Modèle de diffusion à gradient moyenné, comportement asymptotique.

S. Gadat, F. Panloup

Institut de Mathématiques de Toulouse
Rencontres Laboratoire de Physique Théorique, Équipe de Statistiques et Probabilités

Juin 2010

Plan de l'exposé

Introduction

- Dans cet exposé, on étudiera le comporement asymptotique de

$$dx_t = -\left[\frac{1}{k(t)}\int_0^t h(s)\nabla U(x_s)ds\right]dt + \sigma(x_t)dB_t \tag{1}$$

- U est un potentiel coercif strictement positif, $(x_t)_{t\geq 0}$ vit dans $\mathbb{R}^d$.
- B_t est un mouvement brownien d-dimensionnel et σ est une matrice de variance/covariance elliptique.
- Objectif : Comparer le comportement en temps long des solutions de (1) avec des diffusions classiques :

$$d\tilde{x}_t = -\nabla U(\tilde{x}_t)dt + \sigma(\tilde{x}_t)d\tilde{B}_t.$$

Un petit problème de dynamique déterministe

- On considère le mouvement d'une boule pesante posée sur le graphe d'une fonction U.
- La boule est soumise à la pesanteur et à une forte de frottement opposée à son déplacement, le système dissipatif est solution de l'o.d.e.

$$\ddot{x}_t + \lambda \dot{x}_t + \nabla U(x_t) = 0 \tag{2}$$

  et du moment que U est convexe, la solution de (2) converge vers l'unique minimum de U.
- Plus généralement, on peut considérer le système dissipatif

$$\ddot{x}_t + a(t)\dot{x}_t + \nabla U(x_t) = 0 \tag{3}$$

  et étudier le comportement de solutions de (3)
- En réalité, après un changement de temps, on peut vérifier que si y solution de

$$\dot{y}_t = -\frac{1}{k(t)} \int_0^t h(s) \nabla U(y_s) ds,$$

  alors $y \circ \tau$ est solution de (2).

Un petit problème de dynamique déterministe (suite)

- Cas particulier du système déterministe :

$$U(x) = x^2/2, a(t) = 1/t, x_t \quad \text{fonction de Bessel}$$

- Oscillation très lentes lorsque $t \to \infty$ avec phénomène de damping autour du minimum de U.
- De manière plus générale, les systèmes déterministes à mémoire ont une inertie qui implique des oscillations localement.
- D'un point de vue optimisation : cette inertie permet de passer des maximums locaux.

Plan de l'exposé

Notations, hypothèses sur le potentiel U et σ

- U est une fonction de $\mathbb{R}^d$ dans $\mathbb{R}_+^*$ coercive, c'est à dire telle que

$$\lim_{x\to\infty} U(x) = +\infty$$

- On suppose que U satisfait la condition de rappel :

$$\lim \inf_{x\to\infty} \langle x, \nabla U(x)\rangle > 0$$

- U est supposé au moins $\mathcal{C}^3$, et on notera D^2U la matrice

$$(D^2U(x))_{i,j} = (\partial_{x_i}\partial_{x_j}U)(x)$$

- $\sigma(x)$ est une matrice carrée $d \times d$ et $x \to \sigma(x)$ est $\mathcal{C}^\infty$. On note σ^* l'application transposée.
- $\sigma^* D^2 U \sigma \leq CU$ pour une constante C assez grande.

Notations, hypothèses sur le processus stochastique

- On se donne une fonction k croissante, positive, de classe $\mathcal{C}^2$. Typiquement :

$$k(t) = t^{\alpha}, \alpha > 0 \qquad k(t) = e^{t^{\alpha}}, \alpha \geq 1 \qquad k(t) = e^{e^t}$$

- On définit $(B_t)_{t \geq 0}$ un mouvement Brownien standard, et $(x_t)_{t \geq 0}$ est le processus solution de l'eds :

$$dx_t = -\left[\frac{1}{k(t)} \int_0^t \dot{k}(s) \nabla U(x_s) ds\right] dt + \sigma(x_t) dB_t$$

- <u>Problème :</u> $(x_t)_{t \geq 0}$ n'a aucune chance d'être Markovien.
- En posant $y_t = \frac{1}{k(t)} \int_0^t \nabla U(x_s) ds$, on remarque que (x, y) satisfait

$$dx_t = -y_t dt + \sigma(x_t) dB_t \qquad \text{et} \qquad dy_t = r(t)[\nabla U(x_t) - y_t] dt,$$

avec $r(t) = \dot{k}(t)/k(t)$. Le processus $(x_t, y_t)_{t \geq 0}$ est Markovien.

Processus Markovien

- Une fonction importante :

$$r(t) := \frac{\dot{k}(t)}{k(t)}$$

- On définit $(X_t, Y_t)_{t\geq 0}$ solution du processus de descente de gradient moyenné par

$$\begin{cases} dX_t = -Y_t dt + \sigma(X_t) dB_t \\ dY_t = r(t)(\nabla U(X_t) - Y_t) dt. \end{cases} \quad (4)$$

- (X_t, Y_t) est un processus Markovien inhomogène en temps tel que $\forall f \in \mathcal{C}_K^2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_+)$:

$$\mathcal{A}f(x,y,t) = -\langle y, \partial_x f\rangle + r(t)\langle \nabla U(x) - y, \partial_y f\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x) D_x^2 f \sigma(x)\right) + \partial_t f.$$

Existence du processus

Proposition (Existence forte du processus)

Étant donné $T > 0$ et un M.B. $(B_t)_{t\geq 0}$, il existe une unique solution de (6). De plus, si $\mathbb{E}[U(X_0) + Y_0^2] < \infty$, alors $\sup_{t\leq T} \mathbb{E}[U(X_t) + Y_t^2] < \infty$.

Démonstration.

Poser $h(x,y,t) = U(x) + y^2/(2r(t))$ et montrer que

$$\mathcal{A}h(x,y,t) = Tr(\sigma^* D^2 U(x)\sigma) + |y|^2(-1 + \frac{\dot{r}(t)}{2r^2(t)}) \leq C_T h$$

. □

La proposition fournit un contrôle en temps fini de la croissance du processus $(X_t, Y_t)_{t\geq 0}$ (rappel en $|y|^2$) mais est insuffisant pour contrôler finement le comportement en x.

Plan de l'exposé

Rôle de la mémoire (r et k) :

La nature du processus sera radicalement différente selon les variations de r et sa limite $r(t) = \lim_{t\to\infty} r(t)$. Nous étudierons 3 cas :

1. $r_\infty \in]0;+\infty[$: stabilité du processus et comportement non standard.
   Exemple : $k(t) \sim_{t\to\infty} e^{\lambda t}, \lambda > 0$.
2. $r_\infty = +\infty$: stabilité du processus et comportement standard.
   Exemple : $k(t) \sim_{t\to\infty} e^{t^\alpha}, \alpha > 1$.
3. $r_\infty = 0$: instabilité du processus.
   Exemple : $k(t) \sim_{t\to\infty} t^\beta, \beta > 0$.

Le cas standard sera identifié par l'EDS :

$$dS_t = -\nabla U(S_t)dt + \sigma(S_t)d\tilde{B}_t. \tag{5}$$

On notera π une distribution stationnaire des solutions de (5).
Enfin, on fera l'hypothèse sur la mémoire :

$$\lim_{t\to+\infty} \frac{\dot{r}(t)}{r^2(t)} = 0, \text{vrai dès que } k(t) \geq t^\beta, \beta > 0.$$

Hypothèses de rappel U

On distinguera deux types d'hypothèse sur U

- Condition faible

$$(\mathbf{H}_1) \qquad \|\sigma\|^2(D^2U + Id) = O_{x\sim\infty}(\langle x, \nabla U(x)\rangle).$$

Si σ bornée,$(\mathbf{H}_1)$ est satisfaite du moment que $U \sim |x|^p, p > 0$ (croissance en puissance positive de x à l'infini) ou $U(x) \sim \ln(|x| + 1)^\beta$ avec $\beta > 1$.

- Condition forte

$$\exists\, a \in]0;1] \qquad (\mathbf{H}'_a) \begin{cases} (1 + \|\sigma\|^2)(1 + \frac{|\nabla U|^2}{U} + D^2U + D^3U) = O(U(x) \vee |x|^2)^a \\ (\mathrm{U(x)} \vee |x|^2)^a = O(\langle x, \nabla U(x)\rangle) \end{cases}$$

$(\mathbf{H'}_a)$ est vraie pour $U \sim |x|^p, p > 0$ mais plus pour $U(x) \sim \ln(|x| + 1)^\beta$.

- On énoncera des résultats avec $(\mathbf{H}'_a)$ principalement.

Cas stable

On fera l'hypothèse dans cette sous-partie que r possède une limite lorsque $t \to +\infty$ et

$$\lim_{t \to +\infty} r(t) = r_\infty > 0.$$

Fonction de Lyapunov du système ($r_\infty > 0$)

Proposition

On peut trouver ρ solution de l'ode $\rho^2(t) = r(t)\rho(t) + \dot\rho(t)$ telle que $\lim_{t\to\infty} \rho(t)/r(t) = 1$.

Proposition

Il existe $m > 0$ tel que V_m définie par

$$V_m(x,y,t) = U(x) + \frac{|y|^2}{2r(t)} + m\left(|x|^2/2 - \langle x,y\rangle/\rho(t)\right)$$

satisfait $\limsup_{(x,y)\to\infty} \mathcal{A}V_m(x,y,t) = -\infty$

La partie classique $U(x) + \frac{|y|^2}{2r(t)}$ génère un rappel en y tandis que c'est $\langle x, y\rangle$ qui génère le rappel en x.

Compacité des trajectoires ($r_\infty > 0$)

Definition (Mesures d'occupation et d'occupation en moyenne)

Étant donné $z_0 = (x_0, y_0)$, on définit la mesure d'occupation aléatoire :

$$\nu_t^{z_0}(f) = \frac{1}{t}\int_0^t f(X_s^{z_0}, Y_s^{z_0})ds$$

et la mesure d'occupation moyennée :

$$\mu_t^{z_0}(f) = \frac{1}{t}\int_0^t \mathbb{E}[f(X_s^{z_0}, Y_s^{z_0})]ds.$$

On aura besoin d'un peu plus d'hypothèses pour contrôler $\nu_t^{z_0}(f)$ que $\mu_t^{z_0}(f)$.

Compacité des trajectoires (suite) ($r_\infty > 0$)

Théorème

Sous l'hypothèse $(\boldsymbol{H}_1)$, pour tout z_0, la famille de mesures $(\mu_t^z)_{t\geq 0}$ est tendue.
Sous l'hypothèse $(\boldsymbol{H}'_a)$, pour tout z_0, la famille de mesures $(\nu_t^z)_{t\geq 0}$ est tendue.
De plus, si μ_∞ (resp. ν_∞ désigne un point d'accumulation de $(\mu_t^z)_{t\geq 0}$ (resp. de $(\nu_t^z)_{t\geq 0}$), alors :

- *Si $r_\infty = +\infty$, alors μ_∞ (resp. ν_∞) a pour première marginale sur x la mesure de Gibbs associé à U (même mesure stationnaire que pour le processus de diffusion standard $(S_t)t \geq 0$.*
- *Si $0 < r_\infty < +\infty$, μ_∞ (resp. ν_∞) est une distribution invariante pour le processus de Markov homogène :*

$$\begin{cases} dX_t = -Y_t dt + \sigma(X_t) dB_t \\ dY_t = r_\infty(\nabla U(X_t) - Y_t) dt. \end{cases} \tag{6}$$

Cas instable

On fera l'hypothèse dans ce slide que r possède une limite nulle lorsque $t \to +\infty$ et

$$\lim_{t \to +\infty} r(t) = 0.$$

Dans ce cas, on montre que

Théorème

Si U est sous-quadratique : $|\nabla U|^2 = O(U)$, et que $Tr(\sigma^* D^2 U \sigma) \geq \lambda_0 > 0$, alors

$$\limsup_{t \to \infty} r(t)\mathbb{E}[X_t^2] = +\infty.$$

Exemple : Le processus explose lorsque $k(t) = t^\alpha, \forall \alpha \geq 0$ et $U(x) = x^2/2$.

Plan de l'exposé

Hypoellipticité

- Pour la suite de l'exposé, il est nécessaire de l'irréductibilité sur $(X_t, Y_t)_{t\geq 0}$. On supposera U de classe $\mathcal{C}^\infty(\mathbb{R}^d)$.
- C'est non trivial en raison de la dégénérescence de la diffusion sur Y :

$$dY_t = r(t)[\nabla U(X_t) - Y_t]dt.$$

Definition

On définit l'ensemble des points $\mathcal{E}_U$:

$$\mathcal{E}_U = \left\{x \in \mathbb{R}^d, \forall 1 \leq k \leq d : \exists\,(i_1, \ldots i_p)\, \partial_{x_{i_p} \ldots x_{i_1}, x_k} U(x) \neq 0\right\}$$

Théorème (Hypoellipicité)

Si $\sigma^*\sigma \geq \epsilon_0 Id$ et $\dim(\mathbb{R}^d - \mathcal{E}_U) \leq d-1$, alors (X_t^z, Y_t^z) admet une densité $\mathcal{C}^\infty$ par rapport à la mesure de Lebesgue, notée $p_t(z, .)$. Lorsque $r(t) = r > 0$, il y a en plus unicité de la distribution invariante du processus (X_t, Y_t) (qui est Markov homogène).

Identification de la limite dans le cas homogène $0 < r_\infty < +\infty$

Proposition (Identification de la loi stationnaire, cas homogène)

Dans le cas hypoelliptique avec $r(t) = r_\infty \in]0;+\infty[$, il existe une unique probabilité invariante de densité p_∞ strictement positive et $\mathcal{C}^\infty$. De plus, p_∞ satisfait l'équation aux dérivées partielles :

$$\langle y, \partial_x p_\infty \rangle + \frac{1}{2} Tr(\sigma^* D_x^2 p_\infty \sigma) + r_\infty \left(\langle y - \nabla U(x), \partial_y p_\infty \rangle + p_\infty \right) = 0.$$

Dans le cas gaussien ($U(x) = x^2/2$) et σ constant, p_∞ est une mesure gaussienne centrée de matrice de covariance

$$\Sigma^2(r_\infty) = \frac{\sigma^2}{2} \begin{pmatrix} \frac{r_\infty + 1}{r_\infty} & 1 \\ 1 & 1 \end{pmatrix}.$$

Cette proposition décrit la loi limite dans le cas homogène, mais également dans le cas inhomogène lorsque $r(t) \to r_\infty \in]0;+\infty[$.
Dans le cas $\sigma = \sigma Id$ constant, p_∞ satisfait

$$\langle y, \partial_x p_\infty \rangle + \frac{\sigma^2}{2} \Delta p_\infty + r_\infty \left(\langle y - \nabla U(x), \partial_y p_\infty \rangle + p_\infty \right) = 0.$$

Vitesse de convergence, $0 < r_\infty < +\infty$

Théorème

Dans le cas homogène $r(t) = r_\infty$, en supposant la condition de rappel $(\boldsymbol{H}'_a)$, on a pour tout $p \geq 1$:

$$\sup_{\{f, |f| \leq v^p\}} |P_t^{r_\infty}(z_0, f) - p_\infty(f)| \leq C_{a,p,r_\infty} \{v(z_0)\}^p \begin{cases} \exp(-\delta_{p,r_\infty} t) & \text{si } a = 1 \\ t^{-\frac{p+a-1}{1-a}} & \text{si } a \in (0,1). \end{cases}$$

avec $v(z) = U(x) + \frac{|y|^2}{2} + r_\infty \left(\frac{|x|^2}{2} - \frac{\langle x,y \rangle}{r_\infty} \right)$.

- Vitesse de convergence exponentielle lorsque U est essentiellement surquadratique.
- Vitesse géométrique dans le cas où U a une croissance plus faible.
- Pas de résultat lorsque $r(t) \to r_\infty \in]0; +\infty[$.
- Pas de borne inférieure...

Vitesse de convergence, $r_\infty = +\infty$

- On note $\tilde{\sigma}(x,y) = \sigma(y) - \sigma(x)$.
- On note $(\mathbf{AC})$ l'hypothèse :

$$\exists \rho > 0, \forall (x_1, x_2) : \rho |x_2 - x_1|^2 + \frac{1}{2} \mathrm{Tr}\left((\tilde{\sigma}\tilde{\sigma}^*)(x_1, x_2)\right) \leq \langle x_1 - x_2, \nabla U(x_1) - \nabla U(x_2) \rangle$$

- On se place dans le cas hypoelliptique où l'unique distribution stationnaire de S_t est notée π.

Théorème

En supposant que U satisfait l'hypothèse $(\mathbf{AC})$ et est hypoelliptique, alors il existe C_1, C_2 telles que

$$W_2(\mathbb{P}_{X_t}, \pi) \leq C_1 e^{-\rho(t-u)} + \frac{C_2}{k(u)} \int_0^u k(v) dv, \quad \forall u \in [0, t].$$

En optimisant une telle inégalité, on obtient

- *Lorsque* $k(t) = \alpha t^{\alpha - 1} e^{t^\alpha}$, $W_2(\mathbb{P}_{X_t}, \pi) \leq C t^{1-\alpha}$.
- *Lorsque* $k(t) = e^t e^{e^t}$, $W_2(\mathbb{P}_{X_t}, \pi) \leq C e^{-\frac{\rho}{1+\rho} t}$.

Plan de l'exposé

Conclusion

Synthèse :

- Trois régimes distincts identifiés selon le comportement de $r(t) = \frac{\dot{k}}{k}(t)$.
  - $k(t) = t^{\alpha}$ Comportement explosif.
  - $k(t) = e^{e^t}$ Comportement comparable à la diffusion
  $$dS_t = -\nabla U(S_t)dt + \sigma(S_t)dB_t.$$
  - $k(t) = e^{\lambda t}$ Comportement stable mais non standard.
- Vitesses exponentielles ou polynomiales dans les cas stables.

Élargissements :

- Exploiter les vertus exploratoires d'un tel processus pour étudier un recuit simulé $\sigma_t \to 0$ avec une application dédier à l'optimisation.
- Généraliser à des mémoires différentes :
$$dX_t = -\frac{1}{Z_t}\left[\int_{a(t)}^{b(t)} \nabla U(X_s)ds\right] dt + \sigma(X_s)dB_s.$$
- Borne inférieures de vitesses inconnues.
- D'un point de vue de la modélisation, étudier des systèmes à mémoire ayant une écriture comparable.

Conclusion

Merci de votre attention !