Lecture 21 – 2020.07.2 – 12:15 via Zoom

Exam: first oral exam from 27/7-1/8. second oral exam mid september 14/9-25/9. Will prepare a doodle form so that students can sign in.

Boué–Dupuis formula (continued)

We assume that $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ is the canonical d-dimensional Wiener space, i.e. $\Omega = \mathscr{C}^d = C(\mathbb{R}_+, \mathbb{R}^d)$, $X_t(\omega) = \omega(t)$, $\mathbb{P}$ is the law of the Brownian motion and $(\mathscr{F}_t)_{t \geqslant 0}$ is the filtration generated by the canonical process $(X_t)_{t \geqslant 0}$ in particular we have $\mathscr{F}_\infty = \mathscr{F} = \mathscr{B}(\Omega)$. We will also use the notation μ for the Wiener measure $\mathbb{P}$.

Theorem 1. *(Boué–Dupuis formula) For any function $f: \Omega \to \mathbb{R}$ measurable and bounded from below. We have*

$$\log \mathbb{E}_\mu[e^f] = \sup_{u \in \mathbb{H}} \mathbb{E}_\mu \left[f(X + I(u(X))) - \frac{1}{2} \|u(X)\|_{\mathbb{H}}^2 \right]$$

where the supremum on the r.h.s. is taken wrt. all the predictable functions $u: \mathbb{R}_+ \times \Omega \to \mathbb{R}$ such that

$$\|u\|_{\mathbb{H}}^2 = \int_0^\infty |u_s|^2 \mathrm{d}s < \infty, \qquad \mu - a.s. \tag{1}$$

and we write $u(\omega) = u(X(\omega))$ to stress the measurability wrt. the filtratrion $\mathscr{F}$ generated by X and where

$$I(u)(t) = \int_0^t u_s(X) \mathrm{d}s, \qquad t \geqslant 0.$$

We call a function u as above, a drift (wrt. μ).

Large deviations of diffusion

The goal will be now to understand what happens when we have a family of SDEs in $\mathbb{R}^d$ of the form

$$\mathrm{d}X_t^\varepsilon = b(X_t^\varepsilon)\mathrm{d}t + \varepsilon^{1/2}\sigma(X_t^\varepsilon)\mathrm{d}B_t, \qquad X^\varepsilon = x_0 \in \mathbb{R}^d$$

with ε a small parameter and B a d-dimensional BM. Let's assume the coefficient $b: \mathbb{R}^d \to \mathbb{R}^d$, $\sigma: \mathbb{R}^d \to \mathscr{L}(\mathbb{R}^d, \mathbb{R}^d)$ are nice (bounded and Lipshitz) so that we have a strong solution. We would like to understand how the law μ^ε of X^ε looks like as $\varepsilon \to 0$.

Is not difficult to prove that $(\mu^\varepsilon)_\varepsilon$ converges in law (as probability measures on $\mathscr{C}^d = C(\mathbb{R}_+; \mathbb{R}^d)$ with its Borel σ-field) to the Dirac mass μ^0 concentrated on the solution x^0 of the ODE

$$\dot{X}_t^0 = b(X_t^0), \qquad X_0^0 = x_0.$$

(for example proving that $\mathbb{E}[\sup_{t \in [0,T]} |X_t^\varepsilon - X_t^0|^2] \to 0$ and conclude from this).

In large deviations theory one is concerned with the *speed* with which $\mu^\varepsilon \to \mu^0$, namely one would like to quantify this convergence and usually it will happen that this convergence is exponential, in the sense that

$$\mathbb{P}(X^\varepsilon \in A) \approx e^{-r(\varepsilon)C(A)}$$

where $r(\varepsilon) \to \infty$ as $\varepsilon \to 0$ nd it is usually something like $\varepsilon^{-\alpha}$ and $C(A)$ is a constant which depends only on the particular set A.

For example we could ask $A_{\gamma,T,\delta} = \{\omega \in \mathscr{C}^d : \sup_{t\in[0,T]} |\omega(t) - \gamma(t)| < \delta\}$ for given $\gamma \in \mathscr{C}^d$, $\delta > 0$ and $T > 0$. In this case if $\sup_{t\in[0,T]} |\omega(t) - X^0(t)| > \delta$ then $X^0 \notin A_{\gamma,T,\delta}$ and $\mu^\varepsilon(A_{\gamma,T,\delta}) \to 0$. We are going to prove that what will happen is that

$$\varepsilon \log \mu^\varepsilon(A_{\gamma,T,\delta}) = \varepsilon \log \mathbb{P}(X^\varepsilon \in A_{\gamma,T,\delta}) \approx -\inf_{x\in A_{\gamma,T,\delta}} I(x)$$

where I is a function which is only depending on b, σ and on the original problem and is called a *rate function*. They are called large deviations because they happen on an exponential scale. Otherwise stated we have an explitic asymptotic formula for the probability which looks like

$$\mu^\varepsilon(B) \approx e^{-\frac{1}{\varepsilon}\inf_{x\in B} I(x)}$$

Large Deviation Theory is concerned in general in the study of such large fluctuations in a variety of contextes (deviations from the law of large numbers, deviations from the ergodic behaviour, deviations from small noise behaviour like in this case, deviations from the large sample behaviour in statistics).

In order to properly speak about large deviations for the SDEs above we need some standard definitions from large deviation theory.

Definition 2. *A function $I: \mathscr{E} \to [0, +\infty]$ is called a (good) rate function on a Polish space $\mathscr{E}$ if the sets $I^{-1}[0, M] = \{x \in \mathscr{E} : I(x) \leqslant M\} \subset \mathscr{E}$ are compact for all $M < +\infty$.*

In particular, a rate function is always lower semicontinuous.

Definition 3. *Let I be a rate function on a Polish space $\mathscr{E}$ and $(Y^\varepsilon)_{\varepsilon>0}$ a family of random variables with values in $\mathscr{E}$. The this family satisfies the Laplace principle on $\mathscr{E}$ with rate function I (and rate $1/\varepsilon$) if for any function $h \in C_b(\mathscr{E})$ (bounded and continuous) we have*

$$\lim_{\varepsilon\to 0} -\varepsilon \log \mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = \inf_{x\in\mathscr{E}} [I(x) + h(x)]. \tag{2}$$

A Laplace principle is telling us that the law μ^ε of Y^ε is behaving like $e^{-I(x)/\varepsilon}$, in the sense that

$$\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = \int e^{-h(x)/\varepsilon} \mu_\varepsilon(\mathrm{d}x) \approx \int e^{-h(x)/\varepsilon} e^{-I(x)/\varepsilon} \mathrm{d}x = \int e^{-(h(x)+I(x))/\varepsilon} \mathrm{d}x = e^{-\frac{1}{\varepsilon}\inf_{x\in\mathscr{E}}[I(x)+h(x)](1+o(1))}.$$

Definition 4. *A family $(Y^\varepsilon)_{\varepsilon>0}$ satisfies the Large Deviation principle on $\mathscr{E}$ with rate function I (and rate $1/\varepsilon$) if for any open set $A \in \mathscr{E}$ and closed set $B \in \mathscr{E}$ we have*

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(Y^\varepsilon \in A) \geqslant -\inf_{x\in A} I(x),$$

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(Y^\varepsilon \in B) \leqslant -\inf_{x\in B} I(x).$$

Remark 5. Recall that if $\mu^\varepsilon \to \mu$ weakly, then the Portmanteau theorem asserts that for any open set A and closed set B you have

$$\liminf_{\varepsilon\to 0} \mu^\varepsilon(A) \geqslant \mu(A), \qquad \limsup_{\varepsilon\to 0} \mu^\varepsilon(B) \leqslant \mu(B)$$

while if $f \in C_b(\mathscr{E})$ the of course

$$\lim_{\varepsilon\to 0} \int f(x) \mu^\varepsilon(\mathrm{d}x) = \int f(x) \mu(\mathrm{d}x).$$

There are very strong similarities between weak convergence and large deviations.

Theorem 6. *The Laplace principle is equivalent to the Large Deviation principle.*

Proof. Exercise. □

Now we are going to use the Boué–Dupuis formula to prove large deviations for a large class of problems which in particular include the small noise diffusion problem introduced above.

Let $(Y^{\varepsilon})_{\varepsilon>0}$ a family of random variables defined on a Wiener space $(\Omega, \mathscr{F}, \mathbb{W}, X)$ with $\mathbb{W}$ the Wiener measure and taking values in $\mathscr{E}$ which are obtained from X using a family of mappings $\mathscr{G}^{\varepsilon}: \Omega \to \mathscr{E}$ i.e. $Y^{\varepsilon} = \mathscr{G}^{\varepsilon}(X)$.

Let $\mathbb{U}_M \subseteq L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ the subset of elements $u \in L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ such that $\|u\|_{\mathbb{H}} \leqslant M$ and let $\mathscr{U}_M \subseteq L^2_{\mathscr{P}}(\mathbb{R}_{\geqslant 0} \times \Omega; \mathbb{R}^d)$ the subset of drifts $u \in L^2_{\mathscr{P}}(\mathbb{R}_{\geqslant 0} \times \Omega; \mathbb{R}^d)$ such that $\|u\|_{\mathbb{H}} \leqslant M$ holds μ-almost surely, i.e. $u(\cdot, \omega) \in \mathbb{U}_M$ for μ almost every $\omega \in \Omega$.

Note that $\mathbb{U}_M$ is a compact Polish space with respect to the weak topology of $L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ (by Banach-Alaoglu theorem).

We define $J(u)(t) \coloneqq \int_0^t u(s) \mathrm{d}s$ for any $u \in \mathbb{H} = L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ and then $J: L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d) \to C(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d) = \Omega$.

We will make the following assumptions on the family $(\mathscr{G}^{\varepsilon})_{\varepsilon>0}$.

Hypothesis 7. *There exists a measurable mapping $\mathscr{G}^0: \Omega \to \mathscr{E}$ such that the following holds*

a) for every $M < \infty$ and any family $(u^{\varepsilon})_{\varepsilon} \subseteq \mathscr{U}_M$ such that $(u^{\varepsilon})_{\varepsilon}$ converges in law (as a random element of $\mathbb{U}_M$, and with the weak topology of $L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$) to u we have that

$$\mathscr{G}^{\varepsilon}(X + \varepsilon^{-1/2} J(u^{\varepsilon})) \to \mathscr{G}^0(J(u))$$

in law as random variables (on $(\Omega, \mathscr{F}, \mathbb{W})$) with values in $\mathscr{E}$ (of course as $\varepsilon \to 0$).

b) for every $M < \infty$ the set $\Gamma_M \coloneqq \{\mathscr{G}^0(J(u)): u \in \mathbb{U}_M\}$ is a compact subset of $\mathscr{E}$.

For each $x \in \mathscr{E}$ we define

$$I(x) \coloneqq \frac{1}{2} \inf_{u \in \Gamma(x)} \|u\|_{\mathbb{H}}^2 \tag{3}$$

where the infimum is take over the set $\Gamma(x) \subseteq \mathbb{H} = L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ such that $x = \mathscr{G}^0(J(u))$ and is taken to be $+\infty$ if this set is empty.

Lemma 8. *Under the Hypothesis 7 the function I is a rate function.*

Proof. (exercise) □

Theorem 9. *Under the Hypothesis 7 the family $(Y^{\varepsilon} = \mathscr{G}^{\varepsilon}(X))_{\varepsilon>0}$ satisfies the Laplace principle with rate function I as defined in (3) and speed $1/\varepsilon$.*

Proof. We need to show that

$$\lim_{\varepsilon \to 0} -\varepsilon \log \mathbb{E}[e^{-h(Y^{\varepsilon})/\varepsilon}] = \inf_{x \in \mathscr{E}} [I(x) + h(x)]$$

holds for any $h \in C_b(\mathscr{E})$.

Lower bound. By Boué–Dupuis formula we have

$$-\varepsilon \log \mathbb{E}[e^{-h(Y^{\varepsilon})/\varepsilon}] = -\varepsilon \log \mathbb{E}[e^{-h(\mathscr{G}^{\varepsilon}(X))/\varepsilon}] = \inf_u \mathbb{E}\left[h(\mathscr{G}^{\varepsilon}(X + J(u))) + \frac{1}{2}\|\varepsilon^{1/2} u\|_{\mathbb{H}}^2\right]$$

By renaming $u \to \varepsilon^{-1/2} u$ we have

$$-\varepsilon \log \mathbb{E}[e^{-h(Y^{\varepsilon})/\varepsilon}] = \inf_u \mathbb{E}\left[h(\mathscr{G}^{\varepsilon}(X + \varepsilon^{-1/2} J(u))) + \frac{1}{2}\|u\|_{\mathbb{H}}^2\right].$$

Fix $\delta > 0$. Then for any $\varepsilon > 0$ there exists an approximate minimiser u^ε such that

$$-\varepsilon \log \mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] \geqslant \mathbb{E}\left[h(\mathcal{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon))) + \frac{1}{2}\|u^\varepsilon\|_{\mathbb{H}}^2\right] - \delta.$$

This implies in particular that

$$\mathbb{E}\left[\frac{1}{2}\|u^\varepsilon\|_{\mathbb{H}}^2\right] \leqslant \delta - \varepsilon \log \mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] + \|h\|_{C_b(\mathscr{E})} \leqslant \delta + 2\|h\|_{C_b(\mathscr{E})} < \infty,$$

and this bound is independent of ε.
Moreover taking N large enough we can replace u^ε by the stopped process $u_t^{\varepsilon,N} = u_t^\varepsilon \mathbb{1}_{t \leqslant \tau_{\varepsilon,N}}$ with

$$\tau_{\varepsilon,N} := \inf\{t \geqslant 0 : \|u^\varepsilon \mathbb{1}_{[0,t]}\|_{\mathbb{H}} \geqslant N\}.$$

In this case $u_t^{\varepsilon,N} \in \mathscr{U}_N$ and morever we have that

$$\mathbb{P}(u^\varepsilon \neq u^{\varepsilon,N}) \leqslant \mathbb{P}(\|u^\varepsilon\|_{\mathbb{H}} > N) \leqslant \frac{\mathbb{E}[\|u^\varepsilon\|_{\mathbb{H}}^2]}{N^2} \leqslant \frac{2\delta + 4\|h\|_{C_b(\mathscr{E})}}{N^2}$$

uniformly in ε.

□

Next tuesday I will finish this proof and give some applications, to SDEs. And then the rest of the week we will discuss Backward SDE and representations of non-linear PDEs.