Etude et extension des *flots de Ricci, Kähler-Ricci et Calabi* dans le cadre du traitement de l'image et de la géométrie de l'information

F. BARBARESCO[1]
[1]THALES AIR DEFENCE, JOINT RADAR SENSORS UNIT, Project & Engineering Department,
7/9, rue des Mathurins, F-92223 Bagneux, Cedex, FRANCE
frederic.barbaresco@fr.thalesgroup.com

Résumé – La théorie des flots géométriques [1] est au cœur de plusieurs disciplines fondamentales des mathématiques (géométrie différentielle et algèbrique, analyse complexe & globale, EDP) et de la physique mathématique (calcul des variations, relativité générale, variété d'Einstein, théorie des cordes) , mais elle repose avant tout sur le socle de la géométrie Riemannienne dont M.Berger dresse le panorama dans un ouvrage récent [5], son extension à des variétés complexes, la géométrie d'Erich Kähler [4,6], dont la vitalité est vantée par J.P. Bourguignon dans [7] et la théorie des surfaces minimales dont la richesse est synthétisée par F. Hélein dans la postface à l'ouvrage [9]. Le présent article souhaite initier une réflexion sur l'utilisation des flots géométriques intrinsèques, tel le flot de Ricci en traitement d'image pour élaborer de nouveaux opérateurs non-linéaires de filtrages anisotropes, et par le biais de la géométrisation de la théorie de l'information par Chentsov, introduire les flots de Kähler-Ricci et de Calabi pour approfondir la notion de noyaux de diffusion de l'information [3].

Abstract – Geometric Flow Theory [1] is cross fertilized by diverse elements coming from Pure Mathematic (geometry, algebra, analyse, PDE) and Mathematical Physic (calculus of variations, General Relativity, Einstein Manifold, String Theory), but its foundation is mainly based on Riemannian Geometry, as explained by M. Berger in a recent panoramic view [5] of this discipline, its extension to complex manifolds, the Erich Kähler's Geometry [4,6], vaunted for its unabated vitality by J.P. Bourguignon in [7], and Minimal Surface Theory recently synthetized by F. Hélein in the book introduction [9]. This paper would like to initiate seminal studies for applying intrinsic flows, such as Ricci flow in Image Processing to conceive new non-linear operator of anisotropic filtering, and thank to information theory geometrization by Chentsov, by using Kähler-Ricci and Calabi flows to deepen the concept of information diffusion kernel [3].

1 La théorie des flots géométriques

La théorie des flots géométriques est une discipline récente [1] dont les ramifications sont nombreuses dans de multiples branches des mathématiques et de la physique mathématique. Contrairement aux flots classiques extrinsèques (flot de courbure moyenne, flot de Wilmore, flot de Laplace-Beltrami, flot de Gauss,…) maintenant largement utilisés en traitement d'image [2,10], la théorie des flots intrinsèques est basée sur des flots qui n'agissent pas explicitement sur la variété elle-même mais implicitement la déforme de proche en proche en modifiant la métrique différentielle g_{ij} associée à cette variété. Dans le flot de Ricci, la métrique est modifiée proportionnellement au tenseur de Ricci R_{ij} associé :

$$\frac{\partial g_{ij}}{\partial t} = 2\left[-R_{ij} + \frac{1}{n} r . g_{ij}\right] \quad \text{avec} \quad r = \left(\int_{M^n} R d\eta\right) / \left(\int_{M^n} d\eta\right)$$

Le problème peut être formalisé sous forme variationnel en minimisant ,comme le fit D.Hilbert, une fonctionnelle S(g) intégrant la courbure scalaire sur l'ensemble de la variété, à volume constant :

$$S(g) = \int_{M^n} R . \sqrt{\det(g)} . d^n x = \int_{M^n} R . d\eta \quad \text{avec} \quad V_{M^n} = \int_{M^n} d\eta = cste$$

Après un rappel de la théorie du flot de Ricci, nous appliquons cet outil théorique pour définir un nouvel opérateur non-linéaire de filtrage anisotrope d'image, en lieu et place du flot de Laplace-Beltrami, introduit par N. Sochen [2]. Notre approche consite donc à considérer une image *I(x,y)*, comme une variété ouverte orientée de dimension 2 dans un espace $\mathbf{R}^3$: *S[x,y,I(x,y)]*. La métrique de cette variété s'exprime uniquement à partir du gradient de l'image et elle est déformée suivant un flot de Ricci modifié pour lequel la courbure scalaire moyenne, habituellement intégrée sur la variétée complète, est obtenue par une intégration locale de la courbure scalaire au voisinage du point considéré. En dimension 2, toute métrique est d'Einstein $R_{ij} = \frac{R}{2} . g_{ij}$, aussi est-on ramené au flot suivant : $\frac{\partial g_{ij}}{\partial t} = [r - R] g_{ij}$

Ce flot peut s'identifier en dimension 2 à un flot de Gauss, la métrique étant déformée proportionnellement à la différence entre la courbure de Gauss locale et sa moyenne estimée dans son voisinage : $\frac{\partial g_{ij}}{\partial t} = 2[K - K_m] g_{ij}$. En se plaçant dans un système de coordonnées isothermes introduites par G. Darmois et utilisées par A. Lichnerowicz [8], on montre que l'opérateur de filtrage obtenu possède la propriété de conserver les contours (équation de diffusion de la métrique), caractérisés par une courbure de Gauss nulle (c.a.d. une des courbures principales nulle). En revanche, l'avantage par rapport au flot de Laplace-Beltrami, réside dans le fait que les points isolés de gradient fort sont lissés par le flot de Ricci, alors que ceux-ci ne l'étaient pas et bloquaient le flot introduit par N.Sochen dans [2]. L'image est ensuite reconstruite en définissant une fonction de la variable complexe $z = x + iy$ déduite des gradients de l'image :

$$\frac{\partial I}{\partial t} = \mathrm{Re} \int_{z_0}^{z} \frac{\partial \Theta}{\partial t}(z) . dz \quad \text{avec} \quad \frac{\partial \Theta}{\partial t}(z) = \frac{\partial}{\partial t}\left(\frac{\partial I}{\partial x}\right) - i \frac{\partial}{\partial t}\left(\frac{\partial I}{\partial y}\right)$$

Dans le cadre de la géométrie de l'information et des méthodes à noyaux associées («Information Diffusion Kernels » [3]), les lois statistiques classiques (lois gaussiennes, multivarié-gaussiennes, Gamma, elliptiques,…) conduisent à des métriques complexes, pour lesquelles nous utilisons le flot de Khäler-Ricci (extension du flot de Ricci sur des variétés complexes) afin de moyenner un noyau donné sur les densités a posteriori pour obtenir un noyau sur les données $K_t(x,x') = \iint K_t(\theta,\theta') . p(x/\theta) . p(x/\theta') . d\theta . d\theta'$. Nous établissons les comportements asymptotiques de ces lois soumises au flot de Kähler-Ricci. Dans le cas des modèles autorégressifs complexes de moyenne nulle, on retrouve un potentiel de Kähler Φ, tel que $g_{ij} = \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}$,sur le polydisque unité, de la forme $\Phi = \sum_{k=1}^{n-1} \alpha_k . \ln\left[1 - |z_k|^2\right] = \ln K_D(z,z)$, de noyau de Bergman associé $K_D(z,z) = \prod_{k=1}^{n-1} \left(1 - |z_k|^2\right)^{\alpha_k}$, qui fut le cas traité par E. Kähler à l'origine dans son papier précurseur [6] (cas qu'il a appelé ***hyper-Abélien***). En effet, dans le cadre de la géométrie de l'information, le potentiel de Kähler est fourni par l'entropie, qui donne dans le cas autoregressif la même forme que la cas Hyper-Abélien : $\Phi = \sum_{k=1}^{n-1} (n-k) . \ln\left[1 - |\mu_k|^2\right] + n . \ln[\pi . e . P_0]$. Nous verrons que la métrique est d'Einstein en un sens plus large « matriciel » :

$$\boxed{R_{ij} = B^{(n)} g_{ij} \quad \text{avec} \quad R = Tr\left[B^{(n)}\right]}$$

Nous concluons enfin par l'étude du flot plus récent introduit par Calabi [4] et qui conduit à des métriques dites « extrêmes », en déformant directement le potentiel de Kähler proportionnellement à la valeur de la courbure scalaire: $\frac{\partial \Phi}{\partial t} = R_\Phi - \bar{R}$. Ce flot dit de Calabi se déduit également de la minimisation d'une fonctionnelle dépendant dans ce cas du carré de la courbure scalaire intégré sur la variété : $T(g) = \int_M R^2 d\eta$.

2 Définition du flot de Ricci

Le flot de Ricci trouve ses racines dans les travaux de Hilbert dans le cadre de la théorie de la Relativité Générale, pour laquelle il formalisa le problème en utilisant un principe d'Action minimale en utilisant les outils du calcul des variations. Ainsi l'Equation d'Einstein fut dérivée par Hilbert à partir de la minimisation d'une fonctionnelle, habituellement nommée de nos jours « Action de Hilbert ». Soit S cette action : $\boxed{S(g) = \int_{M^n} R.\sqrt{\det(g)}.d^n x = \int_{M^n} R.d\eta}$ avec l'élément de volume : $d\eta = \sqrt{\det(g)}.d^n x$ et la courbure scalaire R définie à partir du tenseur de Ricci $R_{\mu\nu}$: $R = \sum_\mu \sum_\nu g^{\mu\nu}.R_{\mu\nu}$ et g le tenseur de la métrique telle que $g = [g_{ij}]$, $ds^2 = \sum_{i,j} g_{ij} dx_i.dx_j$ et $g^{-1} = [g^{ij}]$.

Le théorème fondamental y est donc donné par Hilbert comme suit :
Soit M^n une variété de dimension $n \geq 3$ et supposons que $S(g)$ possède une dérivée variationnelle nulle pour toutes les variations conformes de la métrique g, conservant le volume $V(g)$ constant, alors la courbure moyenne $R(g)$ est constante sur M^n. De plus, si $S(g)$ possède une dérivée variationnelle nulle pour toutes les variations de g, conservant $V(g)$ constant, alors $R(g)$ est constant et g est une métrique de Einstein ; la courbure de Ricci R_{ij}, vérifie $\boxed{R_{ij} = \frac{1}{n} R.g_{ij}}$, et est proportionnelle à la métrique g_{ij}.

Pour des variétés de dimension ≥ 3, en prenant la trace de la dérivée covariante de l'expression précédente et en utilisant l'identité de Bianchi, on montre que pour une métrique d'Einstein, la courbure scalaire est constante.
Une métrique d'Einstein peut donc être obtenue en déformant une métrique suivant le flot modifié suivant (appelé flot de Ricci) : $\boxed{\frac{\partial g_{ij}}{\partial t} = 2\left[-R_{ij} + \frac{1}{n} R g_{ij}\right]}$. Pour résoudre des problèmes d'existence en temps fini, R. Hamilton a proposé en 1982, le flot suivant qui préserve les volumes :

$$\boxed{\frac{\partial g_{ij}}{\partial t} = 2\left[-R_{ij} + \frac{1}{n} r.g_{ij}\right] \quad \text{avec} \quad r = \left(\int_{M^n} R d\eta\right) / \left(\int_{M^n} d\eta\right)}$$

On remarque qu'il s'agit du même flot que précédemment où la courbure scalaire R a été remplacée par la courbure scalaire moyenne r, pour des problèmes d'existence et de convergence du flot. Il est aisé de démontrer que le volume est conservé par ce dernier flot.

3 Lissage NL par le flot de Ricci

3.1 Rappel sur le flot de Laplace-Beltrami

N. Sochen [2] a étudié en traitement d'image de nouveaux opérateurs de filtrage non-linéaire, appelés « Geometric Filter », basés sur le flot extrinsèque, dit de courbure moyenne et l'opérateur de diffusion de Laplace-Beltrami. Le flot de Beltrami obtenu s'interprète comme un flot de courbure moyenne $\frac{\partial S}{\partial t} = \kappa.\vec{N}$ qui agit sur la surface $S = (x, y, I)$. L'image en niveau de gris I est considérée comme une variété plongée dans un espace $\Re^3$, et sur laquelle la métrique différentielle est définie par :

$$ds^2 = \sum_{i,j} g_{i,j} dx_i dx_j = dx^2 + dy^2 + dI^2 = dx^2 + dy^2 + (\partial_x I.dx + \partial_y I.dy)^2$$

où la métrique est donc : $g = [g_{ij}]_{i,j=1,2} = \begin{bmatrix} 1 + (\partial_x I)^2 & \partial_x I \partial_y I \\ \partial_y I \partial_x I & 1 + (\partial_y I)^2 \end{bmatrix}$

et $g^{-1} = [g^{ij}]_{i,j=1,2} = \frac{1}{1 + (\partial_x I)^2 + (\partial_y I)^2} \begin{bmatrix} 1 + (\partial_y I)^2 & -\partial_y I \partial_x I \\ -\partial_x I \partial_y I & 1 + (\partial_x I)^2 \end{bmatrix}$

avec les notations suivantes que nous conserverons dans la suite : $\partial_x I = \frac{\partial I}{\partial x}$, $\partial_y I = \frac{\partial I}{\partial y}$. L'opérateur de Laplace-Beltrami $\frac{\partial I}{\partial t} = \Delta_g I$ donne directement l'expression du nouvel opérateur non-linéaire $\frac{\partial I}{\partial t} = \frac{1}{\sqrt{1 + |\nabla I|^2}} div\left(\frac{\nabla I}{\sqrt{1 + |\nabla I|^2}}\right)$ qui possède la propriété de lisser le bruit (c'est-à-dire $|\nabla I| << 1 \Rightarrow \frac{\partial I}{\partial t} = div(\nabla I) = \Delta I$) et de conserver les contours (c'est-à-dire $|\nabla I| >> 1 \Rightarrow \frac{\partial I}{\partial t} = 0$).

3.2 Filtrage anisotrope d'image & flot de Ricci

3.2.1 Flot de Ricci agissant sur les gradients de l'image

Nous proposons une approche alternative au flot de Laplace-Beltrami, non plus basée sur le flot de courbure moyenne agissant sur la variété, mais sur un flot de Ricci agissant sur la métrique différentielle de la variété. C'est à dire, partant de la même métrique que précédemment , on va étudier le flot de Ricci ou plus particulièrement la version de Hamilton $\frac{\partial g_{ij}}{\partial t} = 2\left[-R_{ij} + \frac{1}{n} r.g_{ij}\right]$ avec $r = \left(\int_{Voisinage} R d\eta\right) / \left(\int_{Voisinage} d\eta\right)$ avec $d\eta = \sqrt{\det(g)}.d^n x$, que nous adaptons au traitement de l'image en calculant la courbure scalaire moyenne non plus sur l'ensemble de la variété, mais sur un voisinage local. L'opérateur ainsi obtenu sur la métrique g_{ij} permettra par identification avec la dérivée partielle : $\frac{\partial g}{\partial t} = \frac{\partial(\nabla I.\nabla I^T)}{\partial t}$ d'en déduire les opérateurs agissant sur le gradient de l'image. Pour nos développements, il nous faut alors calculer les expressions suivantes pour la nouvelle métrique définie sur l'image, afin d'aboutir à l'expression du tenseur de Ricci et de la courbure scalaire. Premièrement, on détermine le symbole de Christoffel donné par : $\Gamma^\lambda_{\mu\nu} = \frac{1}{2} \sum_\tau g^{\lambda\tau}.\left[\frac{\partial g_{\nu\tau}}{\partial x^\mu} + \frac{\partial g_{\tau\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\tau}\right]$ et le tenseur de Riemann donné par : $R^\lambda_{\mu\nu\tau} = \frac{\partial \Gamma^\lambda_{\mu\tau}}{\partial x^\nu} - \frac{\partial \Gamma^\lambda_{\mu\nu}}{\partial x^\tau} + \sum_\eta \left[\Gamma^\eta_{\mu\tau}.\Gamma^\lambda_{\eta\nu} - \Gamma^\eta_{\mu\nu}.\Gamma^\lambda_{\eta\tau}\right]$.

Le tenseur de Ricci est obtenu à partir du tenseur de Riemann par la relation $R_{\mu\nu} = \sum_\lambda R^\lambda_{\mu\nu\lambda}$ et la courbure scalaire se déduit du tenseur de Riemann par la relation suivante $R = \sum_\mu \sum_\nu g^{\mu\nu}.R_{\mu\nu}$. On déduit pour notre cas d'analyse l'expression du tenseur de Ricci :

$$\boxed{\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \frac{((\partial_{xy} I)^2 - \partial_{xx} I \partial_{yy} I)}{(1 + (\partial_x I)^2 + (\partial_y I)^2)^2} \begin{bmatrix} (1 + (\partial_x I)^2) & \partial_x I \partial_y I \\ \partial_y I \partial_x I & (1 + (\partial_y I)^2) \end{bmatrix}}$$

En remarquant que la courbure scalaire se réduit à : $\boxed{R = \frac{2((\partial_{xy} I)^2 - \partial_{xx} I \partial_{yy} I)}{(1 + (\partial_x I)^2 + (\partial_y I)^2)^2}}$, on vérifie bien que le tenseur d'Einstein pour n=2 , $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu}.R$ est nul $G_{\mu\nu} = 0$. On a alors un flot de Ricci $\boxed{\frac{\partial g_{ij}}{\partial t} = [r - R]g_{ij}}$, dont on déduit sur les gradients le flot :

$$\frac{\partial}{\partial t}\begin{bmatrix}\partial_x I \\ \partial_y I\end{bmatrix} = \left[r - \frac{2\left((\partial_{xy} I)^2 - \partial_{xx} I \partial_{yy} I\right)}{\left(1 + (\partial_x I)^2 + (\partial_y I)^2\right)^2}\right]\begin{bmatrix}\frac{\left(1 + (\partial_x I)^2\right)}{\partial_x I} \\ \frac{\left(1 + (\partial_y I)^2\right)}{\partial_y I}\end{bmatrix}$$

En dimension 2, toutes les métriques sont des métriques d'Einstein, ou de façon équivalente, une métrique d'Einstein n'a pas besoin d'être de courbure scalaire constante, on a alors à remplacer dans l'équation précédente, pour réduire le défaut de courbure la courbure scalaire par la courbure de Gauss K, avec $R = -2K$: $\boxed{\frac{\partial g}{\partial t} = 2[K - K_m]g}$ ou K_m est la courbure de Gauss moyenne. La courbure de Gauss est égale au produit des courbures des directions principales et est donnée par : $K = \frac{R_{1212}}{\det(g)}$ où R_{1212} est une des composantes du tenseur de courbure de Riemann covariant (ou symbole de Riemann de seconde espèce) : $R_{\tau\alpha\lambda\beta} = \sum_{\sigma} g_{\sigma\tau} R^{\sigma}_{\alpha\lambda\beta}$.

3.2.2 Propriétés du flot pour la préservation des contours

Considérons le flot de Ricci non normalisé $\frac{\partial g_{ij}}{\partial t} = -2.R_{ij} \quad \forall i, j$.

Afin d'étudier le comportement du flot suivant les valeurs du gradient de l'image, nous introduisons le système de coordonnées isothermes lié à la métrique, introduit par G. Darmois. Dans un tel système de coordonnées locales $\{x^i\}_{i=1,2}$ isothermes, les quantités suivantes sont nulles : $F^k = \Delta_g x^i = -\sum_{\lambda,\mu} g^{\lambda\mu}\Gamma^k_{\lambda\mu}$ pour $k = 1,2$, avec :

$\Delta_g f = \sum_{\lambda,\mu} g^{\lambda\mu}\left(\partial^2_{\lambda\mu} f - \sum_k \Gamma^k_{\lambda\mu}\partial_k f\right)$ l'opérateur de Laplace-Beltrami.

Etant donnée une variété orientée dans l'espace, elle peut toujours être envisagée comme variété coordonnée dans un système de coordonnées isothermes. Or, ces quantités F^k interviennent d'une manière simple dans les expressions du tenseur de Ricci. Un calcul local permet d'établir [8] : $R_{ij} = -G_{ij} - L_{ij}$ donné par les expressions suivantes :

$$G_{ij} = \frac{1}{2}\sum_{\lambda\mu} g_{\lambda\mu}\partial^2_{\lambda\mu} g_{ij} + H_{ij} \quad \text{et} \quad L_{ij} = \frac{1}{2}\left(\sum_{\mu} g_{i\mu}\partial_j F^{\mu} + \sum_{\mu} g_{j\mu}\partial_i F^{\mu}\right)$$

Les H_{ij} désignent des polynômes par rapport aux $g_{\lambda\mu}$ et $g^{\lambda\mu}$ et à leur dérivées premières. Ainsi en coordonnées isothermes :

$$-2.R_{ij} = \sum_k \sum_l g^{kl}.\frac{\partial^2 g_{ij}}{\partial x^k \partial x^l} + Q_{ij}\left(g^{-1}, \partial g\right)$$

Dans le cas limite $|\nabla I| << 1$ alors $g^{-1} = [g^{ij}]_{i,j=1,2} \approx \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}$ et la métrique est Euclidienne : $ds^2 \approx dx^2 + dy^2$, et la métrique est diffusée isotropiquement $\frac{\partial g_{ij}}{\partial t} = \Delta_g g_{ij}$. Le flot de Ricci diffuse alors la métrique $[g_{ij}]_{i,j=1,2} = Id + \nabla I.\nabla I^T$ sur la variété et donc induit un lissage des gradients : $\frac{\partial\left(\nabla I.\nabla I^T\right)}{\partial t} = \Delta\left(\nabla I.\nabla I^T\right)$. Le flot de Ricci est stoppé $\frac{\partial g_{ij}}{\partial t} = 0$ uniquement aux endroits où $\boxed{\left(\partial^2_{xy} I\right)^2 = \partial^2_{xx} I.\partial^2_{yy} I}$ (équation de Monge-Ampère) pour lesquels le tenseur de Ricci, la courbure scalaire et de Gauss sont nulles. La courbure de Gauss étant égale au produit des courbures principales, cela signifie que l'EDP de diffusion de la métrique est stoppée aux endroits ou l'une des deux courbures principales est nulles (cas des contours de l'image de type : bord d'objet ou coin d'objet). On voit ici une différence avec le Flot de Laplace-Beltrami qui stoppait le flot aux endroits où le gradient était élevé. Dans le cas du flot de Ricci, des points de bruit forts seront lissés alors qu'ils ne le sont pas avec le flot de Laplace-Beltrami.

3.2.3 Reconstruction de l'image à partir des gradients

Considérons la surface $S = (x, y, I)$ et le repère du plan tangent local : $\partial_x S = [1 \quad 0 \quad \partial_x I]^T$, $\partial_y S = [0 \quad 1 \quad \partial_y I]^T$, on définie la fonction : $\Phi(z) = \frac{\partial S}{\partial x} - i\frac{\partial S}{\partial y} = [1 \quad -i \quad (\partial_x I - i\partial_y I)]^T$ avec $z = x + iy$

La surface S est alors reconstruite à une translation prêt, à chaque déformation de la métrique par le flot de Ricci en intégrant dS le long d'un chemin de z_0 à z : $S(x, y) = S(x_0, y_0) + \mathrm{Re}\int_{z_0}^{z}\Phi(z).dz$,

en effet $dS = \frac{\partial S}{\partial x}dx + \frac{\partial S}{\partial y}dy = \mathrm{Re}[\Phi(z)dz]$, et l'intensité de l'image se déduit , en ne conservant que la troisième composante :

$I(x, y) = I(x_0, y_0) + \mathrm{Re}\int_{z_0}^{z}\Theta(z).dz$ avec $\Theta(z) = \frac{\partial I}{\partial x} - i\frac{\partial I}{\partial y}$

le vecteur normal étant donné par $\vec{N} = \Phi(z) \wedge \Phi(z)^* / [i.\Phi(z).\Phi(z)^*]$

On en déduit aussi l'EDP agissant sur I :

$\frac{\partial I}{\partial t} = \mathrm{Re}\int_{z_0}^{z}\frac{\partial\Theta}{\partial t}(z).dz$ avec $\frac{\partial\Theta}{\partial t}(z) = \frac{\partial}{\partial t}\left(\frac{\partial I}{\partial x}\right) - i\frac{\partial}{\partial t}\left(\frac{\partial I}{\partial y}\right)$

4 Flots de Kähler-Ricci et de Calabi en géométrie de l'information

4.1 Notion de noyaux de diffusion de l'information

Dans le cadre de la géométrie de l'information introduite par Rao et axiomatisé par Chentsov, la métrique différentielle est définie dans l'espace des paramètres θ des densités de probabilités $p(x/\theta)$ par :

$ds^2 = \sum_{i,j} g_{ij}(\theta).d\theta_i.d\theta_j^* = d\theta^+.I(\theta).d\theta = Var[d\ln p(x/\theta)]$ avec la métrique $I(\theta) = [g_{ij}]_{i,j}$ donnée par la matrice de Fisher :

$$g_{ij}(\theta) = -\int p(x/\theta).\frac{\partial^2 \ln p(x/\theta)}{\partial\theta_i.\partial\theta_j^*}.dx = -E\left[\frac{\partial^2 \ln p(x/\theta)}{\partial\theta_i.\partial\theta_j^*}\right]$$

Cette métrique est également donnée par le Hessien de l'Entropie :

$$g_{ij}(\theta) = -\frac{\partial^2 H(p)}{\partial\theta_i\partial\theta_j^*} \quad \text{avec} \quad H(\theta) = \int p(x/\theta).\ln p(x/\theta).dx$$

Lafferty [3] moyenne un noyau donné sur les densités a posteriori $\frac{\partial\theta}{\partial t} = \Delta_g\theta$, pour obtenir un noyau sur les données :

$$K_t(x, x') = \iint K_t(\theta, \theta').p(x/\theta).p(x/\theta').d\theta.d\theta'$$

4.2 Potentiel de Kähler et flot de Kähler-Ricci

Nous généralisons l'approche [3] en appliquant le flot de Ricci sur une variété « entropie » dont la métrique est donnée par la matrice de Fisher. Dans notre généralisation le flot de Ricci agit sur des variables complexes, évoluant sur des variétés complexes. Pour ce faire, on utilise l'extension du flot de Ricci à des variétés complexes, dites Kähleriennes. Soit M^n une variété complexe à n dimensions, compacte ou non, munie d'une métrique Kählerienne, que l'on peut représenter localement par sa forme Riemannienne définie positive : $ds^2 = 2\sum_{i,j=1}^{n} g_{i\bar{j}}.dz^i dz^{\bar{j}}$ avec $[g_{i\bar{j}}]_{i,j}$ matrice Hermitienne définie positive. La condition de Kähler s'exprime par l'existence locale d'une fonction potentielle de Kähler Φ, unique à une fonction pluriharmonique près, telle que $\boxed{g_{i\bar{j}} = \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}}$. On définit de même les expressions suivantes des tenseurs associés :

$$\Gamma^k_{ij} = \sum_{l=1}^n g^{k\bar{l}} \frac{\partial g_{i\bar{l}}}{\partial z^j} \text{ et } \Gamma^{\bar{k}}_{\bar{i}\bar{j}} = \sum_{l=1}^n g^{\bar{k}l} \frac{\partial g_{l\bar{i}}}{\partial \bar{z}^j}, R_{i\bar{j}k\bar{l}} = -\frac{\partial^2 g_{i\bar{j}}}{\partial z^k \partial \bar{z}^l} + \sum_{p,q=1}^n g^{p\bar{q}} \frac{\partial g_{i\bar{q}}}{\partial z^k}\frac{\partial g_{p\bar{j}}}{\partial \bar{z}^l}$$

Dans la géométrie de Kähler , le tenseur de Ricci est donnée par l'expression condensée : $\boxed{R_{i\bar{j}} = -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j}}$ dont on déduit la courbure scalaire : $R = \sum_{k,l=1}^n g^{k\bar{l}}.R_{k\bar{l}}$. Le flot de Kähler-Ricci associé y est donné par : $\boxed{\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} + \frac{1}{n} R g_{i\bar{j}}}$

4.3 Potentiel de Kähler associée à l'entropie d'un processus autorégressif complexe

Si l'on considère un vecteur de données complexe X_n de densité de probabilité de Whishart complexe de moyenne nulle :

$$P(X_n/R_n) = (\pi)^{-n}.|R_n|^{-1}.e^{-Tr[\hat{R}_n.R_n^{-1}]} \text{ avec } E[\hat{R}_n] = E[X_n X_n^+] = R_n.$$

Sous l'hypothèse d'un signal suivant un modèle autoregressive complexe, l'expression finale de l'Entropie nous est alors donnée en fonction des coefficients de réflexion $\{\mu_k \in C / |\mu_k| < 1\}$:

$$\boxed{H_n = \ln|R_n| + n.\ln(\pi e) = \sum_{k=1}^{n-1}(k-n).\ln[1-|\mu_k|^2] + n.\ln[\pi.e.P_0]}$$

Si nous revenons à la définition de la métrique de l'information, celle-ci, nous est donnée par le Hessien de l'Entropie :

$$g_{ij}(\theta) = -\frac{\partial^2 H(p)}{\partial\theta_i \partial\theta_j^*} \text{ avec } H(p) = \int p(x/\theta).\ln p(x/\theta).dx$$

En considérant, le vecteur de paramètre suivant :

$$\theta^{(n)} = [P_0 \quad \mu_1 \quad\quad \mu_{n-1}]^T = [\theta_1^{(n)} \ \cdots \ \theta_n^{(n)}]^T$$

et en développant le calcul du Hessien, on trouve : $g_{11} = nP_0^{-1}$,

$$g_{ij} = \frac{(n-i).\delta_{ij}}{(1-|\mu_i|^2)^2} \quad \text{et} \quad \boxed{ds_n^2 = n.\left(\frac{dP_0}{P_0}\right)^2 + \sum_{i=1}^{n-1}(n-i)\frac{|d\mu_i|^2}{(1-|\mu_i|^2)^2}}$$

On en déduit alors une distance au sens de la géometrie de l'information entre deux modèles autoregressifs complexes de moyenne nulle :

$$D(\theta^{(1)},\theta^{(2)}) = \int_{\theta^{(1)}}^{\theta^{(2)}} \sqrt{n(d\ln P_0)^2 + \sum_{i=1}^{n-1}\frac{(n-i)}{4}\left(d\ln\frac{1+r_i}{1-r_i}\right)^2} \quad \text{et } r_i = |\mu_i|$$

On intègre via un automorphisme holomorphique (transf. de Möbius). Si on ne s'intéresse qu'au terme dépendant des μ_i, on est ramené à :

$$\boxed{D(\theta^{(1)},\theta^{(2)}) = \sqrt{\sum_{i=1}^{n-1}\frac{(n-i)}{2}\left(\ln\left[\frac{1+\delta_{i(1)}^{(2)}}{1-\delta_{i(1)}^{(2)}}\right]\right)^2}} \text{ avec } \boxed{\delta_{i(1)}^{(2)} = \left|\frac{\mu_i^{(1)}-\mu_i^{(2)}}{1-\mu_i^{(1)}\mu_i^{(2)*}}\right|}$$

Nous en déduisons la matrice de Ricci : $[R_{k\bar{l}}]_{k,l} = \left[-\frac{\partial^2 \log\det(g_{ij})}{\partial z_k \partial \bar{z}_l}\right]_{k,l}$

$$R_{k\bar{l}} = -2\delta_{kl}(1-|\mu_k|^2)^{-2} \quad \text{pour } k = 2,...,n-1 \text{ et } R_{11} = P_0^{-2}$$

La courbure scalaire étant définie par ailleurs :

$$R = \sum_{k,l} g^{k\bar{l}}.R_{k\bar{l}} \quad \text{soit} \quad R = -2.\left[n^{-1} + \sum_{j=1}^{n-1}(n-j)^{-1}\right] = -2.\left[\sum_{j=0}^{n-1}(n-j)^{-1}\right]$$

La courbure scalaire diverge lorsque n tend vers l'infini. La métrique est d'Einstein en un sens plus large « matriciel » :

$\boxed{R_{ij} = B^{(n)} g_{ij} \text{ avec } R = Tr[B^{(n)}]}$ avec $B^{(n)} = diag\{\cdots,(n-i)^{-1},\cdots\}$

Si nous examinons le flot de Kähler-Ricci dans le cas non normalisé :

$\frac{\partial g_{ij}}{\partial t} = -R_{ij}$, celà donne $\frac{\partial \ln(1-|\mu_i|^2)}{\partial t} = -\frac{1}{(n-i)}$ et $\frac{\partial \ln P_0}{\partial t} = \frac{1}{n}$

$$|\mu_i(t)|^2 = 1-(1-|\mu_i(0)|^2).e^{-\frac{t}{(n-i)}} \xrightarrow[t\to\infty]{} |\mu_i(t)|^2 = 1.$$

Les coefficients de reflexion tendent asymptotiquement vers le cercle unité avec une vitesse dépendante de leur indice.

4.4 Flot de Calabi & géométrie de l'information

La notion de métriques « extrêmes » de Kähler a été introduite par E. calabi dans [4], en introduisant pour montrer leur existence, un nouveau flot, appelé « flot de Calabi ». Considérons la fonctionnelle suivante : $\boxed{\Theta(g) = \int_M R^2 d\eta}$. Les points critiques de $\Theta(g)$ sont appelés les métriques extrêmes. Toute métrique de Kähler-Einstein est également une métrique extrême qui minimise $\Theta(g)$. La solution est donnée par l'EDP parabolique suivante : $\boxed{\frac{\partial\psi}{\partial t} = R_\psi - \bar{R}}$ avec ψ le potentiel de Kähler associé à la métrique de Kähler $g_{i\bar{j}} = \frac{\partial^2\psi}{\partial z^i \partial \bar{z}^j}$, R_ψ la courbure scalaire associée et $\bar{R}$ sa valeur moyenne sur la variété.

Considérons le chemin $\psi(t)$ $(0 \le t < 1)$ et supposons l'existence des potentiels de distorsion de Kähler $\{\phi(s,t) : 0 \le t < 1\}$ soumis à :

$\frac{\partial\phi}{\partial t} = R(\phi) - \bar{R}$ et $\phi(0,t) = \psi(t)$. La longueur du chemin $L(s)$ est alors donnée par : $L(s) = \int_0^1\left(\int_V\left(\frac{\partial\phi(s,t)}{\partial t}\right)^2 .d\eta_{\phi(s,t)}\right)^{\frac{1}{2}}.dt$

Dans le cadre de la géométrie de l'information, ce flot agit sur l'entropie $-H(p)$

$$-\sum_{k=1}^{n-1}(k-n).\frac{\partial\ln[1-|\mu_k|^2]}{\partial t} - n.\frac{\partial\ln[\pi.e.P_0]}{\partial t} = -2.\left[\sum_{k=1}^{n}\frac{1}{n-k} + \frac{1}{n}\right]$$

Ainsi, on obtient : $\frac{\partial\ln(1-|\mu_k|^2)}{\partial t} = -\frac{2}{(n-k)^2}$ et $\frac{\partial\ln P_0}{\partial t} = \frac{2}{n^2}$

Nous obtenons un comportement comparable au flot de Kähler-Ricci à un facteur prêt. L'entropie croit en proportion de la courbure scalaire.

Références

[1] International Conference "Geometric Flows : Theory & Computation", IPAM, UCLA, USA, February 3-4 , 2004

[2] A. Spira, N. Sochen & R. Kimmel, "Geometric Filter, Diffusion Flows and Kernels in Image Processing", Handbook of Geometric Computing, Springer Verlag, Heidelberg, pp.203-230, February 2005

[3] Lafferty J. & Lebanon G., "Diffusion Kernels on Statistical Manifolds", CMU-CS-04-101, work supported by NSF, January 2004

[4] E. Calabi, « Extremal Kähler metrics », Seminar Ann. of Math. Stud., n°102, Princeton University Press, pp.259-290, 1982

[5] M. Berger, « Panoramic View of Riemannian Geometry », Springer, 2004

[6] E. Kähler, « Über eine bemerkenswerte Hermitesche Metrik », Abh. Math. Sem. Hamburg Univ., n°9, pp.173-186, 1932

[7] J.P. Bourguignon, « The Unabated Vitality of Kählerian Geometry », in *Kähler Erich Mathematical Works*, Edited by R. Berndt, Berlin, Walter de Gruyter, ix, 2003

[8] A. Lichnerowicz, « Sur les équations relativistes de la gravitation », bulletin de la SMF, t.80, pp.237-251, 1952

[9] F. Hélein, Postface à "Analyse et géométrie", Coll. Géométrie au XXème siècle : histoire et horizons, J. Konneiher & al Ed.

[10] F. Barbaresco & al ,"Lissage anisotropique de paramètres statistiques par extension de l'opérateur diffusif de Beltrami dans le cadre de la géométrie de l'information », GRETSI'03, Sept. 2003