CHAPITRE V. — SOUS-VARIETES DE $\mathbb{R}^n$

On introduit dans ce chapitre des sous-ensembles particuliers d'espaces vectoriels sur lesquels on peut développer une théorie intéressante du calcul différentiel. Ces ensembles, les sous-variétés, généralisent les graphes de fonctions, comme nous le verrons ci-dessous.

11 Définitions

Soit E un espace vectoriel normé de dimension finie. Un sous-ensemble $\mathrm{X} \subset \mathrm{E}$ est une *sous-variété* de classe $\mathcal{C}^k$, $k \geq 1$, si, pour tout $x \in \mathrm{X}$, il existe un voisinage ouvert U de x dans E, un voisinage ouvert V de 0 dans E, un sous-espace vectoriel fermé $\mathrm{F} \subset \mathrm{E}$ et un difféomorphisme $\varphi : \mathrm{U} \to \mathrm{V}$ tels que

$$\varphi^{-1}(\mathrm{F} \cap \mathrm{V}) = \mathrm{X} \cap \mathrm{U}\,.$$

Cela signifie que X est localement un sous-espace vectoriel quitte à faire un changement de coordonnées approprié. On dit que φ *redresse* X en un sous-espace vectoriel. Cette notion généralise celle de graphes de fonctions, cf. l'exemple 1.

Exemple 1.— On considère dans $\mathbb{R}^n$, muni de la norme euclidienne, le graphe d'une fonction $f : \mathbb{R}^p \to \mathbb{R}^q$, $p + q = n$, de classe $\mathcal{C}^k$:

$$\mathrm{X} = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^q,\ y = f(x)\}\,.$$

L'application $\psi : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^n$ définie par $(x, y) \mapsto (x, f(x) + y)$ est un difféomorphisme sur son image puisque ψ est injective, de classe $\mathcal{C}^k$ et

$$\mathrm{D}_{(x,y)}\psi = \begin{pmatrix} \mathrm{Id}_p & 0 \\ \mathrm{D}_x f & \mathrm{Id}_q \end{pmatrix}$$

De plus, on a $\psi(\mathbb{R}^p \times \{0\}) = \mathrm{X}$, donc l'inverse $\varphi \stackrel{\text{def.}}{=} \psi^{-1}$ convient (à translation près).

Exemple 2.— La sphère unité pour la norme euclidienne

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n,\ \|x\| = 1\}\,.$$

Pour tout $a \in \mathbb{S}^{n-1}$, $a_n \neq 0$, on peut considérer l'application

$$\Phi : (x_1, \ldots, x_{n-1}, x_n) \mapsto \left(x_1, \ldots, x_{n-1}, x_n \pm \sqrt{1 - \sum_{j=1}^{n-1} x_j^2}\right)$$

où on choisit le même signe que a_n et qui est bien définie sur

$$\mathrm{U} = \left\{(x_1, \ldots, x_{n-1}, x_n) \in \mathbb{R}^n,\ \sum_{j=1}^{n-1} x_j^2 < 1\right\}\,.$$

On note $f(x) = \sqrt{1 - \sum_{j=1}^{n-1} x_j^2}$ et on calcule

$$\mathrm{D}_a\Phi = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ \pm\mathrm{D}_a f & 1 \end{pmatrix}$$

On vérifie ainsi que $\Phi : \mathrm{U} \to \mathrm{V} \stackrel{\text{def.}}{=} \Phi(\mathrm{U})$ est un difféomorphisme local au voisinage de a et $\Phi(\mathbb{R}^{n-1} \times \{0\} \cap \mathrm{U}) = \mathbb{S}^{n-1} \cap \mathrm{V}$. Si $a_n = 0$, alors on peut trouver une autre composante a_j non nulle et procéder de la même manière.

L'espace tangent en un point d'une sous-variété est un objet fondamental qui nous permettra de développer le calcul différentiel sur les sous-variétés. Il généralise les tangentes des graphes de fonctions.

Définition 11.1 (espace tangent). — *Soit* X *une sous-variété d'un evn* E *et soit* $x \in \mathrm{X}$*. Le plan tangent* $\mathrm{T}_x\mathrm{X}$ *de* X *au point* x *est l'ensemble des vecteurs* v *de* E *tels que*

$$\lim_{t \to 0} \frac{1}{t} d(x + tv, \mathrm{X}) = 0\,.$$

On rappelle que si $\mathrm{A} \subset \mathrm{E}$ et $x \in \mathrm{A}$ alors

$$d(x.\mathrm{A}) \stackrel{\text{def.}}{=} \inf_{a \in \mathrm{A}} d(x,a) = \inf_{a \in \mathrm{A}} \|a - x\| \,.$$

Lemme 11.2. — *Soit* X *une sous-variété. Pour tout* $x \in \mathrm{X}$*, il existe* $r > 0$ *tel que, pour tout* $y \in \mathrm{B}(a,r)$*, il existe* $z \in \mathrm{X}$ *tel que* $d(y,z) = d(y,\mathrm{X})$*.*

Démonstration. — Prenons $x \in \mathrm{X}$. Il existe un voisinage U de x et V de 0, un sous-espace vectoriel F ainsi qu'un difféomorphisme $\varphi : \mathrm{U} \to \mathrm{V}$ tels que $\varphi^{-1}(\mathrm{F} \cap \mathrm{V}) = \mathrm{X} \cap \mathrm{U}$. Comme U est ouvert, il existe $r > 0$ tel que $\mathrm{B}(x, 3r) \subset \mathrm{U}$. Prenons maintenant $y \in \mathrm{B}(x,r)$. On a $d(y,\mathrm{X}) \leq d(x,y) < r$. Prenons une suite $(z_n)_n$ de X telle que $\lim \|y - z_n\| = d(y,\mathrm{X})$. Remarquons que si $z \in \mathrm{X} \setminus \overline{\mathrm{B}(x,2r)}$, alors $\|y - z\| \geq \|z - x\| - \|x - y\| \geq 2r - r \geq r > d(y,\mathrm{X})$, donc (z_n) appartient à $\mathrm{B}(x,2r)$. Or $\varphi(\overline{\mathrm{B}(x,2r)} \cap \mathrm{X})$ est un compact de F, donc le théorème de Bolzano-Weierstrass nous permet d'extraire une sous-suite $(z_{n_k})_k$ de (z_n) telle que $\varphi(z_{n_k})$ tend vers un point $p \in \mathrm{F} \cap \mathrm{V}$ car $\varphi(\overline{\mathrm{B}(x,2r)}) \subset \mathrm{V}$. Par continuité de φ^{-1}, la sous-suite $(z_{n_k})_k$ tend vers le point $z = \varphi^{-1}(p) \in \mathrm{X}$ et $\|z - y\| = \lim \|z_{n_k} - y\| = d(y,\mathrm{X})$. ■

On a les descriptions suivantes de l'espace tangent en un point.

Proposition 11.3. — *Soient* X *une sous-variété et* $x \in \mathrm{X}$*.*

1. *Un vecteur* $v \in \mathrm{E}$ *est tangent au point* x *si, et seulement si, il existe une application différentiable* $c : \mathrm{I} \to \mathrm{E}$ *définie sur un intervalle contenant* 0 *telle que* $c(\mathrm{I}) \subset \mathrm{X}$, $c(0) = x$ *et* $c'(0) = v$*.*

2. *Si* X *est donnée au voisinage de* x *par un difféomorphisme* $\varphi : \mathrm{V} \to \mathrm{U}$*, avec* $\mathrm{X} \cap \mathrm{V} = \varphi^{-1}(\mathrm{F})$*, alors*

$$\mathrm{T}_x\mathrm{X} = (\mathrm{D}_x\varphi)^{-1}(\mathrm{F})\,.$$

Démonstration. — On considère d'abord X donnée par un difféomorphisme $\varphi : \mathrm{V} \to \mathrm{U}$. Soit $w \in \mathrm{F}$ et posons $c(t) = \varphi^{-1}(tw)$ qui définit un chemin de X. Un développement limité à l'ordre 1 montre

$$c(t) = \varphi^{-1}(tw) = x + t(\mathrm{D}_x\varphi)^{-1}(w) + o(t)$$

donc

$$d(x + t(\mathrm{D}_x\varphi)^{-1}(w), \mathrm{X}) \leq \|x + t(\mathrm{D}_x\varphi)^{-1}(w) - c(t)\| = o(t)$$

et on en déduit que $c'(0) = (\mathrm{D}_x\varphi)^{-1}(w) \in \mathrm{T}_x\mathrm{X}$. Donc $(\mathrm{D}_x\varphi)^{-1}(\mathrm{F}) \subset \mathrm{T}_x\mathrm{X}$.

Faisons l'observation suivante. Si $c : \mathrm{I} \to \mathrm{X}$ vérifie $c(0) = x$ et $c'(0)$ existe, alors $\mathrm{D}_x\varphi(c'(0)) \in \mathrm{F}$. En effet, comme $c(t) \in \mathrm{X}$ pour tout t, on a $\varphi(c(t)) \in \mathrm{F}$ et

$$\mathrm{D}_x\varphi(c'(0)) = (\varphi \circ c)'(0) = \lim_{t \to 0} \frac{1}{t}(\varphi \circ c)(t) \in \mathrm{F}$$

car F est fermé (les sous-espaces vectoriels d'un espace vectoriel normé de dimension finie sont fermés). Du coup, $c'(0) \in (\mathrm{D}_x\varphi)^{-1}(\mathrm{F})$ est tangent à X au point x.

Soit v un vecteur tangent à X au point x. D'après le lemme 11.2, il existe $c(t) \in \mathrm{X}$ tel que $d(x + tv, \mathrm{X}) = \|x + tv - c(t)\| = o(t)$, donc $c(t) = x + tv + o(t)$ ce qui signifie que $c(0) = x$ et $c'(0)$ existe et vaut v. L'observation précédente implique $v \in (\mathrm{D}_x\varphi)^{-1}(\mathrm{F})$. Ceci montre que $(\mathrm{D}_x\varphi)^{-1}(\mathrm{F}) = \mathrm{T}_x\mathrm{X}$, donc établit le point **2.**, mais aussi **1.** car on a vu que tout vecteur $v \in (\mathrm{D}_x\varphi)^{-1}(\mathrm{F})$ s'exprimait comme le vecteur dérivé d'un chemin contenu dans X. ■

12 Différentes présentations des sous-variétés

On donne deux nouvelles définitions équivalentes de sous-variétés. Pour cela, on introduit deux classes de fonctions: les immersions et les submersions.

Définition 12.1 (immersion). — *Une application* $j : \mathrm{U}\,(\subset \mathrm{E}) \to \mathrm{F}$ *de classe* $\mathcal{C}^k$*,* $k \geq 1$*, est une immersion en* $x \in \mathrm{U}$ *si* $\mathrm{D}_x j \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *est injective.*

Remarque 12.2. — *En dimension finie, si* $\mathrm{L} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *est injective, alors* L *admet un inverse à gauche, c'est-à-dire une application linéaire* $\mathrm{L}' \in \mathcal{L}(\mathrm{F},\mathrm{E})$ *tel que* $\mathrm{L}' \circ \mathrm{L} = \mathrm{Id}_\mathrm{E}$*, cf. le lemme 10.1.*

Etant donnée une application différentiable $f : \mathbb{R}^p \to \mathbb{R}^q$, l'application $j : x\,(\in \mathbb{R}^p) \mapsto (x, f(x)) \in \mathbb{R}^{p+q}$ est une immersion car

$$\mathrm{D}j = \begin{pmatrix} \mathrm{Id}_p \\ \mathrm{D}f \end{pmatrix}$$

qui est bien de rang p.

Proposition 12.3. — *Si* X *est une sous-variété de* E *de classe* $\mathcal{C}^k$*,* $k \geq 1$*, et si* $x \in \mathrm{X}$*, alors il existe un voisinage* V *de* x *une immersion* $j : \mathrm{U} \to \mathrm{E}$ *définie au voisinage de l'origine telles que*

$$\mathrm{X} \cap \mathrm{V} = j(\mathrm{U})\,.$$

Démonstration. — On considère un voisinage W de $0 \in \mathbb{R}^n$, un sous-espace vectoriel fermé $\mathrm{F} \subset \mathbb{R}^n$ et un difféomorphisme $\varphi : \mathrm{V} \to \mathrm{W}$ tels que $\varphi^{-1}(\mathrm{F} \cap \mathrm{W}) = \mathrm{X} \cap \mathrm{V}$. On pose $p = \dim \mathrm{F}$ et $q = n - p$. Notons $\psi = \varphi^{-1} : \mathrm{W} \to \mathrm{U}$. Quitte à faire un changement de base de $\mathbb{R}^n$, on peut supposer que $\mathrm{F} = \mathbb{R}^p \times \{0\}$.

On note $j : \mathrm{F} \to \mathrm{E}$ définie par $j = \psi|_\mathrm{F}$, qui est de classe $\mathcal{C}^k$ car j est une restriction de ψ. Comme ψ est un difféomorphisme local, $\mathrm{D}\psi$ est injective, donc $\mathrm{D}j$ aussi. Par construction, $j(\mathrm{F} \cap \mathrm{W}) \subset \mathrm{X}$. ■

On s'intéresse maintenant à la réciproque de cette proposition qui nous fournira une autre manière de définir une sous-variété.

Proposition 12.4. — *Soit* $j : \mathrm{U} \to \mathrm{E}\,(\simeq \mathbb{R}^n)$ *de classe* $\mathcal{C}^k$*, où* U *est un ouvert d'un evn* $\mathrm{G} \simeq \mathbb{R}^p$*. Si* j *est une immersion en* $x \in \mathrm{G}$*, alors il existe un voisinage* V *de* x *tel que* $\mathrm{X} = j(\mathrm{V})$ *soit une sous-variété de* F *et* $\mathrm{T}_{j(x)}\mathrm{X} = \mathrm{Im}\,\mathrm{D}_x j$*.*

Démonstration. — Comme j est une immersion en x on a $p \leq n$. Le rang de j est constant au voisinage de x puisque l'on peut trouver un mineur de $\mathrm{D}_x j$ de taille p non nul. Par continuité de $\mathrm{D}j$ et du déterminant, ce mineur reste non nul au voisinage de x. Comme $\mathrm{rg}\, j \leq p$, le rang ne peut augmenter.

Par conséquent, le théorème du rang constant implique l'existence de difféomorphismes locaux φ et ψ, tels que $\varphi(x) = 0$, $\psi(j(x)) = 0$ et $\psi \circ j \circ \varphi^{-1} = \mathrm{D}_x j$.

Notons $\mathrm{F} = \mathrm{Im}\,\mathrm{D}_x j$. L'application $\psi : \mathrm{W} \to \mathrm{W}'$ est un difféomorphisme sur son image et

$$\psi^{-1}(\mathrm{F} \cap \mathrm{W}') = (\psi^{-1} \circ \mathrm{D}_x j)(\mathrm{V}') = (j \circ \varphi^{-1})(\mathrm{V}') = j(\mathrm{V})\,.$$

Par ailleurs

$$\mathrm{T}_{j(x)}\mathrm{X} = (\mathrm{D}_{j(x)}\psi)^{-1}(\mathrm{F}) = \mathrm{D}_0\psi^{-1} \circ \mathrm{D}_x j(\mathrm{G}) = \mathrm{D}_0(\psi^{-1} \circ \mathrm{D}_x j)(\mathrm{G}) = \mathrm{D}_0(j \circ \varphi^{-1})(\mathrm{G}) = \mathrm{Im}\,\mathrm{D}_x j\,.$$

■

Cette démonstration montre en particulier qu'une immersion est localement injective. Dans ce contexte, on dit que j fournit un *paramétrage local* de X.

Remarque 12.5. — *Si* $j : \mathrm{U}\,(\subset \mathbb{R}^p) \to \mathrm{X}$ *est un paramétrage local et* φ *un redressement de* X *en* $j(0)$*, alors* $\varphi \circ j : \mathrm{U} \to \mathrm{F}$ *est différentiable, et* $\dim \mathrm{F} = p$*. Cela découle du fait que* $j(\mathrm{U}) = \varphi^{-1}(\mathrm{F} \cap \mathrm{W})$*. Donc le théorème d'inversion locale montre que* $\varphi \circ j$ *est un difféomorphisme au voisinage de* 0 *et* $j = \varphi^{-1} \circ (\varphi \circ j)$ *est un homéomorphisme local.*

Du coup, si on a deux paramétrages $j_1, j_2 : \mathrm{U} \to \mathrm{X}$ *alors il existe un difféomorphisme* $\psi : \mathrm{U} \to \mathrm{U}$ *tel que* $j_2 \circ \psi = j_1$*. En effet, on peut écrire*

$$j_1 = j_2 \circ (\varphi \circ j_2)^{-1} \circ (\varphi \circ j_1)\,.$$

Définition 12.6 (submersion). — *Une application* $g : \mathrm{U}\,(\subset \mathrm{E}) \to \mathrm{F}$ *de classe* $\mathcal{C}^k$*,* $k \geq 1$*, est une submersion en* $x \in \mathrm{U}$ *si* $\mathrm{D}_x g \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *est surjective. On dit que* g *est une submersion au-dessus de* $y \in \mathrm{F}$ *si* $g^{-1}(\{y\})$ *est non vide et si* g *est une submersion en chaque point* $x \in g^{-1}(\{y\})$*.*

Remarque 12.7. —

1. *En dimension finie, si* $\mathrm{L} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *est surjective, alors* L *admet un inverse à droite, c'est-à-dire une application linéaire* $\mathrm{L}' \in \mathcal{L}(\mathrm{F},\mathrm{E})$ *tel que* $\mathrm{L} \circ \mathrm{L}' = \mathrm{Id}_\mathrm{F}$*, cf. le lemme 10.1.*

2. *Si* $g : \mathrm{U} \to \mathbb{R}$*, il suffit de vérifier que* $\mathrm{D}_x g \neq 0$ *pour s'assurer que* g *est une submersion en* x*.*

Proposition 12.8. — *Si* X *est une sous-variété de* E *de classe* $\mathcal{C}^k$*,* $k \geq 1$*, et si* $x \in \mathrm{X}$*, alors il existe un voisinage* V *de* x*, une submersion* $g : \mathrm{V} \to \mathbb{R}^q$ *au-dessus de l'origine telle que*

$$\mathrm{X} \cap \mathrm{V} = g^{-1}(\{0\}) \cap \mathrm{V}\,.$$

L'application $g : \mathbb{R}^n \to \mathbb{R}$ définie par $g(x) = \sum x_j^2 - 1$ est une submersion au-dessus de 0 puisque

$$\mathrm{D}_x g = \begin{pmatrix} 2x_1 & \ldots & 2x_n \end{pmatrix}$$

et si $\mathrm{D}_x g = 0$ alors $x = 0$, mais $g(0) \neq 0$. Cela remontre que $\mathbb{S}^{n-1}$ est une sous-variété.

Démonstration. — On considère un voisinage W de $0 \in \mathbb{R}^n$, un sous-espace vectoriel fermé $\mathrm{F} \subset \mathbb{R}^n$ et un difféomorphisme $\varphi : \mathrm{V} \to \mathrm{W}$ tels que $\varphi^{-1}(\mathrm{F} \cap \mathrm{W}) = \mathrm{X} \cap \mathrm{V}$. On note $p = \dim \mathrm{F}$ et $q = n - p$.

Quitte à faire un changement de base de $\mathbb{R}^n$, on peut supposer que $\mathrm{F} = \mathbb{R}^p \times \{0\}$. Notons $\pi : \mathbb{R}^n \to \mathbb{R}^p \times \mathbb{R}^q$ la projection sur $\{0\} \times \mathbb{R}^q$, que l'on identifie à $\mathbb{R}^q$, parallèlement à $\mathbb{R}^p \times \{0\}$. Posons $g = \pi \circ \varphi : \mathbb{R}^n \to \mathbb{R}^q$ qui est de classe $\mathcal{C}^k$. On vérifie que $g(z) = 0$ si et seulement si $\varphi(z) \in \operatorname{Ker} \pi$, soit $\varphi(z) \in \mathrm{F}$. Par conséquent, on a $\mathrm{X} \cap \mathrm{V} = g^{-1}(\{0\}) \cap \mathrm{V}$ pour V assez petit. De plus, $\mathrm{D}g = \pi \circ \mathrm{D}\varphi$. Comme $\mathrm{D}\varphi$ est un isomorphisme et $\operatorname{rg} \pi = q$, on en déduit que $\mathrm{D}g$ est surjective sur V, donc g est bien une submersion. ■

On s'intéresse maintenant à la réciproque de cette proposition.

Proposition 12.9. — *Soit* $g : \mathrm{U} \to \mathrm{G}\,(\simeq \mathbb{R}^q)$ *de classe* $\mathcal{C}^k$*, où* U *est un ouvert d'un evn* $\mathrm{E} \simeq \mathbb{R}^n$*. Si* g *est une submersion au-dessus de* y*, alors* $\mathrm{X} = g^{-1}(\{y\})$ *est une sous-variété et* $\mathrm{T}_a\mathrm{X} = \operatorname{Ker} \mathrm{D}_a g$ *pour tout* $a \in \mathrm{X}$*.*

Démonstration. — Comme g est une submersion au-dessus de y, on a $n \geq q$. Soit $a \in \mathrm{U}$ tel que $g(a) = y$. On remarque que $\operatorname{rg} g$ est constant au voisinage de a, car ce rang est maximal, donc on peut trouver un mineur de $\mathrm{D}_a g$ non nul de taille $q = \dim \mathrm{G}$. Ce mineur restera non nul au voisinage de a par continuité de $\mathrm{D}g$ et du déterminant.

Par conséquent, le théorème du rang constant implique l'existence de difféomorphismes locaux φ et ψ, tels que $\varphi(a) = 0$, $\psi(y) = 0$ et $\psi \circ g \circ \varphi^{-1} = \mathrm{D}_a g$.

Notons $\mathrm{F} = \operatorname{Ker} \mathrm{D}_a g$. L'application $\varphi : \mathrm{V} \to \mathrm{V}'\,(\subset \mathrm{G})$ est un difféomorphisme et $\varphi^{-1}(\mathrm{F} \cap \mathrm{V}') = g^{-1}(\{y\}) \cap \mathrm{V}$ puisque si $x \in \mathrm{V} \cap \mathrm{X}$, alors $\mathrm{D}_a g(\varphi(x)) = (\psi \circ g \circ \varphi^{-1})(\varphi(x)) = \psi(y) = 0$, donc $\varphi(x) \in \mathrm{F}$, et l'on a

$$\psi \circ g \circ \varphi^{-1}(\mathrm{F} \cap \mathrm{V}') = \mathrm{D}_a g(\mathrm{F} \cap \mathrm{V}') = 0 = \psi(y)\,.$$

De plus, on a $g = \psi^{-1} \circ \mathrm{D}_a g \circ \varphi$ donc

$$\operatorname{Ker} \mathrm{D}_a g = \operatorname{Ker} \mathrm{D}_a(\psi^{-1} \circ \mathrm{D}_a g \circ \varphi) = \operatorname{Ker}(\mathrm{D}_a g \circ \mathrm{D}_a \varphi) = (\mathrm{D}_a \varphi)^{-1}(\operatorname{Ker} \mathrm{D}_a g) = (\mathrm{D}_a \varphi)^{-1}(\mathrm{F})\,.$$

■

Notion de dimension. — On remarque que si X est connexe, alors tous les sous-espaces F qui redressent X ont même dimension. Du coup, on définit la *dimension de* X comme la dimension de ces sous-espaces F. Si X est présentée par une immersion $j : \mathrm{U}\,(\subset \mathbb{R}^p) \to \mathrm{X}$, alors $\dim \mathrm{X} = p$ et si X est présentée par une submersion $g : \mathbb{R}^n \to \mathbb{R}^q$ au-dessus d'un point, alors $\dim \mathrm{X} = n - q$. On dit aussi que X est de *codimension* q.

On illustre ces notions en dimension 3. Posons donc $\mathrm{E} = \mathbb{R}^3$.

1. Une sous-variété de dimension 0 est un ensemble localement fini de points.

2. Une sous-variété de dimension 1 est une courbe. Elle peut être donnée par une immersion $j : \mathrm{I}\,(\subset \mathbb{R}) \to \mathrm{E}$, c'est-à-dire telle que $j'(t) \neq 0$; on peut aussi considérer une submersion $g : \mathbb{R}^3 \to \mathbb{R}^2$, ce qui signifie que les différentielles $\mathrm{D}_x g_1$ et $\mathrm{D}_x g_2$ des fonctions coordonnées de g sont linéairement indépendantes.

3. Une sous-variété de dimension 2 est une surface. Elle peut être donnée par une immersion $j : \mathrm{U}\,(\subset \mathbb{R}^2) \to \mathrm{E}$, c'est-à-dire telle que les dérivées partielles $\partial_1 j$ et $\partial_2 j$ sont linéairement indépendantes; on peut aussi considérer une submersion $g : \mathbb{R}^3 \to \mathbb{R}$, ce qui signifie que $\mathrm{D}g \neq 0$.

4. Une sous-variété de dimension 3 est un ouvert de $\mathbb{R}^3$.

Proposition 12.10. — *Si* X_1 *et* X_2 *sont des surfaces de* $\mathbb{R}^3$ *et si, pour tout* $x \in \mathrm{X}_1 \cap \mathrm{X}_2$*, les plans tangents* $\mathrm{T}_x\mathrm{X}_1$ *et* $\mathrm{T}_x\mathrm{X}_2$ *sont différents, alors* $\mathrm{X}_1 \cap \mathrm{X}_2$ *est une courbe et* $\mathrm{T}_x(\mathrm{X}_1 \cap \mathrm{X}_2) = \mathrm{T}_x\mathrm{X}_1 \cap \mathrm{T}_x\mathrm{X}_2$.

Démonstration. — Soit $x \in \mathrm{X}_1 \cap \mathrm{X}_2$ et supposons que X_1 soit définie au voisinage de x par une submersion $g_1 : \mathrm{V}_1 \to \mathbb{R}$ au-dessus de 0 et X_2 par une submersion $g_2 : \mathrm{V}_2 \to \mathbb{R}$ au-dessus de 0. On note $\mathrm{V} = \mathrm{V}_1 \cap \mathrm{V}_2$ et on considère $g : \mathrm{V} \to \mathbb{R}^2$ définie par $g(y) = (g_1(y), g_2(y))$. On a $g(y) = 0$ si et seulement si $g_1(y) = g_2(y) = 0$ donc si et seulement si $y \in (\mathrm{X}_1 \cap \mathrm{X}_2) \cap \mathrm{V}$. Par ailleurs, $\mathrm{D}g_1$ et $\mathrm{D}g_2$ sont des formes linéairement indépendantes puisque les plans tangents sont distincts: si $\lambda \mathrm{D}_y g_1 + \mu \mathrm{D}_y g_2 = 0$, alors, en prenant un élément $v \in \mathrm{T}_x\mathrm{X}_2 \setminus \mathrm{T}_x\mathrm{X}_1$, on obtient $\lambda \mathrm{D}_y g_1(v) = 0$ donc $\lambda = 0$; de même $\mu = 0$. Donc g est une submersion et $\mathrm{X}_1 \cap \mathrm{X}_2$ définit une courbe.

De plus $\operatorname{Ker} \mathrm{D}g = \operatorname{Ker} \mathrm{D}g_1 \cap \operatorname{Ker} \mathrm{D}g_2$. ■

13 Applications différentiables

Soit $\mathrm{X} \subset \mathrm{E}$ une sous-variété de classe $\mathcal{C}^k$, on souhaite définir la différentiabilité de $f : \mathrm{X} \to \mathrm{F}$ et sa différentielle en un point $x \in \mathrm{X}$ qui sera une application linéaire $\mathrm{D}_x f \in \mathcal{L}(\mathrm{T}_x\mathrm{X}, \mathrm{F})$. Nous proposons différentes suggestions selon la manière dont est présentée X, et nous montrerons qu'elles sont toutes équivalentes.

Soient $x \in \mathrm{X}$, V un voisinage ouvert de x dans E et $f : (\mathrm{V} \cap \mathrm{X}) \to \mathbb{R}^n$ une application.

1. On dira que f est différentiable en x si f admet un prolongement $\tilde{f} : \mathrm{V} \to \mathbb{R}^n$ différentiable en x. On notera $\mathrm{D}_x f = \mathrm{D}_x \tilde{f}|_{\mathrm{T}_x\mathrm{X}}$. On dira que f est de classe $\mathcal{C}^m$ sur $\mathrm{X} \cap \mathrm{V}$, $1 \leq m \leq k$, si on peut choisir $\tilde{f}$ de classe $\mathcal{C}^m$ sur V.

2. Si X est donnée par un difféomorphisme local $\varphi : \mathrm{V} \to \mathrm{U}$ tel que $\varphi(x) = 0$ et $\mathrm{X} \cap \mathrm{V} = \varphi^{-1}(\mathrm{F})$, alors on dira que f est différentiable en x si $f \circ \varphi^{-1}|_{\mathrm{F}}$ est différentiable en x. On notera $\mathrm{D}_x f = \mathrm{D}_0(f \circ \varphi^{-1}) \circ \mathrm{D}_x \varphi$ sur $\mathrm{T}_x\mathrm{X}$. On dira que f est de classe $\mathcal{C}^m$ sur V, $1 \leq m \leq k$, si $f \circ \varphi^{-1}|_{\mathrm{F}}$ est de classe $\mathcal{C}^k$ sur $\mathrm{U} \cap \mathrm{F}$.

3. Si X est donnée par un paramétrage local $j : \mathrm{U}\,(\subset \mathbb{R}^p) \to \mathrm{X} \cap \mathrm{V}$, c'est-à-dire une immersion telle que $j(0) = x$, alors on dira que f est différentiable en x si $f \circ j$ est différentiable en 0. On notera $\mathrm{D}_x f = \mathrm{D}_0(f \circ j) \circ (\mathrm{D}_0 j)^{-1}$. On dira que f est de classe $\mathcal{C}^m$ sur V, $1 \leq m \leq k$, si $f \circ j$ est de classe $\mathcal{C}^m$ sur U.

La dernière définition a besoin d'un peu d'explication: la différentielle $\mathrm{D}_0 j \in \mathcal{L}(\mathbb{R}^p, \mathrm{E})$ est injective et son image est $\mathrm{T}_x\mathrm{X}$. Par conséquent, $\mathrm{D}_0 j \in \mathrm{GL}(\mathbb{R}^p, \mathrm{T}_x\mathrm{X})$, donc $(\mathrm{D}_0 j)^{-1}$ désigne l'inverse $(\mathrm{D}_0 j)^{-1} : \mathrm{T}_x\mathrm{X} \to \mathbb{R}^p$.

Théorème 13.1. — *Toutes ces définitions sont équivalentes, et la différentielle ne dépend pas des différents choix.*

Démonstration. — Soit $\tilde{f}$ un prolongement différentiable sur V et prenons $v \in \mathrm{T}_x\mathrm{X}$. Il existe un chemin $c : \mathrm{I} \to \mathrm{X}$ tel que $c(0) = x$ et $c'(0) = v$. L'application $\tilde{f} \circ c$ est donc différentiable et

$$\mathrm{D}_x \tilde{f}(v) = \mathrm{D}_0(\tilde{f} \circ c) = \lim_{t \to 0} \frac{1}{t}(f(c(t)) - f(x))$$

car $c(\mathrm{I}) \subset \mathrm{X}$ et donc $\tilde{f} \circ c = f \circ c$. Donc la différentiabilité de f en x ne dépend pas du prolongement $\tilde{f}$, ni $\mathrm{D}_x f$.

Prenons maintenant un difféomorphisme $\varphi : \mathrm{V} \to \mathrm{U}$. Si $\tilde f$ est différentiable, alors $\tilde f \circ \varphi^{-1}$ est aussi différentiable en 0 et

$$\mathrm{D}_0(\tilde f \circ \varphi^{-1}) = \mathrm{D}_x \tilde f \circ (\mathrm{D}_x\varphi)^{-1}\,.$$

Si on se restreint à F, alors $\tilde f \circ \varphi^{-1} = f \circ \varphi^{-1}$ et $(\mathrm{D}_x\varphi)^{-1}(\mathrm{F}) = \mathrm{T}_x\mathrm{X}$, donc

$$\mathrm{D}_0(\tilde f \circ \varphi^{-1}) \circ \mathrm{D}_x\varphi|_{\mathrm{T}_x\mathrm{X}} = \mathrm{D}_x\tilde f|_{\mathrm{T}_x\mathrm{X}} = \mathrm{D}_x f\,.$$

Réciproquement, supposons $f \circ \varphi^{-1}$ différentiable. Notons $\pi : \mathbb{R}^n \to \mathbb{R}^n$ la projection orthogonale sur F et posons $\tilde f = (f \circ \varphi^{-1}) \circ \pi \circ \varphi$. Il s'agit donc de la composée d'applications différentiables. Du coup

$$\mathrm{D}_x \tilde f = \mathrm{D}_0(f \circ \varphi^{-1}) \circ \pi \circ \mathrm{D}_x\varphi$$

donc

$$\mathrm{D}_x\tilde f|_{\mathrm{T}_x\mathrm{X}} = \mathrm{D}_0(f \circ \varphi^{-1}) \circ \pi \circ \mathrm{D}_x\varphi|_{\mathrm{T}_x\mathrm{X}} = \mathrm{D}_0(f \circ \varphi^{-1}) \circ \mathrm{D}_x\varphi|_{\mathrm{T}_x\mathrm{X}}$$

car $\mathrm{D}_x\varphi(\mathrm{T}_x\mathrm{X}) = \mathrm{F}$ et $\pi|_{\mathrm{F}} = \mathrm{Id}$.

Enfin, si j est une immersion, alors $\tilde f \circ j$ est différentiable et $\mathrm{D}_0(\tilde f \circ j) = \mathrm{D}_x\tilde f \circ \mathrm{D}_0 j$. Comme $\mathrm{D}_0 j : \mathbb{R}^p \to \mathrm{T}_x\mathrm{X}$ est un isomorphisme, on obtient

$$\mathrm{D}_x\tilde f|_{\mathrm{T}_x\mathrm{X}} = \mathrm{D}_0(\tilde f \circ j) \circ (\mathrm{D}_0 j)^{-1}|_{\mathrm{T}_x\mathrm{X}}\,.$$

Si $(f\circ j)$ est différentiable alors on peut construire une extension $\tilde f$ comme pour les redressements puisque j est la restriction de l'inverse d'un redressement sur F. ■

Il existe plusieurs difficultés pour donner du sens aux différentielles d'ordre supérieur, notamment à cause des propriétés de composition. La différentielle de f en x est une application $\mathrm{D}_x f \in \mathcal{L}(\mathrm{T}_x\mathrm{X}, \mathbb{R}^n)$, donc on devrait avoir $\mathrm{D}_x^2 f \in \mathcal{L}^2(\mathrm{T}_x\mathrm{X}, \mathbb{R}^n)$. Pour donner du sens à cette application bilinéaire, on devrait vérifier son invariance par difféomorphisme de sorte qu'elle ne dépende pas du paramétrage local, c'est-à-dire on souhaiterait

$$\mathrm{D}_x^2(f \circ \varphi)(v,w) = \mathrm{D}^2_{\varphi(x)} f(\mathrm{D}_x\varphi(v), \mathrm{D}_x\varphi(w))\,.$$

Or $\mathrm{D}_x(f\circ\varphi)(w) = \mathrm{D}_{\varphi(x)} f \circ \mathrm{D}_x\varphi(w)$ et si on différentie à nouveau, on obtient

$$\mathrm{D}_x^2(f \circ \varphi)(v,w) = \mathrm{D}^2_{\varphi(x)} f(\mathrm{D}_x\varphi(v), \mathrm{D}_x\varphi(w)) + \mathrm{D}_{\varphi(x)} f \circ \mathrm{D}_x^2\varphi(v,w)$$

donc la propriété d'invariance n'est pas préservée. Elle le serait si $\mathrm{D}_{\varphi(x)} f = 0$. Nous verrons au paragraphe suivant que cela suffit pour les points critiques.

14 Extrema liés

On se propose d'étudier les extrema de fonctions différentiables $f : \mathrm{X} \to \mathbb{R}$. Par un paramétrage local, on peut se ramener au cas standard, cf. la proposition 3.12 et le corollaire 5.11. Nous déclinons cette approche lorsque X est donnée par une submersion.

14.1 Calcul à l'ordre 1

On adapte la proposition 3.12 aux sous-variétés.

Théorème 14.1 (extrema liés). — *Soient $\mathrm{X} \subset \mathbb{R}^n$ une sous-variété de dimension p donnée par une submersion $\mathrm{X} = (g = 0)$ et $f : \mathrm{X} \to \mathbb{R}$ une application différentiable; notons $q = n - p$. Si un point x est un extremum local alors, pour toute extension $\tilde f$ de f, il existe des scalaires $\lambda_1, \ldots .\lambda_q$ tels que*

$$\mathrm{D}_x\tilde f = \sum_{j=1}^{q} \lambda_j \mathrm{D}_x g_j$$

où les g_j désignent les fonctions composantes de g. Une fois l'extension $\tilde f$ fixée, les scalaires $\lambda_1, \ldots, \lambda_q$ sont uniques et sont appelés les multiplicateurs de Lagrange.

Démonstration. — Décrivons X au voisinage de x par une immersion $j : \mathrm{U} \to \mathrm{X}$. Dire que x est un extremum local de f revient à dire que 0 est un extremum local de $f \circ j$, donc un point critique: $\mathrm{D}_0(f \circ j) = 0$. Soit $\tilde{f}$ une extension différentiable de f au voisinage de x. On a

$$\mathrm{D}_0(f \circ j) = \mathrm{D}_x \tilde{f} \circ \mathrm{D}_0 j = 0$$

donc

$$\mathrm{T}_x\mathrm{X} = \mathrm{D}_0 j(\mathbb{R}^p) \subset \operatorname{Ker} \mathrm{D}_x \tilde{f}\,.$$

Or $\mathrm{T}_x\mathrm{X} = \operatorname{Ker} \mathrm{D}_x g = \cap \operatorname{Ker} \mathrm{D}_x g_j$, donc $\cap \operatorname{Ker} \mathrm{D}_x g_j \subset \operatorname{Ker} \mathrm{D}_x \tilde{f}$. Le lemme 14.2 ci-dessous d'algèbre linéaire permet de conclure. ∎

Lemme 14.2. — *Soient* E *un espace vectoriel de dimension finie et* $f_1, \ldots, f_p \in \mathrm{E}^*$ *linéairement indépendantes. Alors* $(f_1, \ldots, f_p)$ *est une base du sous-espace* $\mathrm{P} = \{\mathrm{L} \in \mathrm{E}^*,\ \cap \operatorname{Ker} f_j \subset \operatorname{Ker} \mathrm{L}\}$.

Pour ce lemme, on rappelle que si $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ est une base d'un espace vectoriel réel E, alors son dual E^* admet comme base la famille de formes linéaires $\mathcal{B}^* = (\mathrm{e}_1^*, \ldots, \mathrm{e}_n^*)$ où chaque $\mathrm{e}_j^* : \mathrm{E} \to \mathbb{R}$ est définie par

$$\mathrm{e}_j^*(\mathrm{e}_i) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j\,. \end{cases}$$

Si $\mathrm{F} \subset \mathrm{E}$ est un sous-espace de E, son orthogonal $\mathrm{F}^\circ \subset \mathrm{E}^*$ est le sous-espace

$$\mathrm{F}^\circ = \{\mathrm{L} \in \mathrm{E}^* \text{ telle que } \mathrm{L}(x) = 0 \text{ pour tout } x \in \mathrm{F}\}\,.$$

Nous aurons recours au lemme suivant:

Lemme 14.3. — *Avec ces notations, on a* $\dim \mathrm{F} + \dim \mathrm{F}^\circ = \dim \mathrm{E}$.

Démonstration. — Prenons une base $(\mathrm{e}_1, \ldots, \mathrm{e}_p)$ de F que l'on complète en une base $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ de E. Notons $\mathcal{B}^*$ sa base duale. Le lemme sera établi si on montre que F° est engendré par $(\mathrm{e}_{p+1}^*, \ldots, \mathrm{e}_n^*)$.

Prenons $\mathrm{L} \in \mathrm{E}^*$. On écrit $\mathrm{L} = \sum a_j \mathrm{e}_j^*$. Si $\mathrm{L} \in \mathrm{F}^\circ$, alors on a $\mathrm{L}(\mathrm{e}_i) = 0$ pour tout $i \in \{1, \ldots, p\}$. Par conséquent, $\mathrm{L}(\mathrm{e}_i) = a_i = 0$ et donc $\mathrm{L} = \sum_{j=p+1}^n a_j \mathrm{e}_j^*$. Ceci montre $\mathrm{F}^\circ \subset \operatorname{vec}(\mathrm{e}_{p+1}^*, \ldots, \mathrm{e}_n^*)$.

Réciproquement, si $\mathrm{L} \in \operatorname{vec}(\mathrm{e}_{p+1}^*, \ldots, \mathrm{e}_n^*)$, alors $\mathrm{L}(\mathrm{e}_i) = 0$ pour tout $i \in \{1, \ldots, p\}$, donc $\mathrm{L}(x) = 0$ pour tout $x \in \mathrm{F}$. Ceci montre l'inclusion inverse et conclut la démonstration du lemme. ∎

Démonstration du lemme 14.2. — On définit $f : \mathrm{E} \to \mathbb{R}^p$ par $f(x) = (f_1(x), \ldots, f_p(x))$. On a $p = \operatorname{rg}(f_1, \ldots, f_p) = \dim \operatorname{Im} f$ et $\operatorname{Ker} f = \cap \operatorname{Ker} f_j$, donc $\dim \cap \operatorname{Ker} f_j = \dim \mathrm{E} - p$. Par définition P est l'orthogonal de $\operatorname{Ker} f$ donc $\dim \mathrm{P} = \dim \mathrm{E} - \dim \operatorname{Ker} f = p$. Comme les (f_j) sont indépendants et dans P, c'est une base. ∎

Corollaire 14.4. — *Soient* X *une sous-variété,* $f : \mathrm{X} \to \mathbb{R}$ *une application de classe* $\mathcal{C}^k$, $k \geq 1$, *et* $x \in \mathrm{X}$ *un point. Les propositions suivantes sont équivalentes.*

1. *Si* $\tilde{f}$ *est un prolongement différentiable de* f, $\mathrm{T}_x\mathrm{X} \subset \operatorname{Ker} \mathrm{D}_x \tilde{f}$.
2. *Si* X *est donnée par un redressement* φ *sur un sous-espace* F, $\mathrm{F} \subset \operatorname{Ker} \mathrm{D}_0(f \circ \varphi^{-1})$.
3. *Si* X *est donnée par une submersion* g *au-dessus de* 0, $\mathrm{D}_x \tilde{f} = \sum \lambda_j \mathrm{D}_x g_j|_{\mathrm{T}_x\mathrm{X}}$.
4. *Si* X *est donnée par une immersion,* $\mathrm{D}_0(f \circ j) = 0$.

Définition 14.5 (point critique). — *Soient* X *une sous-variété et* $f : \mathrm{X} \to \mathbb{R}$ *une application de classe* $\mathcal{C}^k$, $k \geq 1$. *Un point* $x \in \mathrm{X}$ *est un point critique de* f *si les conditions du corollaire 14.4 sont vérifiées.*

Démonstration. — L'essence du théorème des extrema liés est d'établir les implications **4.** implique **1.** implique **3.**. Si **3.** est vérifié, alors

$$\mathrm{T}_x\mathrm{X} = \operatorname{Ker} \mathrm{D}_x g = \cap \operatorname{Ker} \mathrm{D}_x g_j \subset \operatorname{Ker} \mathrm{D}_x \tilde{f}$$

impliquant ainsi **1.** Comme $\mathrm{T}_x\mathrm{X} = \operatorname{Im} \mathrm{D}_0 j$, **1.** implique aussi $\mathrm{D}_0(f \circ j) = \mathrm{D}_x \tilde{f} \circ \mathrm{D}_0 j = 0$, c'est-à-dire **4.**.

Enfin, l'équivalence avec **2.** se déduit facilement du fait que j est la restriction de l'inverse d'un redressement sur F. ∎

Enfin, on a le corollaire suivant:

Corollaire 14.6. — *Soient* X *une variété de classe* $\mathcal{C}^1$ *et* $f : \mathrm{X} \to \mathbb{R}$ *de classe* $\mathcal{C}^1$*. Si* $x \in \mathrm{X}$ *est un point critique, alors il existe un prolongement* $\widehat{f}$ *de classe* $\mathcal{C}^1$ *au voisinage de* x *tel que* $\mathrm{D}_x\widehat{f} = 0$.

Démonstration. — Soit $\tilde{f}$ un prolongement et notons $\lambda_1, \ldots, \lambda_q$ les multiplicateurs de Lagrange associés à $\tilde{f}$ en x. Posons $\widehat{f} = \tilde{f} - \sum \lambda_j g_j$ et remarquons que cela définit un autre prolongement de f qui vérifie $\mathrm{D}_x\widehat{f} = 0$ par le théorème des extrema liés. ∎

Application au calcul de la distance à une droite. — On se place dans $\mathbb{R}^2$ que l'on munit de la distance euclidienne. On se donne une droite D d'équation $ax + by + c = 0$ avec $(a, b) \neq (0, 0)$ et un point $p = (x_0, y_0)$ et on s'intéresse à $\mathrm{dist}(p, \mathrm{D}) = \inf\{\|p - m\|_2,\ m \in \mathrm{D}\}$. On a

$$\mathrm{dist}(p, \mathrm{D}) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

Pour montrer cela, on définit $g(x, y) = ax + by + c$ de sorte que $\mathrm{D} = g^{-1}(0)$. Montrons que g est une submersion au-dessus de 0. On a $\mathrm{D}g = \begin{pmatrix} a & b \end{pmatrix}$, donc $\mathrm{D}g$ est surjective en tout point puisque $(a, b) \neq (0, 0)$. Par conséquent D est bien une sous-variété définie par une submersion. Posons maintenant $f(x, y) = (x - x_0)^2 + (y - y_0)^2$. On cherche un minimum absolu de f sur D. On a $\mathrm{D}_{(x,y)}f = 2(x - x_0, y - y_0)$. D'après le théorème des extrema liés, pour tout extremum local $m = (x, y) \in \mathrm{D}$, il existe $\lambda \in \mathbb{R}$ tel que $\mathrm{D}_m f = \lambda \mathrm{D}_m g$. Autrement dit, nous devons résoudre le système d'inconnues x, y, λ suivant:

$$\begin{cases} 2(x - x_0) &= \lambda a \\ 2(y - y_0) &= \lambda b \\ ax + by + c &= 0 \end{cases} \quad \text{d'où l'on tire} \quad \begin{cases} x &= x_0 + \lambda a/2 \\ y &= y_0 + \lambda b/2 \end{cases}$$

Du coup, on arrive à

$$ax_0 + by_0 + c + \frac{a^2 + b^2}{2}\lambda = 0 \quad \text{et} \quad \lambda = -2\frac{ax_0 + by_0 + c}{a^2 + b^2}.$$

Nous avons donc un seul extremum local, qui doit être le minimum que l'on recherche. Il vient

$$f(x, y) = \left(\frac{ax_0 + by_0 + c}{a^2 + b^2}\right)^2 a^2 + \left(\frac{ax_0 + by_0 + c}{a^2 + b^2}\right)^2 b^2 = \frac{(ax_0 + by_0 + c)^2}{a^2 + b^2}$$

donc

$$d(p, \mathrm{D}) = \sqrt{\frac{(ax_0 + by_0 + c)^2}{a^2 + b^2}} = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

Application à la diagonalisation des matrices symétriques. — On munit $\mathbb{R}^n$ de la norme euclidienne et on considère la sous-variété

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n,\ \sum x_j^2 = 1\}$$

qui est compacte. On note $g(x) = \sum x_j^2 - 1$. Prenons $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique et posons $f : \mathbb{S}^{n-1} \to \mathbb{R}$ définie par $f(x) = \langle \mathrm{A}x, x\rangle$.

Un point $a \in \mathbb{S}^{n-1}$ est critique s'il vérifie les conclusions du théorème des extrema liés, donc s'il existe $\lambda \in \mathbb{R}$ tel que $\mathrm{D}_a f = \lambda \mathrm{D}_a g$. Or $\mathrm{D}_a g(h) = 2\langle a, h\rangle$ et

$$f(a + h) = f(a) + \langle \mathrm{A}a, h\rangle + \langle \mathrm{A}h, a\rangle + f(h) = f(a) + 2\langle \mathrm{A}a, h\rangle + o(h)$$

donc $\mathrm{D}_a f(h) = 2\langle \mathrm{A}a, h\rangle$. On obtient ainsi $\langle \mathrm{A}a, h\rangle = \lambda\langle a, h\rangle$ pour tout $h \in \mathbb{R}^n$, soit $\langle \mathrm{A}a - \lambda a, h\rangle = 0$. En prenant $h = \mathrm{A}a - \lambda a$, on trouve finalement $\mathrm{A}a = \lambda a$ et

$$f(a) = \langle \mathrm{A}a, a\rangle = \lambda\langle a, a\rangle = \lambda$$

car $a \in \mathbb{S}^{n-1}$.

Les vecteurs propres unitaires de A sont donc les points critiques de f et leurs valeurs propres sont données par la valeur de f.

Supposons maintenant que F est un sous-espace vectoriel invariant par A ($\mathrm{A}(\mathrm{F}) \subset \mathrm{F}$) et que a est un point critique de la restriction de f sur $\mathrm{F} \cap \mathbb{S}^{n-1}$, alors a est aussi un point critique pour la restriction de f sur $\mathbb{S}^{n-1}$. En effet, on peut trouver une famille orthonormale $\mathrm{e}_1, \dots, \mathrm{e}_k$ tels que $x \in \mathrm{F}$ si et seulement si $\langle x, \mathrm{e}_j \rangle = 0$ pour tout j. Du coup, si a est critique alors

$$\langle \mathrm{A}a, h \rangle = \lambda \langle a, h \rangle + \sum_{1 \leq j \leq k} \mu_j \langle \mathrm{e}_j, h \rangle$$

pour tout $h \in \mathbb{R}^n$. Or, comme $a \in \mathrm{F}$, on a aussi $\mathrm{A}a \in \mathrm{F}$, donc, en prenant $h = \mathrm{e}_j$ pour chaque j, on trouve $\mu_j = 0$ et $\langle \mathrm{A}a, h \rangle = \lambda \langle a, h \rangle$, établissant ainsi que a est aussi critique sur $\mathbb{S}^{n-1}$.

Comme la sphère est compacte, f atteint son maximum en un point a, ce qui montre l'existence d'une valeur propre. Supposons que l'on ait construit par réeurrence une famille orthonormée $\mathrm{e}_1, \dots, \mathrm{e}_k$ de vecteurs propres. Notons F_k l'orthogonal de ces vecteurs. Comme A est symétrique, on a bien $\mathrm{A}(\mathrm{F}_k) \subset \mathrm{F}_k$, donc un maximum de f sur $\mathbb{S}^{n-1} \cap \mathrm{F}_k$ sera un vecteur propre de A et on pourra procéder par récurrence pour montrer que A admet une base orthonormée de vecteurs propres.

14.2 Hessienne aux points critiques

Nous avons vu précédemment qu'il n'était en général pas possible de donner du sens à la différentielle seconde d'une application f définie sur une sous-variété. On montre que l'on peut définir proprement la hessienne $\mathrm{H}_f(a)$ en un point critique, ce qui nous permettra d'étendre le corollaire 5.11.

On suppose que X est au moins de classe $\mathcal{C}^2$, $f : \mathrm{X} \to \mathbb{R}$ est aussi de classe au moins $\mathcal{C}^2$ et x un point critique de f. Comme d'habitude, φ, g et j désigneront respectivement un redressement local de X, une submersion au-dessus de 0 et une immersion. De même, on note $\tilde{f}$ un prolongement $\mathcal{C}^2$ de f au voisinage de x.

Proposition 14.7. — *Soit $c : \mathrm{I} \to \mathrm{X}$ un chemin de classe $\mathcal{C}^2$ tel que $c(0) = x$; notons $v \overset{def.}{=} c'(0) \in \mathrm{T}_x\mathrm{X}$. On a*

$$(f \circ c)''(0) = \mathrm{D}_0^2(\tilde{f} \circ \varphi^{-1})(\mathrm{D}_x\varphi(v), \mathrm{D}_x\varphi(v)) = \mathrm{D}_0^2(f \circ j)((\mathrm{D}_0 j)^{-1}(v), (\mathrm{D}_0 j)^{-1}(v))$$

et

$$(f \circ c)''(0) = \left(\mathrm{D}_x^2 \tilde{f} - \sum_{j=1}^{q} \lambda_j \mathrm{D}_x^2 g_j \right)(v, v),$$

où $\lambda_1, \dots, \lambda_q$ sont les multiplicateurs de Lagrange associés au prolongement $\tilde{f}$.

Démonstration. — On écrit $f \circ c = \tilde{f} \circ c$ de sorte que $(f \circ c)'(t) = \mathrm{D}_{c(t)} \tilde{f}(c'(t))$ et

$$(f \circ c)''(0) = \mathrm{D}_x^2 \tilde{f}(v, v) + \mathrm{D}_x \tilde{f}(c''(0)) \,.$$

Or, on a aussi

$$\sum \lambda_j (g_j \circ c)(t) \equiv 0$$

donc en différentiant deux fois, on obtient

$$\sum \lambda_j \mathrm{D}_{c(t)} g_j(c'(t)) \equiv 0$$

puis

$$\sum \lambda_j \mathrm{D}_x^2 g_j(v, v) + \sum \lambda_j \mathrm{D}_x g(c''(0)) \equiv 0 \,.$$

En retranchant cette quantité à $(f \circ c'')(0)$, on obtient par le théorème des extrema liés

$$(f \circ c)''(0) = \mathrm{D}_x^2 \tilde{f}(v, v) - \sum \lambda_j \mathrm{D}_x^2 g_j(v, v) \,.$$

Comme

$$\mathrm{D}_{c(t)} \tilde{f}(c'(t)) = \mathrm{D}_{\varphi(c(t))}(\tilde{f} \circ \varphi^{-1}) \circ \mathrm{D}_{c(t)} \varphi(c'(t))$$

on obtient aussi

$$\begin{aligned}(f\circ c)''(0) &= \mathrm{D}^2_x(\tilde f\circ\varphi^{-1})(\mathrm{D}_0\varphi(v),\mathrm{D}_0\varphi(v)) + \mathrm{D}_0(\tilde f\circ\varphi^{-1})\circ \mathrm{D}^2_x\varphi(v,v)\\ &= \mathrm{D}^2_x(\tilde f\circ\varphi^{-1})(\mathrm{D}_x\varphi(v),\mathrm{D}_x\varphi(v))\end{aligned}$$

car x est un point critique.

Comme $v\in \mathrm{T}_x\mathrm{X}$, on a $\mathrm{D}_x\varphi(v) = (\mathrm{D}_0 j)^{-1}(v)$ et $\tilde f\circ\varphi^{-1}|_{\mathrm{F}} = \tilde f\circ j$. Ceci montre

$$\mathrm{D}^2_0(\tilde f\circ\varphi^{-1})(\mathrm{D}_x\varphi(v),\mathrm{D}_x\varphi(v)) = \mathrm{D}^2_0(f\circ j)((\mathrm{D}_0 j)^{-1}(v),(\mathrm{D}_0 j)^{-1}(v))\,.$$

■

On remarque que $(f\circ c)''(0) = \mathrm{D}^2_x\widehat f(v,v)$ où $\widehat f$ est donné par le corollaire 14.6.

Définition 14.8 (hessienne en un point critique). — *Si* X *est de classe* $\mathcal{C}^2$ *et* $f : \mathrm{X}\to\mathbb{R}$ *aussi avec point critique* x*, on définit la hessienne de* f *en* x *comme la forme bilinéaire symétrique* $\mathrm{H}_f(x) : \mathrm{T}_x\mathrm{X}\times\mathrm{T}_x\mathrm{X}\to\mathbb{R}$ *définie par* $\mathrm{H}_f(x)(v,w) = \mathrm{D}^2_x\widehat f(v,w)$*, où* $\widehat f$ *est un prolongement tel que* $\mathrm{D}_x\widehat f = 0$.

La proposition précédente montre que la hessienne est bien définie et elle donne plusieurs formules pour calculer $\mathrm{H}_f(x)(v,v)$. Le corollaire 5.11 appliqué à $\tilde f\circ j$ implique

Corollaire 14.9 (extrema locaux). — *Soit* $f : \mathrm{X}\to\mathbb{R}$ *une application deux fois différentiables. Si* a *est un maximum local (resp. un minimum local) alors* $\mathrm{D}_a f = 0$ *et* $\mathrm{H}_f(a)$ *est négative (resp. positive). Réciproquement, si* $\mathrm{D}_a f = 0$ *et* $\mathrm{H}_f(a)$ *est définie positive (resp. définie négative) alors* a *est un minimum local strict (resp. maximum local strict).*