Inégalités de courbure-dimension

Jonas Kahn
Édité par Yann Ollivier

18 mai 2004

Nous présentons ici les inégalités de courbure-dimension, introduites par Bakry et Émery. Le problème est le suivant : on se donne une variété riemannienne (M, g) munie aussi d'une mesure μ (supposée absolument continue par rapport à la mesure de volume), et l'on cherche des conditions sous lesquelles l'espace (M, g, μ) serait concentré.

La stratégie adoptée par Bakry et Émery est la suivante : on trouve un processus de diffusion dont la mesure d'équilibre est μ ; alors, une inégalité dite de courbure-dimension, portant sur le générateur infinitésimal du processus de diffusion, garantira lorsqu'elle est vérifiée que (M, g, μ) satisfait une inégalité de Sobolev logarithmique, qui implique la concentration (par un argument simple pour lequel on renvoie à l'exposé d'introduction ; on renvoie à [ABCFGMRS00] ou à ce même exposé pour une définition et une discussion rapide des inégalités de Sobolev logarithmiques). On va donc procéder en dégageant un ensemble de conditions suffisantes pour que cette stratégie fonctionne.

Le cas modèle est celui du processus d'Ornstein-Uhlenbeck sur $\mathbb{R}^n$, présenté dans la deuxième partie de ce texte, et dont la mesure d'équilibre est la gaussienne. L'analyse de l'évolution de l'entropie sous ce processus permet de montrer que $\mathbb{R}^n$ muni de la mesure gaussienne satisfait une inégalité de Solobev logarithmique et est donc concentré. La généralisation présentée ensuite montre que ceci peut s'interpréter en disant qu'en un certain sens, $\mathbb{R}^n$ gaussien est un espace à courbure positive.

Pour une présentation plus détaillée, on renvoie à [Led00] ou [Bak94], qui contiennent eux-mêmes un grand nombre de références.

1 Processus de diffusion sur les variétés

1.1 Définition

On donne ici une définition des processus de diffusion généralisant le noyau de la chaleur bien connu.

Définition 1 – *Un* processus de diffusion *sur une variété mesurée (M, μ) est une famille d'opérateurs $(P_t)_{t\geqslant 0}$ de $L^2(M, \mu) \cap C^\infty(M)$ dans lui-même, tel que ce semi-groupe soit donné par un générateur infinitésimal L :*

$$\lim_{t\to 0} \frac{P_t f - f}{t} = Lf$$

et tel que L soit d'ordre 2 i.e. prenne la forme

$$Lf(x) = \sum g_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum b_i(x) \frac{\partial f}{\partial x_i}(x)$$

où pour tout x, $g_{ij}(x)$ est une forme quadratique définie positive.

D'après cette définition, on a en particulier $P_t 1 = 1$ pour tout t, et $\lim_{t\to 0} P_t f = f$.

L'exemple standard à avoir en tête est bien sûr le semi-groupe de la chaleur sur une variété riemannienne compacte où l'on prend $\mu = \mathrm{vol}$. Le générateur L est alors le laplacien : $L = \Delta$ (on prend comme convention pour le signe du laplacien celle qui donne $\sum \partial^2/\partial x_i^2$ sur $\mathbb{R}^n$).

Si les g_{ij} ne formaient pas une forme quadratique positive, on aurait quelque chose de qualitativement très différent d'une diffusion, qui amplifierait les variations au lieu de les lisser. Si la forme quadratique était positive mais non définie, on aurait un lissage dans certaines directions mais pas toutes.

On n'a pas ajouté de terme d'ordre 0 dans L ; un tel terme correspondrait à de l'apparition ou de la disparition de masse.

Il existe des définitions plus générales ; hors d'une variété, on exprime le fait que le générateur infinitésimal est d'ordre 2 en imposant que, pour toute fonction $\Phi : \mathbb{R}^k \to \mathbb{R}$, on ait la règle de dérivation attendue pour $L(\Phi(f_1, \ldots, f_k))$.

En particulier, les g_{ij} définissent une métrique riemannienne sur la variété M, de sorte qu'on a toujours

$$L = \Delta + X.\nabla$$

où Δ est le laplacien associé à la métrique g, et où X est un champ de vecteurs sur M défini à partir de g et b (les coordonnées de X ne sont pas forcément les b_i à cause des termes de premier ordre apparaissant dans l'expression du laplacien en coordonnées quelconques).

On munit désormais M de la structure riemannienne associée à ce processus.

1.2 Mesures réversibles, existence

On va chercher une mesure de probabilité π invariante par le processus de diffusion. On va en fait chercher une propriété un peu plus forte : une mesure *réversible*, c'est-à-dire que pour toutes fonctions f et g on ait $\int f\, Lg\, d\pi = \int g\, Lf\, d\pi$ (qui vaut alors aussi $-\int \nabla f.\nabla g\, d\pi$, cf. ci-dessous), c'est-à-dire que L agissant sur $L^2(\pi)$ est autoadjoint. Intuitivement, la réversibilité pose que non seulement la mesure π est invariante sous le processus de diffusion, mais qu'en plus, pour tout couple de points infiniment voisins, la masse allant de l'un vers l'autre s'équilibre avec la masse allant de l'autre vers l'un (l'invariance exprime seulement que pour tout point, la masse totale partant vers tous ses voisins est égale à la masse totale reçue de tous les voisins).

En particulier, la masse totale $\int f\, d\pi$ d'une fonction est préservée par le processus ; en effet $\frac{d}{dt}_{|t=0} \int P_t f\, d\pi = \int Lf\, d\pi = \int Lf.1\, d\pi = \int f.L1\, d\pi = 0$.

Cherchons à quoi équivaut la condition $\int f\,Lg\,d\pi = \int g\,Lf\,d\pi$: en écrivant $\pi = \varphi\,d\,\mathrm{vol}$ il vient

$$\begin{aligned}
\int f\,Lg\,d\pi &= \int f\,Lg\,\varphi = \int f\,\Delta g\,\varphi + \int fX.\nabla g\,\varphi \\
&= \int g\Delta\,(f\varphi) + \int fX.\nabla g\,\varphi \\
&= \int g\,\Delta f\,\varphi + 2\int g\nabla f.\nabla\varphi + \int fg\nabla.\nabla\varphi + \int fX.\nabla g\,\varphi \\
&= \int g\,Lf\,\varphi + 2\int g\nabla f.\nabla\varphi - \int \nabla(fg)\,.\nabla\varphi + \int X.(f\nabla g - g\nabla f)\,\varphi \\
&= \int g\,Lf\,\varphi + \int (f\nabla g - g\nabla f)\,.\,(X\varphi - \nabla\varphi)
\end{aligned}$$

En particulier, si on pose $\varphi = e^{-U}$, on a $\nabla\varphi = -\nabla U\,\varphi$ et alors $\int(f\,Lg - g\,Lf)\varphi = \int(f\nabla g - g\nabla f)(X + \nabla U)\varphi$. Ce qui prouve que si[1] $X = -\nabla U$, la mesure $\pi = e^{-U}$ vol est réversible. Si X est un gradient, on a donc construit explicitement une mesure réversible pour le processus de diffusion $L = \Delta + X.\nabla$.

Par la suite, on s'intéressera donc particulièrement au cas $X = -\nabla U$.

Remarquons que dans cette situation, on a $\int f\,Lg\,d\pi = \int f\,\Delta g\,e^{-U} + \int f\,\nabla g.\nabla e^{-U} = -\int \nabla g.\nabla(fe^{-U}) + \int f\,\nabla g.\nabla e^{-U} = -\int \nabla f.\nabla g\,e^{-U}$ et donc

$$\int f\,Lg\,d\pi = \int g\,Lf\,d\pi = -\int \nabla f.\nabla g\,d\pi$$

2 Exemple : le processus d'Ornstein-Uhlenbeck

2.1 Présentation du processus

Ce processus est le processus de diffusion sur $\mathbb{R}^n$ dont la mesure réversible, obtenue par la construction précédente, soit la gaussienne $d\pi = e^{-\|x\|^2/2}dx$. Le gradient de $\|x\|^2/2$ étant x, son générateur infinitésimal est donc $L = \Delta - x.\nabla$. Intuitivement, cela correspond à une diffusion habituelle avec une force de rappel proportionnelle à la distance à l'origine. Le rappel compense la diffusion et la mesure d'équilibre est gaussienne de variance 1.

Il est classique que le mouvement peut s'interpréter comme celui d'une particule dont la position au temps t est $e^{-t}x + \sqrt{1 - e^{-2t}}X$ où x est la position initiale et où X est une variable aléatoire normale centrée dans $\mathbb{R}^n$ dont chaque composante est de variance 1 ; le premier terme est la force de rappel, le second, la diffusion. On a ainsi

$$P_t f(x) = \mathbb{E}_X\left(f\left(e^{-t}x + \sqrt{1 - e^{-2t}}X\right)\right)$$

et en particulier

$$\lim_{t\to\infty} P_t f = \mathbb{E}_X f(X) = \int f\,d\pi$$

On vérifie immédiatement sur l'expression ci-dessus qu'on a l'identité

$$\nabla P_t f = e^{-t}P_t\nabla f$$

et donc $\|\nabla P_t f\| \leqslant e^{-t}P_t\|\nabla f\|$, ce qui nous servira par la suite.

[1] et seulement si, en fait ; $f\nabla g - g\nabla f$ est suffisamment quelconque.

2.2 Concentration de $\mathbb{R}^n$ gaussien

On va maintenant prouver, grâce à ce processus, que l'espace $\mathbb{R}^n$ muni de la mesure gaussienne est concentré, de diamètre observable gaussien 1 (quelle que soit la dimension ! mais à noter qu'un point tiré au hasard selon la gaussienne $\frac{1}{(2\pi)^{n/2}} e^{-\|x\|^2/2} dx$ a une norme de l'ordre de $\sqrt{n}$, et on retrouve ainsi le rapport habituel).

Pour cela, on va passer par le critère de l'inégalité de Sobolev logarithmique présenté dans l'exposé d'introduction. On va donc, pour une fonction f strictement positive sur $\mathbb{R}^n$, évaluer son entropie sous la mesure π et la contrôler par l'intégrale de son gradient.

L'idée est d'étudier l'évolution de l'entropie de $P_t f$ avec le temps. On a

$$\mathrm{Ent}_\pi f = \int_{\mathbb{R}^n} f \log \frac{f}{\int f} \, d\pi = - \int_{\mathbb{R}^n} \int_{s=0}^{\infty} \frac{d}{ds} \left(P_s f \log \frac{P_s f}{\int f} \right) ds \, d\pi$$

or on a par définition $\frac{d}{ds} P_s f = L P_s f$ d'où, par dérivation de $P_s f \log P_s f$,

$$\mathrm{Ent}_\pi f = - \int_{s=0}^{\infty} \int_{\mathbb{R}^n} L P_s f \, (\log P_s f + 1) \, d\pi$$

Or, pour toutes fonctions f et g on a $\int f \, Lg \, d\pi = - \int \nabla f . \nabla g \, d\pi$ car π est la mesure réversible pour le semi-groupe engendré par f, et donc

$$\mathrm{Ent}_\pi f = \int_{s=0}^{\infty} \int_{\mathbb{R}^n} \frac{\|\nabla P_s f\|^2}{P_s f} \, d\pi \leqslant \int_{s=0}^{\infty} \int_{\mathbb{R}^n} e^{-2s} \frac{(P_s \|\nabla f\|)^2}{P_s f} \, d\pi$$

d'après la formule pour $\nabla P_s f$ mentionnée plus haut.

Maintenant on a $(P_s \|\nabla f\|)^2 \leqslant (P_s f) \left(P_s (\|\nabla f\|^2 / f) \right)$ par Cauchy-Schwarz et donc

$$\mathrm{Ent}_\pi f \leqslant \int_{s=0}^{\infty} e^{-2s} \int_{\mathbb{R}^n} P_s \left(\frac{\|\nabla f\|^2}{f} \right) d\pi = \frac{1}{2} \int_{\mathbb{R}^n} \frac{\|\nabla f\|^2}{f}$$

car $\int g \, d\pi = \int P_s g \, d\pi$.

Finalement, en revenant à la définition de l'inégalité de Sobolev logarithmique i.e. en remplaçant f par f^2, on obtient

$$\mathrm{Ent}_\pi f^2 \leqslant 2 \int \|\nabla f\|^2 \, d\pi$$

qui était l'inégalité voulue. On a donc montré la concentration de diamètre gaussien 1.

3 Inégalités de courbure-dimension

Le principe de ce qui suit est d'abstraire les propriétés du processus d'Ornstein-Uhlenbeck que nous avons exploitées, pour dégager des conditions sous lesquelles on peut prouver une inégalité de Sobolev logarithmique, impliquant la concentration.

3.1 Opérateurs du champ

On repart dans la situation plus générale d'un processus de diffusion sur une variété riemannienne donné par le générateur infinitésimal

$$L = \Delta + X.\nabla$$

On définit l'opérateur « carré du champ » par

$$\Gamma(f,g) = \frac{1}{2}\left(L(fg) - g\,Lf - f\,Lg\right) = \nabla f.\nabla g$$

(noter la disparition de X). Son intérêt est que $\Gamma(f,f)$ (souvent abusivement noté $\Gamma(f)$) redonne $\|\nabla f\|^2$ et est défini uniquement à partir de L, ce qui est fort utile dans des contextes plus généraux.

Cette même formule permet de définir l'opérateur « carré du champ itéré » qui ressemblerait à une courbure de L, par

$$\Gamma_2(f,g) = \frac{1}{2}\left(L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(Lf,g)\right)$$

(on notera encore $\Gamma_2(f)$ pour $\Gamma_2(f,f)$).

Si L est simplement le laplacien Δ, rappelons (voir [GHL87]) la formule de Bochner

$$\frac{1}{2}\Delta\left(\|\nabla f\|^2\right) = \|\mathrm{Hess}\, f\|_2^2 + \nabla f.\nabla\Delta f + \mathrm{Ric}(\nabla f, \nabla f)$$

où $\|\mathrm{Hess}\, f\|_2^2$ est la somme des carrés des coefficients de la hessienne dans une base orthonormée (ou encore la norme de $\nabla\nabla f$ dans l'espace des 2-tenseurs munis de la métrique provenant de celle de la variété), et où Ric est le tenseur de Ricci (voir l'exposé de P. Pansu). On obtient alors simplement

$$\Gamma_2(f) = \|\mathrm{Hess}\, f\|_2^2 + \mathrm{Ric}(\nabla f, \nabla f)$$

Si $L = \Delta + X.\nabla$, on a

$$\Gamma_2(f) = \|\mathrm{Hess}\, f\|_2^2 + R_X(\nabla f, \nabla f)$$

où le 2-tenseur R_X est donné par $R_X = \mathrm{Ric} - \mathrm{Sym}\,\nabla X$ ou encore, en coordonnées, $R_X^{ij} = \mathrm{Ric}^{ij} - \frac{1}{2}\left(\nabla^i X^j + \nabla^j X^i\right)$. Ce tenseur peut être considéré comme le tenseur de Ricci associé au processus de diffusion L.

3.2 Présentation des inégalités

Revenons au cas le plus simple où $L = \Delta$ et $\Gamma_2(f) = \|\mathrm{Hess}\, f\|_2^2 + \mathrm{Ric}(\nabla f, \nabla f)$. Remarquons que sur une variété de dimension n on a $\|\mathrm{Hess}\, f\|_2^2 \geqslant \frac{1}{n}(\Delta f)^2$ car le laplacien est la trace de la hessienne (et par Cauchy-Schwarz). Par conséquent, sur une variété de dimension n dont la courbure de Ricci est supérieure à ρ, on aura

$$\Gamma_2(f) \geqslant \frac{1}{n}(\Delta f)^2 + \rho\|\nabla f\|^2$$

Ceci motive la définition générale suivante.

Définition 2 (Inégalité de courbure-dimension) – *On dit que le processus de diffusion L satisfait l'inégalité de courbure-dimension $CD(\rho, n)$ si pour toute fonction f on a*

$$\Gamma_2(f) \geqslant \rho\Gamma(f) + \frac{1}{n}(Lf)^2$$

Plusieurs remarques : d'abord, n n'a a priori rien à voir avec la dimension de la variété (toutefois n est nécessairement au moins cette dimension). Naturellement, ρ est homogène à l'inverse du carré d'une longueur. Ensuite, cette inégalité est *locale* : les dimension et courbure ainsi définies dépendent du point observé. Enfin, il pourrait y avoir a priori un arbitrage à faire entre un meilleur n ou un meilleur ρ (par la suite on se souciera peu de n et on prendra souvent $n = \infty$).

D'après la formule de Bochner, le laplacien sur une variété de dimension n et de courbure de Ricci supérieure à ρ vérifie donc l'inégalité de courbure-dimension $CD(\rho, n)$.

Pour le processus d'Ornstein-Uhlenbeck $L = \Delta - x.\nabla$ sur $\mathbb{R}^n$, appliquant la formule ci-dessus avec R_X, comme Ric est nul sur $\mathbb{R}^n$ et que $\nabla x = \mathrm{Id}$, on obtient que $\Gamma_2(f) = \|\mathrm{Hess}\, f\|_2^2 + \|\nabla f\|^2 \geqslant \|\nabla f\|^2$. Le processus d'Ornstein-Uhlenbeck vérifie donc l'inégalité de courbure-dimension $(1, \infty)$ de courbure 1 (on ne peut pas obtenir de dimension finie : $\|\mathrm{Hess}\, f\|$ contrôle Δf qui n'est pas Lf...). On va voir que cela suffit à impliquer l'inégalité de Sobolev logarithmique que nous avions utilisée plus haut, et donc la concentration de la mesure. Notons que dans ce cas, l'effet de courbure provient vraiment de l'introduction du terme de rappel $-x.\nabla$.

3.3 Inégalités de courbure-dimension et concentration

On en arrive enfin au théorème motivant cet exposé.

Théorème 3 – *Soit $\rho \in \mathbb{R}$ et soit L un processus de diffusion. Alors sont équivalents :*

(i) L satisfait la condition $CD(\rho, \infty)$;

(ii) Pour toute fonction f, pour tout temps t on a

$$\sqrt{\Gamma(P_t f)} \leqslant e^{-\rho t} P_t\left(\sqrt{\Gamma(f)}\right)$$

(iii) Pour toute fonction $f > 0$ on a

$$P_t\left(f^2 \log f^2\right) - P_t(f^2) \log P_t(f^2) \leqslant \frac{2\left(1 - e^{-2\rho t}\right)}{\rho} P_t\Gamma(f)$$

Dans ces expressions, il convient de ne pas perdre de vue que $\Gamma(f)$ est simplement $\|\nabla f\|^2$ dans les cas qui nous intéressent. La deuxième expression est bien sûr simplement la généralisation de l'inégalité $\|\nabla P_t f\| \leqslant e^{-t} P_t \|\nabla f\|$ vérifiée par le processus d'Ornstein-Uhlenbeck.

La troisième inégalité, dont le terme de gauche est une entropie non intégrée de $P_t(f^2)$, et celui de droite la norme du gradient, peut être qualifiée d'inégalité de

Sobolev logarithmique locale (mais une certaine dose d'intégration se cache dans P_t...).

Notons aussi que ces inégalités sont valables pour ρ négatif.

Concernant la démonstration de cette équivalence, on renvoie à la littérature pour des informations précises. L'implication de (i) vers (ii) s'obtient en étudiant les variations de $\psi(s) = e^{-\rho s} P_s \sqrt{\Gamma(P_{t-s} f)}$ en interpolant entre 0 et t. L'implication de (ii) vers (iii) se fait, par analogie avec le cas d'Ornstein-Uhlenbeck, en étudiant les variations de $\psi(s) = P_s \left(P_{t-s}(f^2) \log P_{t-s}(f^2) \right)$. L'implication de (iii) vers (i) passe par un développement limité à l'ordre 2 de l'inégalité de Sobolev logarithmique locale.

À partir de l'inégalité de Sobolev logarithmique locale, par un processus d'intégration en t et en l'espace tout à fait similaire à celui présenté pour le processus d'Ornstein-Uhlenbeck, on obtient (en courbure positive) une inégalité de Sobolev logarithmique ordinaire, et donc la concentration.

Proposition 4 – *Soit $\rho > 0$ et soit L un processus de diffusion vérifiant la condition $CD(\rho, \infty)$, de mesure réversible π. Alors on a pour toute fonction l'inégalité de Sobolev logarithmique*

$$\mathrm{Ent}_\pi f^2 \leqslant \frac{2}{\rho} \int \Gamma(f)\, d\pi$$

et en particulier la variété munie de la mesure π est concentrée de diamètre observable gaussien $1/\sqrt{\rho}$.

En prenant plus précisément en compte la dimension, on peut montrer que l'inégalité $CD(\rho, n)$ entraîne une inégalité de Sobolev de constante $2(n-1)/n\rho$ et donc la concentration de diamètre observable gaussien $\sqrt{(n-1)/n\rho}$ ce qui est légèrement meilleur.

Vu que les variétés à courbure de Ricci positive satisfont l'inégalité de courbure-dimension, ceci généralise le théorème de Lévy-Gromov. La généralisation provient de l'adjonction du terme en gradient, qui produit un phénomène de courbure supplémentaire.

De plus, on peut étendre ceci à des variétés où la courbure de Ricci n'est « pas trop négative », ou plutôt des variétés où la négativité de la courbure à certains endroits n'influe pas trop la croissance des volumes ([Wan97]). Plus précisément, considérons le processus de diffusion donné par $L = \Delta - \nabla U.\nabla$, de mesure invariante $d\pi = e^{-U} d\mathrm{vol}$, et supposons qu'en tout point il vérifie l'inégalité de courbure-dimension $CD(\rho, \infty)$ (où ρ est indépendant du point). Alors, s'il existe un $\varepsilon > 0$ tel que pour un certain x (ou, ce qui est équivalent, pour tout x) on ait

$$\int_y e^{(\max(-\rho,0)+\varepsilon)\,\mathrm{dist}(x,y)^2}\, d\pi(y) < \infty$$

alors il existe une constante C telle que l'espace muni de la mesure invariante π satisfait une inégalité de Sobolev logarithmique de constante C (et donc est gaussiennement concentré).

Références

[ABCFGMRS00] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panoramas et synthèses **10**, SMF (2000).

[Bak94] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, École d'été de probabilités de Saint-Flour, Lecture Notes in Math. **1581**, 1–114, Springer (1994).

[GHL87] S. Gallot, D. Hulin, J. Lafontaine, *Riemannian geometry*, Springer (1987).

[Led00] M. Ledoux, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse **IX**, 305–366 (2000).

[Wan97] F.-Y. Wang, *Logarithmic Sobolev inequalities on noncompact Riemannian manifolds*, Prob. Theory Relat. Fields **109** (1997), 417–424.