Théorème des extremums liées

Julie Parreaux

2018-2019

Références du développement : Avez [1, p.103] ; Objectif agrégation [2, p.20] ; X-ENS Algèbre 1 [5, p.295] ; Rouvière [4, p.380] (les références dépendent de l'orientation de la leçon)

Leçons où on présente le développement : 151 (Dimension et rang) ; 152 (Déterminant) ; 159 (Dualité) ; 214 (TIL, TFI) ; 219 (Extremum).

1 Introduction

Le problème de recherche des extremums liés consiste à maximiser ou minimiser une fonction de plusieurs variables $f(x_1, \dots, x_n)$ lorsque ces variables sont liées par certaines relations. La recherche d'extremums dans un massif montagneux consiste à trouver les sommets, la recherche des extremums liés se rapporteraient à la recherche des cols (de la route) qui passent rarement au sommet de la montagne. Le théorème des extremums liés permet bien souvent de résoudre le problème de la recherche des extremums liés à l'aide des multiplicateurs de Lagrange.

2 Théorème des extremums liés

En fonction des leçons, ce développement n'est pas à réaliser de la même manière. Selon la version du développement présentée, le théorème des extremums liés est plus ou moins démontré.

Schéma du développement (leçons 151 et 214)

Ce qui est important dans ces leçons est de faire apparaître les utilisation de la dimension (et ses calculs) ou la notion de sous-variété. On démontre donc le lemme et le théorème (de manière précise car on manipule la dimension finie et la notion de rang). On se base alors sur la démonstration du théorème donnée par [1, p.103] même si l'énoncé est pris soit dans [2, p.20] ou dans [4, p.343].

— Preuve du lemme

— Preuve du théorème

1. Montrer que M est une sous-variété.
2. Montrer que $T_a(M) = \cap_{i=1}^k \ker(Dg_i(a))$. On pose $T = \cap_{i=1}^k \ker(Dg_i(a))$.
   (a) $T_a(M)$ est un espace vectoriel de même dimension que M.
   (b) Montrons que la dimension de T est $n - k$.
   (c) Montrons que $T_a(M) \subseteq T$.
   (d) On conclut grâce à la dimension.
3. Montrer que $T_a(M) \subseteq \ker(Df(a))$.
4. Appliquer le lemme pour conclure.

Schéma du développement (leçons 152, 159 et 219)

Pour ces leçons, on va traiter le lemme (qui justifie sa présence dans la leçon 159), le théorème et l'application (qui justifie sa présence dans la leçon 152) donnant un exemple d'application du théorème (intéressant dans le cadre de la leçon 219). La preuve du théorème est alors un peu moins poussé, on se base plutôt sur la preuve dans [2, p.20].

- Preuve du lemme
- Preuve du théorème
  1. Dire que M est une sous-variété.
  2. Dire que $T_a(M) = T = \cap_{i=1}^k \ker(Dg_i(a))$.
  3. Montrer que $T_a(M) \subseteq \ker(Df(a))$.
  4. Appliquer le lemme pour conclure.
- Preuve de l'application
  1. Montrer que le maximum de det sur l'ensemble X est atteint et est positif.
  2. Montrer s que si le maximum de det est atteint en $(v_1, \dots, v_n)$ les $(v_i)_i$ forment une base orthogonale de E.
  3. Réciproquement, montrer que si les $(v_i)_i$ forment une base orthogonale de E, $\det(v_1, \dots, v_n) \in \{-1, 1\}$ selon l'orientation de la base

Lemme. *Soient φ, $\varphi_1, \dots, \varphi_n$ des formes linéaires sur un espace vectoriel E de dimension n.*

$$\varphi \in \mathrm{Vect}(\varphi_1, \dots, \varphi_n) \iff \cap_{i=1}^n \ker \varphi_i \subseteq \ker \varphi$$

Démonstration. $\Rightarrow$ Supposons que $\varphi \in \mathrm{Vect}(\varphi_1, \dots, \varphi_n)$, on en déduit qu'il existe des $(\lambda_i)_{i \in [\![1,n]\!]} \in E$ tels que $\varphi = \sum_{i=1}^n \lambda_i \varphi_i$. Soit $x \in \cap_{i=1}^n \ker \varphi_i$, alors pour tout $1 \leq i \leq n$, $\varphi_i(x)$. On a alors $\varphi(x) = 0$ (car $\varphi(x) = \sum_{i=1}^n \lambda_i \varphi_i(x)$) et donc $x \in \ker \varphi$. Donc $\cap_{i=1}^n \ker \varphi_i \subseteq \ker \varphi$.

$\Leftarrow$ **[5, p.295]** Supposons que $\cap_{i=1}^n \ker \varphi_i \subseteq \ker \varphi$. On peut également supposer que φ est non nulle, sans quoi le résultat est évident. Notons F le sous-espace de E^* engendré par les $(\varphi_1, \dots, \varphi_n)$ et D la droite de E^* engendrée par g. On a alors :

$$\begin{array}{rcll}
F^\circ & = & \{x \in E \mid \forall i \in]\!]1,k]\!], \varphi_i(x) = 0\} & \text{(Définition de l'othogonalité au sens de la dualité)} \\
& = & \cap_{i=1}^n \ker \varphi_i & \text{(Définition du noyaux et de l'intersection)} \\
& \subseteq & \ker \varphi & \text{Hypothèse} \\
& = & \{x \in E \mid \varphi(x) = 0\} & \text{(Définition du noyaux)} \\
& = & D^\circ & \text{(Définition de l'othogonalité au sens de la dualité)}
\end{array}$$

En prenant l'orthogonal de F° et D° (qui appartiennent à E) dans E, $(D^\circ)^\perp \subseteq (F^\circ)^\perp$ (car $F^\circ \subseteq D^\circ$ et que la dualité échange l'inclusion).

Comme $\dim E < \infty$, pour tout sous-espace vectoriel de E^*, $(G^\circ)^\perp = G$, donc $D \subseteq F$. On en déduit que $\varphi \in \mathrm{Vect}(\varphi_1, \dots, \varphi_n)$.

□

Théorème (Théorème des extremums liés). *Soient U un ouvert $\mathbb{R}^n$ et $f, g_1, \dots, g_k : U \to \mathbb{R}$ de classe $\mathcal{C}^1$. Posons $M = \{w \in U, g_1(w) = \dots = g_k(w) = 0\}$. Si la restriction de f à M admet un extremum local en $a \in U$, et si la famille $(Dg_i(a))_{i \in [\![1,k]\!]}$ est linéairement indépendante, alors il existe $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ tels que $\sum_{i=1}^k Dg_i(a)\lambda_i = Df(a)$. Les $(\lambda_i)_{i \in [\![1,k]\!]}$ sont appelés les multiplicateurs de Lagrange.*

Démonstration. Conservons les notations du théorème

Étape 1 : montrons que M est une sous-variété [1, p.93, definition 3]

- Posons $g = (g_1, \dots, g_k)$, on a alors $U \cap M = g^{-1}(\{O_{\mathbb{R}^k}\})$ (comme image réciproque de 0).

— M est une sous-variété si et seulement si $\forall x \in M$, $Dg(x)$ est surjective (dans la troisième définition des sous-variétés, l'hypothèse sur $U \cap M$ est déjà vérifié (remarque précédente)).

— $\forall x \in M$, $Dg(x) = \begin{pmatrix} Dg_1(x) \\ \vdots \\ Dg_k(x) \end{pmatrix}$. La famille $(Dg_i(a))_{i \in I}$ est libre, elle forme alors une base de $\mathbb{R}n - k$ et donc $\forall x \in M$, $Dg(x)$ est surjective car elle est de rang $-k$.

Donc M est une sous-variété de dimension $n - k$.

Étape 2 : montrons que $T_a(M) = \cap_{i=1}^k \ker(Dg_i(a))$ **[1, p.98, théorème 2.3]** Notons $T = \cap_{i=1}^k \ker(Dg_i(a))$. Montrons alors que $T_a(M) = T$.

Étape a : $T_a(M)$ **est un espace vectoriel de même dimension que** M **[1, p.98, théorème 2.2]** D'après la première définition des plans tangents (carte locale), $D\varphi(a)(T_a(M)) = \mathbb{R}^d \times \{0\}$ (composition par $D\varphi(a)$). Donc $T_a(M)$ est un espace vectoriel de dimension $d = \dim M = n - k$.

Étape b : montrons que la dimension de T **est** $n - k$ Posons $g = (g_1, \ldots, g_k)$. On a $T = \ker g$ et $\mathbb{R}^k = \mathrm{Im} g$ car g est surjective. Donc par le théorème du rang appliqué à g : $\dim \mathbb{R}^n = \dim(\mathrm{Im} g) + \dim(\ker g)$. On en déduit que $n = \dim R^k + \dim T$, soit $\dim T = n - k$.

Étape c : montrons que $T_a(M) \subseteq T$ Soit $v \in T_a(M)$.

— Il existe un intervalle ouvert I contenant 0 et une application différentielle $\gamma : I \to M$ telle que $\gamma(0) = a$ et $\gamma'(0) = v$ (définition du plan tangent).

— Pour tout $i \in [\![1, k]\!]$ et tout $t \in I$, on a $g_i(\gamma(t)) = 0$ (car $\forall t \in I$ $\gamma(t) \in M$ et par définition de γ).

— Pour tout $i \in [\![1, k]\!]$ et tout $t \in I$, $Dg_i(\gamma(t))(\gamma'(t)) = 0$ (car $Dg_i(\gamma(t)) = 0$ et on applique le théorème de la différentielle composée).

— On évalue en 0 et on obtient que pour tout $i \in [\![1, k]\!]$, $Dg_i(\gamma(0))(\gamma'(0)) = 0 = Dg_i(a)(v)$. Donc $v \in T$.

Donc, comme $\dim T_a(M) = \dim T$ et que $T_a(M) \subseteq T$, on en conclut que $T_a(M) = T$.

Étape 3 : montrons que $T_a(M) \subseteq \ker(Df(a))$ **[1, p.98, théorème 2.4]** Soit $v \in T_a(M)$. Alors, il existe un intervalle ouvert I contenant 0 et une application différentielle $\gamma : I \to M$ telle que $\gamma(0) = a$ et $\gamma'(0) = v$ (définition du plan tangent).

— Pour tout $t \in I$, $f|_M \circ \gamma(t) = f \circ \gamma(t)$ (car pour tout $t \in I$, $\gamma(t) \in M$).

— Cette fonction admet un extremum pour $t = 0$ (car $\gamma(0) = a$ et a est un extremum de f par hypothèse, si $t \neq 0$, on obtient une contradiction).

— On en déduit que $t \mapsto Df(\gamma(t))(\gamma'(t))$ s'annule en 0 car $f(\gamma(t))$ admet un extremum en 0. Donc $Df(a)(v) = 0$ (évaluation en 0) et $v \in \ker Df(a)$.

En particulier $\cap_{i=1}^k \ker(Dg_i(a)) \subseteq \ker D(f)$. On peut alors appliquer le lemme, ce qui nous permet de conclure. □

Application (Inégalité de Hadamard [4, p.380])**.** *Soit $E = \mathbb{R}^n$ muni du produit scalaire usuel noté $< . >$. Pour tout $x_1, \ldots, x_n \in E$, on a l'inégalité suivante :*

$$|\det(x_1, \ldots, x_n)| \leq \|x_1\| \ldots \|x_n\|$$

avec égalité si et seulement si les $(x_i)_{1 \leq i \leq n}$ forment une base orthogonale de E.

Démonstration. Notons $X = \{(v_1, \ldots, v_n) \in E^n, \|v_1\| = \cdots = \|v_1\| = 1\}$.

Étape 1 : montrons que le maximum de det **sur l'ensemble** X **est atteint et est positif**

— X, produit des sphère unité de E, est un compacte de E^n. De plus, det est continue sur E^n (et même de classe $\mathcal{C}^\infty$ car c'est une fonction polynomiale). Donc, la fonction det atteint son maximum sur X.

— Notons e_i les vecteurs de la base canonique de $\mathbb{R}^n$. Comme $(e_1, \dots, e_n) \in X$ (car leur norme vaut 1 par définition) et que $\det(e_1, \dots, e_n) = 1$ (définition du déterminant de la base canonique), le maximum de det sur X vaut au moins 1.

Étape 2 : montrons que si le maximum de det **est atteint en** $(v_1, \dots, v_n)$ **les** $(v_i)_i$ **forment une base orthogonale de** E

— Ce maximum est un extremum de det lié par les conditions

$$\begin{cases} g_1(v_1, \dots, v_n) \\ \quad\vdots \\ g_n(v_1, \dots, v_n) \end{cases} \quad \text{où } g_i(v_1, \dots, v_n) = \|v_i\|^2 - 1$$

— $\forall (v_1, \dots, v_n) \in X$ et $\forall (h_1, \dots, h_n) \in E^n$, $Dg_i(v_1, \dots, v_n)(h_1, \dots, h_n) = 2 \langle v_i, h_i \rangle$. Elle sont donc indépendantes comme formes linéaires.

— X est une sous-variété de $E^n = \mathbb{R}n^2$ de dimension $n^2 - n$.

— Le théorème des extremums liés nous assure l'existence de λ_i tels que $D\det(v_1, \dots, v_n)(h_1, \dots, h_n) = \sum_{i=1}^n \lambda_i \langle v_i, h_i \rangle$, pour tout $h_i \in E$.

— Par la linéarité de det, pour tout v_i, on a $D\det(v_1, \dots, v_n)(0, \dots, 0, h_i, 0, \dots, 0) = \det(v_i, \dots, v_{i-1}, h_i, v_{i+1}, \dots, v_n)$. Donc $D\det(v_1, \dots, v_n)(h_1, \dots, h_n) = \sum_{i=1}^n \lambda_i \langle v_i, h_i \rangle$ est équivalent à $\det(v_i, \dots, v_{i-1}, h_i, v_{i+1}, \dots, v_n) = \lambda_i \langle v_i, h_i \rangle$ (on somme les différentielles partielles qui sont continues).

— On en déduit que si $h_i = v_i$, $\lambda_i = f(v_1, \dots, v_n) > 0$ et $v_i.v_j = 0$.

Étape 3 : réciproquement, montrons que si les $(v_i)_i$ **forment une base orthogonale de** E, $\det(v_1, \dots, v_n) \in \{-1, 1\}$ **selon l'orientation de la base**

— Par les étapes 1 et 2, le maximum de det sur X vaut 1 et est atteint en $(v_1, \dots, v_n)$ si et seulement si les $(x_i)_{1 \le i \le n}$ forment une base orthogonale de E.

— Par homogénéité en chacun des v_i, $\det(v_1, \dots, v_n) \le \|v_1\| \dots \|v_n\|$. On en déduit le résultat si les v_i sont tous nuls ou s'ils forment une base orthogonale normalisée.

□

3 Compléments autours des extremums liés

Le théorème d'inversion locale

Le théorème d'inversion local consiste en la résolution de l'équation $y = f(x)$ en inversant f, soit $x = f^{-1}(y)$ [4, p.187]. Avant d'énoncer le théorème, rappelons la définition d'un $\mathcal{C}^k$-difféomorphisme.

Application $\mathcal{C}^k$-difféomorphisme [3, p.707]

Définition (Difféomorphisme global)**.** Soient E et F deux espaces vectoriels normés, U est un ouvert de E et V un ouvert de F. Soit $k \in \mathbb{N}^* \cup \{+\infty\}$.

1. Une application $f : U \to V$ est un $\mathcal{C}^k$-difféomorphisme global (ou simplement $\mathcal{C}^k$-difféomorphisme) si f est bijective de U dans V, et si f^{-1} est $\mathcal{C}^k$ sur V.
2. On dit que U et V sont $\mathcal{C}^k$-difféomorphe s'il existe un $\mathcal{C}^k$-difféomorphisme de U dans V.

Définition (Difféomorphisme local)**.** Soient E et F deux espaces vectoriels normés, U est un ouvert de E et $k \in \mathbb{N}^* \cup \{+\infty\}$.

1. Une application $f : U \to F$ est un $\mathcal{C}^k$-difféomorphisme local en $x_0 \in U$ s'il existe un voisinage ouvert V de x_0 et un voisinage W de $f(x_0)$ tels que $f|_V : V \to W$ est un $\mathcal{C}^k$-difféomorphisme.
2. Une application $f : U \to F$ est un $\mathcal{C}^k$-difféomorphisme local U, si f est un $\mathcal{C}^k$-difféomorphisme local en tout point de U.

Les variantes du théorème d'inversion locale

Théorème (Théorème d'inversion locale)**.** *Soient U un ouvert de $\mathbb{R}^n$, a un point de U, et $f : U \to \mathbb{R}^n$ une application de classe $\mathcal{C}^1$. On suppose que la matrice jacobienne $Df(a)$ est inversible (i.e.* $\det Df(a) \neq 0$*). Il existe alors un ouvert V contenant a (et contenu dans U) et un ouvert W contenant $b = f(a)$, tels que f (restreinte à V) soit un difféomorphisme de classe $\mathcal{C}^1$ de V sur $W = f(V)$.*

Idée de la preuve. On peut développer $f(x)$ à l'ordre un puis appliqué l'inverse de $Df(a)$. □

Démonstration. Utilisation d'un point fixe : on se ramène à un problème de point fixe en posant $F(x) = x - Df(a)^{-1}(f(x) - y)$. La solution sera donc une limite d'une suite récurrente définie par F. On conclut en vérifiant la régularité de la solution obtenue. □

Remarque : On peut se placer dans des Banach en supposant que $df(a)$ est bijective de E dans F.

Interprétation : On a ainsi : ($x \in V$ et $y = f(x)$) $\Leftrightarrow$ ($y \in W$ et $x = f^{-1}(y)$) en notant $f^{-1} : W \to V$ la réciproque de f restreinte à V. L'équation $y = f(x)$ admet alors une unique solution dans V, mais elle peut en avoir d'autre en dehors de V (on peut avoir $V \neq U$).

Remarque : On a $D\left(f^{-1}\right)(y) = Df(x)^{-1}$ pour $x \in V$.

En pratique, l'application f se présente sous la forme $f(x) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))$ et l'hypothèse à vérifiée est

$$\det Df(a) = \det \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} (a_1, \dots, a_n) \neq 0$$

Notons que cette hypothèse est nécessaire par la dérivée de composition.

Exemple ([4, p.202]) : Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ tel que $f : (x, y) \mapsto (s, p) = (x + y, xy)$. Alors $U = \{x > y\}$ sont les ouverts connexes maximaux tels que $f : \Omega \to f(\Omega)$ soit un difféomorphisme.

Contre-exemple ([4, p.204]) : Soit $f(x) = x + x^2 \sin \frac{\pi}{2}$ si $x \neq 0$ et $f(0) = 0$ est dérivable mais pas inversible localement.

Remarque : On obtient le même résultat en prenant $\mathcal{C}^k$ à la place de $\mathcal{C}^1$.

Théorème (Théorème d'inversion globale)**.** *Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ de classe $\mathcal{C}^1$. On suppose que f est injective sur U et que, pour tout $u \in U$ la matrice jacobienne $Df(u)$ est inversible (i.e.* $\det Df(u) \neq 0$*). Alors, l'ensemble image $f(U)$ est un ouvert de $\mathbb{R}^n$ et f est un $\mathcal{C}^1$-difféomorphisme de U sur $f(U)$.*

Remarque : L'hypothèse d'injectivité est nécessaire. En pratique la montrer revient souvent à calculer un inverse. Ce théorème perd alors tout intérêt.

Contre-exemple ([4, p.204]) : Soit $f(x, y) = (x^2 - y^2, 2xy)$ est une difféomorphisme local au voisinage de tout point de U mais pas un difféomorphisme global.

Application (Inverse globale [4, p.221]) : Soient $k > 0$ une constante et $f : \mathbb{R}^n \to \mathbb{R}^n$ une application de classe $\mathcal{C}^1$ supposée k-dilatante, i.e. $|f(x) - f(y)| \geq k|x - y|$, pour tout $x, y \in \mathbb{R}^n$. Alors, f est un difféomorphisme global de $\mathbb{R}^n$ sur lui-même.

Démonstration.
— f est injective : $f(x) = f(y)$ entraîne (via les normes) que $x = y$.
— $f(\mathbb{R}^n)$ est un fermé de $\mathbb{R}^n$: caractérisation séquentielle.
— $Df(x)$ inversible pour tout x : calcul des normes de la définition.
— $f(\mathbb{R}^n)$ est un ouvert de $\mathbb{R}^n$: Inversion locale au voisinage d'un point.
— Conclure : connexité de $\mathbb{R}^n$ montre f surjective et on conclut par le théorème d'inversion globale.

□

Remarque : Ce théorème est un cas particulier du théorème d'Hadamard–Lévy. L'hypothèse k-dilatante entraîne que f est propre et que $Df(x)$ est inversible en tout point.

Théorème (Théorème d'inversion version holomorphe). *Soient U un ouvert connexe de $\mathbb{C}$ et $f : U \to \mathbb{C}$ une fonction holomorphe. On suppose que f est injective sur U. Alors, l'image $f(U)$ est un ouvert de $\mathbb{C}$ et f est un difféomorphisme holomorphe de U sur $f(U)$, i.e. f^{-1} est holomorphe sur $f(U)$.*

Remarque : Il est remarquable qu'on est besoin d'aucune hypothèse sur la dérivée de f : l'injectivité de celle-ci suffit à montrer que f' ne s'annule pas. Sinon, le développement en série entière de f indiquerait qu'elle se comporte comme x^k, $k \geq 2$ qui n'est pas injective.

Application (Inversion de la fonction holomorphe [4, p.234]) : Soit f une fonction holomorphe sur le disque unité $|z| < 1$. On suppose $f(0) = 0$, $f'(0) = 1$ et $|f'(z)| < M$ pour $|z| < 1$. Alors, f est un difféomorphisme holomorphe du disque $|z| < R$ sur un ouvert contenant le disque $|w| < \frac{R}{2}$.

Démonstration.
— Il existe $R \in]0,1[$ tel que $|f'(z) - 1| \leq \frac{|z|}{R}$ pour $|z| < 1$: on applique le principe du maximum à $\frac{f'(z)-1}{z}$ et on choisi $R = \frac{1}{(M+1)}$.
— f injective sur le disque $|z| \leq R$: compare les valeurs de $f(z) - z$ en deux points.
— $|f(z) - z| \leq \frac{|z|^2}{2R}$ pour $|z| < 1$: calcul d'intégrale
— Conclure : application du théorème de Rouché à la fonction $z \mapsto f(z) - w$ puis le théorème d'inversion locale version holomorphe.

□

Application en analyse Nous donnons ici quelques applications de ce théorème. Ces applications viennent de problèmes d'analyses.

Théorème (Théorème de changement de coordonnées). *Soient $f_1, \dots, f_n$ des fonctions numériques de classes $\mathcal{C}^1$ au voisinage d'un point a de $\mathbb{R}^n$. Les relations $u_1 = f_1(x_1, \dots, x_n), \dots, u_n = f_n(x_1, \dots, x_n)$ définissent un changement de coordonnées sur un voisinage de a si et seulement si le déterminant du Jacobien n'est pas nul, c'est-à-dire, si les différentielles $Df_1(a), \dots, Df_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$.*

Démonstration. Reformulation du théorème d'inversion locale. □

Exemple ([4, p.71]) : Coordonnées polaires

Corollaire. *Soient $f_1, \dots, f_p$ des fonctions numériques de classes $\mathcal{C}^1$ au voisinage d'un point a de $\mathbb{R}^n$. On peut les compléter par des fonctions $f_{p+1}, \dots, f_n$ en un changement de coordonnées au voisinage de a sur $\mathbb{R}^n$ si et seulement si les différentielles $Df_1(a), \dots, Df_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$.*

Remarque : Ce résultat est une généralisation du théorème de la base incomplète en algèbre linéaire. Il permet de simplifier un problème en changeant les coordonnées de fonctions qui y jouent un rôle important (résolution d'équations aux dérivées partielles).

Proposition (Perturbation de l'identité [3, p.714]). *Soient E un espace de Banach et $g : E \to E$ une application $\mathcal{C}^1$. S'il existe $M > 0$ tel que, pour tout $x \in E$, $|d_x g|_{\mathcal{L}(E)} \leq M$, alors, pour tout $\epsilon \in \left]0, \frac{1}{M}\right[$, l'application $f_\epsilon = I_E + \epsilon g$ est un $\mathcal{C}^1$-difféomorphisme de E dans E.*

Démonstration.
— f_ϵ est bijective : on se ramène à un résultat des points fixe et on montre que $z - \epsilon g(z)$ est contractante via le théorème des accroissements finis.
— f_ϵ est bijective : on applique le théorème d'inversion locale à f_ϵ.

□

Proposition (Isométrie de $\mathbb{R}^{n*}$ [3, p.715]). *Soient $n \geq 2$ et $f : \mathbb{R}^{n*} \to \mathbb{R}^{n*}$ une application $\mathcal{C}^1$ telle que $\forall x \in \mathbb{R}^{n*}$, $Df(x) \in GL_n(\mathbb{R})$ et $|f(x)|_2 = |x|$. Alors f est surjective.*

Démonstration.
— On se restreint au cas où l'isométrie est définie sur la sphère S_r de centre l'origine et de rayon r qui est connexe.
— $f(S_r)$ est fermé : image d'un compact.
— $f(S_r)$ est ouvert : théorème d'inversion locale.

□

Théorème (Lemme de Morse à deux variables [4, p.330]). *Soit U un ouvert de $\mathbb{R}^2$ tel que $0 \in U$. Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la forme quadratique $D^2(f)(0,0)$ est non-dégénérée.*

1. *Si la signature de $D^2(f)(0,0)$ est $(2,0)$ alors il existe u et v tels que $f(x,y) - f(0,0) - df(0,0)(x,y) = u(x,y)^2 + v(x,y)^2$.*

2. *Si la signature de $D^2(f)(0,0)$ est $(1,1)$ alors il existe u et v tels que $f(x,y) - f(0,0) - df(0,0)(x,y) = u(x,y)^2 - v(x,y)^2$.*

Démonstration. — Développement de Taylor à l'ordre un avec reste intégrale.

— On effectue une réduction de Gauss et par positivité du déterminant (via la signature du cas 1), on obtient les fonction u et v. On conclut avec le théorème d'inversion locale.

— Si la signature est $(+,-)$ on se ramène au cas précédent, si elle est $(+,+)$ on décompose par Gauss. Sinon (elle est $(-,-)$) et on applique ce qui précède à $-f$.

□

Application ([4, p.341]) : Étude affine locale d'une surface via le lemme de Morse : décompose selon la signature.

Application en algèbre et en géométrie Nous donnons ici quelques applications de ce théorème. Ces applications viennent de problèmes d'algèbre.

Théorème (Réduction des formes quadratiques [4, p.209])**.** *On note E l'espace des matrices réelles $n \times n$ et S le sous-espace des matrices symétriques. On fixe $A_0 \in S$ inversible. Soit $\varphi : E \to S$ l'application définie par $\varphi(M) =^t MA_0M$. Alors, il existe un voisinage V de A_0 dans S et une application de $A \mapsto M \in V$ de classe $\mathcal{C}^1$ telle que $A)^tMA_0M$.*

Démonstration. — L'application φ est de classe $\mathcal{C}^1$ car polynomiale de noyau de $D\varphi(I)$ est les matrices H telles que A_0H soit antisymétrique.

— Décomposition des matrices en somme d'une matrice symétrique et une matrice antisymétrique : supplémentaire du noyau de $D\varphi(I)$ est dans F. On restreint alors φ à F et on applique le théorème d'inversion locale.

□

Remarque : Pour toute forme quadratique suffisamment proche d'une forme quadratique non dégénérée, elle lui est équivalente (on s'y ramène par changement de base). Elles ont en particulier la même signature. Donc les matrice de même signature donnée forment un ouvert de S.

De plus, en appliquant ce résultat à $A_0 = I$, on voit que tout matrice A symétrique suffisamment proche de l'identité admet une racine carré symétrique.

Application (Lemme de Morse à n variables [4, p.354]) : Soit U un ouvert de $\mathbb{R}^2$ tel que $0 \in U$. Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la forme quadratique $D^2(f)(0,0)$ est non-dégénérée de signature $(p, n-p)$ et que $Df(0) = 0$. Alors, il existe un difféomorphisme $x \mapsto u = \varphi(x)$ entre deux voisinage de l'origine, de classe $\mathcal{C}^1$, tel que $\varphi(0) = 0$ et $f(x) - f(0) = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2$.

Idée de la preuve. — Formule de Taylor à l'ordre un avec reste intégral : existence d'une matrice Q symétrique.

— Réduction des formes quadratiques : décomposition de $Q(x) =^t M(x)Q(0)M(x)$.

— Théorème d'inversion local sur la matrice M.

□

Le TIL transforme un système quelconque (de classe $\mathcal{C}^1$) de deux équations à deux inconnues se discute comme un système linéaire de rang le rang de la matrice jacobienne.

Application (Système de deux équations à deux inconnues [4, p.211]) : Soient U un ouvert de $\mathbb{R}^2$ et $\varphi : U \to \mathbb{R}^2$ une application de classe $\mathcal{C}^1$ telle que $\varphi(x,y) = (f(x,y), g(x,y))$. Étudions le système d'équation $f(x,y) = u$, $g(x,y) = v$ où u et v sont données.

— Si $D\varphi$ est de rang 2 en un point de U, alors il existe un ouvert V (donné par le TIL) tel qu'il existe une unique solution.

— Si $D\varphi$ est de rang 1 en tout point de U, alors le système possède soit une infinité de solution, soit aucune.

— Si $D\varphi$ est de rang 0 en tout point de U, alors le système possède une solution si et seulement si $u = A$ et $v = B$; tout point $(x,y) \in U$ est alors solution.

Théorème (Surjectivité de l'exponentielle de matrice [6, p.48])**.** *Soit $A \in \mathcal{M}_n(\mathbb{C})$, on a* $\exp(\mathbb{C}[A]) = \mathbb{C}\times[A]$.

Corollaire. *1.* $\exp(\mathcal{M}_n(\mathbb{C})) = GL_n(\mathbb{C})$

2. $\exp(\mathcal{M}_n(\mathbb{R})) = \{A^2,\ A \in GL_n(\mathbb{R})\}$

3. $GL_n(\mathbb{C})$ ne possède pas de sous-groupe arbitrairement petit.

Le théorème des fonctions implicites

Le théorème des fonctions implicites consiste en la résolution de l'équation $f(x,y) = 0$ en exprimant y comme une fonction en c, soit $y = \varphi(x)$ [4, p.187]. Avant d'énoncer le théorème, rappelons la définition d'un $\mathcal{C}^k$-difféomorphisme.

Les variantes du théorème des fonctions implicites [4, p.192]

Théorème (Théorème des fonctions implicites)**.** *Soient U un ouvert de $\mathbb{R}^n \times \mathbb{R}^p$, (a,b) un point de U, et $f : (x,y) \mapsto f(x,y)$ une application de classe $\mathcal{C}^1$ de U dans $\mathbb{R}^p$. On suppose que $f(a,b) = 0$ et que la matrice jacobienne $D_y f(a,b)$, formée des dérivées partielles par rapport à y, est inversible, i.e.* $\det D_y f(a,b) \neq 0$.

Alors, l'équation $f(x,y) = 0$ peut être résolue localement par rapport aux variables y : il existe V (voisinage ouvert de a dans $\mathbb{R}^n$), W (voisinage ouvert de b dans $\mathbb{R}^p$), avec $V \times W \subset U$ et $D_y f(x,y)$ inversible pour tout $(x,y) \in V \times W$, et une unique application $\varphi : V \to W$ telle que $(x \in V, y \in W$ et $f(x,y) = 0) \Leftrightarrow (x \in V$ et $y = \varphi(x))$. De plus φ est de classe $\mathcal{C}^1$ sur V.

Idée de la preuve. On développe $f(x,y)$ à l'ordre un. L'équation ainsi obtenu est admet une unique solution si et seulement si le jacobien en y est inversible. On peut alors calculer cette différentielle. □

Démonstration. On se ramène au théorème d'inversion locale que nous appliquons sur $g(x,y) = (x, f(x,y))$. La fonction implicite φ est donnée par la deuxième composante de φ calculée à l'aide d'un point fixe. □

Remarque : Les théorème d'inversion locale et des fonctions implicites sont équivalents : dans l'autre sens on prend la fonction implicite de $f(x) - y$ qui donne l'inverse.

Définition. On appelle φ la fonction implicite définie par f au voisinage du point de départ (a,b). En particulier, $\varphi(a) = b$.

Remarque : On peut obtenir une version $\mathcal{C}^k$ de ce théorème en remplaçant les $\mathcal{C}^1$ par des $\mathcal{C}^k$. De même, on obtient une version holomorphe en remplaçant les $\mathbb{R}^n$ par des $\mathbb{C}$ et les $\mathcal{C}^1$ par des fonctions holomorphes.

Interprétation : Ce théorème se traduit géométriquement par le fait que la courbe définie implicitement par l'équation $f(x,y) = 0$ peut être localement perçu comme le graphe d'une fonction φ. Ce résultat est bien local car si on agrandit trop W l'unicité de φ peut être compromise par la présente d'un deuxième point d'abscisse a. Sous les bonnes hypothèses, on peut avoir deux fonctions implicites pour chacun des points.

Pratique : On a un système d'équations tel qu'un certain nombre de variables soit considéré comme des inconnues (les autres sont données). On souhaite exprimer ce système uniquement en fonction des inconnues selon une hypothèse de départ. Il nous suffit alors de regarder la non nullité du déterminant des dérivées partielles des fonctions données par rapport aux inconnues. On peut ainsi appliquer le théorème.

*Différentielle de la fonction implicite*L'existence de la différentielle de la fonction implicite est donnée par le théorème. Son calcul est un calcul de dérivée composé et de la résolution d'un système de Crammer.

Exemple ([4, p.237]) : Folium de Descartes

Remarque : Le théorème des fonctions implicites, le théorème d'inversion locale, le théorème de Cauchy–Lipschitz et le théorème du point fixe de Picard sont équivalents (les uns permettent de montrer les autres).

Applications dans les mathématiques et autres

Application (Résolution approchée d'une équation) : Donnons une valeur approchée de la solution réelle de l'équation : $x^7 + 0.99x - 2.03 = 0$. On trouve $x \simeq 1.005$. Pour ce faire, on applique le théorème des fonctions implicites à la fonction $f(x) = x^7 + x - 2$ (via la différentielle on a une précision à l'ordre un).

Application : Soit $f(x,\epsilon) = (x-a)(b-x) + \epsilon x^3$ où $a < b$ sont fixés et ϵ est un paramètre. Donnons un développement asymptotique de l'intégrale $I(\epsilon) = \int_{x_1(\epsilon)}^{x_2(\epsilon)} \frac{dx}{\sqrt{f(x,\epsilon)}}$ où x_1 et x_2 sont deux racines. On a alors $I(\epsilon) = \pi + \frac{3\pi}{4}(a+b)\epsilon + O(\epsilon^2)$.

- On applique le théorème des fonctions implicites au voisinage de $(a,0)$ ou $(b,0)$ pour trouver un développement asymptotique de f à l'ordre deux.
- Changement de variable avec la demi-somme et la demi-différence des deux racines.

Remarque : Le théorème des fonctions implicites joue un rôle important en physique : il a été l'origine de la première considération de la théorie de la relativité générale. Il permet également d'expliquer Kepler.

Les sous-variétés

Les sous-variétés sont une notion permettant d'uniformiser la théorie autours des objets appelés surface (parabole, ellipse, cylindre, tore, ...) [4, p.197]. Elle se concentre sur l'aspect local des sous-ensembles de $\mathbb{R}^n$: ce sont des sous-espaces affines propres qui ont été transformés.

Sous-variétés

Définition. Soient V un sous-ensemble de $\mathbb{R}^n$, $a \in V$, et $d \in \mathbb{N}$. On dit que V est lisse en a, de dimension d, s'il existe un difféomorphisme F de classe $\mathcal{C}^1$ d'un voisinage ouvert U de a dans $\mathbb{R}^n$ sur le voisinage ouvert $F(U)$ de 0 dans $\mathbb{R}^n$, qui transforme V en un sous-espace vectoriel de dimension d : $F(V \cap U) = V' \cap F(U)$ où $V' = \mathbb{R}^d \times \{0\} \subset \mathbb{R}^n$.

On dit que V est une sous-variété de dimension d de $\mathbb{R}^n$ si V est lisse (de dimension d) en chacun de ces points.

Remarque : Une sous-variété se ramène localement à un sous-espace vectoriel par simple changement de coordonnées. Par définition, ces notions (sous-variété, lisse) sont invariantes par difféomorphisme. On peut étendre la définition pour un $\mathcal{C}^k$-difféomorphisme.

Théorème (Théorème des sous-variétés). *Soient V un sous-ensemble de $\mathbb{R}^n$, $a \in V$ et d un entier naturel. Les quatre propriétés suivantes sont équivalentes :*

1. *(carte locale) V est lisse en a, de dimension d (il existe un $\mathcal{C}^p$-difféomorphisme local $\varphi : U \to \mathbb{R}^n$ tel que $\varphi(V \cap U) = \varphi(U) \cap \left[\mathbb{R}^k \times \{0\}^{n-k}\right]$).*
2. *(équation) Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$ et $n-d$ fonctions $f_i : U \to \mathbb{R}$ de classe $\mathcal{C}^1$, telles que $(x \in V \cap U) \Leftrightarrow (x \in U$ et $f_1(x_1, \dots, x_n) = 0, \dots, f_n(x_1, \dots, x_n) = 0)$ et les différentielles $Df_1(a), \dots, Df_{n-d}(a)$ sont indépendantes (existence d'une fonction surjective telle que son image réciproque en 0 soit $U \cap V$).*
3. *(graphe) Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, un voisinage ouvert U' de $(a_1, \dots, a_d)$ dans $\mathbb{R}^d$ et $n-d$ fonctions $g_i : U' \to \mathbb{R}$ de classe $\mathcal{C}^1$, telles que (après permutation éventuelle des coordonnées) $(x \in V \cap U) \Leftrightarrow \begin{cases} (x_1, \dots, x_d) \in U' \text{ et} \\ x_{d+1} = g_1(x_1, \dots, x_d), \dots, x_n = g_{n-d}(x_1, \dots, x_d) \end{cases}$ (consiste en un changement de coordonnées donné par les g_i).*
4. *(nappe paramétrée) Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, un voisinage ouvert Ω de (0 dans $\mathbb{R}^d$ et n fonctions $\varphi_i : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^1$, telles que l'application $\varphi : u = (u_1, \dots, u_d) \mapsto x = (\varphi_1(u), \dots, \varphi_n(u))$ soit un homéomorphisme de de Ω sur $V \cap U$, avec $a = \varphi(0)$ et que la matrice jacobienne $D\varphi(0)$ soit injective (de rang r) (φ est bijective et bi-continue dont la différentielle en 0 est injective).*

Idée de la preuve. **Implication** $1 \Rightarrow 2$ On pose $f_i = x_{d+1}(\varphi(x))$. On montre l'indépendance linéaire via le calcul d'une composée de différentielle.

Implication $2 \Rightarrow 3$ Application du théorème des fonctions implicites pour résoudre le système d'équations $f_1(x) = \cdots = f_{n-d}(x) = 0$ par rapport à $x_{d+1}, \dots, x_n$.

Implication $3 \Rightarrow 1$ On pose $\varphi : x = (x_1, x_2) \mapsto (x_1, x_2 - u(x_1))$ où u est le vecteur composé des g_i. On conclut en appliquant le théorème de la différentielle des fonctions composées.

Implication $3 \Rightarrow 4$ On pose $\varphi : \tilde{z} \mapsto (z_0 + \tilde{z}, u(z_0 + \tilde{z}))$ où u est le vecteur composé des g_i.

Implication $4 \Rightarrow 3$ Application du théorème d'inversion locale pour inverser $(\varphi_1, \dots, \varphi_d)$ ce qui nous donne $u_1, \dots, u_d$ fonctions de $x_1, \dots, x_d$. Ensuite on les reporte dans $(\varphi_{d+1}, \dots, \varphi_n)$ pour avoir $x_{d+1}, \dots, x_n$ fonction de $x_1, \dots, x_d$. □

Espace tangent

Définition. Soient V un sous-ensemble de $\mathbb{R}^n$ et $a \in V$. Un vecteur $v \in \mathbb{R}^n$ est dit tangent en a à V s'il existe une fonction dérivable $\gamma : I \to \mathbb{R}^n$ où I est un intervalle ouvert autours de 0 telle que $\gamma(I) \subset V$, $\gamma(0) = a$ et $\gamma'(0) = v$.

Théorème. *Si V est lisse en a, de dimension d, ses vecteurs tangents en a forment un sous-espace vectoriel de dimension d, appelé espace vectoriel tangent en a à V, noté $T_a V$.*

Démonstration. Ce résultat n'a rien d'évident avec cette définition. Si on utilise la caractérisation par équation, on obtient le résultat par noyau d'une application linéaire. Le point difficile de la preuve "classique" de ce résultat est la stabilité par addition [3, p.749]. La fonction ainsi définie nous permet également de donnée une base. □

Remarque : On utilise souvent le plan tangent affine, parallèle au précédent mais passant par a.

Théorème (Caractérisation du plan tangent). *Soient V un sous-ensemble de $\mathbb{R}^n$, $a \in V$ et d un entier naturel. En utilisant les notations du théorème des sous variétés, les quatre propriétés suivantes sont équivalentes :*

1. *(carte locale)* $T_aV = d\varphi(x_0)^{-1}\left(\mathbb{R}^k \times \{0\}^{n-k}\right)$
2. *(équation)* $T_aV = \ker(Df(a))$
3. *(graphe)* $T_aV = \{(h, Dg(a_1, \ldots, a_d)), h \in \mathbb{R}^n\}$
4. *(nappe paramétrée)* $T_aV = D\varphi(O)(\mathbb{R}^n)$

Démonstration. **Carte locale** On construit des chemins dans $\mathbb{R}^k \times \{0\}^{n-k}$

Nappe paramétrée $\beta : t \mapsto j^{-1} \circ \gamma(t)$ est un chemin. $\gamma : t \mapsto \varphi(tw)$ où $v = D\varphi(0)(w)$.

Graphe et équation On utilise la preuve du théorème des sous-variétés.

□

Références

[1] A. Avez. *Calcul différentiel*. Masson, 1983.

[2] V. Beck, J. Malik, and G. Peyré. *Objectif Agrégation*. H et K, 2004.

[3] J.-P. Marco ; P. Thieullen ; J.-A.Weil. *Mathématiques pour la L2*. Pearson education, 2007.

[4] F. Rouvière. *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*. Cassini, 3^{me} edition, 2009.

[5] S.Nicolas S. Francinou, H. Gianella. *Oraux X-ENS, Algèbre 1*. Cassini.

[6] M. Zavidovique. *Un max de maths*. Calvage & Mounet, 2013.