n$, lorsque $f : \mathcal{O} \to \mathbb{R}$ est différentiable on a :

$$df : m \mapsto \sum_{i=1}^{n} \partial_i f(m) dx_i.$$

2. L'ensemble $\mathcal{D}(\mathcal{O}, \mathbb{R})$ des fonctions différentiables sur $\mathcal{O}$ est un espace vectoriel et $f \mapsto df$ est linéaire.

1.6 Opérations sur les fonctions de classe C^1

Théorème 3

1. *L'ensemble $C^1(\mathcal{O}, \mathbb{R})$ des applications de classes C^1 de $\mathcal{O}$ dans $\mathbb{R}$est un espace vectoriel : toute combinaison linéaire de deux fonctions C^1 sur $\mathcal{O}$ l'est encore.*
2. *Le produit, le quotient (si...) de deux fonctions C^1 est encore une fonction C^1.*
3. *Soient $f : \mathcal{O} \to \mathbb{R}$ et $\varphi : \mathbb{R} \to \mathbb{R}$ deux fonctions de classe C^1. Alors $\varphi \circ f$ est de classe C^1. De plus pour tout m dans $\mathcal{O}$ on a :*

$$\partial_i(\varphi \circ f)(m) = \varphi' \circ f(m) \times \partial_i f(m)$$

Démonstration.

1. Soient f et g dans $C^1(\mathcal{O}, \mathbb{R})$ et $\alpha \in \mathbb{R}$. On pose $h = \alpha f + g$ et on considère un point $m_0 = (a_1, \ldots, a_n)$ de $\mathcal{O}$.
   Soit $i \in \{1, \ldots, n\}$. Notons f_i, g_i et h_i les i-ème applications partielles en m_0 de f, g et h respectivement

   On a $h_1 = \alpha f_i + g_i$, donc h_i est bien dérivable en a_i, comme somme de fonctions dérivables en a_i, avec :

   $$h_i'(a_i) = \alpha \partial_i f(m_0) + \partial_i(m_0).$$

   Ainsi la dérivée partielle de h par rapport à la i-ème variable en m_0 est :

   $$\partial_i h(m_0) = \alpha \partial_i f(m_0) + \partial_i(m_0).$$

   Ceci étant valable quelque soit le choix initial de m_0, on a l'égalité de fonctions

   $$\partial_i fh = \alpha \partial_i f + \partial_i.$$

   Cette dernière fonction est bien continue comme somme de fonctions continues.
   On peut donc conclure que h est de classe c^1 sur $\mathcal{O}$.

2. Soient f et g dans $C^1(\mathcal{O}, \mathbb{R})$ et $\alpha \in \mathbb{R}$. On pose $h = fg$ et on considère un point $m_0 = (a_1, \ldots, a_n)$ de $\mathcal{O}$.
   Soit $i \in \{1, \ldots, n\}$. Notons f_i, g_i et h_i les i-ème applications partielles en m_0 de f, g et h respectivement

   On a $h_1 = f_i g_i$, donc h_i est bien dérivable en a_i, comme somme de fonctions dérivables en a_i, avec :

   $$\begin{aligned} h_i'(a_i) &= f_i'(a_i) g_i(a_i) + f_i(a_i) g_i'(a_i) \\ &= \partial_i f(m_0) \times g(m_0) + f(m_0) \times \partial_i g(m_0). \end{aligned}$$

   On a donc $\partial_i h = \partial_i f \times g + f \times \partial_i g$ et cette dernière application est continue sur $\mathcal{O}$...

3. Soit $m_0 = (a_1, \ldots, a_n) \in \mathcal{O}$. Soit $i \in \{1, \ldots, n\}$. Notons $h = \varphi \circ f$ ainsi que f_i et h_i les i-ème applications partielles en m_0 de f, et h respectivement.
   On a $h_i = \varphi \circ f_i$ donc h_i est dérivable en a_i par composition avec :

   $$h_i'(a_i) = \varphi \circ f_i(a_i) \times f_i'(a_i),$$

   ce qui amène de suite la formule proposée pour $\partial_i(\varphi \circ f)(m_0)$...

□

Corollaire 3.1
Les fonctions polynomiales, les fractions rationnelles sont des fonctions de classe C^1.

1.7 Dérivées selon un chemin, dérivées directionnelles

Théorème 4
Soit $f : \mathcal{O} \to \mathbb{R}$. On considère une courbe paramétrée $c : I \to \mathcal{O}$ de classe C^1 où I est un intervalle de $\mathbb{R}$. Si f est de classe C^1 alors $g = f \circ c$ est de classe C^1 sur I avec :

$$g'(t) = \langle \nabla f(c(t)), c'(t) \rangle .$$

Remarque 1.6

1. Avec les notations du théorème 4, en terme de différentielle, on a donc :

   $$g'(t) = df_{c(t)}(c'(t)).$$

2. **Un cas particulier essentiel : dérivées directionnelles**.
   Soient $m_0 \in \mathcal{O}$, $v \in \mathbb{R}^n \setminus \{0\}$ et $f : \mathcal{O} \to \mathbb{R}$. Puisque $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$, l'application $c : t \mapsto m_0 + tv$ est à valeur dans $\mathcal{O}$ pour $|t|$ petit.
   Il existe donc un intervalle ouvert I de $\mathbb{R}$ centré en 0 tel que l'application :

   $$g_v : t \mapsto f(m_0 + tv)$$

   soit correctement définie. Lorsque f est de classe C^1 alors g_v est de classe C^1 sur I avec :

$$\boxed{g'_v(t) = \langle \nabla f(m_0 + tv), v \rangle .}$$

Le nombre $g'_v(0) = \langle \nabla f(m_0), v \rangle$ est la dérivée directionnelle de f en m_0 dans la direction de v.

3. **Interprétation géométrique du gradient**.
On a donc (notation ci-dessus), $g'_v(0) = \langle \nabla f(m_0), v \rangle$. Ainsi, $|g_v(0)|$ est maximal lorsque v et $\nabla f(m_0)$ sont colinéaires (par CSB) : « la fonction f varie le plus en quittant m_0 dans la direction de $\nabla f(m_0)$ ».

4. Si on note $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$, la dérivée directionnelle de f en m dans la direction de e_i est, sous réserve d'existence, la dérivée partielle $\partial_i f(m)$.

Corollaire 4.1
Lorsque Ω est un ouvert convexe de $\mathbb{R}^n$, alors f est constante sur Ω si et seulement si pour tout $m \in \Omega$ on a $\nabla f(m) = 0$.

Démonstration.
• On suppose que f est constante sur Ω. Soit $m \in \Omega$. Soit $v \in \mathbb{R}^n$. Il existe alors $\alpha > 0$ tel que la fonction

$$g_v : t \mapsto f(m + tv)$$

soit correctement définie sur $]-\alpha, \alpha[$. Cette fonction est de classe C^1 (d'après le théorème 4) avec, pour tout $t \in]-\alpha, \alpha[$:

$$g'_v(t) = \langle \nabla f(m + tv), v \rangle .$$

Comme g_v est une fonction constante, sa dérivée est nulle et en particulier $g'_v(0) = 0$: le vecteur $\nabla f(m)$ appartient donc à l'orthogonal de $\mathbb{R}^n$, il est donc nul.

• Supposons que ∇f soit la fonction nulle sur Ω. Soit $m_0 \in \Omega$. Soit $m \in \omega$. On pose $v = m - m_0$. Par convexité de Ω, pour tout $t \in [0, 1]$ on a $m_0 + tv \in \Omega$, et on considère la fonction $g_v : [0, 1] \to \mathbb{R}$ définie par $g_v(t) = f(m_0 + tv)$.
Cette fonction est de classe C^1 sur $[0, 1]$ de dérivée nulle par hypothèse : elle est constante. Ainsi $g_v(0) = g_v(1)$ i.e. $f(m_0) = f(m)$. □

2 Calcul différentiel à l'ordre 2

2.1 Définition des dérivées partielles secondes

• Soit $f : \mathcal{O} \to \mathbb{R}$ qui est de classe C^1. On a donc n fonctions de plusieurs variables :

$$\partial_i f : m \mapsto \partial_i f(m)$$

Lorsque la fonction $\partial_i f : \mathcal{O} \to \mathbb{R}$ admet en m une dérivée partielle par rapport à la *j-ième* variable on note :

$$\partial^2_{j,i} f(m) = \partial_j \left(\partial_i f \right)(m)$$

On parle de dérivée partielle seconde...

• Lorsque f admet en tout m de $\mathcal{O}$ des dérivées partielles secondes, on dispose de n^2 applications :

$$\partial^2_{i,j} f : \mathcal{O} \to \mathbb{R}$$

2.2 Fonctions de classe C^2

• Une fonction $f : \mathcal{O} \to \mathbb{R}$ est dite de classe C^2 lorsque qu'elle admet des dérivées partielles secondes en tout point de $\mathcal{O}$ et que les fonctions dérivées partielles secondes sont continues sur $\mathcal{O}$.

• **Opérations sur les fonctions de classe C^2**. Les énoncés du théorème 3 et du corolaire 3.1 s'adaptent...On retient notamment que les fonctions polynomiales et les fractions rationnelles sont de classe C^2 (adapter les énoncé).

Théorème 5 (Lemme de Schwarz)
Soit $f : \mathcal{O} \to \mathbb{R}$ de classe C^2. Alors la matrice Hessienne de f en tout point de $\mathcal{O}$ est symétrique, i.e. pour tout i et j dans $\{1, \dots, n\}$ on a l'égalité de fonctions sur $\mathcal{O}$:

$$\partial^2_{i,j} f = \partial^2_{j,i} f$$

2.3 Matrice Hessienne

Définition 7 *Soit $f : \mathcal{O} \to \mathbb{R}$ de classe C^2. Soit m dans $\mathcal{O}$. La matrice hessienne de f en m est :*

$$H_f(m) = \left[\partial^2_{i,j} f(m) \right]_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} .$$

On note aussi $\nabla^2 f(m)$ la matrice hessienne de f en m.

Remarque 2.1

1. Selon le lemme de Schwarz, c'est une matrice symétrique. On notera q_m « la forme quadratique associée » à la matrice $H_f(m)$ i.e. pour tout X dans $\mathcal{M}_{n,1}(\mathbb{R})$:
$$q_m(X) = X^\top H_f(m) X.$$

2. **Dérivées directionnelles secondes.** Soit $f : \mathcal{O} \to \mathbb{R}$ de classe C^2. Soient m_0 dans $\mathcal{O}$ et v dans $\mathbb{R}^n$.
Il existe donc un intervalle ouvert I de $\mathbb{R}$ centré en 0 tel que l'application :
$$g_v : t \mapsto f(m_0 + tv)$$
soit correctement définie.
Alors g_v est de classe C^2 sur I et $g''(t) = X^T H_f(m_0 + tv)X$ où X est la colonne des coordonnées de v dans la base canonique.

   Démontrons ce dernier résultat. On sait que g_v est de classe C^1 avec :
$$g'_v(t) = \langle \nabla f(m_0 + tv), v\rangle = \sum_{i=1}^{n} v_i \underbrace{\partial_i f(m0 + tv)}_{=h_i(t)}.$$

   Chaque fonction h_i est de classe C^1 avec :
$$h'_i(t) = \langle \nabla \partial_i f(m_0 + tv), v\rangle = \sum_{j=1}^{n} v_j \partial^2_{j,i} f(m_0 + tv).$$

   Par somme g'_v est donc C^1 avec : $g'_v(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j \partial^2_{j,i} f(m_0 + tv)$.

   La dérivée directionnelle seconde de f en m_0 dans la direction de v est alors :
$$g''_v(0) = X^T H_f(m_0) X.$$

2.4 Développement limité du second ordre

Théorème 6
Soit $f : \mathcal{O} \to \mathbb{R}$ de classe C^2. Alors f admet un développement limité à l'ordre 2 en tout point m_0 de $\mathcal{O}$ i.e. il existe une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}$ nulle et continue en 0 telle que pour tout v vérifiant $m_0 + v \in \mathcal{O}$:
$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v\rangle + \frac{1}{2} q_{m_0}(v) + \|v\|^2 \varepsilon(v)$$

Autre expression. Si on confond le gradient avec sa ligne des coordonnées dans la base canonique on peut réécrire le DL à l'ordre 2 sous la forme :
$$f(m_0 + v) = f(m_0) + \nabla f(m_0) X + \frac{1}{2} X^T H_f(m_0) X + \mathrm{o}(\|v\|^2),$$
où X est la colonne des coordonnées de v dans la base canonique de $\mathbb{R}^n$.

Démonstration. Soit $m_0 \in \mathcal{O}$. Il existe $\alpha > 0$ tel que $B(m_0, \alpha) \subset \mathcal{O}$. Pour $m \in B(m_0, \alpha)$, qui est convexe, on pose $v = m - m_0$. La formule de Taylor-Lagrange entre 0 et 1 donne pour la fonction $g_v : t \mapsto f(m_0 + tv_m)$, qui est de classe C^2 sur $[0, 1]$, l'existence de $\theta_v \in [0, 1]$ tel que :
$$\underbrace{g_v(1)}_{f(m)} = \underbrace{g_v(0)}_{f(m_0)} + \underbrace{g'_v(0)}_{\langle \nabla f(m_0), v\rangle} + \frac{1}{2} g''_v(\theta_v).$$

Il reste à montrer que $g''_v(\theta_v) \underset{v\to 0}{=} q_{m_0}(v) + \mathrm{o}(\|v\|^2)$.
Or pour tout $t \in [0, 1]$ on a : $g''_v(t) = q_{m_0+tv}(v) = X^\top \nabla^2 f(m_0 + tv) X$ où X est la colonne de v dans la base canonique de $\mathbb{R}^n$.
Mais, en notant $\|\ \|_*$ une norme sur $\mathcal{M}_n(\mathbb{R})$, par Cauchy-Schwarz et continuité de l'application $X \mapsto MX$ sur $\mathcal{M}_{n,1}(\mathbb{R})$ (où M est une matrice), on obtient :
$$\begin{aligned} |g''_v(\theta_v) - g''_v(0)| &= \left|X^\top \left(\nabla^2 f(m_0 + \theta_v v) - \nabla^2 f(m_0)\right) X\right| \\ &\leqslant \left\|\nabla^2 f(m_0 + \theta_v v) - \nabla^2 f(m_0)\right\|_* \|v\|^2 \end{aligned}$$

Fixons $\varepsilon > 0$. Comme la fonction $m \mapsto \nabla^2(m)$ est continue sur $\mathcal{O}$, donc en m_0, il existe $\eta \in]0, \alpha[$ tel que si $m \in B(m_0, \eta)$ on ait :
$$\left\|\nabla^2 f(m_0 + \theta_v v) - \nabla^2 f(m_0)\right\|_* \leqslant \varepsilon.$$

Si donc $\|v\| \leqslant \eta$ il vient $|g''_v(\theta_v) - g''_v(0)| \leqslant \varepsilon \|v\|^2$, ce qui implique bien que $g''_v(\theta_v) \underset{v\to 0}{=} q_{m_0}(v) + \mathrm{o}(\|v\|^2)$. □

3 Extrema

3.1 Généralités

• Le vocabulaire : extrema locaux ou relatifs, globaux, extrema locaux stricts...

• Rappels dans le cas $n = 1$.

Théorème (condition nécessaire d'extremum local).
Soient I un intervalle de $\mathbb{R}$, $f : I \to \mathbb{R}$, $a \in I$ ***non borne de*** *I.*
Si f admet en a un extremum relatif alors $f'(a) = 0$.

Démonstration. On suppose que f admet en a un minimum local. Si t est proche de a avec $t > a$ on a donc :

$$\frac{f(t) - f(a)}{t - a} \leqslant 0.$$

Avec $t \to a^+$ on obtient $f'(a) \leqslant 0$.
En considérant le cas $t < a$ on obtient cette fois $f'(a) \geqslant 0$. □

3.2 Condition nécessaire du premier ordre

Théorème 7 (Condition nécessaire du premier ordre)
Soit $f : \mathcal{O} \to \mathbb{R}$ une fonction de classe C^1. Si f admet en m_0 un extremum local alors m_0 est un point critique de f i.e. $\nabla f(m_0) = 0$.

Démonstration. On suppose que f admet en m_0 un extremum local.
Soit $v \in \mathbb{R}^n$. Il existe $\alpha > 0$ tel que pour tout $t \in]-\alpha, \alpha[$ on ait $c(t) = m_0 + tv \in \mathcal{O}$.
La fonction $g_v : t \mapsto f(m_0 + tv)$ est dérivable sur $]-\alpha, \alpha[$ et admet un extremum local en $t = 0$, donc $g_v'(t) = 0$ i.e. $\langle \nabla f(m_0), v \rangle = 0$.
Ceci étant vrai pour tout $v \in \mathbb{R}^n$, $\nabla f(m_0) \in (\mathbb{R}^n)^\perp = \{0\}$. □

3.3 Conditions du second ordre

Théorème 8 (Condition nécessaire du second ordre)
Soit $f : \mathcal{O} \to \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. Si f admet un minimum local en m_0 alors la matrice hessienne de f en m_0 est positive i.e. si q_{m_0} est la forme quadratique associée à $H_f(m_0)$ on a $q_{m_0}(v) \geqslant 0$ pour tout v dans $\mathbb{R}^n$.

Démonstration. On suppose que f admet un minimum local en m_0. On sait déjà (théorème 7) que $\nabla f(m_0) = 0$. Soit $v \in \mathbb{R}^n$. Pour t petit $m_0 + tv \in \mathcal{O}$ et le développement limité à l'ordre 2 de f en m_0 donne une fonction $\varepsilon : \mathcal{O} \to \mathbb{R}$ nulle et continue en $m8$ telle que pour tout $t \in]-\alpha, \alpha[$ où $\alpha > 0$ on ait :

$$0 \leqslant f(m_0 + tv) - f(m_0) = \frac{1}{2} q_{m_0}(tv) + t^2 \|v\| \varepsilon(m_0 + tv)$$

ce qui amène, puisque $q_{m_0}(tv) = t^2 q_{m_0}(v)$, lorsque $t \neq 0$:

$$0 \leqslant \frac{1}{2} q_{m_0}(v) + \|v\| \varepsilon(m_0 + tv).$$

On fait alors tendre t vers 0 pour obtenir : $q_{m_0}(v) \geqslant 0$. □

Théorème 9 (Condition suffisante du second ordre)
Soient $f : \mathcal{O} \to \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. On suppose que :

- *m_0 est un point critique de f.*
- *La matrice hessienne $H_f(m_0)$ appartient à $S_n^{++}(\mathbb{R})$ i.e. ses valeurs propres sont strictement positives.*

Alors f admet un minimum local strict en m_0.

Démonstration. Le développement limité à l'ordre 2 de f en m_0 donne une fonction $\varepsilon : \mathcal{O} \to \mathbb{R}$ nulle et continue en m_0 telle que pour tout $m \in \mathcal{O}$ on ait, en notant $v_m = m - m_0$:

$$f(m) - f(m_0) = \frac{1}{2} q_{m_0}(tv_m) + \|v_m\| \varepsilon(m)$$

Or pour tout $w \in \mathbb{R}^n$ on a $q_{m_0}(w) \geqslant \lambda \|w\|^2$ où λ est la plus petite des valeurs propres de $H_f(m_0)$.
Il vient donc, pour tout $m \in \mathcal{O}$:

$$f(m) - f(m_0) \geqslant \|v_m\|^2 \left(\frac{\lambda}{2} + \varepsilon(m) \right).$$

Comme $\varepsilon(m) \underset{m \to m_0}{\longrightarrow} 0$, il existe $\alpha > 0$ tel que $\|v_m\| \leqslant \alpha$ implique $\left(\frac{\lambda}{2} + \varepsilon(m) \right) \geqslant 0$.
Ainsi f admet en m_0 un minimum local. □

Corollaire 9.1
Soient $f : \mathcal{O} \to \mathbb{R}$ de classe C^2 et $m_0 \in \mathcal{O}$. On suppose que m_0 est un point critique de f.

1. *Si les valeurs propres de $\nabla^2 f(m_0)$ sont strictement positives alors f admet en m_0 un minimum local strict.*
2. *Si les valeurs propres de $\nabla^2 f(m_0)$ sont strictement négative alors f admet en m_0 un maximum local strict.*

3. *Si $\nabla^2 f(m_0)$ admet une valeur propre strictement négative et une valeur propre strictement positive, alors f n'admet pas en m_0 d'extremum relatif.*

Procédure pour la recherche d'extrema sur un ouvert.

1. On vérifie que l'on travaille sur un ouvert et on justifie du caractère C^1 ou C^2 de la fonction.
2. On cherche les points critiques en résolvant le système donné par $\nabla f(m) = 0$.
3. • On regarde en chacun des points critiques la hessienne de f. **Une** idée est de regarder ses valeurs propres. La condition nécessaire du second ordre permet d'éliminer les mauvais candidats.
   • On peut aussi faire une réciproque en étudiant (localement) le signe de $f(m) - f(m_0)$ ou de $f(m+h) - f(m)$.

Bien noter que tout cela ne donne qu'un résultat local ; il faut utiliser d'autres arguments pour avoir des résultats **globaux** .

Le cas où $n = 2$

On calcule le déterminant de la hessienne en m_0 :

— Si ce déterminant est strictement positif, il y a en m_0 un minimum local strict lorsque la trace est positive et un maximum local strict lorsque la trace est négative.
— Si ce déterminant est strictement négatif, il n'y a pas d'extremum local en m_0 : on a un point selle (ou col).
— Dans les autres cas, on ne peut rien dire : on peut regarder la différence $f(m) - f(m_0)\ldots$

3.4 Compléments sur les extrema

3.4.1 Passer des extrema locaux aux extrema globaux

Je vais présenter une situation modèle qui peut s'adapter à divers cas. On considère une fonction $f : \mathbb{R}^n \to \mathbb{R}$, disons de classe C^1 qui admet un unique point critique m_0.

On suppose que pour tout $v \in \mathbb{R}^n$ on a $f(v) \geqslant \varphi(\|v\|)$ où $\varphi : \mathbb{R}^+ \to \mathbb{R}$ vérifie : $\varphi(t) \underset{t\to+\infty}{\longrightarrow} +\infty$.

Notons $\alpha = f(m_0)$. Comme $\varphi(t) \underset{t\to+\infty}{\longrightarrow} +\infty$, il existe $R > 0$ tel que pour tout $t \geqslant R$ on ait $\varphi(t) \geqslant 2|\alpha|$. Ainsi pour tout $v \in \mathbb{R}^n$ tel que $\|v\| \geqslant R$ on a :

$$f(v) \geqslant 2|\alpha|.$$

La restriction g de f à la boule fermée $B_f(0,R)$ est continue sur le fermé borné $B_f(0,R)$ de l'espace vectoriel $\mathbb{R}^n$ de dimension finie : elle admet donc un minimum global en un point p_0 de $B_f(0,R)$. Ainsi, pour tout $m \in B_f(0,R)$ on a :

$$f(m) = g(m) \geqslant g(p_0) = f(p_0).$$

De plus, pour tout m hors de la boule $B_f(0,R)$, on a :

$$f(m) \geqslant 2|\alpha| \geqslant f(m_0),$$

et comme $m_0 \in B_f(0,R)$, il vient $f(m) \geqslant f(p_0)$.

Ainsi f admet en p_0 un minimum global sur $\mathbb{R}^n$. Comme c'est un ouvert et que f est de classe C^1, p_0 est un point critique de f donc $p_0 = m_0$, par unicité du point critique.

En conclusion f admet en m_0 un minimum global.

3.4.2 Une condition suffisante de minimum global

Soient Ω un **ouvert convexe** de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ de classe C^2. Soient $m_0 \in \Omega$, $m \in \Omega$ et $v = m - m_0$.
Par convexité de Ω, la fonction

$$g_v : t \mapsto f(m_0 + tv)$$

est correctement définie sur $[0,1]$.
Puis on a vu que g_v est de classe C^2 sur $[0,1]$ et $g''(t) = X^T H_f(m_0 + tv)X$ où X est la colonne des coordonnées de v dans la base canonique.
Une intégration par parties (ou Taylor-intégral à l'ordre 1) donne alors :

$$g(1) = g(0) + g'(0) + \int_0^1 (1-t)g''(t)\,\mathrm{d}t.$$

Cela se traduit par :

$$f(m_0 + v) = f(m_0) + \langle \nabla f(m_0), v\rangle + \int_0^1 (1-t)X^T H_f(m_0 + tv)X\,\mathrm{d}t.$$

Ainsi, dès que la matrice hessienne $H_f(*)$ est dans $S_n^{++}(\mathbb{R})$ pour tout $* \in \Omega$, on peut dire que f admet en m_0 un minimum global strict.

Remarque 3.1
Pour $\mathcal{O}$ convexe, la condition $H_f(*)$ est dans $S_n^{++}(\mathbb{R})$ pour tout $* \in \mathcal{O}$ est équivalente au fait que f est convexe.

4 Le cas des fonctions d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

Cadre supplémentaire

Dans tout ce paragraphe

- n est un entier supérieur (ou égal) à 1 ;
- $\mathbb{R}^n$ et $\mathbb{R}^p$ sont munis de leurs structures euclidiennes canoniques respectives ;
- $\mathcal{O}$ désigne un ouvert non vide de $\mathbb{R}^n$;
- Pour $f : \mathcal{O} \to \mathbb{R}^p$ on écrit :

$$f = (f_1, \dots, f_p),$$

où pour tout $i \in \{1, \dots, p\}$, $f_i : \mathcal{O} \to \mathbb{R}$.

- On confondra le gradient d'une fonction $g : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 avec la matrice ligne de $\mathcal{M}_{1,n}(\mathbb{R})$ dont les coefficients sont les coordonnées du gradient.

4.1 Fonction de classes C^1 d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$, matrice jacobienne.

Définition 8 *On dit que f est de classe C^1 lorsque chaque application f_i pour $i \in \{1, \dots, p\}$ est de classe C^1.*
Lorsque $m_0 \in \mathcal{O}$ la matrice jacobienne de f en m_0 est l'élément de $\mathcal{M}_{p,n}(\mathbb{R})$, notée $J_f(m_0)$, « dont la i-ème ligne est le gradient $\nabla f_i(m_0)$ » :

$$J_f(m_0) = \begin{pmatrix} \partial_1 f_1(m_0) & \partial_2 f_1(m_0) & \cdots & \partial_n f_1(m_0) \\ \partial_1 f_2(m_0) & \partial_2 f_2(m_0) & \cdots & \partial_n f_2(m_0) \\ \vdots & \vdots & & \vdots \\ \partial_1 f_p(m_0) & \partial_2 f_p(m_0) & \cdots & \partial_n f_p(m_0) \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R}).$$

Lorsque $n = p$ le déterminant de $J_f(m_0)$ s'appelle le jacobien de f en m_0.

Théorème 10

1. ***La règle des chaînes.***
   Soient $g : \mathcal{O} \to \mathbb{R}^p$ et $h : \mathbb{R}^p \to \mathbb{R}$ des fonctions de classe C^1. Alors $h \circ g : \mathcal{O} \to \mathbb{R}$ est de classe C^1 avec, pour tout $m \in \mathcal{O}$:

$$\nabla(h \circ g)(m) = \nabla h(g(m)) \times J_g(m),$$

   i.e. pour tout $i \in \{1, \dots, n\}$:

$$\partial_i(h \circ g)(m) = \sum_{j=1}^{p} \partial_j h(g(m)) \times \partial_i g_j(m).$$

2. ***Le théorème des fonctions composées.*** *Soient $q \geqslant 1$ entier, Ω un ouvert de $\mathbb{R}^q$, $g : \Omega \to \mathcal{O}$ et $f : \mathcal{O} \to \mathbb{R}^p$ des fonctions de classe C^1.*
   Alors la fonction $f \circ g : \Omega \to \mathbb{R}^p$ est de classe C^1 et pour tout $m \in \Omega$ on a :

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

Démonstration.

1. On écrit $g = (g_1, \dots, g_p)$. Soit $m_0 \in \mathcal{O}$. On écrit $m_0 = (a_1, \dots, a_n)$. La première application partielle de $h \circ g$ en m_0, correctement définie sur un intervalle de la forme $]a_1 - \alpha, a_1 + \alpha[$ où $\alpha > 0$, est :

$$h_1 : t \mapsto h(g_1(t, a_2, \dots, a_n), \dots, g_p(t, a_2, \dots, a_n)).$$

   On pose pour $t \in]a_1 - \alpha, a_1 + \alpha[$:

$$c(t) = (g_1(t, a_2, \dots, a_n), \dots, g_p(t, a_2, \dots, a_n)).$$

   Comme c est de classe C^1, le théorème 4 s'applique. La fonction h_1 est dérivable en a_1 avec :

$$h_1'(a_1) = \langle \nabla h(c(0)), c'(0) \rangle,$$

   ce qui donne $\partial_1(h \circ g)(m_0) = \sum_{j=1}^{n} \partial_j h(m_0) \times \partial_1 g_j(m_0)$.
   On procède de la même manière pour les autres dérivées partielles.

2. Résulte de suite du point 1 et de la définition de la jacobienne.

□

4.2 Le point de vue de la différentielle

Définition 9 *Soient $m_0 \in \mathcal{O}$ et $f : \mathcal{O} \to \mathbb{R}^p$. On dit que f admet un développement limité à l'ordre 1 en m_0 lorsqu'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ et une fonction $\varepsilon : \mathcal{O} \to \mathbb{R}^p$ continue et nulle en m_0 telle que pour tout $m \in \mathcal{O}$ on ait :*

$$f(m) = f(m_0) + L(m - m_0) + \|m - m_0\| \varepsilon(m).$$

Théorème 11
Soit $f : \mathcal{O} \to \mathbb{R}^p$.

1. *Lorsque f admet en $m_0 \in \mathcal{O}$ un développement limité à l'ordre 1 alors l'application $L : \mathbb{R}^n \to \mathbb{R}^p$ de la définition 9 est unique : c'est la différentielle de f en m_0 et on la notera $T_{m_0}(f)$ ou df_{m_0} ou...*
2. *On suppose que f est de classe C^1 sur $\mathcal{O}$. La fonction f admet alors un développement limité en tout point m de $\mathcal{O}$ et la matrice de l'application linéaire $T_{m_0}(f)$ dans le couple des bases canoniques est la matrice jacobienne de f en m_0.*

Remarque 4.1
Le théorème 10 se réécrit en terme de différentielles de la manière suivante.
Soient $q \geqslant 1$ entier, Ω un ouvert de $\mathbb{R}^q$, $g : \Omega \to \mathcal{O}$ et $f : \mathcal{O} \to \mathbb{R}^n$ des fonctions de classe C^1.
Alors la fonction $f \circ g : \Omega \to \mathbb{R}^n$ est de classe C^1 et pour tout $m \in \Omega$ on a :

$$d(f \circ g)_m = df_{g(m)} \circ df_m.$$

5 Difféomorphismes et inversion locale

5.1 Difféomorphismes

Définition 10 *Soient V un ouvert de $\mathbb{R}^n$ et W un ouvert de $\mathbb{R}^p$. On dit que $f : V \to W$ est C^1-difféomorphisme lorsque f est bijective, de classe C^1 avec f^{-1} de classe C^1.*

Remarque 5.1

1. L'application $f = \begin{pmatrix} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^3 \end{pmatrix}$ n'est pas un C^1-difféomorphisme.
2. Avec les notations de la définition 10, lorsque f est un difféomorphisme, en notant $g = f^{-1}$, on a :
$$\begin{cases} f \circ g = \mathrm{Id}_W \\ g \circ f = \mathrm{Id}_V \end{cases}.$$
Ainsi par le théorème des fonction composées, si $x \in V$, alors :
$$\begin{cases} df_x \circ dg_{f(x)} = \mathrm{Id}_{\mathbb{R}^p} \\ dg_{f(x)} \circ df_x = \mathrm{Id}_{\mathbb{R}^n} \end{cases}.$$
Ainsi df_x est une application linéaire bijective, ce qui force $n = p$.
3. $x \mapsto \dfrac{x}{1 - \|x\|}$ est un difféomorphisme de la boule ouverte $B(0,1)$ sur $\mathbb{R}^n$.
4. Le carré $]-1,1[^2$ est difféomorphe à $\mathbb{R}^2$ via :
$$(x,y) \mapsto \left(\frac{x}{1-x^2}, \frac{y}{1-y^2} \right).$$

Théorème 12
Soient V un ouvert de $\mathbb{R}^n$, W un ouvert de $\mathbb{R}^p$, $f : V \to W$ un homéomorphisme de classe C^1. Les propriétés suivantes sont équivalentes.

(1) f est un C^1-difféomorphisme.
(2) $(\forall x \in V)(df_x \in \mathrm{Isom}(E,F))$

Démonstration. • (1) $\Rightarrow$ (2). Soit x dans V. On note $g = f^{-1}$ et $y = f(x)$. Comme $g \circ f = \mathrm{Id}_V$, il vient : $dg_y \circ df_x = \mathrm{Id}_E$ et on a de même $df_x \circ dg_y = \mathrm{Id}_F$, ce qui prouve (2).

• (2) $\Rightarrow$ (1) (rapidement...). On doit prouver que $g = f^{-1}$ est de classe C^1.
Mais g est différentiable en tout point de W (écrire l'identité de la différentielle) et si $y \in W$ on a $dg_y = \left(df_{g(y)}\right)^{-1}$.
Ainsi la différentielle dg est une composée de df, de g et de $u \mapsto u^{-1}$ qui sont toutes continues : dg est continue. □

5.2 Le théorème d'inversion locale

Théorème 13 (Inversion locale)
Soient $a \in U$ un ouvert de $\mathbb{R}^n$ et $f \in C^1(U, \mathbb{R}^n)$.
On suppose que df_a est inversible.
Il existe alors un voisinage ouvert V de a dans $\mathbb{R}^n$, inclus dans U et W un voisinage ouvert de $f(a)$ dans $\mathbb{R}^n$ tels que $f|_V^W$ soit un C^1-difféomorphisme.

Définition 11 *Soient U un ouvert de $\mathbb{R}^n$ et $f \in C^1(U, \mathbb{R}^n)$. On dit que f est un difféomorphisme local (ou **application étale**) lorsque pour tout $x \in U$ l'application linéaire df_x est inversible.*

Exemple 5.1 (Les coordonnées polaires)
L'application

$$f = \begin{pmatrix} \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta) \end{pmatrix}$$

est de classe C^1 par composantes.
Pour $(r,\theta) \in \mathbb{R}_+^* \times \mathbb{R}$, le déterminant jacobien de f en (r,θ) est :

$$DJ_f(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r > 0,$$

et ainsi $df_{(r,\theta)}$ est une application linéaire inversible.
Il en résulte que f est un difféomorphisme local. La réciproque locale $g = (g_1, g_2)$ étant définie par :

$$\begin{cases} g_1(u,v) & = g_1(r\cos\theta, r\sin\theta) = r = \sqrt{u^2+v^2} \\ g_2(u,v) & = g_2(r\cos\theta, r\sin\theta) = \theta \text{ défini selon...} \end{cases}$$

5.3 Applications importantes

Théorème 14 (Théorème de l'image ouverte)
Une application étale est ouverte.

Démonstration. Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une application étale. Soit Ω un ouvert de U. On veut montrer que $f(\Omega)$ est un ouvert de $\mathbb{R}^n$. Soit $y \in f(\Omega)$. Il existe $x \in U$ tel que $f(x) = y$. Comme df_x est inversible, on peut appliquer le théorème d'inversion locale : il existe V ouvert de U contenant x tel que f induise un difféomorphisme $g : V \to g(V)$. Comme g^{-1} est continue, $g(V)$ est un ouvert de $\mathbb{R}^n$ et $x \in g(V) \subset f(\Omega)$, ce qui démontre que $f(\Omega)$ est un ouvert de $\mathbb{R}^n$. □

Théorème 15 (Théorème d'inversion globale)
Une application étale injective est un difféomorphisme sur son image.

Démonstration. Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une application étale injective. On sait d'après le théorème 14 que $V = f(U)$ est un ouvert de $\mathbb{R}^n$.
Selon le théorème d'inversion locale $g = f^{-1}$ est localement de classe C^1, donc est localement continue donc est continue, de sorte que f est un homéomorphisme sur son image : on applique alors le théorème 12. □