1. Motivation from sampling.

$\mathbb{E}[f(X)] = \int f(x)\, p(x)\, dx$

How do we draw a sample from $X \sim \mu \leftrightarrow p(x)$ given?

change of measure

$= \int f(x) \frac{p(x)}{q(x)} q(x)\, dx$

$= \mathbb{E}_q [f(X) \frac{p(X)}{q(X)}]$

- We can sample a few distributions eg. Bernoulli, Gaussian.

$\rightarrow$ draw samples from $q(x)$, then assign weight $w(x) = \frac{p(x)}{q(x)}$.

$(X^m, w^m) \longrightarrow \approx \frac{1}{M} \sum_{m=1}^{M} f(X^m) w^m$

- How to sample a trajectory $X_{[0,T]}$ with a given distribution $\mathbb{P}^X$?
  - sample from $\mathbb{Q}^X$ defined by an SDE, then assign weight $\Delta\Delta$

2. Motivation from time series classification.

- Given a time series sampled from one of two known diffusions. How to classify it?

1> Given a data X^i sampled from $f_\theta(x)$, either $\theta = \theta_0$ or θ_1; Determine X^i is sample from θ_1 or θ_0.

$\Rightarrow$ Hypothesis testing $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta = \theta_1 \end{cases}$ choose a rejection set R: $\begin{cases} \text{accept } H_0 \text{ if } X^i \notin R \\ \text{reject } H_0 \text{ otherwise.} \end{cases}$

☆ Likelihood ratio test. $l(\theta_0, \theta_1 | X) = \log \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}$

$R_k = \{ X : l(\theta_0, \theta_1 | X) \geq k \}$

Neyman-Pearson Lemma, LRT is the uniformly most powerful test.

2>. Generalize the above to time series. What is the likelihood ratio?

Girsanov Theorem on change of measure.

1. A motivation from inference (other: importance density)

<u>Problem 1</u>: Given data $\{X^{(m)}\}_{m=1}^{M}$, $X^{(m)} \sim X$ with pdf u_{θ_0}, where $\{u_\theta : \theta \in \Theta\}$ are pdfs

To do: estimate θ_0.

Maximal likelihood estimator (MLE)

$$\hat{\theta}_M = \underset{\theta \in \Theta}{\arg\max} \underbrace{\frac{1}{M}\sum_{m=1}^{M} \ell(\theta, X^{(m)})} \qquad \log u_\theta(X^{(m)})$$

Fact: If $\mathbb{E}[\ell(\theta, X)]$ has ! maximizer. Then $\hat{\theta} - \theta_0 \sim \frac{1}{\sqrt{M}} N(0, \sigma^2)$ as $M \uparrow \infty$.

(need suitable conditions on Fisher information matrix.)

<u>Problem 2</u>. Given Data $\{X_{[0,T]}^{(m)}\}_{m=1}^{M}$: sample paths of $dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t$.

ToDo: estimate b. · If $\sigma \equiv 0$, then $\dot{X}_t = b(X_t)$ $\quad \mathcal{E}(b) = \int_0^T \|\dot{X}_t - b(X_t)\|^2 dt$

① . A "pdf" for $X_{[0,T]}$ (and then take log)? $\quad p(X_{[0,T]})$ No.

· Discrete-time approximation? $\quad X_{t_1}, \cdots, X_{t_L} \quad p(X_{t_1}, \cdots X_{t_L})$; $\log u(X_{t_1}, \cdots, X_{t_L})$

$\downarrow L \uparrow \infty \ (\Delta t \downarrow 0)$

$\xrightarrow{\text{Markov}} \quad p(x_1 \cdots x_L) = p(x_2|x_1)\, p(x_3|x_2) \cdots p(x_L|x_{L-1}) \qquad ?.?$

$\xrightarrow{\text{EM}} \quad X_{t_{l+1}} \approx X_{t_l} + \Delta t\, b(X_{t_l}) + \sqrt{\Delta t}\, \sigma(X_{t_l}) W_n$

$$p(x_{l+1}|x_l) \simeq N(x_l + \Delta t\, b(x_l), \Delta t\, \sigma(x_l)^2) = \frac{1}{\sqrt{2\pi\Delta t}\,\sigma(x_l)} \exp\left(-\frac{|x_{l+1} - x_l - \Delta t\, b(x_l)|^2}{2\Delta t\, \sigma(x_l)^2}\right)$$

$$\Rightarrow \log p(x_1 \cdots x_L) \approx \sum_{l=1}^{L-1}\left[-\cancel{\log(\sqrt{2\pi\Delta t}\,\sigma(x_l))} - \frac{|X_{l+1} - X_l - \Delta t\, b(X_l)|^2}{2\Delta t\, \sigma(X_l)^2}\right]$$

$\downarrow L \uparrow \infty$

negative-log

drop it, (it is indpt. of b) $\quad$ why ok? $\quad \sigma = 1$

$$\int_0^T \left(\cancel{\left|\frac{dX}{dt}\right|^2} + b(X_t)^2 - 2b(X_t)\frac{dX_t}{dt}\right) dt \longleftarrow -\int_0^T \frac{|dX_t - b(X_t)dt|^2}{2\,dt\,\sigma(X_t)^2} = -\frac{1}{2}\int_0^T \frac{\left|\frac{dX_t}{dt} - b(X_t)\right|^2}{\sigma(X_t)^2} dt$$

$$= \int_0^T b(X_t)^2 dt - 2\int_0^T b(X_t)\, dX_t$$

But $\frac{dX}{dt}$ D.N.E. a.s. !!!

$= \mathcal{E}(b)$ a function (functional) of b. ✓✓

Example OU : $dX_t = \theta X_t dt + dB_t$

$$X_{t+\Delta t} = e^{\theta \Delta t} X_t + \underbrace{\int_t^{t+\Delta t} e^{\theta(t+\Delta t - s)} dB_s}_{N(0,\sigma^2) \quad \sigma^2 = \frac{1}{2\theta}(1-e^{2\theta \Delta t})}$$

$$Y_{n+1} = (1+\theta \Delta t) Y_n + \sqrt{\Delta t}\, N(0,1)$$

$$\downarrow$$

$$l(\theta, X_{t_1:t_L}) = \sum_{l=1}^{L-1} \frac{|X_{t_{l+1}} - (1+\theta\Delta t) X_{t_l}|^2}{2\Delta t}$$

$$\Rightarrow \hat{\theta} = \left(\sum_{l=1}^{L-1} X_{t_l}^2\right)^{-1} \left(\sum_{l=1}^{L} \frac{X_{t_{l+1}} - X_{t_l}}{\Delta t} X_{t_l}\right)$$

$\theta < 0$ $\quad t_L \uparrow \infty, \downarrow$

$$\mathbb{E}[X_t^2]^{-1} \mathbb{E}\left[X_t^2 \frac{e^{\theta \Delta t} - 1}{\Delta t}\right] = \frac{e^{\theta \Delta t} - 1}{\Delta t} = \theta + O(\Delta t^2)$$

But $\Delta t = \frac{T}{L}$, $L \uparrow \infty$

$$\lim_{L\uparrow\infty} l(\theta, X_{t_1:t_L}) = \lim_{L\uparrow\infty} \sum_{l=1}^{L-1} \frac{|X_{t_{l+1}} - X_{t_l} - \theta \Delta t X_{t_l}|^2}{2\Delta t}$$

$$= \int_0^T \frac{|dX_t - \theta X_t dt|^2}{2\, dt} = \frac{1}{2}\int_0^T \left|\frac{dX_t}{dt} - \theta X_t\right|^2 dt \qquad = \infty$$

② A change of measure.

· Recall pdf of X : $\quad u_\theta(x) = \dfrac{dP_\theta(x)}{dP_L(x)}$ $\longrightarrow \mathbb{P}(X\in D) = \int_D dP_\theta(x)$; $dP_L(x) \longrightarrow$ Lebesgue measure on $\mathbb{R}^n$

We can also use any other measure $P_*(dx) = u_*(x)dx$, with u_0 known

$$e^{l(\theta, X)} = \frac{u_\theta(x)}{u_*(x)} = \frac{dP_\theta(x)}{dP_*(x)}$$

· Back to the process $X_{t_1:t_L}$ (AR1)

$$e^{l(\theta, X_{t_1:t_L})} = \frac{dP_\theta(x_1:x_L)}{dP_*(x_1:x_L)} = \frac{C_1 e^{-\frac{1}{2\Delta t}\sum_{l=1}^{L-1} |X_{l+1} - X_l - \theta\Delta t X_l|^2}}{C_2\, e^{-\frac{1}{2\Delta t}\sum_{l=1}^{L-1} |X_{l+1} - X_l|^2}}$$

Gaussian increments

$$\Rightarrow \quad l(\theta, X_{t_1:t_L}) = C_{\Delta t} - \frac{1}{2\Delta t} \sum_{l=1}^{L-1} \left[-(X_{l+1} - X_l)\Delta t X_l + \Delta t^2 X_l^2\right]$$

$$= C_{\Delta t} - \frac{1}{2} \sum_{l=1}^{L-1} \left[X_l^2 \Delta t - 2X_l(X_{l+1} - X_l)\right]$$

$$\xrightarrow{\Delta t = \frac{T}{L} \downarrow 0} \quad l(\theta, X_{[0,T]}) = \lim_{L\uparrow\infty} l(\theta, X_{t_1:t_L}) = -\frac{1}{2}\left(\int_0^T X_t^2 dt - 2\int_0^T X_t\, dX_t\right)$$

· Question: what is the limit of the measure $\mathbb{P}_*(X_1:X_L)$ as $L\uparrow\infty$?

$\downarrow$

finite L: Gaussian process $X_{t_1:t_L}$ with indpt increment

$X_{t_{l+1}} - X_{t_l} \sim N(0, t_{l+1} - t_l)$

i.e. Brown motion at discrete times.

· Question 2 : can we use other measure ? Yes, any reference measure s.t. $\mathbb{P}_\theta \ll \mathbb{P}_*$

For computation, some measure works better. (similar in importance sampling).

11/1 Girsanov Thm

From previous examples $\left.\frac{dQ}{dP}\right|_{\mathcal{F}_T}(X_{[0,t]})$

- MLE likelihood ratio
- Importance sampling
- Time series classification

$$dX_t = b(t, X_t)\,dt + \sigma(X_t)\,dB_t$$

$$dY_t = b_2(t, Y_t)\,dt + \sigma_2(Y_t)\,dB_t^2$$

$\sigma_1 = \sigma_2$ $\quad \frac{dQ^X}{dP^Y}(X_{[0,t]}) = \exp\left(\int_0^t \Delta b \cdot dX_s - \frac{1}{2}\int_0^t \left(|b_1|^2 - |b_2|^2\right) ds\right)$ ← How to get it from the 3 Thms?

- When the ratio exists? → Radon-Nikodym
  ($\mathbb{P}$, t changing)
- what is the essential? → change of meas.
- Can B & B^2 be different? ✓ distn.
  σ_1 & σ_2 ? No.
- other applications? weak sdu.

3.8 Girsanov Theorem

Girsanov theorem says that a Brownian motion with drift $B_t+\lambda t$ can be seen as a Brownian motion without drift, with a change of probability. We first discuss changes of probability by means of densities.

Suppose that $L \geq 0$ is a nonnegative random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $E(L) = 1$. Then,

$$\boxed{Q(A) = E\left(\mathbf{1}_A L\right)}$$

defines a new probability. In fact, Q is a σ-additive measure such that

$$Q(\Omega) = E(L) = 1.$$

We say that L is the *density* of Q with respect to P and we write

$$\frac{dQ}{dP} = L.$$

The expectation of a random variable X in the probability space $(\Omega, \mathcal{F}, Q)$ is computed by the formula

$$E_Q(X) = E(XL).$$

The probability Q is absolutely continuous with respect to P, that means,

$$P(A) = 0 \Longrightarrow Q(A) = 0.$$

If L is strictly positive, then the probabilities P and Q are *equivalent* (that is, mutually absolutely continuous), that means,

$$P(A) = 0 \iff Q(A) = 0.$$

The next example is a simple version of Girsanov theorem.

Example 9 Let X be a random variable with distribution $N(m, \sigma^2)$. Consider the random variable

$$L = e^{-\frac{m}{\sigma^2}X + \frac{m^2}{2\sigma^2}}.$$

which satisfies $E(L) = 1$. Suppose that Q has density L with respect to P. On the probability space $(\Omega, \mathcal{F}, Q)$ the variable X has the characteristic function:

$= \mathbb{E}[e^{itx} \frac{dQ}{dP})$

$$\begin{aligned} E_Q(e^{itX}) &= E(e^{itX}L) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2} - \frac{mx}{\sigma^2} + \frac{m^2}{2\sigma^2} + itx} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2} + itx} dx = e^{-\frac{\sigma^2 t^2}{2}}, \end{aligned}$$

so, X has distribution $N(0, \sigma^2)$.

Let $\{B_t, t \in [0, T]\}$ be a Brownian motion. Fix a real number λ and consider the martingale

$$\boxed{L_t = \exp\left(-\lambda B_t - \frac{\lambda^2}{2}t\right).} \tag{24}$$

2. Girsanov Theorem (theory)

1>. a Lévy characterization of Bm

2>. absolute continuity of measures

3>. Girsanov formulas.

1> Theorem 8.6.1 (The Lévy characterization of Bm)

David Nualart;s note:

Let $X=(X_1(t),\cdots,X_n(t))$ be a cts process on $(\Omega,\mathcal{H},Q)$ with values in $\mathbb{R}^n$. Then

(a) $X(t)$ is a Bm w.r.t. Q $\Leftrightarrow$

(b) (i) $X(t)$ is a mG wrt Q and its own filtration

(ii) $X_i(t)X_j(t)-\delta_{ij}\cdot t$ is a mG wrt Q & $\mathcal{F}_t^X$. $\forall i,j\in\{1,\cdots,n\}$.

Lemma 8.6.2 (Bayes' rule) $\qquad \mathbb{P}(A|B)=\frac{\mathbb{P}(AB)}{\mathbb{P}(B)}=\frac{P(B|A)P(A)}{P(B)}$

Let μ & ν be two probab. measures on $(\Omega,\mathcal{G})$ st $d\nu(\omega)=f(\omega)\,d\mu(\omega)$, $f\in L^1(\mu)$.

Let $X\in L^1(\nu)$, $\mathcal{H}\subseteq\mathcal{G}$. Then $\qquad \mathbb{E}_\nu[X|\mathcal{H}]\cdot\mathbb{E}_\mu[f|\mathcal{H}]=\mathbb{E}_\mu[fX|\mathcal{H}] \qquad (*)$

"Proof" $\forall H\in\mathcal{H}$, $\quad \mathbb{E}_\mu[1_H\,\mathbb{E}_\nu[X|\mathcal{H}]f]=\int_H \mathbb{E}_\nu[X|\mathcal{H}]f(\omega)d\mu(\omega)=\int_H \mathbb{E}_\nu[X|\mathcal{H}]\,d\nu(\omega)$

$\mathbb{E}_\mu[\mathbb{E}_\mu[1_H\mathbb{E}_\nu[X|\mathcal{H}]f|\mathcal{H}]] \qquad =\int_H Xf\,d\mu \leftarrow \;=\int_H X\,d\nu$

$=\mathbb{E}_\mu[1_H\,\mathbb{E}_\nu[X|\mathcal{H}]\,\mathbb{E}_\mu[f|\mathcal{H}]] \quad \xleftrightarrow[\forall H\in\mathcal{H}]{} \quad =\mathbb{E}_\mu[fX1_H] \qquad \Rightarrow (*)$ by def.

2> Absolute continuity of measures

$(\Omega,\mathcal{F},P)$ a probab. space, $\{\mathcal{F}_t\}_{t\geq0}$ a filtration. Q another probab. meas. on $\mathcal{F}_T$.

Q is absolutely continuous w.r.t. $P|_{\mathcal{F}_T}$ if $P(H)=0\Rightarrow Q(H)=0$, $\forall H\in\mathcal{F}_T$. "$Q\ll P$."

Radon-Nikodym Theorem: $Q\ll P\Leftrightarrow\exists$ $\mathcal{F}_T$-measurable r.v. Z_T st. $dQ(\omega)=Z_T(\omega)dP(\omega)$

$$\frac{dQ}{dP}=Z_T \text{ on } \mathcal{F}_T. \qquad \text{Radon-Nikodym derivative.}$$

Lemma 8.6.3. Sps $Q\ll P|_{\mathcal{F}_T}$ with $\frac{dQ}{dP}=Z_T$ on $\mathcal{F}_T$. Then $Q|_{\mathcal{F}_t}\ll P|_{\mathcal{F}_t}$ $\forall t\in[0,T]$, and

$Z_t:=\frac{dQ|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}}$ is a martingale w.r.t. $\mathcal{F}_t$ and P. $\quad$ (# $\Leftarrow \mathbb{E}_P[Z_T|\mathcal{F}_t]=Z_t$ $\Leftarrow$ by Z_t def)

Proof: $\mathcal{F}_t\subseteq\mathcal{F}_T\Rightarrow Q\ll P$ on $\mathcal{F}_t$; $\quad\forall F\in\mathcal{F}_t$, $\mathbb{E}_P[1_F\,\mathbb{E}_P[Z_T|\mathcal{F}_t]]=\mathbb{E}_P[1_F Z_T]=\mathbb{E}_Q 1_F=\mathbb{E}_P[1_F Z_t]$

Girsanov Thm I. $\underline{dY_t = a(t,\omega)dt + dB_t\,;\ Y_0 = 0 \in \mathbb{R}^n;}$ B_t: $\mathbb{R}^n$-valued Bm

Then, $\{Y_t\}_{0\le t\le T}$ is a Bm wrt. Q s.t. $\frac{dQ}{dP} = M_T$ if

$$M_t = \exp\left(-\int_0^t a(s,\omega)dB_s - \frac{1}{2}\int_0^t a^2(s,\omega)ds\right) \text{ is a mG wrt. } \{\mathcal{F}_t^B\} \,\&\, P.$$

[Proof see next page.]

- A sufficient condition (Novikov) $\mathbb{E}_P \exp\left(\frac{1}{2}\int_0^t a^2(s,\omega)ds\right) < \infty$.
- Since $M_T(\omega) > 0$ a.s., we have $\mathbb{P} \ll Q$ too. $\Rightarrow$ P & Q are equivalent i.e.

$$\mathbb{P}(Y_{t_1}\in F_1;\cdots;Y_{t_k}\in F_k) > 0 \Leftrightarrow Q(Y_{t_1}\in F_1;\cdots;Y_{t_k}\in F_k) > 0$$

$$\Leftrightarrow P(B_{t_1}\in F_1,\cdots,B_{t_k}\in F_k) > 0.$$

- Example $a(t,\omega) = a(t)$ deterministic.

Girsanov Theorem II. $\underline{dY_t = \beta(t,\omega)\,dt + \theta(t,\omega)\,dB_t\,;}$ $\beta\in\mathbb{R}^n$, $\theta\in\mathbb{R}^{n\times m}$, $B_t\in\mathbb{R}^m$

If $\exists$ $u(t,\omega)\in W_{\mathcal{H}}^{B}$ and $\alpha(t,\omega)\in W_{\mathcal{H}}^{Y}$ s.t. $\underline{\theta(t,\omega)\,u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega)}$.

Assume $\boxed{M_t = \exp\left(-\int_0^t u(s,\omega)dB_s - \frac{1}{2}\int_0^t u^2(s,\omega)ds\right)}$ is a mG w.r.t. $\mathcal{F}_t^Y$ & P,

then $\hat{B}_t = \int_0^t u(s,\omega)ds + B_t$ is a Bm w.r.t. Q s.t. $\frac{dQ}{dP} = M_T$ on $\mathcal{F}_T^B$

and $dY_t = \alpha(t,\omega)\,dt + \theta(t,\omega)\,d\hat{B}_t$.

Proof: $\xrightarrow{\text{Thm I}}$ Q is a probab. meas. on $\mathcal{F}_T^B$ and $\hat{B}_t$ is a Bm w.r.t. Q.

$$dY_t = \beta(t,\omega)dt + \theta(t,\omega)\left(d\hat{B}_t - u(t,\omega)dt\right) = \alpha(t,\omega)\,dt + \theta\, d\hat{B}_t. \quad \#.$$

Girsanov III (for Ito diffusion $\Leftrightarrow$ process)

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t\,, \qquad t\in T,\quad X_0 = x$$

$$dY_t = [\gamma(t,\omega) + b(Y_t)]\,dt + \sigma(Y_t)\,dB_t,\quad t\in T,\quad Y_0 = x$$

$b\in\mathbb{R}^n$, $\sigma\in\mathbb{R}^{n\times m}$ linear growth, Lipschitz. B_t: $\mathbb{R}^m$-Bm.

Assume $\gamma\in W_H^n$, $\exists$ $u(t,\omega)\in W_H^m$ s.t. $\sigma(Y_t)\,u(t,\omega) = \gamma(t,\omega)$.

$\sigma u = b_Y - b_X$

Define M_t, Q & $\hat{B}_t$ as above, and assume M_t is a mG wrt. $\mathcal{F}_t^B$ & P.

Then Q is a probab. meas. $dY_t = b(Y_t)\,dt + \sigma(Y_t)\,d\hat{B}_t$

i.e. the Q-law of Y_t^x is the same as the P-law of X_t^x.

$$\frac{dP_Y}{dQ} = \frac{dP_X}{dP}$$

$$\Rightarrow \frac{dP_Y}{dP_X} = \frac{dQ}{dP}$$

$$= \exp\left(-\int_0^t \frac{b_Y - b_X}{\sigma}dB - \frac{1}{2}\int_0^t u^2 ds\right)$$

Proof: Direct application of Theorem II: $\theta(t,\cdot) = \sigma(Y_t)$, $\beta(t,\cdot) =$

$\Rightarrow b_r = 0, \qquad \frac{dP_r}{dP_r} = \exp\left(\int_0^t b_x\, dB + \frac{1}{2}\int_0^t u^2 ds\right)$

Proof of Thm I:

Since M_t is a mG, we have $Q(\Omega) = \mathbb{E}_Q 1 = \mathbb{E}_P[M_T] = \mathbb{E}_P[M_0] = 1. \Rightarrow Q$ is a probab. meas.

WLOG, assume $a(s,\omega)$ is bdd (otherwise consider $a \wedge k$ first, and then send $k \to \infty$)

In view of Levy's characterization of Bm, we need to verify that

(i) $Y_t = (Y_1(t), \cdots, Y_n(t))$ is a mG w.r.t. Q

(ii) $Y_i(t) Y_j(t) - \delta_{ij} \cdot t$ is a mG w.r.t. Q, $\forall\, i, j$.

To verify (i): let $K(t) = M_t Y_t$. By Ito's formula,

$$d e^{Z_t} = e^{Z_t}\left(dZ_t + \frac{1}{2}\langle dZ\rangle_t\right) \Rightarrow dM_t = M_t(-a\, dB_t)$$

$$\begin{aligned}
dK_i(t) &= M_t\, dY_i(t) + Y_i(t)\, dM_t + dM_t\, dY_i(t) \\
&= M_t(\cancel{a_i dt} + dB_i(t)) + Y_i(t) M_t(-a \cdot dB_t) + \cancel{M_t}\underbrace{(-a\, dB_t)\, dB_i(t)}_{\to\, \cancel{(-a_i dt)}} \\
&= M_t(\underbrace{dB_i(t) - Y_i(t)\, a\, dB_t}) = M_t\, \gamma^{(i)} dB_t \\
&= dB_i(t) - Y_i(t) \sum_{j=1}^n a_j(t)\, dB_j(t) = (\underbrace{\vec{e}_i - Y_i(t)\, a}) dB.
\end{aligned}$$

Hence, $K_i(t)$ is a mG w.r.t. P, so By the Lemma (Bayes rule), we get $\forall\, t > s$

$$\mathbb{E}_Q[Y_i(t) \mid \mathcal{F}_s] = \frac{\mathbb{E}_P[M_t Y_i(t) \mid \mathcal{F}_s]}{\mathbb{E}_P[M_t \mid \mathcal{F}_s]} = \frac{\mathbb{E}[K_i(t) \mid \mathcal{F}_s]}{M_s} = \frac{K_i(s)}{M_s} = Y_i(s),$$

i.e., $Y_i(t)$ is a mG w.r.t. Q. This proves (i).

The proof of (ii) is similar.

Applications / examples

Example 8.6.7 Let $dY(t) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} dt + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}$ ($dY = \beta dt + \theta dB_t$)

Let $\alpha(t,\omega) = 0$ in Thm II, $\theta u = \beta$ $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

Then, for Q: $\frac{dQ}{dP} = e^{-3B_1(T) + B_2(T) - 5T}$; $M_t = \exp(-\int_0^t u\, dB_s - \frac{1}{2}\int_0^t u^2 ds)$

$d\hat{B}_t = \begin{pmatrix} 3 \\ -1 \end{pmatrix} dt + dB(t)$; $= \exp(-3dB_1(t) + dB_2(t) - 5t)$

we have $\hat{B}_t$ is a Bm w.r.t. Q, and $dY_t = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} d\hat{B}_t$.

Application of Thm III: weak soln to SDE.

Sps Y_t is a known weak or strong soln. to $dY_t = b(Y_t)dt + \sigma(Y_t) dB_t$ (1) $b, a: \mathbb{R}^n \to \mathbb{R}^n$

We wish to find a weak soln. to $dX_t = a(X_t)dt + \sigma(X_t) dB_t$. (2) $\sigma: \mathbb{R}^n \to \mathbb{R}^{n \times m}$

Assume $\exists\ u_0\ \mathbb{R}^n \to \mathbb{R}^n$ s.t. $\sigma(y) u_0(y) = b(y) - a(y)$, $\forall y \in \mathbb{R}^n$. ($u_0 = \sigma^{-1}(b-a)$ if $\sigma^{-1} \exists$)

Then if $u_0(Y_t(\omega))$ satisfies the Novikov condition, $\mathbb{E}[\exp(\frac{1}{2}\int_0^T u^2 ds)] < \infty$,

we have $\hat{B}_t = \int_0^t u(s,\omega) ds + B_t$ is a Bm w.r.t. Q $\frac{dQ}{dP} = M_t = \exp(-\int_0^t u\, dB_s - \frac{1}{2}\int_0^t u^2 ds)$

and $dY_t = a(Y_t)dt + \sigma(Y_t) d\hat{B}_t$.

That is, $(Y_t, \hat{B}_t)$ is a weak soln. to (2).

Example: To construct a weak soln. to $dX_t = a(X_t)dt + \sigma dB_t$; $X_0 = x \in \mathbb{R}^n$. $\sigma \in \mathbb{R}$ (4)

Start from $dY_t = \sigma dB_t$; $Y_0 = x$. $u_0 = \sigma^{-1}(b-a) = -\sigma^{-1} a$

Fix $T < \infty$ and put $\frac{dQ}{dP} = M_T$ on $\mathcal{F}_T^B$ $M_t = \exp(+\int_0^t \sigma^{-1} a(Y_s) dB_s - \frac{1}{2}\int_0^t |\sigma^{-1} a(Y_s)|^2 ds$

Then, $\hat{B}_t = -\int_0^t \sigma^{-1} a(Y_s) ds + B_t$ is a Bm w.r.t. Q

and $dY_t = \sigma dB_t = a(Y_s)ds + \sigma d\hat{B}_t$, i.e. $(Y_t, \hat{B}_t)$ is a weak soln to (4).

Application: likelihood of data $X_{[0,T]}$; $X_0 = x$;

$$\frac{dP_a}{dP_0} = \exp\left(\int_0^t \frac{a(X_s)}{\sigma^2} dX_s - \frac{1}{2}\int_0^t \frac{a(X_s)^2}{\sigma^2} ds \right).$$

X $\quad dX_t = a(X_t)dt + \sigma dB_t \quad \to \mathbb{P}_a$

X^0 $\quad dX_t = \sigma dB_t \quad \to \mathbb{P}_0$

X^b $\quad dX_t = b(X_t)dt + \sigma dB_t \quad \to \mathbb{P}_b$

$$\frac{dP_a}{dP_b} = \exp\left(\int_0^t \frac{(a-b)}{\sigma^2}(X_s) dX_s - \frac{1}{2}\int_0^t \frac{a^2-b^2}{\sigma^2}(X_s) ds \right). \quad (*)$$

$\frac{p_x}{p_0}$ $\quad -\sum \frac{|\Delta X - a\,dt|^2}{2\,dt}$ $\quad = \frac{1}{2}\sum \left[\frac{(dX)^2}{dt} - 2\langle a, dX\rangle + a^2 dt\right] = \Box + \langle a, dX\rangle - \frac{1}{2}a^2 dt$

WLOG $\sigma = 1$. ① Q-law of X^0 = P-law of X^a $\quad \frac{dP_0}{dQ} = \frac{dP_a}{dP}$

$$\frac{dP_a}{dP_0} = 1\Big/\frac{dQ}{dP} = \exp\Big(+\int_0^t a(X_s)\underbrace{dB_s}_{dX_s - a(X_s)ds} + \frac{1}{2}\int_0^t a^2(X_s)ds\Big) = \exp\Big(+\int_0^t a(X_s)\,dX_s - \frac{1}{2}\int_0^t a(X_s)^2 ds\Big)$$

② Q-law of X^b = P-law of X^a $\quad \frac{dP_b}{dQ} = \frac{dP_a}{dP}$ $\quad \sigma u = a - b$

$$\Rightarrow \frac{dP_a}{dP_b} = 1\Big/\frac{dQ}{dP} = \exp\Big(\int_0^t \frac{a-b}{\sigma} dB_s + \frac{1}{2}\int_0^t \left(\frac{b-a}{\sigma}\right)^2 ds\Big)$$

$$= \exp\Big(\int_0^t \frac{a-b}{\sigma}(X_s)\, dX_s - \frac{1}{2}\int_0^t (a^2-b^2)\, ds\Big)$$

$(a-b)(dX_s - a\,ds) + \frac{1}{2}(a-b)^2$
$= (a-b)dX_s + ds\left[-a^2+ab + \frac{1}{2}(a-b)^2\right]$
$-\frac{1}{2}(a^2-b^2)$

$\frac{dP_a}{dP_0}\frac{dP_0}{dP_b} \Rightarrow (*)$

With $b(X)$ as the true drift (i.e. X_t from b), we get a loss functional (log-likelihood)

$$\ell_{X_{[0,t]}}(a) = \log \frac{dP_a}{dP_0} = \int_0^t a(X_s)\,dX_s - \frac{1}{2}\int_0^t a(X_s)^2 ds$$

Setting 1: Data: multi-trajectory: $\{X^{(m)}_{[0,T]}\}_{m=1}^M$ $\quad$ setting 2: Ergodic, $T \nearrow \infty$, $M=1$

$$L_M(a) = \frac{1}{M}\sum_{m=1}^M \ell_{X^m_{[0,T]}}(a) \to \mathbb{E}\Big[\int_0^T a(X_s)\underbrace{dX_s}_{b_*(X_s)dt + \sigma dB_s} - \frac{1}{2}\int_0^t a(X_s)^2 ds\Big] = L_\infty(a)$$

$$\hat{a} = \operatorname*{argmax}_{a \in \mathcal{H}} L_M(a)$$

$a = \sum_{i=1}^n c_i \varphi_i$, then

$$L_M(a) = \frac{1}{M}\sum_{m=1}^M \int_0^t \sum_i c_i \varphi_i(X_s^m)\,dX_s^m - \frac{1}{2}\sum_{ij} c_i c_j \varphi_i \varphi_j(X_s)\,ds$$

$$= b^T c - \frac{1}{2} c^T A c$$

$\Rightarrow \nabla L_M(c) = Ac - b$

$c = A^{-1}b \in \mathbb{R}^n$

① Is A inverstible?

as $n \nearrow \infty$?

$$= \mathbb{E}\int_0^T [a(X_s)b_*(X_s) - \frac{1}{2}a(X_s)^2]\,ds$$

$$= \int_{\mathbb{R}^d}\underbrace{\int_0^T p_s(x)}\left[a(x)b_*(x) - \frac{1}{2}a(x)^2\right] ds\, dx$$

$$= \int_{\mathbb{R}^d}\left[a(x)b_*(x) - \frac{1}{2}a(x)^2\right] \rho_T(x)\,dx$$

$$\langle \nabla L_\infty(a), \phi\rangle = \lim_{\varepsilon \downarrow 0} \frac{L_\infty(a+\varepsilon\phi) - L_\infty(a)}{\varepsilon}$$

$$= \mathbb{E}\int_0^T \phi(X_s)\left[b(X_s) - a(X_s)\right] ds, \ \forall \phi \in L^2(\rho_T)$$

$\nabla L_\infty(a) = b - a$ in $L^2(\rho_T)$

$\Rightarrow a = b$ is the ! soln in $L^2(\rho_T)$

Inference and nonparametric regression, ML.

$\{x_i, y_i\}_{i=1}^{M} \longrightarrow Y = f(X) + \varepsilon.$

$\mathcal{E}_M(f) = \frac{1}{M}\sum_{m=1}^{M} |y_i - f(x_i)|^2$

$\downarrow M \to \infty$

$Y = \mathbb{E}[Y|X] + \varepsilon$

$\hat{f}_M = \underset{f \in \mathcal{H}}{\operatorname{argmin}}\ \mathcal{E}_M(f)$

$\mathcal{E}_\infty(f) = \frac{1}{2}\mathbb{E}[|Y - f(X)|^2]$

$Y = Y_X + Y_\perp$

$\downarrow M \to \infty$

$\hat{f}_\infty(x) = \mathbb{E}[Y|X=x]$ $\Leftarrow$

$= \frac{1}{2}\int |\hat{y}(x) - f(x)|^2 \rho_X(x)dx = \frac{1}{2}\langle f, f\rangle_{L^2(\rho)} - \langle f, \hat{y}\rangle + |y|^2$

$\nabla \mathcal{E}_\infty(f) = f_\infty - \hat{y}_\infty$

$\Rightarrow \hat{f}_\infty = I^{-1}\hat{y}_\infty.$

- Back to the inference of SDE.

- Inference for interacting particle system

$$dX_t^i = \frac{1}{N}\sum_{j\neq i} \varphi(|X^j - X^i|)\frac{X^j - X^i}{|X^j - X^i|}dt + \sigma dB_t^i$$

Given $\{\vec{X}^{(m)}_{[0,T]}\}_{m=1}^{M}$, to estimate $\varphi: \mathbb{R}^+ \to \mathbb{R}$

$$d\vec{X}_t = f_\varphi(\vec{X}_t)dt + \sigma d\vec{B}_t$$

$\ell(\vec{X}_{[0,T]}|\varphi) = -\int_0^T \langle f_\varphi(\vec{X}_s), d\vec{X}_s\rangle + \frac{1}{2}\int_0^T |f_\varphi(\vec{X}_t)|^2 dt$ (negative log

$\mathcal{E}_M(\varphi) = \frac{1}{M}\sum_{m=1}^{M} \ell(\vec{X}^{(m)}_{[0,T]}|\varphi) \to \mathcal{E}_\infty(\varphi) = -\frac{1}{2}\mathbb{E}[\int_0^T \langle f_\varphi(\vec{X}_s), d\vec{X}_s\rangle - \frac{1}{2}\int_0^T |f_\varphi(\vec{X}_t)|^2 dt]$

$= \langle \varphi, L_G\varphi\rangle - 2\langle \varphi^D, L_G\varphi\rangle + C^D.$