Anaël Marit ENS Rennes

LEMME DE MORSE

- 157, 171, 206, 214, 215, 218 –

—

On sait, de par le théorème de Sylvester, qu'une forme quadratique réelle q admet toujours une base q-orthogonale de l'espace dans laquelle la matrice de q a pour forme :

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1}$$

En d'autres termes, pour un vecteur x qu'on écrirait $x = \sum x_1 e_1 + \dots x_d e_d$ dans une telle base, on obtiendrait :

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 \tag{2}$$

On cherche maintenant à s'intéresser à des objets un peu plus sauvages que des formes quadratiques : des fonctions deux fois différentiables. Si f est une telle fonction, qu'on suppose par ailleurs nulle et de différentielle nulle en 0, la formule de Taylor-Young donne :

$$\forall x \in \mathbb{R}^d,\ f(x) = \mathrm{d}^2 f(0) \cdot (x, x) + o(\|x\|) \tag{3}$$

Ce qui semble indique qu'au voisinage de 0, f se comporte sensiblement comme sa hessienne (qui est une forme quadratique d'après le lemme de Schwarz). On est alors en droit de se demander s'il existe toujours un changement de coordonnées, qu'on ne cherchera plus linéaire (ce serait bien trop restrictif !) mais C^1, pour lequel l'équation (2) tienne encore. On va voir ici qu'en rajoutant quelques hypothèses sur f, c'est effectivement le cas !

On considère un entier d.

Etant donnée une application f n-fois différentiable en a de $\mathbb{R}^d$ dans lui-même, on notera $\mathrm{d}^n f(a)$ sa différentielle n-ième, qui s'identifie à une application n-linéaire. Dans le cas particulier $n = 2$, la différentielle seconde s'identifie à une forme bilinéaire symétrique de matrice dans la base canonique de $\mathbb{R}^d$ $H_f(a)$, de sorte qu'on puisse écrire :

$$\mathrm{d}^2 f(a) \cdot (x, x) = {}^t x \cdot H_f(a) \cdot x \tag{4}$$

où l'on assimile x avec le vecteur colonne de ses coordonées dans la base canonique. Nous allons démontrer ici :

Théorème 1 (lemme de Morse ([2], exercice 114)). *Soit $f : U \to \mathbb{R}^d$ une application de classe C^3, avec U un ouvert de $\mathbb{R}^d$. On fait les hypothèses suivantes :*

1. *$f(0) = 0$ et $\mathrm{d}f(0) = 0$.*
2. *La forme quadratique $x \mapsto \mathrm{d}^2 f(0) \cdot (x, x)$ est non dégénérée. On notera $(p, d-p)$ sa signature.*

Il existe V et W deux voisinages ouverts de l'origine et $u : V \to W$ un C^1-difféomorphisme envoyant 0 sur 0 tel que, en notant (u_i) les composantes de u (i), on ait :

$$\forall x \in V,\ f(x) = u_1(x)^2 + \cdots + u_p(x)^2 - u_{p-1}(x)^2 - \cdots - u_d(x)^2 \tag{5}$$

Réduction des formes quadratiques et calcul différentiel

Avant de se lancer à l'assaut du lemme de Morse, on va chercher à mieux comprendre la réduction des formes quadratiques sous le prisme du calcul différentiel, car on l'aura compris, c'est bien au croisement de ces deux domaines que se trouve notre résultat.

Lemme 2 ([2], exercice 66). *Soit $A_0 \in GL_d(\mathbb{R}) \cap S_d(\mathbb{R})$ une matrice symétrique inversible. Il existe V un voisinage ouvert de A_0 dans $S_d(\mathbb{R})$ et $\Phi : V \to GL_d(\mathbb{R})$ un C^∞ difféomorphisme tel que :*

$$\forall A \in V,\ A = {}^t\Phi(A) \cdot A_0 \cdot \Phi(A) \ \wedge\ \Phi(A) = I_d \tag{6}$$

Remarque : La forme matricielle de ce lemme peut faussement donner l'impression qu'il s'agit d'un résultat d'algèbre linéaire, mais celui-ci se comprend beaucoup mieux en utilisant le langage des formes quadratiques. En effet, une matrice symétrique peut s'interpréter comme la matrice d'une forme quadratique plutôt que d'une application linéaire. Ici, on encadre la matrice A_0 entre une matrice inversible et sa transposée, ce qui correspond au changement de bases pour des formes quadratiques, et non pour des applications linéaires. Ce qu'exprime ce lemme, c'est que les matrices « proches » de A_0 (qui sont dans le voisinage V) sont congruentes à A_0, c'est-à-dire qu'elles représentent la même forme quadratique dans des bases différentes, et que par ailleurs le changement de base peut se faire de manière lisse. ♦

Démonstration. Pas de grand mystère, ce lemme ne cache pas le chemin de sa démonstration : la solution va venir du théorème d'inversion locale. On considère l'application :

$$\begin{aligned} \psi : GL_d(\mathbb{R}) &\to S_d(\mathbb{R}) \\ M &\mapsto {}^tM \cdot A_0 \cdot M \end{aligned} \tag{7}$$

Cette application est C^∞, car toutes ses composantes sont polynomiales en les coefficients de M. Par ailleurs, on calcule sa différentielle à vue :

$$\forall M \in GL_d(\mathbb{R}),\ \forall H \in \mathbb{M}_d(\mathbb{R}),\ \mathrm{d}\psi(M) \cdot H = {}^tH \cdot A_0 \cdot M + {}^tM \cdot A_0 \cdot H \tag{8}$$

(i). Par « composantes de u », on entend les fonctions (u_i) telles que, si $(e_1, \ldots, e_d)$ est la base canonique de $\mathbb{R}^d$, alors $u_i = e_i^* \circ u$. En d'autres termes, $u(x) = \sum u_i(x) e_i$.

On s'intéresse au noyau de cette différentielle en l'identité. Prenons $H \in \mathbb{M}_d(\mathbb{R})$. On a :

$$\begin{aligned}
H \in \operatorname{Ker}(\mathrm{d}\psi(I_d)) &\iff {}^tH \cdot A_0 = -A_0 \cdot H && (9)\\
&\iff {}^t(A_0 \cdot H) = -A_0 \cdot H \qquad \text{car } A_0 \text{ est symétrique} && (10)\\
&\iff H \in A_0^{-1}\mathfrak{A}_d(\mathbb{R}) && (11)
\end{aligned}$$

avec $\mathfrak{A}_d(\mathbb{R})$ l'ensemble des matrices antisymétriques de taille d. Or, on a la classique décomposition $\mathbb{M}_d(\mathbb{R}) = S_d(\mathbb{R}) \oplus \mathfrak{A}_d(\mathbb{R})$. Le théorème du rang implique donc que $\mathrm{d}\psi(I_d)$ est surjective. Par ailleurs, si W est un supplémentaire de $A_0^{-1}\mathfrak{A}_d(\mathbb{R})$ dans $\mathbb{M}_d(\mathbb{R})$, $\mathrm{d}[\psi(I_d)_{|W}]$ est inversible. Ainsi, par théorème d'inversion locale, $\psi_{|W}$ est un C^∞-difféomorphisme local au voisinage de I_d. Puisque $\psi(I_d) = A_0$, on obtient l'application Φ en prenant la réciproque de $\psi_{|W}$ là où elle est définie. □

Le lemme de Morse

Allons-y pour la réduction plus générale de f ! Comme f est C^3, elle est C^2 et on peut utiliser la formule de Taylor avec reste intégal à l'ordre 2 :

$$\forall x \in \mathbb{R}^d,\ f(x) = \cancelto{0}{f(0)} + \cancelto{0}{\mathrm{d}f(0) \cdot x} + \int_0^1 (1-\theta)\mathrm{d}^2 f(\theta x) \cdot (x, x)\mathrm{d}\theta \tag{12}$$

Histoire de se raccrocher plus facilement aux formes quadratiques, passons en écriture matricielle, qui pour une fois s'avérera bien utile :

$$\forall x \in \mathbb{R}^d,\ f(x) = {}^tx \cdot \underbrace{\int_0^1 (1-\theta) H_f(\theta x)\mathrm{d}\theta}_{=:Q(x)} \cdot x \tag{13}$$

Or l'application Q est à valeurs dans les matrices symétriques, car définie comme une intégrale de matrices symétriques. Comme f est de classe C^3, elle est par ailleurs C^1 en vertu du théorème de différentiation sous l'intégrale (ii) : si K est un compact de $\mathbb{R}^d$, l'application $(\theta, x) \mapsto (1-\theta)H_f(\theta x)$ est continûment différentiable par rapport à x, et l'application $(\theta, x) \mapsto (1-\theta)\mathrm{d}_x[H_f(\theta x)]$ (à valeurs dans $\mathscr{L}_c(\mathbb{R}^d, S_d(\mathbb{R}))$) est dominée uniformément sur K, en norme d'opérateur, par une constante, qui est intégrable sur $[0, 1]$.

Considérons un voisinage $\tilde{V}$ de $Q(0) = \frac{1}{2}H_f(0)$ (matrice inversible car $H_f(0)$ est supposée non dégénérée) et une fonction Φ C^∞ de V dans $GL_d(\mathbb{R})$ fournie par le lemme précédent. Comme Q est continue, il existe un voisinage V de l'origine telle que si $x \in V$, alors $Q(x) \in \tilde{V}$, et on a alors :

$$\forall x \in V,\ Q(x) = {}^t\Phi(Q(x)) \cdot Q(0) \cdot \Phi(Q(x)) \tag{14}$$

On peut alors appliquer le théorème d'inertie de Sylvester à $H_f(0) = 2Q(0)$: il existe B une matrice inversible telle que :

$$Q(0) = {}^t\left(\frac{1}{\sqrt{2}}B\right) \cdot \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \cdot \left(\frac{1}{\sqrt{2}}B\right) \tag{15}$$

(ii). J'ai l'impression que cette étape est souvent balayée par un rapide « f est C^3 donc il est trivial que Q est C^1, mais la réalité me semble un peu plus subtile que ça. J'ai ajouté en annexe ledit théorème de différentiation qui n'est pas si connu que cela, histoire d'avoir bien en tête les outils du calcul différentiel qui sont invoqués.

Posons alors :

$$\begin{aligned} u : V &\to \mathbb{R}^d \\ x &\mapsto \frac{1}{\sqrt{2}} B \cdot \Phi(Q(x)) \cdot x \end{aligned} \tag{16}$$

u est de classe C^1 et $u(0) = 0$. On a de plus :

$$\forall x \in V,\ f(x) = {}^t x \cdot Q(x) \cdot x = {}^t u(x) \cdot \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \cdot u(x) \tag{17}$$

Ce qui est exactement l'expression souhaitée. Il reste à montrer que u induit un C^1 difféomorphisme d'un voisinage de l'identité sur un autre voisinage de l'identité. On va encore une fois appliquer le théorème d'inversion locale (iii) :

$$\begin{aligned} \forall h \in V,\ u(0+h) &= \frac{1}{\sqrt{2}} B \cdot (\Phi(Q(0)) + \mathrm{d}[\Phi \circ Q](0) \cdot h + o(\|h\|)) \cdot h \quad (18) \\ &= u(0) + B \cdot I_d \cdot h + O(\|h\|^2) \quad (19) \end{aligned}$$

d'où u a pour jacobienne en 0 la matrice B, inversible par hypothèse. Le théorème d'inversion locale s'applique, ce qui achève la démonstration du lemme de Morse.

Remarque : Si on augment la régularité de f, la régularité de Q augmente également. Ainsi, si f est supposée C^{2+n} avec $n \in \mathbb{N}^*$, la même preuve permet d'affirmer que u peut être choisi C^n. ♦

Annexe : différentiation sous l'intégrale

On a l'habitude de dérivée sous l'intégrale lorsqu'on manipule des fonctions définies comme intégrale dépendant d'un paramètre réel, voire complexe. Mais ici, notre paramètre est un vecteur de $\mathbb{R}^d$, et donc ce n'est pas le même théorème qui s'applique (même si en dimension finie on peut s'y ramener) ! En fait, le bon critère est le suivant :

Théorème 3 (Différentiation sous l'intégrale ([1], théorème 12.13)). *Soit E et G deux e.v.n. de dimension finie et soit $(X, \mathfrak{F}, \mu)$ un espace mesuré. Soit $f : E \times X \to G$ telle que :*

1. $\forall v \in E,\ f(x, -) \in L^1(\mu)$
2. *pour presque tout $x \in X$, $f(-, x)$ est de classe C^1.*
3. *pour tout compact K de E, il existe $g_K \in L^1(\mu)$ telle que pour tout $(v, x) \in K \times X$, $|||\mathrm{d}_x f(v, x)||| \leqslant g_k(x)$.*

Alors l'application :

$$F : v \mapsto \int_X f(v, x) \mathrm{d}\mu(x) \tag{20}$$

est de classe C^1 sur E et on a :

$$\forall v \in E,\ \forall h \in E,\ \mathrm{d}F(v) \cdot h = \int_X \mathrm{d}_x f(v, x) \cdot h \, \mathrm{d}\mu(x) \tag{21}$$

(iii). Cette étape là me parait également négligée dans la référence...

Démonstration. En dimension finie, on s'épargne l'abstraction du calcul différentiel en revenant aux dérivées partielles et en utilisant l'équivalence des normes. En effet, pour tout $x \in X$, $f(-,x)$ admet des dérivées partielles selon toutes ses coordonnées, et en exploitant l'équivalence de la norme d'opérateur avec d'autres normes sur $\mathscr{L}_c(E,G)$ où la dépendance aux coordonnées est plus explicite, on obtient que toutes les dérivées partielles sont dominées uniformément sur les compacts de E par des applications L^1 sur X. Le théorème de dérivation sous l'intégrale s'applique alors et F admet toutes ses dérivées partielles, et celles-ci sont continues et s'obtiennent en dérivant directement sous l'intégrale. Ceci implique que F est C^1 et qu'on peut intervertir différentielle et intégrale. □

Remarque : Le théorème reste vrai en dimension infinie. Il faut alors faire une démonstration directe, sans utiliser la dérivation sous l'intégrale, en exploitant l'inégalité des accroissements finis. Ce n'est pas spécialement plus compliqué, mais c'est beaucoup plus abstrait, alors j'ai préféré me limiter ici à la dimension finie, qui nous suffit amplement. ♦

Références

[1] Jean-Yves Ouvrard. *Probabilités 2.* Cassini, 2000. isbn : 978-2-84225-144-4.

[2] François Rouvière. *Petit guide du calcul différentiel.* Cassini, 2015.