

Inversion locale et extrema liés

I. Difféomorphismes

Dans tout le chapitre, E et F désignent des espaces vectoriels normés de dimension finie, le plus souvent $\mathbb{R}^n$. Les résultats restent vrais dans des espaces de Banach, avec la même démonstration. Une application est dite de classe C^k si elle est k fois différentiable, avec une différentielle k -ième continue.

Définition :

Soient U un ouvert de E et V un ouvert de F . Une application $f:U \rightarrow V$ est un **C^k -difféomorphisme** ($k \geq 1$) si f est bijective, de classe C^k , et si sa réciproque $f^{-1}:V \rightarrow U$ est aussi de classe C^k .

Un difféomorphisme conserve donc toute la structure différentielle. En particulier, sa différentielle est inversible en chaque point. En effet, en différentiant $f^{-1} \circ f = \text{id}_U$ et $f \circ f^{-1} = \text{id}_V$, on obtient :

Propriété :

Si $f:U \rightarrow V$ est un C^1 -difféomorphisme, alors pour tout $x \in U$, $Df(x)$ est un isomorphisme de E sur F et

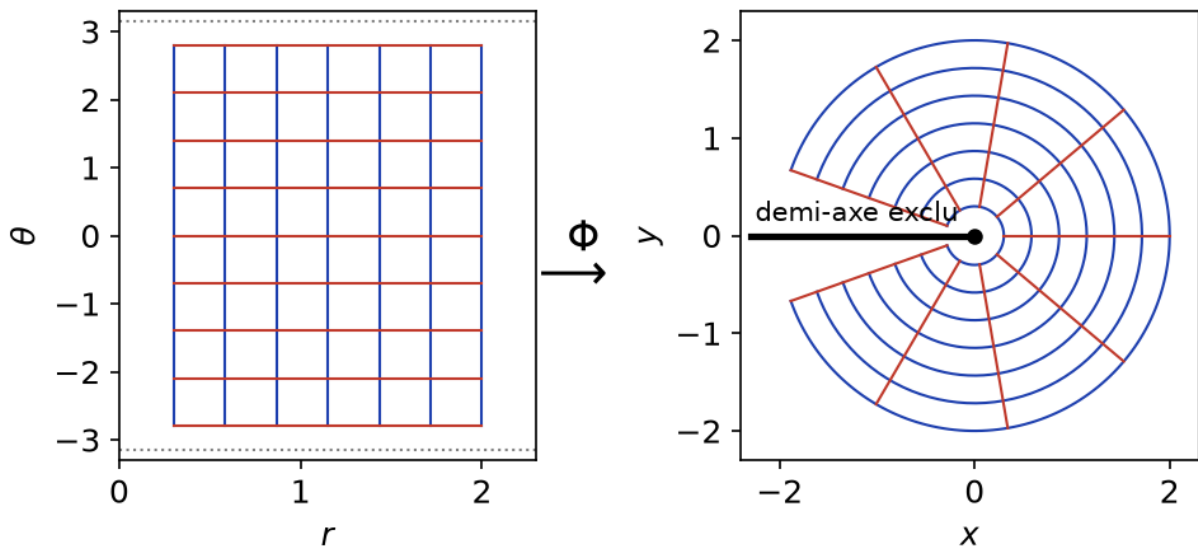
$$D(f^{-1})(f(x)) = (Df(x))^{-1}.$$

Par conséquent, $\dim E = \dim F$.

Exemple :

L'application $x \rightarrow x^3$ est une bijection continue de $\mathbb{R}$ sur $\mathbb{R}$, de réciproque continue. Cependant, ce n'est pas un difféomorphisme : sa dérivée s'annule en 0 , et sa réciproque $y \rightarrow y^{1/3}$ n'est pas dérivable en 0 . Autrement dit, un homéomorphisme de classe C^∞ n'est pas forcément un difféomorphisme.

L'exemple fondamental est celui des coordonnées polaires. L'application $\Phi(r, \theta) = (r \cos \theta, r \sin \theta)$ envoie la bande $]0, +\infty[\times]-\pi, \pi[$ sur le plan privé du demi-axe $\mathbb{R}_- \times \{0\}$. Comme le montre la figure ci-dessous, elle transforme un quadrillage rectangulaire en un réseau de cercles et de rayons.

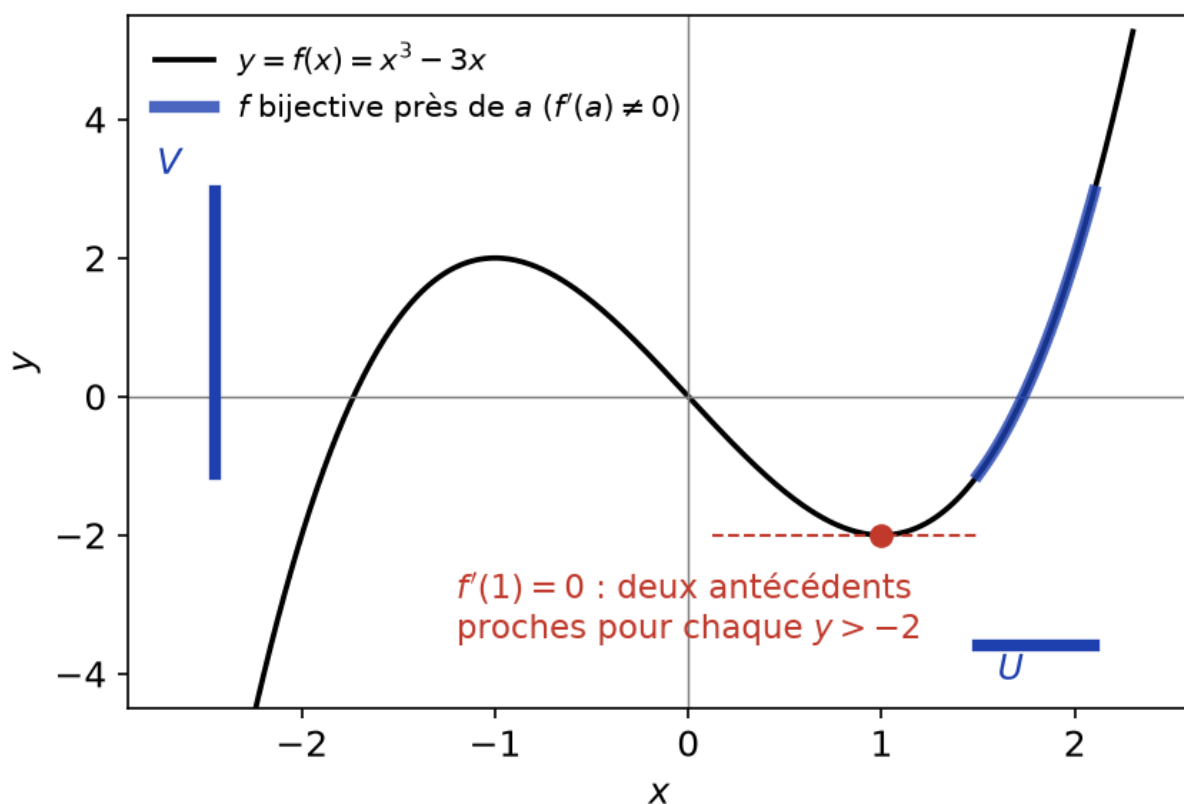


Nous verrons plus loin que Φ est un C^∞ -difféomorphisme sur son image. Son jacobien vaut $\det D\Phi(r, \theta) = r$, qui ne s'annule jamais sur la bande.

II. Le théorème d'inversion locale

La question est la suivante : si $Df(a)$ est inversible, f est-elle inversible près de a ? En dimension 1, la réponse est classique. Si $f'(a) \neq 0$, alors f' garde un signe constant près de a , donc f est

strictement monotone. En revanche, en un point où f' s'annule, la fonction peut prendre deux fois la même valeur, comme le montre la figure ci-dessous.



1. Un lemme de point fixe

La démonstration repose sur le théorème du point fixe de Banach. Rappelons-le, car il fournit à la fois l'existence et l'unicité de l'antécédent.

Théorème :

(Point fixe de Banach) Soient X un espace métrique complet et $\varphi : X \rightarrow X$ une application k -lipschitzienne avec $0 \leq k < 1$. Alors φ admet un unique point fixe.

On utilise aussi l'inégalité des accroissements finis. Si h est C^1 sur un convexe C et si

$|Dh(x)| \leq k$ pour tout $x \in C$, alors h est k -lipschitzienne sur C .

2. Énoncé et démonstration

Théorème :

(Inversion locale) Soient U un ouvert de $\mathbb{R}^n$, $f:U \rightarrow \mathbb{R}^n$ de classe C^k ($k \geq 1$) et $a \in U$. Si $Df(a)$ est inversible, il existe un ouvert $U_0 \ni a$ et un ouvert $V_0 \ni f(a)$ tels que f induise un C^k -difféomorphisme de U_0 sur V_0 .

Démonstration :

D'abord, on se ramène au cas $a = 0$, $f(0) = 0$ et $Df(0) = I$. Il suffit de remplacer f par $x \rightarrow Df(a)^{-1}(f(a+x) - f(a))$.

Posons $h(x) = x - f(x)$. Alors $Dh(0) = 0$. Par continuité de Df , il existe $r > 0$ tel que $|Dh(x)| \leq \frac{1}{2}$ sur la boule fermée $\overline{B}(0,r)$. Donc h est $\frac{1}{2}$ -lipschitzienne sur cette boule, et $|h(x)| \leq \frac{1}{2}|x|$ car $h(0) = 0$.

Ensuite, fixons y avec $|y| < r/2$ et posons $\varphi_y(x) = y + h(x)$. On a $f(x) = y$ si et seulement si $\varphi_y(x) = x$. De plus, $|\varphi_y(x)| \leq r/2 + r/2 = r$, donc φ_y envoie la boule fermée, qui est complète, dans elle-même. Elle y est $\frac{1}{2}$ -contractante : elle a donc un unique point fixe x . Ainsi, y a un unique antécédent dans $\overline{B}(0,r)$.

Posons $V_0 = B(0,r/2)$ et $U_0 = B(0,r) \cap f^{-1}(V_0)$, qui est ouvert. Alors $f:U_0 \rightarrow V_0$ est bijective. Notons g sa réciproque. Pour $x, x' \in U_0$, on a $|x - x'| \leq |f(x) - f(x')| + \frac{1}{2}|x - x'|$. Par conséquent, g est 2 -lipschitzienne, donc continue.

Enfin, $Df(x) = I - Dh(x)$ est inversible sur U_0 , car $|Dh(x)| < 1$. Soient $y = f(x)$ et

$y + k = f(x + \ell)$. On a $|\ell| \leq 2|k|$ et $k = Df(x)\ell + o(\ell)$. Donc $\ell = Df(x)^{-1}k + o(k)$: g est différentiable et $Dg(y) = Df(g(y))^{-1}$. Comme l'inversion est C^∞ sur $GL_n(\mathbb{R})$, cette formule montre par récurrence que g est de classe C^k .

Corollaire :

Si $f : U \rightarrow \mathbb{R}^n$ est C^1 et si $Df(x)$ est inversible pour tout $x \in U$, alors f est une application ouverte : l'image de tout ouvert est un ouvert.

3. Inversion globale

L'inversion locale ne dit rien de l'injectivité sur tout U . Par exemple, $(x, y) \rightarrow (e^x \cos y, e^x \sin y)$ a une différentielle inversible partout, mais elle est 2π -périodique en y . Il faut donc ajouter l'injectivité.

Théorème :

(Inversion globale) Soit $f : U \rightarrow \mathbb{R}^n$ de classe C^k sur un ouvert U . Si f est injective et si $Df(x)$ est inversible pour tout $x \in U$, alors $V = f(U)$ est ouvert et f est un C^k -difféomorphisme de U sur V .

Démonstration :

D'après le corollaire, V est ouvert. Ensuite, $f : U \rightarrow V$ est bijective. Soit $y_0 = f(x_0) \in V$. Par inversion locale, f induit un difféomorphisme de $U_0 \ni x_0$ sur $V_0 \ni y_0$. Sur V_0 , la réciproque globale coïncide avec cette réciproque locale, grâce à l'injectivité. Elle est donc C^k au voisinage de chaque point.

Méthode :

Pour montrer que $f : U \rightarrow f(U)$ est un difféomorphisme :

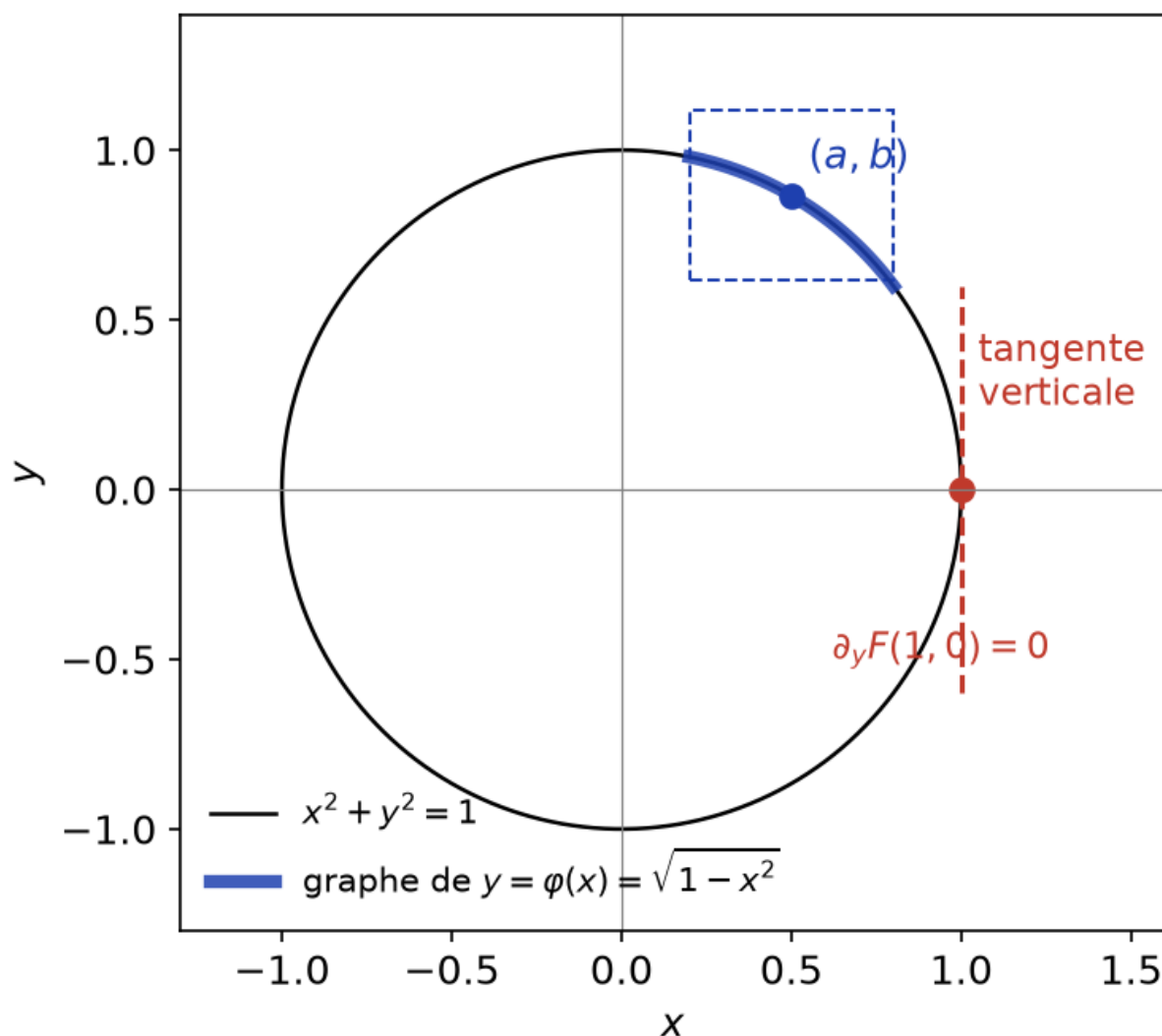
- vérifiez que f est de classe C^1 ;
- calculez le jacobien et montrez qu'il ne s'annule pas sur U ;
- prouvez l'injectivité, souvent en résolvant $f(x) = y$;
- concluez par le théorème d'inversion globale.

Exemple :

Pour les coordonnées polaires, Φ est C^∞ et $\det D\Phi(r,\theta) = r > 0$. De plus, Φ est injective : r est le module de $\Phi(r,\theta)$, puis θ est l'unique argument dans $] -\pi, \pi[$. Donc Φ est un C^∞ -difféomorphisme de la bande sur $\mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$.

III. Le théorème des fonctions implicites

On considère une équation $F(x,y) = 0$, avec $x \in \mathbb{R}^p$ et $y \in \mathbb{R}^q$. On veut exprimer y en fonction de x . Par exemple, le cercle $x^2 + y^2 = 1$ n'est pas un graphe, mais il l'est localement près de tout point où la tangente n'est pas verticale. La figure ci-dessous illustre cette situation.



Théorème :

(Fonctions implicites) Soient W un ouvert de $\mathbb{R}^p \times \mathbb{R}^q$, $F: W \rightarrow \mathbb{R}^q$ de classe C^k ($k \geq 1$) et $(a, b) \in W$ tel que $F(a, b) = 0$. On suppose que la différentielle partielle $D_y F(a, b)$ est inversible. Alors il existe des voisinages ouverts V de a et W_0 de b , et une application $\varphi: V \rightarrow W_0$ de classe C^k , tels que :

$$\forall (x, y) \in V \times W_0, F(x, y) = 0 \Leftrightarrow y = \varphi(x).$$

De plus, $D\varphi(x) = -D_y F(x, \varphi(x))^{-1} \circ D_x F(x, \varphi(x))$.

Démonstration :

Posons $\Phi(x,y) = (x, F(x,y))$. Sa différentielle en (a,b) s'écrit par blocs

$\begin{pmatrix} p & 0 \\ D_x F & D_y F \end{pmatrix}$. Elle est donc inversible. Par inversion locale, Φ est un difféomorphisme d'un voisinage de (a,b) sur un voisinage de $(a,0)$. Sa réciproque a la forme

$\Psi(u,v) = (u, \psi(u,v))$, puisque Φ conserve la première composante. Alors $F(x,y) = 0$ équivaut à $\Phi(x,y) = (x,0)$, c'est-à-dire $y = \psi(x,0)$. On pose $\varphi(x) = \psi(x,0)$. Enfin, on différentie l'identité $F(x,\varphi(x)) = 0$ pour obtenir la formule annoncée.

Méthode :

Pour étudier une courbe $F(x,y) = 0$ près d'un point (a,b) :

- vérifiez que $F(a,b) = 0$ et que $\partial_y F(a,b) \neq 0$;
- concluez à l'existence de φ de classe C^1 avec $\varphi(a) = b$;
- dérivez $F(x,\varphi(x)) = 0$: $\varphi'(x) = -\partial_x F / \partial_y F$;
- dérivez une seconde fois pour obtenir φ'' , donc la position de la courbe par rapport à sa tangente.

Exemple :

Soit $F(x,y) = x^2 + y^2 - 1$ et (a,b) avec $b > 0$. On a $\partial_y F = 2b \neq 0$. Donc $y = \varphi(x)$ près de a , et $\varphi'(x) = -x/\varphi(x)$. Ici, on reconnaît $\varphi(x) = \sqrt{1-x^2}$. En revanche, en $(1,0)$, on a $\partial_y F = 0$: c'est x qui s'exprime en fonction de y .

IV. Sous-variétés de $\mathbb{R}^n$ et espace tangent

Les fonctions implicites permettent de définir proprement les courbes et les surfaces lisses. On adopte ici l'approche élémentaire par équations.

Définition :

Une partie M de $\mathbb{R}^n$ est une **sous-variété de dimension** d si, pour tout $a \in M$, il existe un ouvert $U \ni a$ et $g: U \rightarrow \mathbb{R}^{n-d}$ de classe C^1 tels que $M \cap U = g^{-1}(\{0\})$ et $Dg(a)$ soit surjective. Pour $d = n - 1$, on parle d'**hypersurface**.

Par le théorème des fonctions implicites, une sous-variété est, près de chacun de ses points, le graphe d'une application C^1 de d coordonnées vers les $n-d$ autres.

Définition :

Le **vecteur tangent** à M en a est tout vecteur $\gamma'(0)$, où $\gamma:]-\varepsilon, \varepsilon[\rightarrow M$ est C^1 avec $\gamma(0) = a$. L'ensemble de ces vecteurs est l'**espace tangent** $T_a M$.

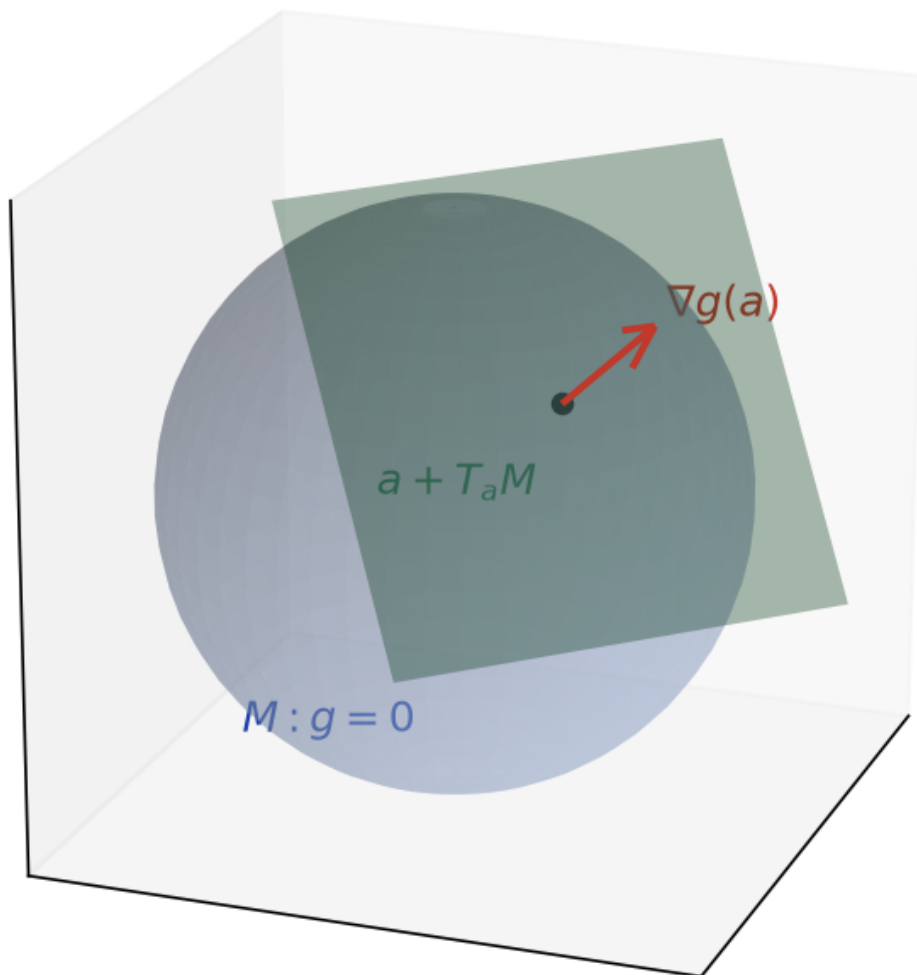
Théorème :

Avec les notations de la définition, $T_a M = \ker Dg(a)$. C'est un sous-espace vectoriel de dimension d . Pour une hypersurface $g = 0$, on a $T_a M = \nabla g(a)^\perp$.

Démonstration :

D'abord, si γ est tracée sur M , alors $g(\gamma(t)) = 0$. En dérivant en 0 , on obtient $Dg(a)\gamma'(0) = 0$. Donc $T_a M \subset \ker Dg(a)$, qui est de dimension d par le théorème du rang. Réciproquement, écrivons M localement comme le graphe $\{(u, \varphi(u))\}$. Pour tout $v \in \mathbb{R}^d$, la courbe $t \rightarrow (u_0 + tv, \varphi(u_0 + tv))$ donne le vecteur tangent $(v, D\varphi(u_0)v)$. Ces vecteurs forment un espace de dimension d . Par inclusion et égalité des dimensions, on conclut.

Ainsi, pour la sphère unité, $g(x) = |x|^2 - 1$ et $\nabla g(a) = 2a$. Le plan tangent en a est donc l'orthogonal de a , comme le montre la figure ci-dessous.



Méthode :

Pour déterminer l'espace tangent à $M = \{g = 0\}$ en a :

- vérifiez que $Dg(a)$ est surjective, c'est-à-dire que les gradients des composantes sont libres ;
- écrivez $T_a M = \{h : Dg(a)h = 0\}$;
- pour le plan tangent affine, prenez $a + T_a M$.

Attention :

La condition de rang est indispensable. Le cône $x^2 + y^2 - z^2 = 0$ a un gradient nul à l'origine. D'ailleurs, ce n'est pas une sous-variété en ce point : privé de l'origine, il a deux composantes connexes près de 0 .

V. Extrema libres

Soient U un ouvert de $\mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}$. On dit que f admet un minimum local en a si $f(x) \geq f(a)$ pour tout x voisin de a .

Théorème :

(Condition du premier ordre) Si f est différentiable en a et admet un extremum local en a , alors $Df(a) = 0$. On dit que a est un **point critique**.

La réciproque est fausse. En effet, $x^2 - y^2$ a un point critique en 0 sans extremum. On examine donc la hessienne $H_f(a) = (\partial_i \partial_j f(a))$, symétrique par le théorème de Schwarz.

Théorème :

(Condition du second ordre) Soit f de classe C^2 et a un point critique.

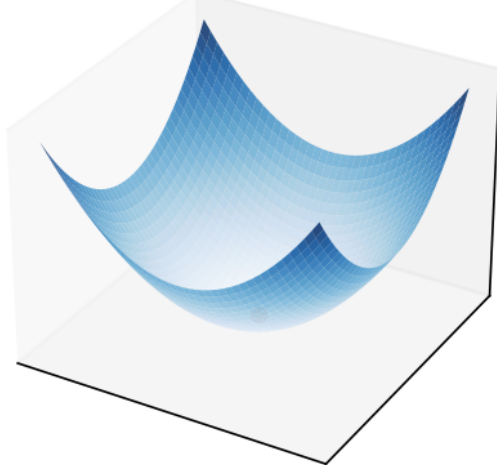
- Si $H_f(a)$ est définie positive, f a un minimum local strict en a .
- Si f a un minimum local en a , alors $H_f(a)$ est positive.
- Si $H_f(a)$ a une valeur propre strictement positive et une strictement négative, a est un point selle.

Démonstration :

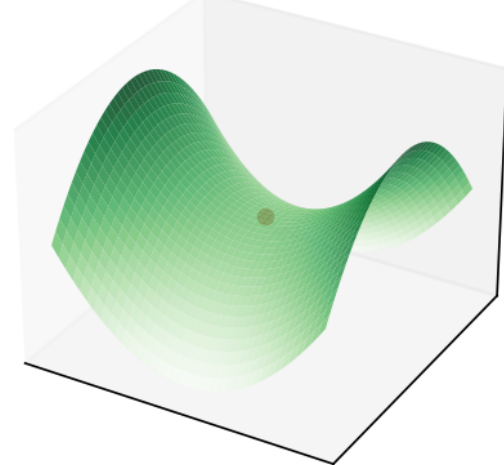
Par la formule de Taylor-Young, $f(a+h) = f(a) + \frac{1}{2}h^T H_f(a)h + o(|h|^2)$. Si $H_f(a)$ est définie positive, sa plus petite valeur propre α est strictement positive, et $h^T H_f(a)h \geq \alpha|h|^2$. Donc $f(a+h) - f(a) \geq \frac{\alpha}{4}|h|^2 > 0$ pour $h \neq 0$ assez petit. Pour les deux autres points, on se restreint à une droite $t \rightarrow a + tv$, avec v vecteur propre.

En dimension 2, on note $r = \partial_{xx}f$, $s = \partial_{xy}f$, $t = \partial_{yy}f$. Si $rt - s^2 > 0$, c'est un extremum : minimum si $r > 0$, maximum si $r < 0$. Si $rt - s^2 < 0$, c'est un point selle. Enfin, si $rt - s^2 = 0$, on ne peut pas conclure. La figure ci-dessous compare un minimum et un point selle.

$z = x^2 + y^2$: minimum



$z = x^2 - y^2$: point selle



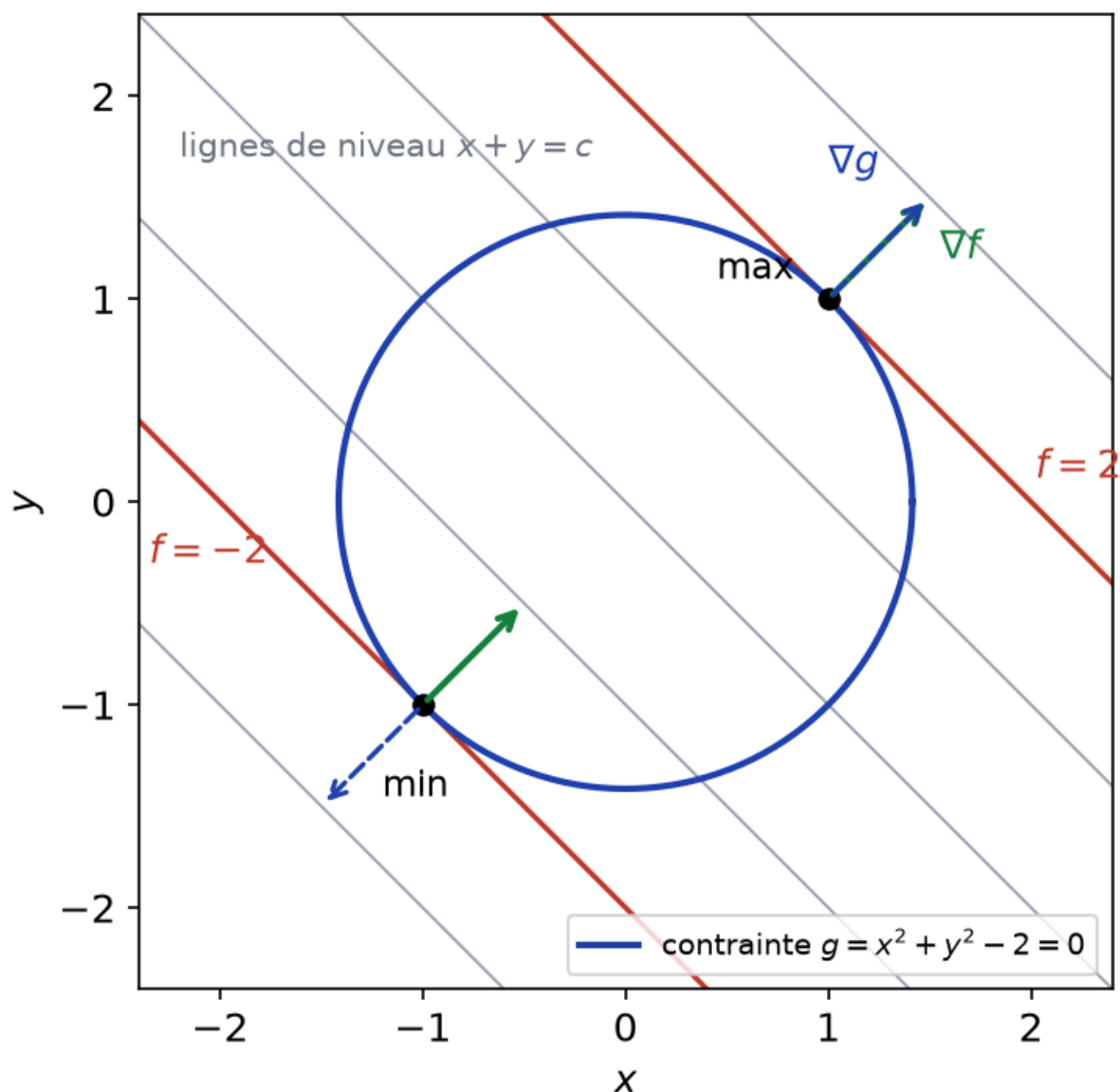
Exemple :

Soit $f(x,y) = x^2 + xy + y^2 - 3y$. Le gradient $(2x + y, x + 2y - 3)$ s'annule en $(-1, 2)$. Ensuite, $r = 2$, $s = 1$, $t = 2$, donc $rt - s^2 = 3 > 0$ et $r > 0$. C'est un minimum local, qui vaut $f(-1,2) = -3$.

VI. Extrema liés et multiplicateurs de Lagrange

On cherche maintenant les extrema de f restreinte à une contrainte

$M = \{g_1 = 1, \dots, g_p = 0\}$. Au point cherché, la ligne de niveau de f est tangente à la contrainte. Autrement dit, les gradients de f et de g sont colinéaires, comme le montre la figure ci-dessous.



Théorème :

(Multiplicateurs de Lagrange) Soient U un ouvert de $\mathbb{R}^n$, $f, g_1, \dots, g_p : U \rightarrow \mathbb{R}$ de classe

C^1 , et $M = \{x \in U : g_1(x) = \dots = g_p(x) = 0\}$. Soit $a \in M$ tel que $Dg_1(a), \dots, Dg_p(a)$ soient linéairement indépendantes. Si $f|_M$ admet un extremum local en a , il existe des réels $\lambda_1, \dots, \lambda_p$ tels que

$$\nabla f(a) = \lambda_1 \nabla g_1(a) + \dots + \lambda_p \nabla g_p(a).$$

Démonstration :

Grâce à l'indépendance des $Dg_i(a)$, M est une sous-variété près de a , et $T_a M = \bigcap_i \ker Dg_i(a)$. Soit $v \in T_a M$ et γ une courbe tracée sur M avec $\gamma(0) = a$, $\gamma'(0) = v$. La fonction $f \circ \gamma$ a un extremum local en 0 , donc $Df(a)v = 0$. Ainsi, $\ker Df(a)$ contient $\bigcap_i \ker Dg_i(a)$. Par un lemme classique d'algèbre linéaire, $Df(a)$ est alors combinaison linéaire des $Dg_i(a)$.

Méthode :

Pour trouver les extrema de f sous la contrainte $g = 0$:

- vérifiez que les ∇g_i sont libres en tout point de M , ou traitez à part les points où ils ne le sont pas ;
- résolvez le système $\nabla f = \sum \lambda_i \nabla g_i$, $g = 0$;
- si M est compact, f atteint ses bornes : comparez les valeurs aux points candidats ;
- sinon, concluez par convexité ou par une étude directe.

Exemple :

Cherchons les extrema de $f(x,y) = x + y$ sur le cercle $g(x,y) = x^2 + y^2 - 2 = 0$. Le gradient $\nabla g = (2x, 2y)$ ne s'annule pas sur le cercle. La condition $(1,1) = \lambda(2x, 2y)$ donne $x = y$, puis $x = \pm 1$. Le cercle est compact, donc f atteint ses bornes. Par conséquent, le maximum vaut $f(1,1) = 2$ et le minimum $f(-1,-1) = -2$.

VII. Fonctions convexes et minimisation

La convexité transforme la condition nécessaire $\nabla f(a) = 0$ en condition suffisante. C'est pourquoi elle joue un rôle central en optimisation.

Définition :

Soit C un convexe de $\mathbb{R}^n$. Une fonction $f : C \rightarrow \mathbb{R}$ est **convexe** si $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ pour tous $x, y \in C$ et $t \in [0,1]$. Elle est **strictement convexe** si l'inégalité est stricte pour $x \neq y$ et $t \in]0,1[$.

Proposition :

Soit f différentiable sur un ouvert convexe U .

- f est convexe si et seulement si $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ pour tous $x, y \in U$.
- Si f est C^2 , f est convexe si et seulement si $H_f(x)$ est positive en tout point. Si $H_f(x)$ est définie positive partout, f est strictement convexe.

Théorème :

Soit f convexe et différentiable sur un ouvert convexe U .

- Si $\nabla f(a) = 0$, alors a est un minimum global de f sur U .
- Si f est strictement convexe, elle admet au plus un point de minimum.
- Si $U = \mathbb{R}^n$, f continue et $f(x) \rightarrow +\infty$ quand $|x| \rightarrow +\infty$, alors f atteint son minimum.

Démonstration :

Le premier point découle de l'inégalité de la proposition avec $x = a$. Pour le deuxième, si $a \neq b$ sont deux minima de valeur m , alors $f(\frac{a+b}{2}) < m$, ce qui est absurde. Pour le troisième, f est supérieure à $f(0)$ hors d'une boule fermée B . Or f atteint son minimum sur le compact B , et ce minimum est global.

Exemple :

Soient A symétrique définie positive et $b \in \mathbb{R}^n$. La fonction $J(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$ a pour gradient $Ax - b$ et pour hessienne A . Elle est donc strictement convexe. Son unique minimum est atteint en $x^* = A^{-1}b$, et $J(x^*) = -\frac{1}{2} \langle A^{-1}b, b \rangle$. Par exemple, le problème des moindres carrés $\min |Ax - b|^2$ conduit aux équations normales $A^T Ax = A^T b$.

Remarque :

Si f est convexe et les contraintes g_i affines, la condition de Lagrange devient suffisante. En effet, la fonction $f - \sum \lambda_i g_i$ est alors convexe, de gradient nul au point trouvé. Elle y atteint donc son minimum global, et elle coïncide avec f sur la contrainte.

Ce qu'il faut retenir

- Un difféomorphisme a une différentielle inversible en tout point, mais un homéomorphisme lisse n'est pas toujours un difféomorphisme.
- Inversion locale : si $Df(a)$ est inversible, f est un difféomorphisme local près de a . La preuve repose sur le point fixe de Banach.
- Inversion globale : injectivité et différentielle inversible partout donnent un difféomorphisme sur l'image.

- Fonctions implicites : si $D_y F(a,b)$ est inversible, $F(x,y) = 0$ équivaut localement à $y = \varphi(x)$, avec $D\varphi = -(D_y F)^{-1} D_x F$.
- Une sous-variété est localement définie par $g = 0$ avec $Dg(a)$ surjective. Son espace tangent est $\ker Dg(a)$.
- Un extremum libre est un point critique. La hessienne définie positive donne un minimum strict ; une hessienne de signature mixte, un point selle.
- Lagrange : en un extremum lié régulier, ∇f est combinaison linéaire des ∇g_i . La condition est seulement nécessaire.
- Pour une fonction convexe, tout point critique est un minimum global, unique si la convexité est stricte.

Questions fréquentes sur inversion locale et extrema liés

Quelle est la différence entre inversion locale et inversion globale ?

L'inversion locale ne demande l'inversibilité de la différentielle qu'en un point, et ne donne une réciproque que sur un voisinage de ce point. L'inversion globale suppose en plus l'injectivité sur tout l'ouvert, et la différentielle inversible partout. On obtient alors un difféomorphisme de l'ouvert sur son image.

Comment savoir sur quelle variable appliquer le théorème des fonctions implicites ?

On cherche la variable par rapport à laquelle la dérivée partielle de F ne s'annule pas au point étudié. Par exemple, si $\partial_y F(a,b) \neq 0$, l'équation définit y en fonction de x près de (a,b) . Si seule $\partial_x F$ est non nulle, c'est x qui s'exprime en fonction de y .

La condition de Lagrange suffit-elle pour trouver un extremum lié ?

Non, elle est seulement nécessaire. Elle fournit des points candidats, qu'il faut ensuite départager. On conclut souvent par compacité de la contrainte, en comparant les valeurs de f aux points candidats, ou par convexité quand f est convexe et les contraintes affines.

Pourquoi la convexité simplifie-t-elle la minimisation ?

Pour une fonction convexe différentiable, tout point critique est un minimum global. De plus, la stricte

convexité garantit l'unicité du minimum. Il suffit donc de résoudre $\nabla f(a) = 0$ sans étudier la hessienne point par point.

Pour aller plus loin en L3

- Les énoncés : [exercices de maths en L3 sur inversion locale et extrema liés](#)
- À maîtriser avant : [Calcul différentiel dans les espaces normés](#), [Complétude, espaces de Banach et point fixe](#)
- Chapitre précédent : [Calcul différentiel dans les espaces normés](#)
- Chapitre suivant : [Équations différentielles et théorème de Cauchy-Lipschitz](#)
- Tester vos connaissances : [QCM de maths en L3 par chapitre](#)
- Le sommaire : [tous les chapitres de maths de L3](#) et la [licence de maths de L1 à L3](#)