

Espaces de formes

Hugo Malamut, Noé Sotto

sous la direction de Jean Feydy

Introduction

Dans le domaine de l'imagerie médicale, on dispose aujourd'hui de grandes quantités d'images dont on aimerait pouvoir tirer des informations informatiquement. Ainsi, à partir d'une banque de données d'images de coeurs ou de cerveaux, on aimerait pouvoir situer celui du patient pour diagnostiquer ou non une maladie : est-il "proche" d'un cerveau "normal", est-il ou non "dans la moyenne" ?

On voit que le premier pas pour pouvoir donner des informations statistiques sur différentes formes est de définir une notion de proximité, et donc de distance sur l'ensemble des formes. Pour que les statistiques aient un sens, il faut que la distance soit adaptée au problème précis, par exemple que des cerveaux jugés proches par des médecins aient une distance faible. Mais ce n'est pas le seul enjeu : en effet, l'objectif est de traiter un grand nombre de données. Le calcul de la distance doit donc être rapide.

Enfin, il ne faut pas oublier que le but est de traiter des données numériques discrétisées, et c'est pourquoi on considère que l'espace étudié est celui des nuages de points, correspondant aux formes discrétisées. On l'appellera aussi espace de "landmarks" dans la suite.

Dans un premier temps, nous verrons comment définir un grand nombre de distances naturelles sur l'espace des landmarks en le munissant d'une structure de variété riemannienne et en utilisant une distance de plus court chemin. Puis nous décrirons une famille plus restreinte de variété riemannienne bien adaptée au problème des landmarks, à savoir les cométriques à noyau. Enfin nous calculerons la courbure de cette variété pour estimer si les éléments statistiques qui correspondent pour certains à des approximations linéaires peuvent avoir du sens. Tout au long, nous introduirons également les outils informatiques permettant de calculer rapidement la distance entre les points et autres informations significatives sur l'espace.

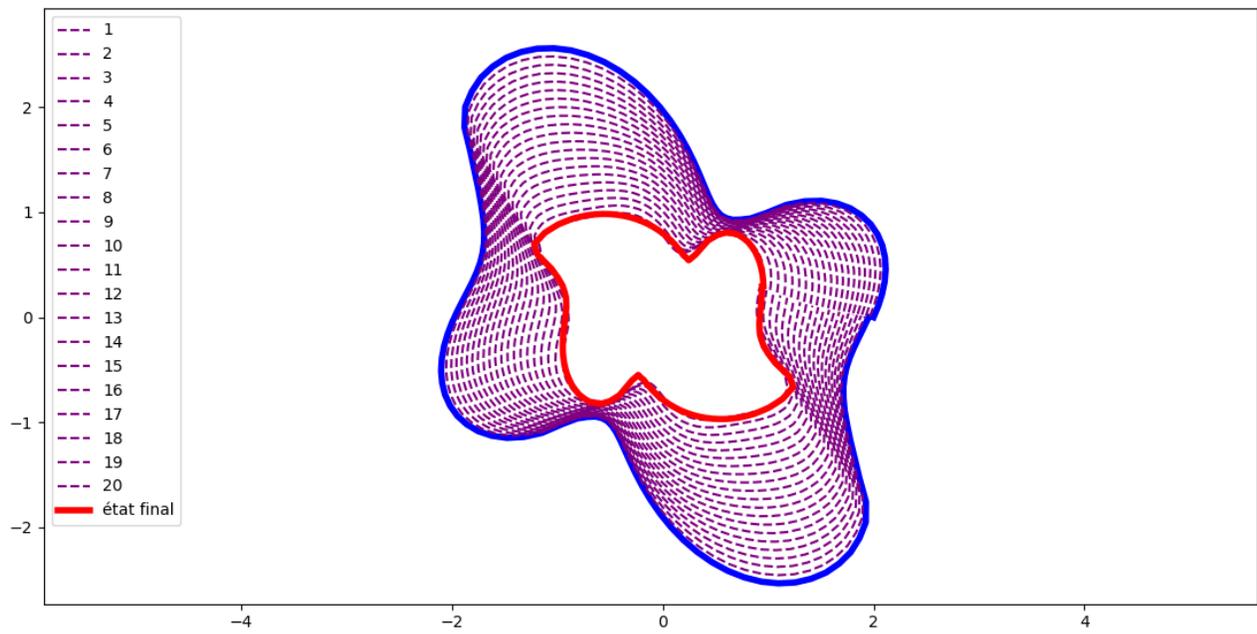


FIGURE 1 – Déformation progressive d'une forme en une autre dans l'espace muni d'une cométrie à noyau

Table des matières

1 Métrique riemannienne et géodésiques	4
1.1 Métrique riemannienne, distance de chemins	4
1.2 Equation hamiltonienne d'une géodésique	8
1.3 Calcul algorithmique d'une géodésique, schéma différentiel et pytorch	10
2 Un certain type de métrique : la métrique à noyau	15
2.1 Quelle métrique pour quelle déformation ?	15
2.2 Interpolation continue, normes RKHS et difféomorphismes	18
2.3 Intérêt algorithmique, premiers exemples	20
3 Etude de l'espace des landmarks dans des cas simples, courbure	24
3.1 Introduction à la courbure et triangles	25
3.2 Exemples de courbure de l'espace des landmarks	29

Chapitre 1

Métrie riemannienne et géodésiques

1.1 Métrie riemannienne, distance de chemins

Présentation On cherche donc à construire une distance entre deux nuages de points, ie une distance sur $(\mathbb{R}^d)^N$. On va se placer dans un formalisme assez large, qui permet de construire un large éventail de distances. Certaines, on le verra, sont particulièrement adaptées pour modéliser une distance entre des images de cerveaux par exemple.

L'idée est d'introduire une norme locale : à tout point $q_0 \in (\mathbb{R}^d)^N$, on associe une "norme locale" $\|\cdot\|_{q_0}$ qui mesure les distances autour de q_0 . On peut alors définir la longueur d'un chemin $q : [0, 1] \rightarrow (\mathbb{R}^d)^N C^1$ par morceaux en posant :

$$l(q)^2 = \int_0^1 \|dq(t)\|_{q(t)}^2 = C(q)$$

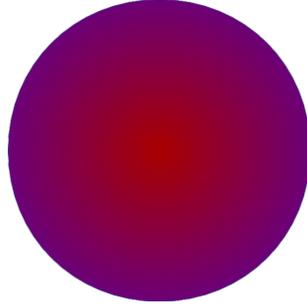
On peut voir $C(q)$ comme l'énergie nécessaire pour transporter $q(0)$ jusqu'à $q(1)$ en passant par le chemin q .

Dans la suite de la partie, on se placera dans le cas général $\mathbb{R}^n := (\mathbb{R}^d)^N$.

Remarque. (*intérêt d'une distance de chemins*) On peut calculer cette longueur en discrétisant l'intégrale, ce qui n'est pas le cas de toutes les distances. Par exemple, si on prend comme distance sur la sphère la corde entre deux points, la discrétisation ne converge pas vers la distance.

Exemple. (*disque de Poincaré*)

$$\|\cdot\|_x = \frac{1}{1 - \|x\|^2} \|\cdot\|$$



Métrie sur un espace euclidien, cométrie sur son dual

Définition 1. (métrie) On appelle métrie sur \mathbb{R}^n une matrice symétrique définie positive g définissant un nouveau produit scalaire $\langle gv, w \rangle$ et donc une nouvelle norme :

$$\|v\|_g^2 = \langle gv, v \rangle$$

Lemme 1.1.1. (cométrie) Si g est une métrie sur \mathbb{R}^n , $K := g^{-1}$ est la métrie sur le dual, appelée la cométrie.

On appellera les éléments du dual les moments.

L'application $v \rightarrow gv$ est une isométrie de (\mathbb{R}^n, g) vers (\mathbb{R}^n, K) .

Démonstration.

$$|\langle p, v \rangle| = \langle g^{-1/2}p, g^{1/2}v \rangle \leq \langle g^{-1/2}p, g^{-1/2}p \rangle \langle g^{1/2}v, g^{1/2}v \rangle = \langle p, g^{-1}p \rangle \langle g^{1/2}v, g^{1/2}v \rangle = \langle p, Kp \rangle \|v\|_g^2$$

donc $\|p\|_\star \leq \langle p, Kp \rangle$, or il y a égalité avec $v = Kp$, donc $\|p\|_\star = \langle p, Kp \rangle$ \square

Remarque. Si $p = gv$, $\|v\|_g^2 = \langle g_q v, v \rangle = \langle p, v \rangle = \langle p, K_q p \rangle = \|p\|_{K_q}^2$

Métrie riemannienne

Définition 2. (métrie riemannienne) On appelle métrie riemannienne sur une variété M une application qui associe à tout point de la variété une métrie sur l'espace tangent en ce point.

$$g : q \in M \rightarrow g_q \in T_q M$$

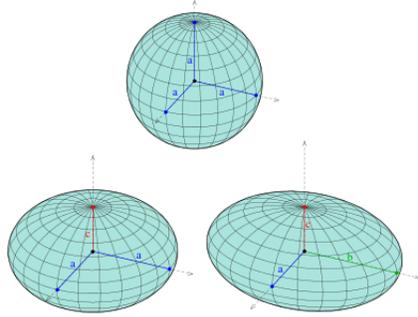
Dans la suite, on prendra g C^1 .

Remarque. g_q est la métrie qui s'applique aux vitesses en q : elle pénalise certaines directions de vitesses et en favorise d'autres, comme on va le voir dans la suite.

Influence du choix de la métrique sur le coût de déplacement g_q est symétrique définie positive, donc diagonalisable à valeurs propres strictement positives $\lambda_1 \leq \dots \leq \lambda_n$, dans une certaine base orthonormale (e_1, \dots, e_n) :

$$g_q = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

et $p = g_q v$



Si par exemple v est entièrement portée par e_i , alors $p = \lambda_i v$. Ainsi, si la vitesse v (qu'on prend normée) est portée par e_1 , le moment nécessaire pour mettre en mouvement une particule à la vitesse v est de norme plus petite que dans le cas où v est portée par e_n (pour un coût de λ_1 contre un coût de λ_n supérieur). On peut se représenter p comme le carburant nécessaire pour créer une vitesse v (en q), carburant qui aurait un coût de λ_i (ou encore une efficacité de $\frac{1}{\lambda_i}$) dans la direction e_i . On peut aussi voir p , élément du dual, comme une mesure qui accorde un poids variable à chaque direction.

Longueur d'un chemin, plus court chemin

Définition 3. (longueur d'un chemin) Soit $q : [0, 1] \rightarrow MC^1$ par morceaux un chemin sur une variété. On définit la longueur de q par :

$$l(q)^2 = \int_0^1 \|\dot{q}(t)\|_{q(t)}^2 dt =: C(q)$$

On peut alors munir M d'une structure métrique en posant :

$$d(a, b) = \inf_{q: a \rightarrow b} l(q)$$

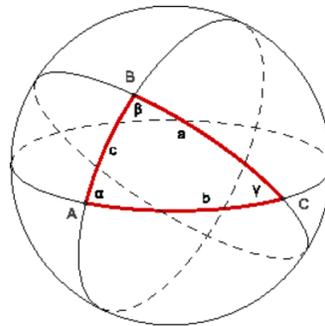
Remarque. On verra un chemin entre a et b comme une déformation de a à b , dont on cherche à minimiser le coût de transport, c'est-à-dire sa longueur.

Définition 4 (plus court chemin). On appelle plus court chemin a et b un chemin dont la longueur réalise $d(a,b)$. C'est un chemin de coût de transport minimal.

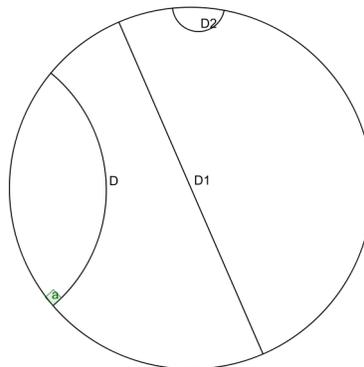
Exemple. Pour la distance euclidienne, la droite est l'unique géodésique entre deux points.



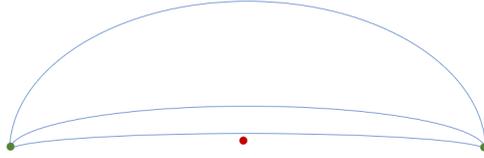
Sur la sphère : il peut y avoir plusieurs géodésiques entre deux points.



Sur le disque de Poincaré :



Dans le plan privé d'un point : il peut ne pas exister de géodésique entre deux points (défaut de complétude).



Une de nos préoccupations majeures dans la suite sera de réussir à calculer effectivement ces géodésiques.

1.2 Equation hamiltonienne d'une géodésique

On va établir une équation différentielle vérifiée par les plus courts chemins, ce qui permettra de les calculer algorithmiquement.

Ici, on travaille sur U , ouvert de \mathbb{R}^n .

Equation sur le fibré tangent On pose

$$E : (q, v) \in \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \frac{1}{2} \|v\|_q^2 = \frac{1}{2} \langle g_q v, v \rangle$$

qui représente, par analogie avec la physique, "l'énergie cinétique". E est C^1 et on a :

$$C(q) = 2 \int_0^1 E(q_t, \dot{q}_t) dt$$

Lemme 1.2.1. (formule de Christoffel)

$$\frac{d}{dt} \partial_v E(q, \dot{q}) = \partial_q E(q, \dot{q})$$

Démonstration. Soit q un plus court chemin entre deux points a et b de U . Par définition, q est un point critique de la fonctionnelle de coût.

On se ramène en dimension 1 en posant $f(x) = C(q + xQ)$, où Q est fixé tel que $Q(0) = Q(1) = 0$ et x suffisamment petit. On vérifie simplement que les hypothèse de convergence dominée sont vérifiées, d'où f dérivable et

$$0 = f'(0) = 2 \int_0^1 (\langle \partial_q E(q_t, \dot{q}_t), Q_t \rangle + \langle \partial_v E(q_t, \dot{q}_t), \dot{Q}_t \rangle) dt$$

En intégrant par parties le deuxième terme dans l'intégrale (le crochet est nul grâce aux conditions aux limites de Q), on obtient :

$$\int_0^1 (\langle \partial_q E(q_t, \dot{q}_t), Q_t \rangle - \langle \frac{d}{dt} \partial_v E(q_t, \dot{q}_t), Q_t \rangle) dt = 0$$

Par densité des chemins s'annulant en 0 et en 1 dans l'ensemble des chemins, on obtient :

$$\frac{d}{dt} \partial_v E(q, \dot{q}) = \partial_q E(q, \dot{q})$$

□

Equation sur le fibré cotangent On pourrait développer cette équation, mais la formule résultante est très compliquée. On remarque plutôt que :

$$\partial_v E(q, v) = g_q v = p$$

On fait alors le changement de variable $(q, p) := (q, g_q v)$.

Remarque. *Le couple de variables vit maintenant dans le fibré cotangent, qu'on appelle aussi l'espace des phases.*

Avec ces nouvelles variables, on redéfinit

$$H(q, p) = \frac{1}{2} \langle p, K_q p \rangle = E(q, K_q p)$$

Alors :

$$\begin{aligned} \partial_q H(q, p) &= -\partial_q E(q, K_q p) \\ \partial_p H(q, p) &= K_q p \end{aligned}$$

On obtient alors :

Théorème. *(équations de Hamilton)*

$$\begin{aligned} \dot{q} &= +\partial_p H(q, p) \\ \dot{p} &= -\partial_q H(q, p) \end{aligned}$$

On utilisera ces équations pour calculer des plus courts chemins algorithmiquement.

Remarque (flot hamiltonien). *Le couple (q, p) suit le gradient de H tourné de $-\frac{\pi}{2}$, le gradient symplectique de H :*

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = R_{-\frac{\pi}{2}} \nabla \tilde{H} \begin{pmatrix} q \\ p \end{pmatrix}$$

$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix}$ est orthogonal à ∇H tout le long du chemin, donc q suit une "équi-potentielle" : $H(q, p)$ est constant le long d'un plus court chemin.

On a donc

$$C(q) = \|\dot{q}(0)\|_{q(0)}^2$$

On réutilisera cette expression pour le calcul algorithmique du coût minimal.

Remarque. *La réciproque est fautive : un chemin vérifiant les équations de Hamilton n'est pas forcément un plus court chemin entre ses deux extrémités. On peut seulement affirmer qu'il est localement un plus court chemin, c'est-à-dire une géodésique.*

1.3 Calcul algorithmique d'une géodésique, schéma différentiel et pytorch

On a donc obtenu une équation différentielle qu'il suffit de suivre pour tracer les géodésiques de l'espace. En intégrant cette équation, grâce par exemple à un schéma d'Euler, cette équation permet d'explicitier les géodésiques numériquement. C'est essentiel pour comprendre la dynamique dans l'espace de forme, pour avoir une idée de l'évolution naturelle d'une forme selon la métrique choisie. Mais pour pouvoir extraire de réelles statistiques d'une base de données, il faudrait également pouvoir calculer le plus court chemin entre deux points, et donc la distance qui les sépare. Dans ce cas, on pourra par exemple calculer des moyennes de Fréchet d'un ensemble de landmark, c'est à dire trouver le landmark minimisant la somme des distances au carrés avec cette ensemble. Mais pour pouvoir appliquer cette méthode, il faut un cadre algorithmique adapté : il ne faut pas perdre de vue que l'étude des nuages de points s'inscrit dans un cadre de traitement de données nécessairement informatique. Si les algorithmes sont inefficaces (trop lents pour traiter le nombre de données suffisant), l'intérêt est nul. Nous présenterons donc ensuite le module pyTorch de python que nous avons utilisé pour effectuer l'ensemble de nos expériences numériques.

Schéma différentiel utilisé pour tracer un plus court chemin

Tirer un segment Soit $q \in \mathbb{R}^n$ une position et $p \in \mathbb{R}^n$ un vecteur cotangent en q . On veut tirer le segment d'extrémité q relatif à ce vecteur cotangent. On a les équations pour suivre une géodésique :

$$dq = \partial_p H dt$$

$$dp = -\partial_q H dt$$

Nous avons utiliser un simple schéma d'Euler explicite. On a $H(q, p) = \|p\|_q^2 = pK_q p$, et donc on suppose avoir implémenté une fonction H quadratique en p et lisse en q . Finalement, cette routine s'écrit comme indiqué sur la figure ?? . On a enfin le résultat suivant, très pratique

Proposition 1.3.1. *La longueur du segment q de moment cotangent p_0 à l'extrémité q_0 et tiré sur un temps 1 est $l(q) = \|p_0\|_{q_0}$*

Démonstration. H est une constante du mouvement le long d'une géodésique, on a donc pour un le segment

$$E(q) = \int_0^1 \|p(t)\|_{K_{q(t)}}^2 dt = 2H(q_0, p_0) = \|p_0\|_{q_0}^2$$

Cela montre que la longueur du segment tracé est bien de $\|p_0\|_{q_0}$ □

Trouver une géodésique Soient $(q_i, q_f) \in (\mathbb{R}^n)^2$ deux nuages de points. On cherche maintenant à tracer un segment allant de q_i à q_f . Grâce à la formulation hamiltonienne, le problème se reformule en un problème d'équation différentielle avec conditions aux limites : un segment q doit vérifier

$$\begin{cases} dq = \partial_p H dt \\ dp = -\partial_q H dt \\ q(0) = q_i \\ q(1) = q_f \end{cases}$$

Les deux premières équations formant une équation d'ordre 2 sur q . Une stratégie classique pour résoudre ce type problème est de le transformer en un problème d'optimisation sur un problème de Cauchy-Lipschitz. Pour un moment p cotangent en q_i , on peut tracer le segment géodésique correspondant sur un temps 1. On obtient alors un certain point $q(1)$. Si $q(1) = q_f$, alors on aura réussi à tracer un segment de q_i à q_f . Il faut donc optimiser p pour avoir $q(1) = q_f$. On pose donc $cost(p) = \|q(1) - q_f\|^2$ où la norme est la norme euclidienne entre points de \mathbb{R}^n . On peut alors différentier $cost(p)$ et effectuer une descente de gradient, pour la minimiser. On a alors espoir de faire tendre la distance entre $q(1)$ et q_f vers 0. En effet on a le résultat suivant :

Proposition 1.3.2. *Soit $(p_n)_{n \in \mathbb{N}}$ une suite de moments cotangents en q_i et $q_n(t)$ les segments correspondants. Si $cost(p_n) \rightarrow 0$ alors la distance géodésique entre $q_n(1)$ et q_f tend vers 0*

Démonstration. Comme $\|q_n(1) - q_f\|^2$ tend vers 0, $q_n(1)$ est une suite bornée. Soit K un compact convexe de \mathbb{R}^n contenant les $(q_n)_n$ et q_f . Soit γ_n le segment euclidien de $q_n(1)$ à q_f : $\gamma_n(t) = (1-t)q_n(1) + tq_f$. On a

$$E(\gamma_n) = \int_0^1 \langle g_{\gamma_n(t)} \dot{\gamma}_n(t), \dot{\gamma}_n(t) \rangle dt$$

Mais $\|\dot{\gamma}_n(t)\|_2 = \|q_f - q_n(1)\|_2$ et comme K est convexe, $\gamma_n(t) \in K$ et donc $\|g_{\gamma_n(t)}\|_2 \leq \|g\|_{\infty, K}$ (g est bien bornée sur K par continuité sur un compact). Par inégalité triangulaire et inégalité de Cauchy-Schwarz, on a donc

$$E(\gamma_n) \leq \|g\|_{\infty, K} \|q_f - q_n(1)\|_2^2$$

Or, par définition, la distance géodésique entre $q_n(1)$ et q_f , $d(q_n(1), q_f)$ est plus petite que la longueur de tous chemin les joignant, en particulier γ_n . Et donc

$$0 \leq d(q_n(1), q_f) \leq \sqrt{\|g\|_{\infty, K}} \|q_f - q_n(1)\|_2$$

Et donc, par comparaison, $d(q_n(1), q_f) \rightarrow 0$. □

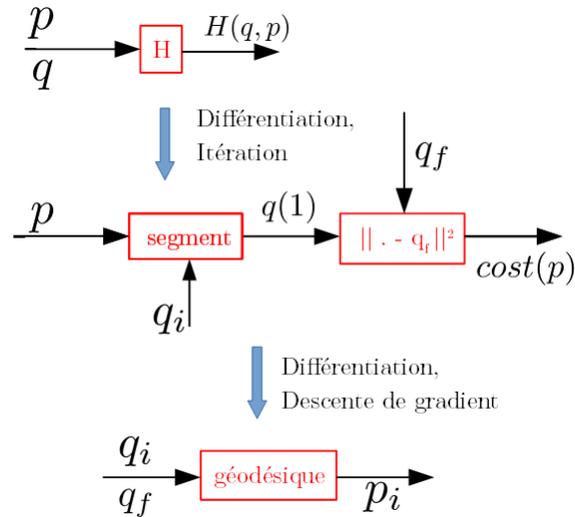


FIGURE 1.1 – Schéma numérique complet pour obtenir une géodésique entre deux points. On a besoin de quatre routines

Procédure entière On est donc parvenu à mettre en place un schéma algorithmique pour trouver un segment géodésique entre deux points. On peut le récapituler avec la figure 1.1

Le principal défaut de cette procédure est tout d'abord qu'on applique une descente de gradient à une fonction compliquée dont on ignore toute propriété, typiquement de convexité. On a donc aucune garantie que la descente de gradient aboutisse bien à un minimum global, ou même local. De plus, cette procédure pose aussi des problèmes plus techniques : on va devoir différencier une fonction compliquée, comprenant notamment la fonction appelée dans la figure segment qui nécessite de nombreuses itérations du schéma d'Euler pour être précise. Pour limiter la complexité, plusieurs conditions sont nécessaires :

- pouvoir calculer rapidement $H(q, p)$ en tout point
- avoir une procédure de différenciation rapide et précise

Nous allons voir que PyTorch est un outil bien adapté pour ces deux points.

Utilisation de PyTorch

Pour effectuer le calcul de H Un grand intérêt de PyTorch est qu'il permet d'utiliser les cartes graphiques des ordinateurs afin d'effectuer de nombreux calculs simples en parallèle. Une application naturelle est le calcul tensoriel,

puisque la multiplication de deux matrices, par exemple, se décompose en calculs simples, à savoir le calcul indépendant de chaque case. On sait que la fonction H s'écrit $H(q, p) = \langle K_q p, p \rangle$. Les multiplications par p seront très rapides sous PyTorch. Reste en suspens le calcul de K_q , qui dépendra de la métrique locale choisie.

Différentiation automatique

privilégier la différentiation composée Pour effectuer numériquement une différentiation, la méthode classique est de discrétiser :

pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $x \mapsto f(x)$, la différentielle de f est alors approximée par

$$df \simeq \left(\frac{f(x + dx_j) - f(x)}{dx_j} \right)_{1 \leq j \leq n}$$

où les dx_j sont choisis suivant la direction i et "suffisamment" petits.

Cette méthode est assez peu adaptée à PyTorch car les calculs des $f(x + dx_j)$ sont souvent trop compliqués pour être effectués en parallèle. À la place, PyTorch utilise la dérivation composée pour calculer le gradient de façon exacte. Si f est une composée, notée $f = h \circ g$, on a la formule

$$d_x(h \circ g) = d_{g(x)}h \circ d_xg$$

En notant ∂g la transposée de dg , on obtient pour le gradient

$$\nabla f = \partial_x g \circ \partial_{g(x)} h$$

En fait, si les valeurs des paramètres x et g(x) sont connus et le remplissage des matrices de type ∂g est simple, le calcul pourra être rapide sous PyTorch.

parcours aller-retour Le principe de la différentiation automatique se déroule en deux temps : dans un premier temps, la fonction f est évaluée, et au passage est créé un graphe de relation entre les variables d'entrée et la valeur de sortie. Ce graphe est composé d'opération élémentaire dont PyTorch connaît la différentielle. Dans le même temps, ce graphe est aussi rempli avec les valeurs utiles pour le calcul du gradient (x et g(x)) dans notre exemple de composée. Dans un second temps, si la différentielle est appelée, le graphe est parcouru en sens inverse en multiplication matricielle. La figure ?? résume cette technique pour $f = h \circ g$ et la figure 1.2 donne un exemple de graph.

Ainsi, la différentiation automatique permet de réduire la complexité. Le gain est variable en fonction des opérations utilisées. Un grand avantage est aussi d'avoir des formules exactes pour la différentielle. La contrepartie évidente est une perte en complexité spatiale : stocker un tel graphe des liaisons a évidemment un coût spatial important.

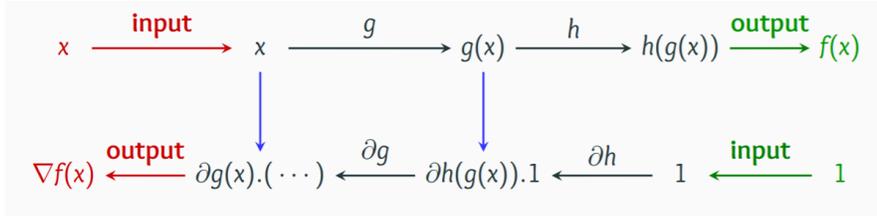


FIGURE 1.2 – Un exemple de graph créé par l'appel d'une fonction

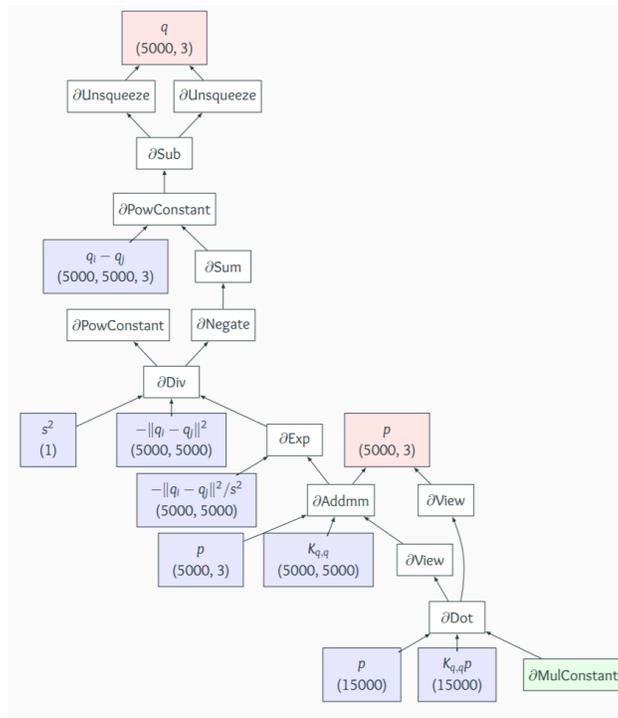


FIGURE 1.3 – Action de la différentiation automatique. Le parcours aller (en haut) conserve les données utiles pour le parcours retour (en bas)

Chapitre 2

Un certain type de métrique : la métrique à noyau

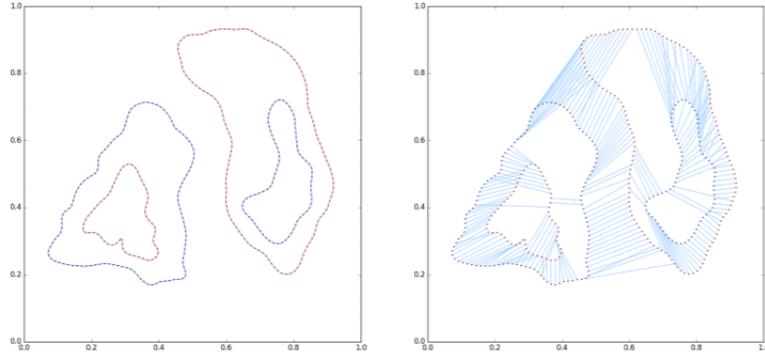
2.1 Quelle métrique pour quelle déformation ?

Le choix de la métrique Maintenant qu'on a vu dans un cadre général comment tirer des géodésiques sur une variété riemannienne, on s'intéresse plus précisément au cas de la variété dont un point représente une image, modélisée par un nuage de point, qu'on appelle "landmark". Il s'agit de choisir une métrique pour laquelle les géodésiques modélisent des déformations naturelles d'images, et qui est bien adaptée à des calculs algorithmiques. Mais alors, que signifie "avoir des formes proches" ?

Étudions quels sens cette expression peut avoir dans des situations différentes.

Si on s'intéresse à l'ensemble des triangles du plan, on voudrait dire que deux triangles sont proches s'ils sont globalement la même forme. Ainsi, on s'intéresse à la structure, rigide, des triangles, et on s'autorise donc des déformations qui conservent parfaitement la structure : les similitudes. De telles déformations considèrent un nuage de points comme un bloc d'un seul tenant.

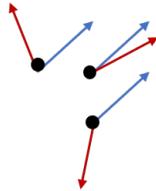
Deuxième situation : on interprète maintenant nos nuages de points de départ et d'arrivée comme des mottes de terre (départ) à transporter dans des trous (arrivée). Un ouvrier doit transporter chaque motte, et on veut que l'ensemble des ouvriers se fatigue le moins possible. Il s'agit alors simplement de minimiser la distance euclidienne entre les deux nuages de points. De telles déformations, qui considèrent tous les points comme indépendants les uns des autres, relient les points par des droites (géodésiques pour la distance euclidienne) : elles sont extrêmement souples mais déforment complètement la structure de la forme initiale. Ainsi, deux points proches peuvent être envoyés sur des points très éloignés.



On voit que selon les cas, les mêmes genres de déformations pour passer d'un nuage de points à l'autre ne sont pas adaptées. Dans notre cas, on cherche un entre-deux : des déformations assez souples mais qui ne déchirent pas complètement la structure. Dans l'idée, on voudrait que nos déformations envoient des points proches sur des points qui sont également proches.

Construction d'une métrique aux bonnes propriétés Soit (q_1, \dots, q_N) un amas de N points de l'espace, affectés de vitesses $v = (v_1, \dots, v_N)$.

Si la largeur de l'amas est inférieure à une certaine longueur caractéristique l , on aimerait forcer les vitesses v_i à être proches. En supposant que les points de l'amas sont tous à une distance d des uns des autres, pour faire apparaître une longueur caractéristique, on peut utiliser un paramètre gaussien $a = e^{-\frac{d^2}{l^2}}$.



Soit $F = \{(v_1, \cdot, v)_1, v_1 \in \mathbb{R}^d\}$. Si v est dans la direction de F , cela signifie que la vitesse appliquée à chaque point est la même. On veut donc que F soit un sous-espace propre de la métrique de valeur propre $f(a)$ petite (carburant de grande efficacité selon u , faible coût), et que F^\perp soit un sous-espace propre de valeur propre $g(a)$ grande (carburant de faible efficacité, grand coût).

Plus précisément, on peut souhaiter que

$$\lim_{a \rightarrow 0} f(a), g(a) = 1$$

(pas d'interaction quand les points sont éloignés)

$$\lim_{a \rightarrow 1} g(a) = +\infty$$

$$\lim_{a \rightarrow 1} f(a) = \frac{1}{N}$$

(coût minimal, de déplacement d'un point usuellement)

Le plus simple est alors de prendre, dans la base de $F \oplus F^\perp$, en écrivant par blocs de dimension d :

$$g_q = \begin{pmatrix} \frac{1}{1-a(n-1)} & 0 & (0) \\ 0 & \frac{1}{1-a} & \\ & & \ddots \\ (0) & & & \frac{1}{1-a} \end{pmatrix}$$

On s'intéresse surtout à la cométrie, qui est celle qui apparaît dans les équations de Hamilton :

$$K_q = \begin{pmatrix} 1 - a(n-1) & 0 & (0) \\ 0 & 1 - a & \\ & & \ddots \\ (0) & & & 1 - a \end{pmatrix}$$

ce qui se réécrit dans la base canonique :

$$K_q = \begin{pmatrix} 1 & a & \cdots & a \\ a & 1 & & \vdots \\ \vdots & & \ddots & a \\ a & \cdots & a & 1 \end{pmatrix}$$

C'est une expression particulièrement simple de la cométrie, qui modélise le comportement souhaité pour un amas de points à distance d les uns des autres.

Définition 5. (*cométrie à noyau*) Pour une répartition quelconque de points dans l'espace, on définit

$$K_q = (k(q_i - q_j))$$

par blocs de dimension d , avec

$$k(x) = e^{-\frac{\|x\|^2}{t^2}}$$

On appelle K_q une cométrie à noyau.

Remarque. Cependant, il n'est à priori pas évident que K_q reste définie positive : on le montrera dans la partie suivante.

Bilan sur le comportement de la métrique à noyau Des points proches peuvent difficilement s'éloigner, et sont contents d'avancer dans des directions proches : ils font ce qu'on peut appeler du "covoiturage". A la limite, il suffit d'un coût de 1 pour transporter N points au même endroit, au lieu d'un coût de N pour N points indépendants.

2.2 Interpolation continue, normes RKHS et difféomorphismes

Interpolation en champs de vitesses

Remarque (introductive). On peut réécrire :

$$k(q_i - q_j) = \langle \delta_{q_i}, k \star \delta_{q_j} \rangle$$

Lemme 2.2.1. (interpolation) Soit $\pi = \sum p_i \delta_{q_i}$. Alors par linéarité :

$$(K_q p)_i = \langle \delta_{q_i}, k \star \pi \rangle = (k \star \pi)(q_i)$$

En posant $u = k \star \pi$, on obtient :

$$v_i = u(q_i)$$

Remarque. On a ainsi défini un champ de vecteur $u \in C^\infty$, défini sur tout l'espace, qui coïncide avec les v_i en les q_i . C'est l'interpolation naturelle du vecteur $v = (v_0, \dots, v_N)$ sur l'espace entier. On a :

$$\forall x \in \mathbb{R}^d, u(x) = \sum k(x - q_i) p_i$$

La convolution de π par k rend interdépendants les q_i , selon le rapport entre la distance qui les sépare et la longueur caractéristique de k . Dans tous les cas, $\lim_{\infty} u(x) = 0$.

π s'interprète comme une mesure sur les champs de vitesses, ce qu'on va formaliser dans la suite.

Définition 6. On définit

$$H = \left\{ u \in C^\infty(\mathbb{R}^d, \mathbb{R}^d), \int \frac{1}{\hat{k}(\omega)} \|\hat{u}(\omega)\|^2 d\omega < +\infty \right\}$$

Proposition 2.2.1 (espace dual).

$$H^* = \left\{ \pi \in \mathcal{S}'(\mathbb{R}^d), \int \hat{k}(\omega) \|\hat{\pi}(\omega)\|^2 d\omega < +\infty \right\}$$

Définition 7 (espace RKHS). Si $\forall x \in \mathbb{R}^d, \delta_x \in H^*$, on dit que H est un RKHS (espace de Hilbert à noyau reproduisant). Dans ce cas, les diracs sont des mesures sur l'espace des champs de vitesses. En particulier,

Remarque. H est un RKHS $\iff \hat{k} \in L^1$: k gaussienne donne bien un RKHS. Les diracs sont des mesures sur l'espace des champs de vitesses. En particulier, $\pi = \sum p_i \delta_{q_i}$ est une mesure.

Extension de la métrique sur le RKHS

Lemme 2.2.2. (K_q est définie positive) Si $\hat{k} > 0$, pour $\pi = \sum p_i \delta_{q_i}$ et $u = k \star \pi$, on a :

$$\langle p, K_q p \rangle = \langle \pi, k \star \pi \rangle_{L^2} = \langle \hat{\pi}, \hat{k} \hat{\pi} \rangle_{L^2} = \left\langle \frac{1}{\hat{k}} \hat{u}, \hat{u} \right\rangle_{L^2} > 0$$

K_q est alors définie positive.

Définition 8. Si $\hat{k} > 0$, on peut munir H de la norme :

$$\|u\|_H^2 = \left\langle \frac{1}{\hat{k}} \hat{u}, \hat{u} \right\rangle_{L^2}$$

De manière analogue à la dimension finie, la cométrique sur le dual vaut :

$$\|\pi\|_{H^*}^2 = \langle \hat{\pi}, \hat{k} \hat{\pi} \rangle_{L^2}$$

Remarque. Cette norme étend la métrique à noyau aux champs de vitesses interpolés. Elle donne une nouvelle interprétation de la métrique à noyau : cette norme pénalise les hautes fréquences contenues dans le champ de vitesses, et donne un coût faible aux champs qui contiennent uniquement des basses fréquences. Ainsi, un champ qui applique des vitesses opposées à des points proches contiendra des hautes fréquences et sera fortement pénalisé, comme on l'a vu précédemment.

Flot du champ de vitesses et plus court chemin Maintenant qu'on a étendu la métrique à l'espace des champs de vitesses, on va s'intéresser aux champs de vitesses de coût minimaux qui envoient un nuage de points sur un autre.

Définition 9. Soit $u \in L^1([0, 1], H)$ un champ de vecteur dépendant du temps, soit $q = (q_0, \dots, q_N)$ un nuage de points de départ, $q' = (q'_0, \dots, q'_N)$ un nuage de points d'arrivée.

On dit que u transporte q sur q' si q' est l'image de q par le flot en $t = 1$ de u . On appelle coût de u $\|u\| = \int_0^1 \|u(t)\| dt$.

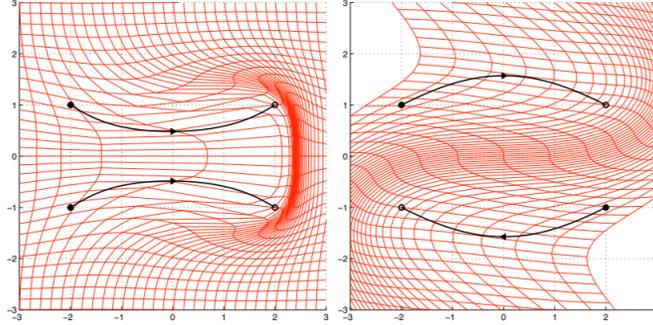
Lemme 2.2.3. Soit $V = \text{Vect}(\delta_{q_i})$.

A la décomposition orthogonale $H^* = V \oplus^\perp V^\perp$ correspond la décomposition orthogonale $H = (k \star V) \oplus^\perp \{u, \forall i, u(q_i) = 0\}$ (pour les produits scalaires respectifs).

Démonstration. H et H^* munis de leurs produits scalaires respectifs sont des espaces de Hilbert, d'où l'existence de l'orthogonal d'un sous-espace de dimension finie. \square

Corollaire. Soit $v = (v_0, \dots, v_N)$ des vitesses appliqués aux points $q = (q_0, \dots, q_N)$. Soit $p = g_q v := (p_0, \dots, p_N)$ les moments correspondants, $\pi = \sum p_i \delta_{q_i}$ la mesure associée dans H^* , et $u = k \star \pi$.

Parmi les champs de vitesses qui interpole v en q , u est le champ de norme minimale.



Corollaire. Soit $Q(t) = (Q_0(t), \dots, Q_N(t))$ un plus court chemin entre q et q' dans l'espace des nuages de points.

On note $p(t) = g_q \dot{Q}(t)$ le moment qui guide la trajectoire de Q .

Soit $\pi(t) = \sum p_i(t) \delta_{Q_i(t)}$ la mesure associée, et $u(t) = k \star \pi(t)$ le champ de vitesses qui lui correspond. Alors :

$$\forall t \in [0, 1], u(t) = \dot{Q}(t)$$

et u est le champ de vitesses (variable) de coût minimal qui envoie q sur q' .

Il a le même coût que le chemin Q .

2.3 Intérêt algorithmique, premiers exemples

Avantage numérique des cométries à noyau

Au niveau algorithmique, en section 1.3, on avait remarqué qu'un des principaux freins en terme de complexité était l'évaluation de $H(q, p) = \langle K_q p, p \rangle$. En fait, à aucun autre moment, on a besoin d'utiliser la métrique. Nul part n'apparaît l'expression de g_q . On sait déjà que les multiplications par p sont rapides sous PyTorch. Reste donc le remplissage de la matrice K_q . Or c'est là que l'on voit tout l'intérêt des cométries à noyau : pour calculer K_q , pas besoin d'inverser à chaque fois la matrice g_q . Au contraire, on dispose d'une expression directe de la cométrie : $K_q = (k(q_i - q_j))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$. Cette expression donne une formule simple pour chaque case : PyTorch peut donc remplir toutes les cases en parallèle en utilisant la carte graphique. C'est en fait une raison majeure du succès des cométries à noyau dans le domaine : elles permettent une vitesse d'exécution inégalable par d'autres formules. Elles ont donc le mérite de donner des résultats en grande dimension et pour un grand nombre de données. Reste à savoir si ces résultats sont fiables et traduisent correctement la réalité ou pas.

Premiers exemples : déformations d'un cercle

Pour commencer à manipuler cette métrique riemannienne, quoi de mieux que de l'appliquer à une forme simple ressemblant un peu à un cerveau : un cercle.

Déformation d'un cercle en ellipse et translation On considère un cercle comme point de départ et une ellipse décalée comme point d'arrivée et on veut voir comment notre métrique déforme l'un en l'autre de manière géodésique. On obtient le résultat de la figure 2.1

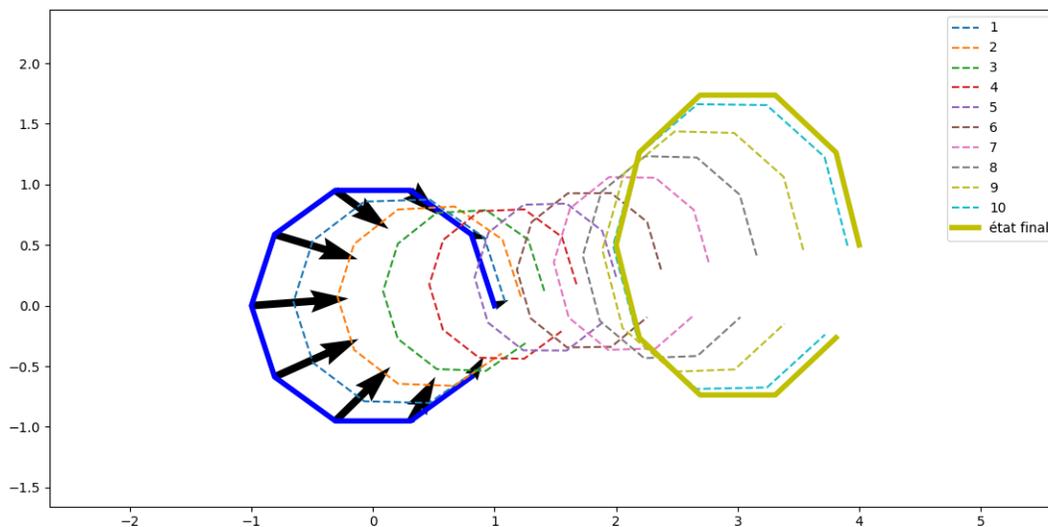


FIGURE 2.1 – Géodésique entre un cercle et une ellipse : on a représenté chaque forme intermédiaire obtenue par discrétisation du temps

On peut faire plusieurs remarques :

- La cométrique respecte bien la régularité des formes : la déformation est progressive, et en tout point intermédiaire de la géodésique, on obtient une forme régulière, qui "ressemble" à une ellipse.
- On observe un phénomène de "covoiturage" : les points ont tendance à se rapprocher pendant le trajet. Ce phénomène s'explique et peut permettre de montrer un cas de courbure et sera analysé plus avant dans la partie 3. On peut dire que c'est un effet indésirable : il serait préférable que

l'agrandissement du cercle en ellipse soit progressif plutôt qu'il y ait une phase de rétrécissement.

Déformation d'un cercle par rotation On peut aussi déformer le cercle en le faisant tourner : c'est un des autres aspects de la batterie d'algorithmes que nous avons mis sur pied : pour une position initiale et un moment initial donné, on peut suivre la géodésique obtenue. En prenant des moments tangents au cercle initial, on obtient comme géodésique la figure 2.2

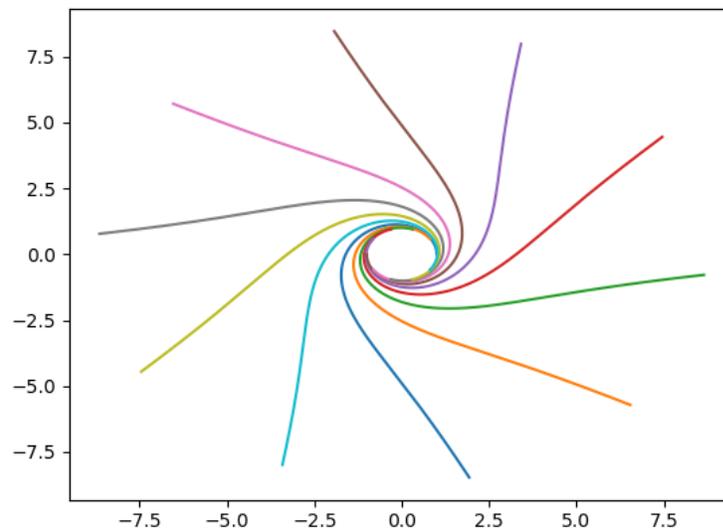


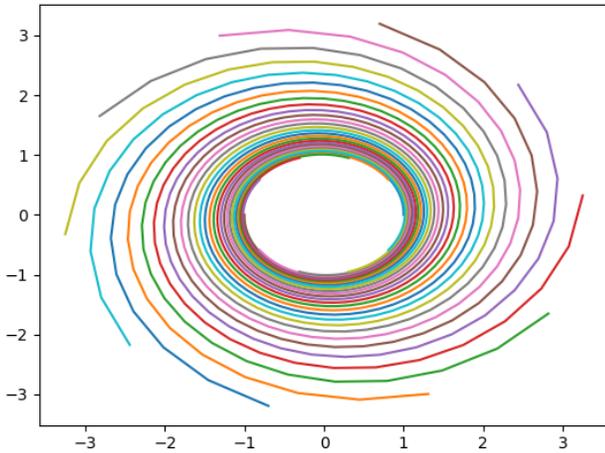
FIGURE 2.2 – Le nuage initial est formé de dix points en cercle de rayon 1, avec des moments égaux en norme et tangents au cercle. on a comme noyau $\sigma = 5$, on a représenté la trajectoire de chaque point

On peut analyser les comportements des points :

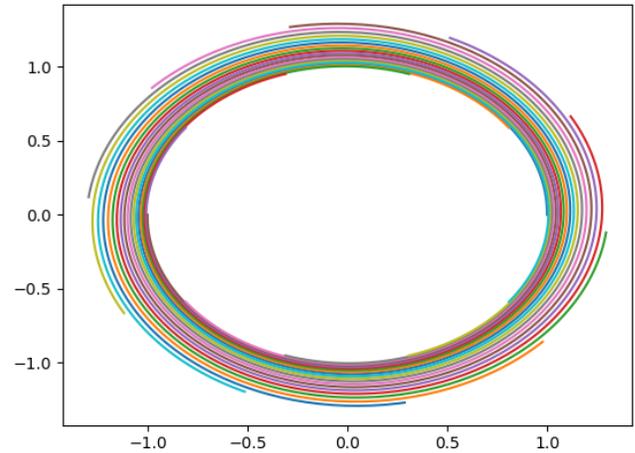
- À l'infini, les points n'interagissent pas et partent donc en ligne droite
- Quand ils sont proches, les points ont tendance à avoir des vitesses orthoradiales : cela permet en chaque point d'avoir une vitesse proche de celle de ses deux proches voisins sur le cercle.

On peut aussi comparer les résultats obtenus suivant les paramètres donnés en entrée comme sur la figure 2.3.

Défaut du schéma d'Euler On voit que si n_t change, l'allure reste globalement la même, mais le moment où les points diverge est plus rapidement atteint si n_t est petit. C'est un phénomène qui montre qu'il faut se méfier des résultats obtenus, qui peuvent être faussés par des imprécisions numériques. Nous verrons plus loin, en introduisant la courbure, dans quelles situations il faut particulièrement se méfier de ce phénomène. Ici, tant que les points restent assez serrés, on a tout de même des situations très proches, et d'ailleurs on le voit mal, mais même dans la situation où il y a une différence notable, il y a beaucoup de tours où l'on peut considérer que les deux figures coïncident.



(a) $n_t = 200$



(b) $n_t = 500$

FIGURE 2.3 – Défaut numérique : sur les images 2.3a et 2.3b on a la même situation initiale avec le même noyau $\sigma = 30$. Seul le paramètre n_t varie

Chapitre 3

Etude de l'espace des landmarks dans des cas simples, courbure

Moyenne de Fréchet Maintenant que nous avons doté l'espace des landmarks d'une métrique riemannienne qui fait sens, et que nous avons les moyens pratiques de tracer des géodésiques entre les landmarks, la prochaine étape est de calculer des grandeurs statistiques sur ces formes afin d'en tirer des informations. Supposons disposer d'un ensemble de données, par exemple des formes de cerveau, sous forme d'une famille de landmarks $(q_a)_{a \in A}$. On a étiqueté ces landmarks avec diverses informations : âge, pathologie... Un premier but pourrait être de voir si les landmarks de mêmes étiquettes se ressemblent. Pour cela, une méthode très efficace est d'utiliser une moyenne de Fréchet et l'atlas associé :

Définition 10. *On appelle **moyenne de Fréchet** de la famille $(q_a)_{a \in A}$ un landmark m minimisant $\sum_{a \in A} d(m, q_a)^2$. Si ces distances sont atteintes par des plus courts chemins, l'**atlas** associé est une famille $(p_a)_{a \in A}$ de vecteurs cotangents en m qui permettent de tirer des plus courts chemins vers chacun des q_a .*

On remplace alors les q_a par les p_a , qui sont bien plus simples à étudier puisqu'ils sont plongés dans un espace pré-hilbertien. On peut par exemple un hyper-plan séparant certaines étiquettes, trouver une direction privilégiée des pathologies d'un certain type...

Limitation par la courbure Mais cette méthode a ses limites. En fait, on est en train de remplacer notre modèle par un problème linéaire : on fait une approximation linéaire. Or on risque d'y perdre beaucoup d'information : suivant la situation, de façon imagée, c'est comme si on essayait d'enrober une sphère avec une feuille de papier. On voit de façon instinctive qu'une caractéristique de la variété risque de s'opposer à cette approximation, ou plutôt à sa capacité à

représenter la réalité. Cette caractéristique, on le comprend avec la sphère, c'est la courbure.

3.1 Introduction à la courbure et triangles

Avant d'aller plus loin, il faut définir ce que l'on entend par courbure dans l'espace des landmarks. Déjà, en dimension 2, il y a plusieurs type de courbure. Pour les surfaces, nous nous intéressons à la courbure de Gauss. En dimension supérieure, la définition est plus délicate, on peut introduire un tenseur de courbure, mais nous nous contenterons d'appliquer la courbure 2-dimensionnelle à certains plans grâce à la notion de courbure sectionnelle.

Courbure sur les surfaces

Lien entre courbure et triangles géodésiques

Jusqu'ici, les outils qu'on a mis en place pour travailler sur l'espace de forme sont avant tout des outils servant à tracer des géodésiques. Or il existe un lien assez étroit entre géodésique et courbure gaussienne. C'est ce lien que nous exploiterons en pratique pour calculer des courbures sectionnelles dans l'espace des landmarks. Une façon de faire le lien entre courbure gaussienne et géodésique est d'étudier les triangles géodésiques :

Définition 11. *Un triangle géodésique est la donnée de 3 points (sommets du triangle) et de 3 segments géodésiques les reliant 2 à 2.*

Dans la suite, on confondra souvent ce triangle et la surface qu'il délimite. Dans un plan, on sait que pour un tel triangle, la somme des angle est égale à π . Cette propriété est modifiée par la courbure. La figure 3.1 illustre ce phénomène

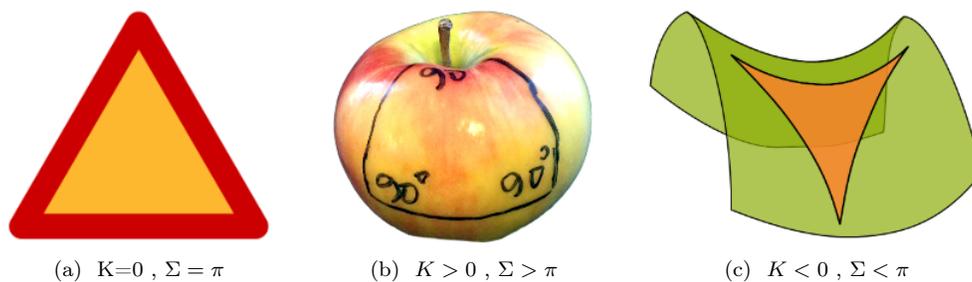


FIGURE 3.1 – La somme Σ des angles des triangle varie

Il est donc intéressant d'introduire l'angle entre deux vecteurs : naturellement, on a envie d'écrire

$$\langle g_q v_1, v_2 \rangle = \|v_1\|_q \|v_2\|_q \cos(v_1, v_2)$$

Définition 12. Soit $q \in \mathbb{R}^{Nd}$ une position et $v_1, v_2 \in (\mathbb{R}^{Nd})^2$ deux vecteurs tangents en q , On définit l'**angle**(v_1, v_2) par

$$(v_1, v_2) = \arccos\left(\frac{\langle g_q v_1, v_2 \rangle}{\|v_1\|_q \|v_2\|_q}\right)$$

Remarque. On définit de même l'angle entre vecteurs cotangents (en remplaçant g_q par K_q), et si pour $i = 1, 2$, $p_i = g_q v_i$, alors $(p_1, p_2) = (v_1, v_2)$

Démonstration. C'est immédiat, car $\langle g_q v_1, v_2 \rangle = \langle K_q p_1, p_2 \rangle$ □

Pour un triangle géodésique, il est de plus visiblement intéressant d'étudier la différence entre la somme des angles et π :

Définition 13. Soit T un triangle d'angles intérieurs positifs α, β, γ . On définit le **défaut** de T , noté $\delta(T)$ par

$$\delta(T) = (\alpha + \beta + \gamma) - \pi$$

On remarque que le défaut est une propriété additive des triangles : En effet, découpons le triangle T en deux triangles T_1 et T_2 comme indiqué sur la figure 3.2.

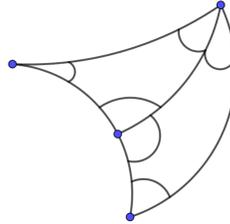


FIGURE 3.2 – On découpe un triangle en deux sous triangles

On a $\delta(T) = (\alpha + (\beta + \theta) + \varphi) - \pi$ et $\beta + \gamma = \pi$. En additionnant ces relations, on obtient

$$\delta(T) = \delta(T_1 + T_2) = \delta(T_1) + \delta(T_2)$$

On voit que le défaut se comporte assez comme une mesure signée, et on a donc l'espoir d'avoir une forme intégrale $\delta(T) = \int_T \kappa(q) dA(q)$

En fait, une telle forme existe, grâce au théorème de Gauss-Bonnet :

Théorème (Gauss-Bonnet (admis)). Soit M une surface compacte dotée d'une métrique riemannienne donnant une forme d'aire dA . On note ΓM la frontière de M , κ la courbure sur M , k la courbure géodésique sur ΓM et $\chi(M)$ la caractéristique d'Euler de M . Alors

$$\int_M \kappa dA + \int_{\Gamma M} k ds = 2\pi \chi(M)$$

On peut appliquer ce théorème à un triangle géodésique et on obtient :

Proposition 3.1.1. *Soit S une surface dotée d'une métrique riemannienne donnant une forme volume dA . Soit T un triangle géodésique de S . Si T est simplement connexe, on a l'égalité*

$$\delta(T) = \int_T K(q)dA(q)$$

Démonstration. Comme T est une surface à bord simplement connexe, $\chi(T) = 1$. De plus, sans entrer dans les détails de la définition de la courbure géodésique, on a les propriétés suivantes :

- $k = 0$ sur les géodésiques
- au voisinage $]-\varepsilon, +\varepsilon[$ d'un angle α , $\int_{-\varepsilon}^{\varepsilon} k ds \sim (\pi - \alpha)$.

Cela permet de conclure. □

Exemples de surfaces de courbure non nulle

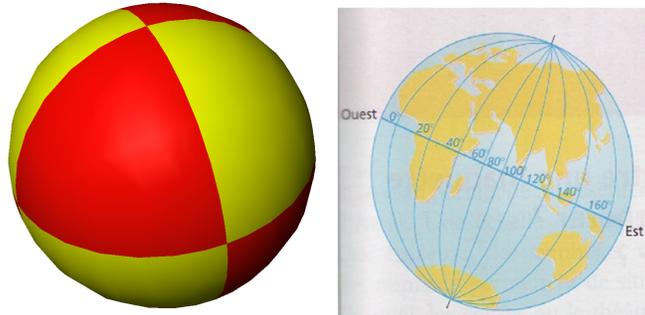
Maintenant que l'on a vu le lien avec le tracé géodésique, on peut regarder des exemples d'espaces de courbure non nulle, afin d'observer les effets potentiels de la courbure positive et de la courbure négative sur nos données et nos programmes informatiques.

Un exemple de courbure positive : la sphère La sphère est l'exemple par excellence de courbure positive : sa courbure est constante, égale à $\frac{1}{R^2}$ où R est son rayon. Les géodésiques sont les grands arcs de cercle de rayon R . On en déduit deux propriétés caractéristiques de la courbure positive :

- La somme des angles d'un triangle est supérieure à π comme vu précédemment.
- Deuxièmement, il existe des **points conjugués**, ie reliés par deux segments géodésiques (intrinsèques) distincts.

L'existence de points conjugués est notamment un problème pour utiliser une moyenne de Fréchet : par exemple, deux points tout proches de l'un des points peuvent avoir des moments très différents dans l'atlas associé à l'autre point conjugué.

Un exemple de courbure négative : le disque de Poincaré De même, l'exemple typique de la courbure négative est le disque de Poincaré déjà introduit comme exemple typique de variété riemannienne. Dans ce disque, les géodésiques sont les cercles perpendiculaires au cercle frontière. Comme attendu, la somme des angles d'un triangle fait donc moins de $\frac{\pi}{2}$ – attention, l'aire doit être calculée dans la métrique locale.



(a) L'intersection de la sphère et (b) Tous les méridiens sont des d'un octant définit un triangle dont géodésiques reliant les deux pôles tous les angles sont droits

FIGURE 3.3 – Illustrations des propriétés de la courbure positive : des triangles dont la somme des angles dépasse π (3.3a) et des points conjugués (3.3b)

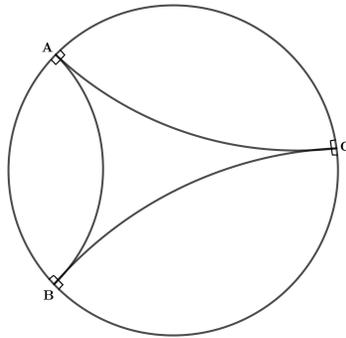


FIGURE 3.4 – ABC est un triangle limite dont tous les angles sont nuls. Les sommets n'appartiennent pas au disque mais on peut approcher ces triangles en approchant les points et alors les angles tendent vers 0

Une caractéristique très embêtante du disque de Poincaré est que des segments lancés du milieu dans des directions proches se retrouvent vite très éloignés puisque les distances sont dilatées sur les bords. De nouveau, les informations fournies par un atlas à la moyenne de Fréchet auront peu de sens si des moments proches correspondent à des landmarks très éloignés. Une autre conséquence très fâcheuse de ce phénomène est une forte instabilité numérique puisqu'une légère déformation entraîne de grands changements.

Courbure sectionnelle dans l'espace des landmarks

Courbure sectionnelle

La courbure dans un espace de dimension supérieure à 3 n'est pas un objet évident. C'est pourquoi nous nous ramènerons au plan en utilisant une courbure sectionnelle :

Définition 14. Soit $q \in \mathbb{R}^n$ une position et P un plan de l'espace cotangent en q . Les géodésiques passant par q et ayant leur moment dans ce plan P forment une sous-variété de dimension 2 de l'espace ambiant, (qu'on appellera ensuite surface géodésique). La **courbure sectionnelle** au point q relatif au plan P , notée est alors la courbure de cette surface au point q .

Calcul explicite et triangles Comme on peut facilement tracer des segments, on peut aussi tracer des triangles et calculer leurs angles. On va donc utiliser cette méthode pour calculer la courbure en un point. On a pour cela le résultat suivant.

Proposition 3.1.2. Soit $q \in \mathbb{R}^n$ un landmark, et P un plan de l'espace cotangent en q . Soit $(p_1, p_2) \in P^2$ deux moments non colinéaires. Soit $(T_\epsilon)_{\epsilon > 0}$ des triangles de sommet q et de côtés issus de q de longueur ϵ et de moment initiaux p_1 et p_2 . On a, en notant $\kappa(q)$ pour la courbure sectionnelle de q relatif à P ,

$$\kappa(q) = \lim_{\epsilon \rightarrow 0} \frac{2\delta(T_\epsilon)}{\epsilon^2 |\sin(p_1, p_2)|}$$

Démonstration. On a $T_\epsilon \rightarrow_{\epsilon \rightarrow 0} q$. Donc d'après notre définition de la courbure, $\frac{\delta(T_\epsilon)}{A(T)} \rightarrow K(q)$ où l'aire de T est l'aire de la surface délimitée par T sur la surface géodésique. Il s'agit donc simplement de démontrer $A(T) \sim \frac{1}{2} \|p_1\| \|p_2\| |\sin(p_1, p_2)| \epsilon^2$. Or quand ϵ tend vers 0, la métrique tend en tout point de T_ϵ vers la métrique en q . À l'ordre 1, chacun des segments du triangle tend donc vers un segment euclidien et donc T tend vers le triangle euclidien de cotés adjacents $\epsilon \frac{p_i}{\|p_i\|}$. Or ce triangle a pour aire $\frac{1}{2} |\sin(p_1, p_2)| \epsilon^2$. Ainsi $\frac{2\delta(T_\epsilon)}{\epsilon^2 |\sin(p_1, p_2)|} \rightarrow \kappa(q)$

□

3.2 Exemples de courbure de l'espace des landmarks

Nous allons maintenant passer à des exemples dans l'espaces des landmarks pour se familiariser.

Situation de courbure négative

Un exemple typique est celui d'un nuage de points avec un seul moment initial non nul, celui du centre.

Nous avons repris cet exemple : Tout d'abord, la géodésique suivie par le landmark s'explique assez bien en terme d'interaction des points : la mise en mouvement du point central a tendance à favoriser le mouvement des autres points dans le même sens, au prorata de la proximité de ces derniers. Cela amène donc une sur-concentration des points dans le sens de la vitesse et un éparpillement dans l'autre sens.

En prenant deux moments distincts, on obtient les situations finales données en figure 3.5a. Nous avons enfin tracé le dernier côté du triangle correspondant : le tracé en est donné en figure 3.5b On peut aussi calculer le défaut de ce triangle : on obtient $\delta = -0.13$. On a donc bien $\delta < 0$ et donc une courbure négative.

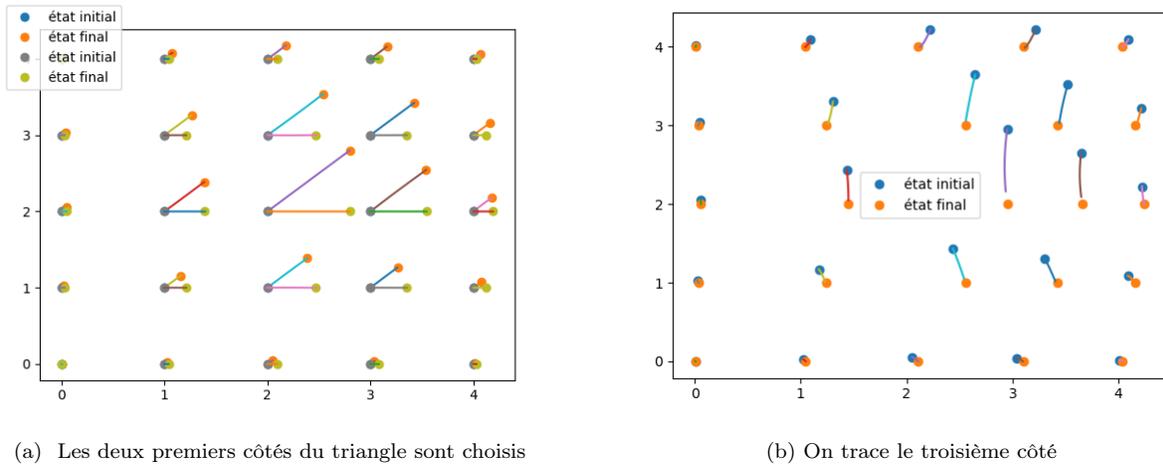


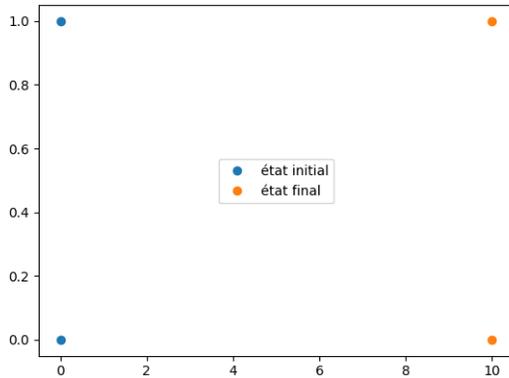
FIGURE 3.5 – Exemple de points conjugués

Analyse des géodésiques Intuitivement, cela s'explique, car pour passer de la situation finale 1 à la situation finale 2, il s'agit de séparer des points très proches. Comme c'est cher, mieux vaut donc emprunter pour commencer un chemin qui les sépare alors qu'ils bougent dans des sens proches : on a donc intérêt à revenir sur nos pas vers la situation initiale. C'est une propriété classique de la courbure négative en lien avec les triangles fins.

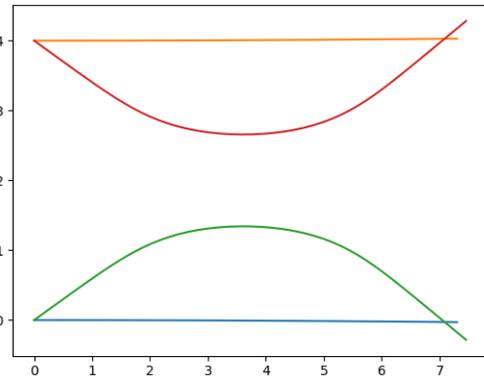
Situation de courbure positive

Pour exhiber une situation de courbure positive, des points conjugués sont très intéressants. Pour cela, on va exploiter un phénomène de covoiturage. On se place dans l'espace à 2 dimensions et avec seulement 2 points. Les situations

initiale et finale sont décrites par la figure 3.6a . Les deux géodésiques les reliant sont donnés figure 3.6b



(a)



(b) Deux géodésiques différentes passent par ces points

FIGURE 3.6 – Exemple de points conjugués

Analyse des géodésiques La première géodésique est instinctive : quoi de mieux pour relier des points que les droites, surtout si elles sont parallèles et si les déplacements dans le même sens sont moins coûteux. Pour la deuxième géodésique, on a affaire à un phénomène de convoiturage : les phases de rapprochement et d'éloignement sont certes coûteuses car les vitesses sont alors dans des sens relativement opposés, mais la phase centrale est avantageuse. Pour expliquer cela, le relèvement en difféomorphisme est parlant : si les points sont assez éloignés, le difféomorphisme les transportant est la somme des deux difféomorphismes capables de les transporter un par un, alors que si les points sont assez proches, un seul de ces difféomorphismes suffit (dans le cas où ils sont confondus) ou presque. On peut donc diviser quasiment le coût par 2. C'est pourquoi on parle de convoiturage, puisque les deux points sont conduits par le même difféomorphisme. Si la période de convoiturage est assez longue, cette géodésique est donc nécessairement plus intéressante, car les coûts des phases initiale et finale seront compensés.

Calculs effectifs de courbure

Enfin, nous avons réussi à calculer la courbure dans des cas précis de forte symétrie du problème.

Symétrie par rapport à un axe Dans l'espace des landmarks à deux points à deux dimension, il y a un plan particulier qui peut être intéressant : celui qui rassemble les situations symétriques par rapport à l'axe des abscisses (par exemple). Nous avons tracé la courbure observée sur cet axe en figure 3.7

Symétrie par rotation On peut aussi penser à considérer les deux points étant symétriques par rapport au centre, et cette situation s'inscrit dans une situation plus générale : on peut multiplier le nombre de points et considérer le plan où les points sont toujours placés en polygone régulier, ie quand ils sont tous obtenus par rotation d'angle $\frac{2\pi}{N}$ (où N est leur nombre) à partir d'un seul d'entre eux. La courbure est décrite en figure 3.8.

On peut faire les mêmes réflexions basiques pour les deux cas :

- Comme les longueurs explosent en $\rho = 0$, les résultats ne sont plus fiables près de ce point
- En $\rho \rightarrow +\infty$, les points n'interagissent plus et la courbure tend vers celle d'un espace euclidien : 0

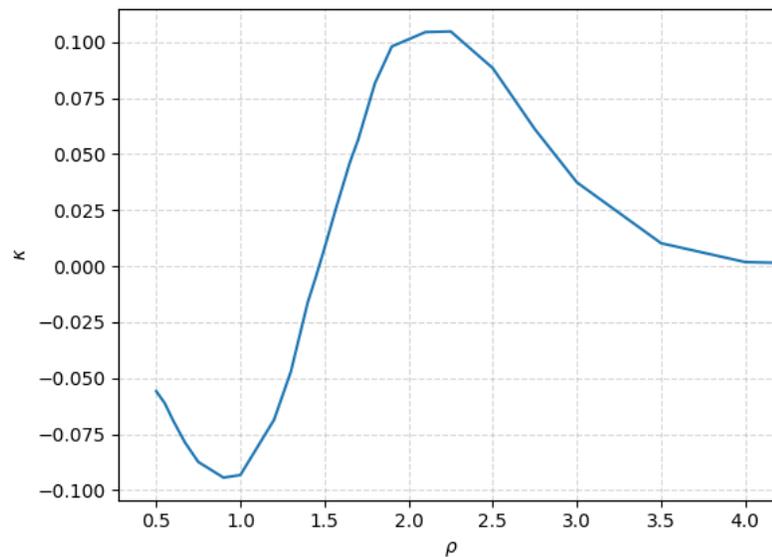


FIGURE 3.7 – Courbure en fonction de $\rho = |q_1 - q_2|$, le noyau est gaussien avec $\sigma = 1$

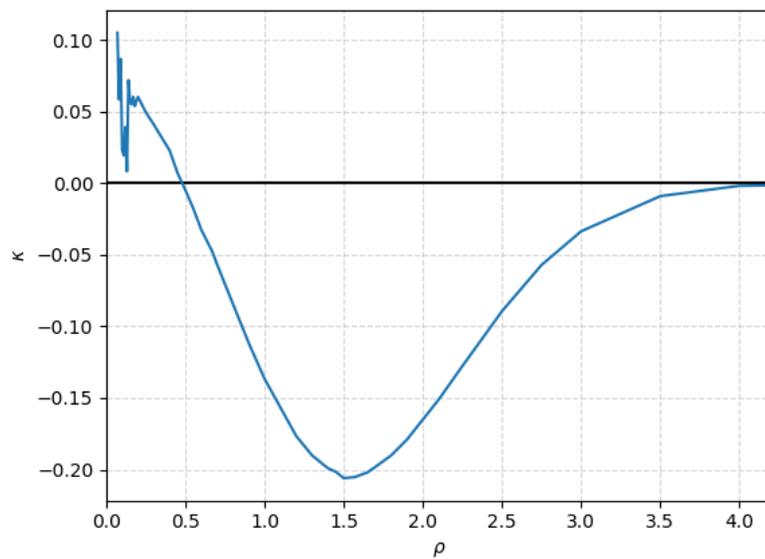


FIGURE 3.8 – Courbure en fonction de ρ , le module de tous les points. Le noyau est gaussien avec $\sigma = 1$

Conclusion

Bilan On a donc réussi à trouver une distance de plus court chemin sur les nuages de points, qui modélise à la fois des déformations naturelles de nos images, et qui est facile à calculer, et donc utilisable en pratique dans les laboratoires médicaux : c'est la distance issue d'une cométrie à noyau. Le développement d'outils informatiques récents, en particulier de la différenciation automatique avec pytorch, a permis de mettre à la portée des laboratoires des techniques jusqu'ici très coûteuses.

Pour approfondir Le premier axe d'approfondissement concerne le choix de la fonction à noyau k , car le comportement de la métrique en dépend. Cependant, tant qu'on reste dans le cadre des cométries à noyaux, on va toujours assister aux comportements de type "covoiturage" qu'on a décrit précédemment, et qui ne sont pas les comportements les plus naturels. Pour éviter cela, on peut s'autoriser à prendre des cométries qui s'en inspirent, mais un peu plus compliquées.

Deuxièmement, on pourrait approfondir l'étude de l'attache aux données. L'attache aux données caractérise, lors du calcul des géodésiques, la proximité que l'on exige entre le point d'arrivée de la géodésique et le point d'arrivée cherché. S'il semble d'abord que ces deux points doivent idéalement être les plus proches possibles, on peut en fait souhaiter garder une certaine liberté sur cet écart. Cela permet d'obtenir algorithmiquement une géodésique assez fidèle mais de coût raisonnable, ce qui revient à éliminer les hautes fréquences, le bruit. Une autre façon de faire est de minimiser directement l'écart entre le point d'arrivée de la géodésique, mais auquel on a rajouté un paramètre proportionnel au coût de la géodésique. Si ce paramètre est nul, comme dans notre cas, on dit qu'on fait du matching exact.

Enfin, il faudrait s'intéresser au problème de l'étiquetage des nuages de points qu'on étudie. En effet, dans toute notre théorie, on cherche à matcher des nuages de points vus comme des n -uplets, numérotés. Or, si on voit ces nuages de points comme la discrétisation de courbes, il n'y a priori aucune raison qu'ils possèdent un étiquetage préalable. Une solution est de voir les formes de départ et d'arrivée comme des répartition de masse, des mesures. Le problème devient alors un problème de transport de masse : il s'agit de trouver un difféomorphisme qui

transporte une mesure sur une autre. Les idées qu'on a développées dans l'étude des RKHS s'appliquent ici aussi.

Bibliographie

- [1] J. FEYDY *Cours de culture mathématique Fondations, Analyse, Géométrie*, Ulm 2017-2018, ch 4-5-6 et Applications de Jean Feydy, Chapitre Cours de culture mathématique Fondations, Analyse, Géométrie et Applications Ch 4 - 5 - 6
- [2] M. MICHELI, P. W. MICHOR, and D. MUMFORD. *Sectional curvature in terms of the cometric, with applications to the Riemannian manifolds of landmarks*, SIAM Journal on Imaging Sciences , 394–433, 2012
- [3] J. FEYDY *Automatic differentiation for applied mathematics*