EDS avec petit bruit

Grandes déviations

Soit $\{\mu_\varepsilon\}_{\varepsilon>0}$ une famille de mesures sur un espace polonais $(\Omega, \mathcal{F})$ (espace métrique complet et séparable muni de sa tribu Borelienne $\mathcal{F}$)

Définition 1. *La famille $\{\mu_\varepsilon\}_{\varepsilon>0}$ satisfait un principe de grandes déviations (PDG) ssi il existe $v(\varepsilon)$ et $I\colon \Omega \to \mathbb{R}_+ \vee \{+\infty\}$ tels que*

1. *$v(\varepsilon) \to 0 + \infty$ pour $\varepsilon \to 0$*

2. *Pour tout $\ell \geqslant 0$ l'ensemble $\mathcal{I}(\ell) = \{I \leqslant \ell\}$ est compact dans Ω.*

3. *Pour tout $F \in \mathcal{F}$ on a*

$$-\inf_{\omega\in F^\circ} I(\omega) \leqslant \liminf_{\varepsilon\to 0} \frac{1}{v(\varepsilon)} \log \mu_\varepsilon(F) \leqslant \limsup_{\varepsilon\to 0} \frac{1}{v(\varepsilon)} \log \mu_\varepsilon(F) \leqslant -\inf_{\omega\in \bar{F}} I(\omega)$$

où F° et $\bar{F}$ sont respectivement l'intérieur et l'aderence de F.

La fonction I est appelé le fonctionnel d'action et v la vitesse du PDG.

Si on a que $-\inf_{\omega\in F^\circ} I(\omega) = -\inf_{\omega\in\bar{F}} I(\omega) = I(F)$ alors $\mu_\varepsilon(F) = \exp(-v(\varepsilon)[I(F) + o(1)])$ pour $\varepsilon \to 0$.

Ces formulations évitent des problèmes liées au fait que d'un côté on a des objets qui viennent de la théorie de la mesure, dans l'autre, des objets topologiques. Considérons le cas très simple d'une famille Gaussienne: $\mu_\varepsilon \sim \mathcal{N}(0, 1/\varepsilon^2)$, donc on voudrait dire que

$$\varepsilon^2 \log \mu_\varepsilon(A) = \int_A e^{-x^2/2\varepsilon^2} \frac{\mathrm{d}x}{\sqrt{2\pi}\varepsilon} \simeq -\inf_{x\in A} \frac{x^2}{2} \qquad \text{pour } \varepsilon \to 0$$

par (une sorte de) méthode de Laplace. Mais cela a peu de sens car si on prend $A = \{a\}$ alors à gauche on a $-\infty$ et a droite on a $-a^2/2$... La formulation qui s'est avérée être solide (et pas toujours fausse) est celle que l'on a donné.

Exemple 2. [Cramer] Si $(X_k)_{k\geqslant 1}$ est une suite de v.a. iid et $\mathbb{E}e^{\alpha X_1} < +\infty$ pour $\alpha \in V$ voisinage de 0, alors les lois μ_n des moyennes empiriques $\bar{X}_n$ satisfont un PDG avec vitesse n et fonctionnel d'action

$$I(x) = \sup_\alpha [\alpha x - \log \mathbb{E} e^{\alpha X_1}]$$

(transformé de Cramer).

Lemme 3. *Des conditions suffisantes pour la validité du PDG sont:*

1. *$\mathcal{I}(\ell)$ est compact dans Ω pour tout $\ell \geqslant 0$.*

2. *Pour tout $\omega_0 \in \Omega$ tel que $I(\omega_0) < +\infty$ et tout $\delta > 0$*

$$\liminf_{\varepsilon\to 0} \frac{1}{v(\varepsilon)} \log \mu_\varepsilon(d(\omega, \omega_0) < \delta) \geqslant -I(\omega_0)$$

3. Pour tous $\delta > 0$ et $\ell > 0$:

$$\limsup_{\varepsilon \to 0} \frac{1}{v(\varepsilon)} \log \mu_{\varepsilon} (d(\omega, \mathcal{I}(\ell)) \geqslant \delta) \leqslant -\ell.$$

Démonstration. (Exercice) Montrons par exemple que (1),(3) impliquent la borne supérieure du PDG pour tout fermé F. On peut supposer que $I(F) > 0$. Soit $0 < \ell < I(F)$ alors $F \cap \mathcal{I}(\ell) = \emptyset$ et puisque $\mathcal{I}(\ell)$ est compact il existe $\delta > 0$ tel que $F \subseteq \{\omega: d(\omega, \mathcal{I}(\ell)) \geqslant \delta\}$ d'où

$$\limsup_{\varepsilon \to 0} \frac{1}{v(\varepsilon)} \log \mu_{\varepsilon}(F) \leqslant \limsup_{\varepsilon \to 0} \frac{1}{v(\varepsilon)} \log \mu_{\varepsilon} \{\omega: d(\omega, \mathcal{I}(\ell)) \geqslant \delta\} \leqslant -\ell. \qquad \square$$

Grandes déviations pour le mouvement Brownien

Soit $(B_t)_{t \geqslant 0}$ un mouvement Brownien à valeurs dans $\mathbb{R}^d$, issu de zéro. Pour tout $T > 0$ on considère l'espace $\Omega = C([0, T]; \mathbb{R}^d)$ muni de la norme uniforme $\| \cdot \|$ et l'espace de Cameron-Martin

$$H = \left\{ h \in \Omega: h(t) = \int_0^t \dot{h}(s) \mathrm{d}s \text{ pour quelque } \dot{h} \in L^2([0, T]; \mathbb{R}^d) \right\}$$

on défini le bon fonctionnel d'action

$$I(\omega) = \begin{cases} \frac{1}{2} \int_0^T |\dot{w}_s|^2 \mathrm{d}s & \text{si } \omega \in H \\ +\infty & \text{sinon} \end{cases}$$

Lemme 4. *Pour tout $\ell \geqslant 0$ l'ensemble $\mathcal{I}(\ell)$ est compact dans Ω.*

Démonstration. (Exercice) $\square$

Lemme 5. *Pour tous $\delta \geqslant 0$ et $c > 0$*

$$\mathbb{P}\left(\sup_{s \in [0, c]} |B_s| \geqslant \delta \right) \leqslant 2 d \exp\left(-\frac{\delta^2}{2 d^2 c} \right)$$

Démonstration. Une conséquence de la majoration exponentielle pour les martingales continues. $\square$

Remarque 6. [Majoration exponentielle] Soit $(M_t)_{t \geqslant 0}$ une martingale locale continue, nulle en 0. Pour tous $T > 0, k > 0$ et $c > 0$ on a

$$\mathbb{P}(\langle M \rangle_T \leqslant k, \sup_{0 \leqslant s \leqslant T} |M_s| \geqslant c) \leqslant 2 \exp(-c^2 / 2k).$$

Théorème 7. *[Schilder, 1966] Pour tout $\varepsilon > 0$ soit μ_{ε} la loi de $(\varepsilon B_t)_{t \in [0, T]}$ sur l'espace canonique Ω muni de sa tribu Borelienne. La famille $(\mu_{\varepsilon})_{\varepsilon > 0}$ vérifie un PDG avec vitesse $1/\varepsilon^2$ et fonctionnel d'action I.*

Démonstration. D'abord la borne supérieure. Pour tout $n \geqslant 0$ soit B^n l'approximation polygonale de B sur $[0, T]$:

$$B_t^n = B_{kT/n} + \frac{n}{T}\left(t - \frac{kT}{n} \right) \left(B_{(k+1)T/n} - B_{kT/n} \right) \qquad \text{si } t \in [kT/n, (k+1)T/n[$$

pour tout $k=0,1,\ldots,n-1$. On a donc $B^n \in H$ et

$$I(B^n) = \frac{n}{2T}\sum_{k=0}^{n-1} |B_{(k+1)T/n} - B_{kT/n}|^2$$

la v.a. $2I(B^n)$ suit une loi du χ^2 à nd degrés de liberté, de plus

$$\sup_{t\in[0,T]} |B_t - B_t^n| \leqslant 2 \sup_{k=0,\ldots,n-1} \sup_{t\in[kT/n,(k+1)T/n[} |B_t - B_{kT/n}|.$$

Soient $\ell > 0$ et $\delta > 0$. On a

$$\mathbb{P}(d(\varepsilon B, \mathcal{I}(\ell)) \geqslant \delta) \leqslant \mathbb{P}(d(\varepsilon B, \varepsilon B^n) \geqslant \delta) + \mathbb{P}(\varepsilon B^n \notin \mathcal{I}(\ell))$$

et

$$\mathbb{P}(d(\varepsilon B, \varepsilon B^n) \geqslant \delta) \leqslant \sum_{k=0}^{n-1} \mathbb{P}\left(\sup_{t\in[kT/n,(k+1)T/n[} |B_t - B_{kT/n}| \geqslant \frac{\delta}{2\varepsilon}\right) \leqslant 2nd \exp\left(- \frac{n\delta^2}{8d^2 T\varepsilon^2}\right)$$

d'après le lemme 5. De plus

$$\mathbb{P}(\varepsilon B^n \notin \mathcal{I}(\ell)) = \mathbb{P}(\varepsilon^2 I(B^n) > \ell) = \mathbb{P}(\xi_1^2 + \cdots + \xi_{nd}^2 > 2\ell/\varepsilon^2) \leqslant \exp - \frac{2c\ell}{\varepsilon^2} + nd \log \mathbb{E} e^{c\xi_1^2}$$

pour tout $0 \leqslant c < 1/2$, où $\xi_1, \ldots, \xi_{nd}$ sont des v.a. iid $\mathcal{N}(0,1)$. Donc

$$\mathbb{P}(d(\varepsilon B, \mathcal{I}(\ell)) \geqslant \delta) \leqslant 2nd \exp\left(- \frac{n\delta^2}{8d^2 T\varepsilon^2}\right) + \exp\left(- \frac{2c\ell}{\varepsilon^2} + nd \log \mathbb{E} e^{c\xi_1^2}\right)$$

et choisissant convenablement $n \propto c\ell$ on obtient $\limsup_{\varepsilon\to 0} \varepsilon^2 \mathbb{P}(d(\varepsilon B, \mathcal{I}(\ell)) \geqslant \delta) \leqslant -2c\ell$ pour tout $c < 1/2$ et donc aussi pour $c = 1/2$.

Montrons maintenant la borne inf. Soient $\omega_0 \in H$ et $\delta > 0$. D'après la formule de Cameron-Martin

$$\mathbb{P}(\|\varepsilon B - \omega_0\| < \delta) = \mathbb{E}\left[1_{\{\|\varepsilon B\| < \delta\}} e^{-\varepsilon^{-1}\int_0^T \dot{\omega}_s \mathrm{d}B_s - (1/2\varepsilon^2)\int_0^T |\dot{\omega}_s|^2 \mathrm{d}s}\right]$$

$$\geqslant_{(\text{Jensen})} \mathbb{P}(\|\varepsilon B\| < \delta) \exp\left(- \frac{I(\omega_0)}{\varepsilon^2} - \frac{1}{\varepsilon}\mathbb{E}\left[\int_0^T \dot{\omega}_s \,\mathrm{d}B_s; \|\varepsilon B\| < \delta\right]\right)$$

$$= \mathbb{P}(\|\varepsilon B\| < \delta) \exp\left(- \frac{I(\omega_0)}{\varepsilon^2}\right)$$

donc $\liminf_{\varepsilon\to 0} \varepsilon^{-2}\, \mathbb{P}(\|\varepsilon B - \omega_0\| < \delta) \geqslant -I(\omega_0)$ car $\mathbb{P}(\|\varepsilon B\| < \delta) \to 1$ pour $\varepsilon \to 0$. □

Grandes déviations pour les EDS Browniennes

Soit X^ε la solution de

$$\mathrm{d}X_t^\varepsilon = b(X_t^\varepsilon)\mathrm{d}t + \varepsilon \mathrm{d}B_t, \qquad X_0^\varepsilon = x$$

on note μ_ε la loi de $(X_t^\varepsilon)_{t\in[0,T]}$ sur l'espace canonique Ω^x. On définit l'application $F\colon \Omega \to \Omega$ tel que $X_t^\varepsilon = F(\varepsilon W)$.

Théorème 8. *La famille* $(\mu_\varepsilon)_{\varepsilon>0}$ *satisfait un PDG avec vitesse* $1/\varepsilon^2$ *et fonctionnel d'action*

$$J(\varphi) = \inf \{I(\omega) : \omega \in H, \varphi = F(\omega)\}$$

ou l'infimum est $+\infty$ *si l'ensemble est vide.*

Démonstration. On prouve d'abord la borne sup. Comme dans le théorème de Schilder on considère des approximations B^n de B. Soit $X^{\varepsilon,n} = F(\varepsilon B^n)$ alors

$$\mathbb{P}(d(X^\varepsilon, \mathcal{I}(\ell)) > \delta) \leqslant \mathbb{P}(d(X^\varepsilon, X^{\varepsilon,n}) > \delta, I(\varepsilon B^n) \leqslant \ell) + \mathbb{P}(I(\varepsilon B^n) > \ell)$$

Il est facile de voir qu'il existe une constante C_T telle que $\|F(x) - F(y)\| \leqslant C_T \|x - y\|$, donc

$$\mathbb{P}(d(X^\varepsilon, X^{\varepsilon,n}) > \delta, I(\varepsilon B^n) \leqslant \ell) \leqslant \mathbb{P}(d(B, B^n) > \delta / C_T \varepsilon) \leqslant n\, d \exp\left(- \frac{n \delta^2}{8 d^2 C_T^2 T \varepsilon^2} \right)$$

et on peut conclure la preuve par le même argument utilisé plus en haut. Pour la borne inf, soit $\varphi = F(\theta)$, on a $\mathbb{P}(\|X^\varepsilon - \varphi\| < \delta) \geqslant \mathbb{P}(\|\varepsilon B - \theta\| < \delta / C_T)$ et donc

$$\liminf_{\varepsilon \to 0} \frac{1}{\varepsilon^2} \log \mathbb{P}(\|X^\varepsilon - \varphi\| < \delta) \geqslant - I(\theta).$$

□

Pour des EDS plus générales (avec un coefficient de diffusion devant $\mathrm{d}W$) on a aussi un PDG mais la preuve devient plus élaborée. Si X^ε résoudre

$$\mathrm{d}X_t^\varepsilon = b(X_t^\varepsilon)\mathrm{d}t + \varepsilon \sum_{i=1}^m \sigma_i(X_t^\varepsilon)\mathrm{d}W_t^i, \qquad X_0^\varepsilon = x$$

où b, σ_1, ..., σ_m sont des champs de vecteurs Lipshitziens sur $\mathbb{R}^d$. Si on considère l'application $\Phi = H \to H_x$ définie par

$$\Phi(h)_t = x + \int_0^t [b(\Phi(h)_s) + \sigma(\Phi(h)_s)\dot{h}_s]\mathrm{d}s$$

alors la famille $(\mu_\varepsilon)_{\varepsilon>0}$ satisfait un PDG avec vitesse $1/\varepsilon^2$ et fonctionnel d'action

$$J(\varphi) = \inf \{I(h) : h \in H, \varphi = \Phi(h)\}.$$

Si la matrice $a(x) = \sigma\sigma^T(x)$ est uniformément positive alors on peut récrire le fonctionnel d'action dans une forme plus explicite:

$$J(\varphi) = \frac{1}{2} \int_0^T \langle \dot{\varphi}_s - b(\varphi_s), a(\varphi_s)^{-1}(\dot{\varphi}_s - b(\varphi_s)) \rangle \mathrm{d}s.$$

où on considère que $J(\varphi) = +\infty$ si $\dot{\varphi} \notin L^2([0,T])$.