1. **Moment from samples.**

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

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

   - We can sample a few distributions, e.g., Bernoulli, Gaussian.

   \[ \rightarrow \text{draw samples from } p(x), \text{ then assign weight } w(x) = \frac{p(x)}{p_0(x)}, \frac{p(x)}{p_0(x)} \]

   - How to sample a trajectory \( X_{[0,T]} \) with a given distribution \( p^a \)?

     - Sample from \( \mathcal{G}^a \) defined by an SDG, then assign weight \( \alpha \sim \alpha \).

2. **Moment from time series classification.**

   - Given a time series sampled from one of two known data-generating processes, how to classify \( x \)?

   1. Given a data \( x \) sampled from \( f_0(x) \), either \( \theta = \theta_0 \) or \( \theta_1 \); determine \( x \) is sampled from \( \theta_0 \) or \( \theta_1 \).

   \[ \Rightarrow \text{Hypothesis testing } H_0: \theta = \theta_0, \quad H_1: \theta = \theta_1 \]

   - Choose a rejection set \( R \); accept \( H_0 \) if \( x \notin R \) reject \( H_0 \) otherwise.

   - **Likelihood ratio test:**

     \[ l(\theta, \theta_1; x) = \log \frac{f_1(x)}{f_0(x)} \]

     \[ R_k = \{ x : l(\theta, \theta_1; x) > k \} \]

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

2. Generalize the above to time series. What is the likelihood ratio?
1. A motivation from inference (other: importance density)

**Problem 1.** Given data \( X^{(n)}_{m=1} \), \( X^{(m)} \sim X \) with pdf \( \theta_0 \), where \( \{\theta_0 : \theta \in \Theta\} \) are pdfs

To do: estimate \( \theta_0 \).

**Maximum likelihood estimate (MLE)**

\[
\hat{\theta}_M = \arg \max_{\theta \in \Theta} \frac{1}{n} \sum_{m=1}^{n} l(\theta, X^{(m)}) \quad \text{log} \ U_0(X^{(m)})
\]

Fact: If \( E[l(\theta, X)] \) has 1 maximizer, then \( \hat{\theta} - \theta_0 \sim \frac{1}{n} N(0, \sigma^2) \) as \( n \to \infty \).

(need suitable conditions on Fisher information matrix.)

**Problem 2.** Given data \( X^{(m)}_{t \in [0,T]} \) \( m = 1 \): sample paths of \( dX_t = b(X_t) \, dt + \sigma(X_t) \, dB_t \).

To do: estimate \( b \).

- If \( \delta \equiv b \), then \( \dot{X}_t = b(X_t) \) \( \delta(b) = \sqrt{\|b(X_t)\|^2} \).
- A "pdf" for \( X^{(m)}_{t \in [0,T]} \) (and then take log)? \( p(X^{(m)}_{t \in [0,T]}) \) \( \mathcal{N} \).

- Discrete-time approximation? \( X_t, \ldots, X_{t+1} \)

**Markov**

\[
p(X_t, \ldots, X_{t+1}) = p(X_{t+1} | X_t) \cdot p(X_t | X_{t-1}) \cdot \cdots \cdot p(X_0)
\]

**EM**

\[
x_{t+1} \approx X_t + \delta \, b(X_t) + \sigma \, \text{G}(X_t) \, W_t
\]

\[
p(X_{t+1} | X_t) \approx N(X_t + \delta \, b(X_t), \sigma^2 \text{G}(X_t))
\]

\[
\log p(X_t, \ldots, X_{t+1}) \approx \sum_{t=1}^{T} \left[ - \frac{1}{2} \log(\text{det} \Sigma) - \frac{1}{2} \frac{(X_{t+1} - X_t - \delta \, b(X_t))^2}{\sigma^2 \text{G}(X_t)} \right]
\]

\[
\int_0^T \left[ (\frac{dX_t}{dt})^2 + b(X_t)^2 - 2 \frac{b(X_t) \, dB_t}{dt} \right] \, dt \quad \text{why ok?} \quad \sigma = 1
\]

\[
= \int_0^T b(X_t)^2 \, dt - 2 \int_0^T b(X_t) \, dB_t
\]

\[
= E(b) \quad \text{a function (functional) of } b \quad \sqrt{}
\]

But \( \frac{dB_t}{dt} \) D.N.E. a.s. ! !!
Example \( \text{OU} : \) 
\[ dX_t = \theta (X_t - \mu) dt + dB_t \]
\[ X_{t+\Delta t} = e^{\theta \Delta t} X_t + \sqrt{\frac{\theta}{2}} \left( e^{\theta \Delta t} - 1 \right) dB_{t+\Delta t} \]

\[ Y_{t+1} = \left( 1 + \theta \Delta t \right) Y_t + \sqrt{\Delta t} N(0,1) \]
\[ \mathbf{b} = \frac{\partial}{\partial Y_t} \mathbb{E} [ -\log f(Y_{t+1}) ] \]
\[ b(\theta, X_t; t) = \frac{\partial}{\partial Y_t} \mathbb{E} [ -\log f(Y_{t+1}) ] \]
\[ \Rightarrow \theta = \frac{\mathbf{b}}{\mathbf{b}^T \mathbf{b}} = \frac{\mathbf{b}}{\mathbf{b}^T \mathbf{b}} \]
\( \theta > 0 \) \( \Rightarrow \mathbf{b} \sim \mathcal{N} \)
\[ \mathbb{E} \left[ X_t^2 \right] = \mathbb{E} \left[ e^{\theta X_t} \right] = \frac{e^{\theta^2}}{\theta^2} = \theta + O(\Delta t^2) \]

\[ \lim_{\Delta t \to 0} \int_0^T \frac{dx_t - \theta \Delta t}{\Delta t} = \int_0^T \frac{dx_t - \theta X_t}{\Delta t} dt = 0 \]

Q) A change of measure.
Recall pdf of \( X \):
\[ \mathcal{L}(x) = \frac{dP_n(x)}{dx} \rightarrow \text{Lebesgue measure on } \mathbb{R}^n \]
We can also use any other measure \( P_c(dx) = u_0(x) dx \), with \( u_0 \) known
\[ \mathcal{L}(x) = \frac{dP_n(x)}{dP_c(x)} = \frac{dP_n(x)}{u_0(x)} \]

- Back to the process \( X_{t\in[0,T]} \) (AR1)
\[ \mathcal{L}(x) = \frac{dP_n(x)}{dP_c(x)} = \mathcal{N}(X_t, \sigma^2) \]
\[ \Rightarrow \mathcal{L}(x) \sim \mathcal{N}(X_t, \sigma^2) \]
\[ \mathbb{E} \left[ X_t^2 \right] = \mathbb{E} \left[ e^{\theta X_t} \right] = \frac{e^{\theta^2}}{\theta^2} = \theta + O(\Delta t^2) \]

- Question: what is the limit of the measure \( P_c(X_t; x) \) as \( \Delta t \to 0 \)?

- Question 2: can we use other measure? Yes, any reference measure \( P \), \( P_c \ll P \)

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

- Gaussian process \( X_{t\in[0,T]} \) with white noise:
\[ X_{t\in[0,T]} \sim \mathcal{N}(0, \sigma^2) \]
- Brown motion at discrete times.
From previous examples:

\[ \frac{d\theta}{dp} \left| \frac{1}{F_T} \right| \]

MIE likelihood ratio

Importance sampling

Time series classification

- When the ratio exists? \( \Rightarrow \) Random walk
  \( \Rightarrow \) (IP, + changing)
- What is the essential? \( \Rightarrow \) Change of mean
- Can \( \beta \) be different? \( \checkmark \) doesn't matter.
  \( \beta_1 \) & \( \beta_2 \) \( \Rightarrow \) No.
- Other applications? \( \Rightarrow \) weak side.

\[ \mathbb{E}_x \left[ (X_{\text{out}}) \right] = \exp \left( \int_0^T \left( \frac{d}{dt} X_t - \frac{1}{2} \left( \sigma^2 + \left( \mu - \frac{\gamma}{\beta} \right)^2 \right) \right) dt \]

How to get it from the 3 Thms?
3.8 Girsanov Theorem

Girsanov’s 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,

\[
Q(A) = E(1_A L)
\]

defines a new probability. In fact, \( Q \) is a \( \sigma \)-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,

\[
P(A) = 0 \implies Q(A) = 0.
\]

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’s theorem.

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

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

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:

\[
E_Q(e^{itX}) = E(e^{itX}L) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2} - \frac{m^2}{2\sigma^2} + itx} dx
\]

\[
= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{e^2}{2\sigma^2} + itx} dx = e^{-\frac{e^2}{2}},
\]

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

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

\[
L_t = \exp \left( -\lambda B_t - \frac{\lambda^2}{2} t \right).
\]
2. Girsanov Theorem (theory)

17. a Levy characterization of Bm

22. absolute continuity of measures

33. Girsanov formulas

David Nualart's note:

Theorem 8.6.1 (The Levy characterization of Bm)

Let \( X=(X_t), t \in \mathbb{R}_+ \) be a cta process on \((\Omega, \mathcal{F}, \mathbb{P})\) with values in \(\mathbb{R}^n\). Then

(a) \( X_t \) is a Bm w.r.t. \( \mathbb{P} \) \iff

(b) \( X_t \) is an mG w.r.t. \( \mathbb{Q} \) and its own filtration

(i) \( X_t \) is a mG w.r.t. \( \mathbb{Q} \) & \( F^X_t \), \( \forall \mathcal{F}_t \in \{\ldots, \mathcal{F}_t \} \)

Lemma 8.6.2 (Bayes' rule)

Let \( \mu \) & \( \nu \) be two prob. measures on \((\Omega, \mathcal{F})\) s.t. \( d\nu = \mathbb{E}[f] d\mu, \quad f \in \mathcal{L}^1(\mu) \).

Let \( X \in L^1(\nu) \), \( H \in \mathcal{G} \). Then

\[
\mathbb{E}_\nu[X|H] \mathbb{E}_\nu[f|H] = \mathbb{E}_\nu[fX|H]
\]

Proof: \( \forall H \in \mathcal{G} \), \( \mathbb{E}_\nu[fX|H] = \int_H \mathbb{E}_\nu[fX|\mathcal{F}_t] d\nu w = \int_H \mathbb{E}_\nu[fX|\mathcal{F}_t] d\nu w = \mathbb{E}_\nu[\mathbb{E}_\nu[fX|H] 1_H]
\]

\[
= \mathbb{E}_\nu[\mathbb{E}_\nu[fX|H] 1_H] \quad \forall H \in \mathcal{G} \Rightarrow (1) \text{ by def.}
\]

27. Absolute continuity of measures

\( \Omega, \mathcal{F}, \mathbb{P} \) a prob. space. \( L^0 \) a field. \( \mathbb{Q} \) another prob. measure on \( \mathcal{F} \).

\( \mathbb{Q} \) is absolutely continuous wrt. \( \mathbb{P} \) if \( \mathbb{Q}(H)=0 \Rightarrow \mathbb{P}(H)=0, \forall H \in \mathcal{F} \). "\( \mathbb{Q} \ll \mathbb{P} \)"

Radon-Nikodym Theorem:

\( \mathbb{Q} \ll \mathbb{P} \iff \exists \mathcal{F}-measurable r.v. Z \in \mathcal{L}_1^+ \text{ s.t. } \mathbb{Q}(H) = \int_H Z d\mathbb{P} \)

\[
\frac{d\mathbb{Q}}{d\mathbb{P}} = Z \quad \text{on } \mathcal{F} \quad \text{Radon-Nikodym derivative.}
\]

Lemma 8.6.3. Sps. \( \mathbb{Q} \ll \mathbb{P} \) on \( \mathcal{F} \) with \( \frac{d\mathbb{Q}}{d\mathbb{P}} = Z \). Then \( \mathbb{Q} \mathbb{E}_\mathbb{F}(\mathcal{F}_t) \ll \mathbb{P} \) on \( \mathcal{F}_t \), and

\[
Z_t = \frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} \text{ as a martingale wrt. } \mathcal{F}_t \text{ and } \mathbb{P}.
\]

Proof: \( \mathbb{F}_t \in \mathcal{F}_t \Rightarrow \mathbb{Q} \ll \mathbb{P} \text{ on } \mathcal{F}_t \); \( \forall \mathcal{F}_t \subseteq \mathcal{F}_t \), \( \mathbb{E}_\mathbb{F}[1_{\mathcal{F}_t} 1_{\mathcal{F}_t}] = \mathbb{E}_\mathbb{F}[1_{\mathcal{F}_t}] = \mathbb{E}_\mathbb{F}[1_{\mathcal{F}_t}]
\]
Girsanov Theorem I
\[ dY_t = a(t, \omega) dW_t + dB_t, \quad Y_0 = 0, \quad B_t : \mathbb{R}^n - \text{valued Bm} \]
Then, \( \{Y_t\}_{t \geq 0} \) is a Bm w.r.t. \( \mathbb{Q} \) s.t. \( \frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{M}_T \) if
\[ \mathcal{M}_T = \exp \left( -\int_0^T a(t, \omega) dB_t - \frac{1}{2} \int_0^T a(t, \omega)^2 ds \right) \] is a m. G w.r.t. \( \mathbb{F}_T \) & \( \mathbb{P} \).

[Proof: see next page.]

A sufficient condition (Nowak):
\[ \mathbb{E}_{\mathbb{P}} \exp \left( \frac{1}{2} \int_0^T a(t, \omega)^2 ds \right) < \infty. \]

So if \( \mathcal{M}_T(\omega) > 0 \) a.s., we have \( \mathbb{P} \ll \mathbb{Q} \) \( \mathbb{Q} \) a.e. \( \Rightarrow \mathbb{P} \) & \( \mathbb{Q} \) are equivalent i.e.

\[ P(\{Y_t, \alpha \in \mathcal{F}_t, \ldots, Y_T, \alpha \in \mathcal{F}_T\} > 0) \equiv Q(\{Y_t, \alpha \in \mathcal{F}_t, \ldots, Y_T, \alpha \in \mathcal{F}_T\} > 0) \equiv P(B_t, \alpha \in \mathcal{F}_t, \ldots, B_T, \alpha \in \mathcal{F}_T) > 0. \]

Example
\[ a(t, \omega) = a(t) \text{ deterministic.} \]

Girsanov Theorem II
\[ dY_t = \beta(t, \omega) dt + a(t, \omega) dB_t, \quad \beta \in \mathbb{R}^n, \quad a \in \mathbb{R}^{n \times n}, \quad B_t : \mathbb{R}^n - \text{Bm} \]
If \( \exists \) \( u(t, \omega) \in W^{1,1}_T \) and \( v(t, \omega) \in W^{1,1}_T \) st. \( \Theta(t, \omega) u(t, \omega) = \beta(t, \omega) - a(t, \omega) \frac{dW_t}{d\mathbb{P}} \).

Assume
\[ \mathcal{M}_T = \exp \left( -\int_0^T a(t, \omega) dB_t - \frac{1}{2} \int_0^T u(t, \omega)^2 ds \right) \] is a m. G w.r.t. \( \mathbb{F}_T^B \) & \( \mathbb{P} \), then
\[ \mathcal{B}_t = \int_0^t u(s, \omega) ds + B_t \] is a Bm w.r.t. \( \mathbb{Q} \) s.t. \( \frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{M}_T \) on \( \mathbb{F}_T^B \).

and
\[ dY_t = a(t, \omega) dB_t + \Theta(t, \omega) d\mathcal{B}_t. \]

Proof: Similar to Theorem I, \( \mathcal{Q} \) is a prob. meas. on \( \mathbb{F}_T^B \) and \( \mathbb{B}_t : \mathbb{R}^n - \text{Bm w.r.t.} \mathbb{Q} \).

Girsanov III (for Ito diffusion \( X \) process)
\[ dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad t \in [0, T], \quad X_0 = x, \quad b \in \mathbb{R}^n, \quad \sigma \in \mathbb{R}^{n \times n} \text{ matrix-valued}, \quad W_t : \mathbb{R}^n - \text{Brownian}. \]

Assume \( X \in W^{1,1}_T \), \( \Theta(t, \omega) u(t, \omega) = \sigma(t, \omega) dW_t \).

\begin{align*}
\mathcal{M}_T & = \exp \left( -\int_0^T b(t, \omega) dt - \frac{1}{2} \int_0^T \sigma(t, \omega)^2 u(t, \omega)^2 dt \right) \text{ is a m. G w.r.t. \( \mathbb{F}_T^B \) & \( \mathbb{P} \).}
\end{align*}

Define \( \mathcal{B}_t, \mathcal{Q} \) & \( \mathbb{B}_t \) as above, and assume \( \mathcal{M}_T \) is a m. G w.r.t. \( \mathbb{F}_T^B \) & \( \mathbb{P} \).

Then \( \mathcal{Q} \) is a prob. meas. \( \frac{d\mathcal{Q}}{d\mathbb{P}} = \exp \left( \int_0^T -\frac{1}{2} \sigma(t, \omega)^2 u(t, \omega)^2 dt \right) \).

Then \( \mathcal{Q} \)-law of \( X^\mathcal{Q}_T \) is the same as the \( \mathbb{P} \)-law of \( X^\mathbb{P}_T \).

Proof: Direct application of Theorem II.
\[ \Theta(t, \omega) = \sigma(t, \omega), \quad \frac{d\mathbb{Q}}{d\mathbb{P}} = \exp \left( \int_0^T b(t, \omega) dt - \frac{1}{2} \int_0^T \sigma(t, \omega)^2 u(t, \omega)^2 dt \right). \]
Proof of Theorem 1:

Since $M_t$ is a $M_b$, we have $Q(w) = E_{\omega} 1 = E_p [ M_T ] = E_{\omega} [ M ] = 1 \Rightarrow Q$ is a probability measure.

WLOG, assume $a(s,w)$ is bold (otherwise consider a $N_b$ first, and then send it off).

In view of Lévy's characterization of $\text{BM}$, we need to verify that

(i) $Y_t = (Y_t(b), \ldots, Y_n(t))$ is a $M_b$ w.r.t. $Q$.

(ii) $Y_t(b)$ is $\delta_t$-additive w.r.t. $Q$, for $a \neq b$.

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

$$d K_t = M_t d Y_t(t) + Y_t(t) d M_t + M_t d Y_t(t)$$

$$= M_t (d X_t + d B(t)) + Y_t(t) M_t (-a d B(t)) + M_t Y_t(t) dB(t)$$

$$= M_t (d X_t - Y_t(t) a dt) + M_t Y_t(t) dB(t)$$

$$= dB_t(t) - Y_t(t) a dt + M_t Y_t(t) dB(t)$$

Hence, $K_t(t)$ is a $M_b$ w.r.t. $P$, so by the Lemma (Bayes rule), we get for $a \neq b$

$$E_{Q} [ Y_t(b) | F_s ] = \frac{E_p [ M_t Y_t(b) | F_s ]}{E_p [ M_t | F_s ]} = \frac{E_p [ K_t(t) | F_s ]}{E_p [ M_t | F_s ]} = \frac{K_s(t)}{M_s} = Y_s(t),$$

i.e., $Y_t(b)$ is $\delta_t$-additive w.r.t. $Q$. This proves (ii).

The proof of (ii) is similar.
Example 8.6.7 Let \( dY_t = \left( \begin{array}{c} \frac{2}{4} \ dt + \left( \begin{array}{c} 1 \ 1 \\ 0 \end{array} \right) \left( \frac{dB_t}{dt} \right) \end{array} \right) \), \( dY_t = \beta \ dt + \sigma \ dB_t \).

Let \( d(t, w) = 0 \) in Thm II, \( \theta U = \beta \) \( \left( \begin{array}{c} 1 \ 1 \\ 0 \end{array} \right) \left( \begin{array}{c} U_t \ U_t \end{array} \right) = \left( \begin{array}{c} \frac{2}{4} \ \frac{3}{4} \end{array} \right) \), \( \left( \begin{array}{c} U_t \ U_t \end{array} \right) = \left( \begin{array}{c} -1 \ 1 \end{array} \right) \)

Then, for \( \alpha \):

\[
\frac{d\beta t}{dp} = e^{-3B_t} + B_t - 5t
\]

\[
M_t = \exp \left(-\int_0^t u \ dB_t - \frac{1}{2} \int_0^t u^2 ds \right)
\]

\[
d\beta t = \left( \begin{array}{c} -1 \ \frac{3}{4} \end{array} \right) dt + dB_t
\]

we have \( \beta_t \) is a BM w.r.t. \( \alpha \), and \( dY_t = \left( \begin{array}{c} 1 \ 1 \\ 0 \end{array} \right) \frac{dB_t}{dt} \).

Application of Thm III: weak soln to SDE.

Let \( Y_t \) be a known weak or strong soln. to

\[
dY_t = b(Y_t) \ dt + \sigma(Y_t) \ dB_t \hspace{1cm} b, \ a: \mathbb{R}^n \rightarrow \mathbb{R}^n
\]

We wish to find a weak soln. to

\[
d\beta t = a(\beta_t) \ dt + \sigma(\beta_t) \ dB_t
\]

Assume \( \exists \ U_t (\mathbb{R}^n \rightarrow \mathbb{R}^n) \) s.t. \( \sigma(Y_t) U_t(Y_t) = b(Y_t) - a(Y_t) \), \( b, \ y \in \mathbb{R}^n \).

\( u_o = \sigma^{-1}(b-a) \) \( \sigma^{-1} \in \mathbb{R}^n \)

Then if \( u(Y_t) \) satifies the Novikov condition,

\[
\mathbb{E}[ \exp \left( \frac{1}{2} \int_0^t u^2 ds \right) ] < \infty
\]

we have \( \beta_t = \int_0^t u(s, w) ds + \beta_t \) a BM w.r.t. \( \beta \)

and

\[
\frac{d\beta_t}{d\beta} = M_t = \exp \left(-\int_0^t u \ dB_t - \frac{1}{2} \int_0^t u^2 ds \right)
\]

That is, \( (Y_t, \beta_t) \) is a weak soln to (e).

Example: To construct a weak soln. to

\[
\frac{dX_t}{d\beta} = a(X_t) \ dt + \sigma \ dB_t \hspace{1cm} X_0 = \xi \in \mathbb{R}^n \hspace{1cm} \mathbb{R} \in \mathbb{R}
\]

Start from \( d\beta_t = \sigma \ dB_t \); \( \beta_t = \beta \).

For \( T < \tau \) and put

\[
\frac{d\beta_t}{d\beta} = M_T = \exp \left( + \int_0^B \sigma \ dt + \int_0^B \sigma \ dB_t \right)
\]

Then,

\[
\beta_t = -\int_0^t \sigma \ dB_t + \beta_t \text{ is a BM w.r.t. } \beta
\]

and

\[
d\beta_t = \sigma \ dB_t = a(X_t) ds + \sigma \ dB_t
\]

That is, \( (Y_t, \beta_t) \) is a weak soln to (e).
Application: likelihood of data $X_{[0,T]}: X \sim X^a$.

\[
\frac{dP_a}{dP_0} = \exp \left( \int_0^T \frac{a(x_s)}{\sigma x_s} dx_s - \frac{1}{2} \int_0^T \frac{a(x_s)^2}{\sigma^2 x_s^2} dx_s \right).
\]

\[
x \sim a(x_s) dx_s + \sigma dB_s \rightarrow \mathbb{P}_a
\]

\[
x \sim a(x_s) dx_s \rightarrow \mathbb{P}_0
\]

\[
x^b \sim b(x_s) dx_s + \sigma dB_s \rightarrow \mathbb{P}_b
\]

\[
\frac{dP_a}{dP_b} = \exp \left( \int_0^T \frac{a(x_s)}{\sigma b(x_s)} dx_s - \frac{1}{2} \int_0^T \frac{a(x_s)^2}{\sigma^2 b(x_s)^2} dx_s \right).
\]

\(\sigma = 1\).

\(\nabla \phi = 0\).

\[
\frac{dP_a}{dP_b} \cdot \frac{dP_b}{dP_0} \cdot \frac{dP_0}{dP_a} = 1
\]

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

\[
L_{X_{[0,T]}(a)} = \log \frac{dP_a}{dP_0} = \int_0^T a(x_s) dx_s - \frac{1}{2} \int_0^T a(x_s)^2 dx_s
\]

Setting 1: Data: multi-trajectory, $X_{[0,T]}$, $N$.

Setting 2: Ergodic, $T \to \infty$.

\[L_T(a) = \frac{1}{N} \sum_{i=1}^{N} L_{X_{[0,T]}(a)} \rightarrow \mathbb{E} \left[ \int_0^T a(x_s) dx_s - \frac{1}{2} \int_0^T a(x_s)^2 dx_s \right] = L_T(a)
\]

\[\hat{a} = \arg \max_{a \in \mathbb{R}} L_T(a), \quad \Rightarrow \quad \mathbb{E} \left[ \int_0^T a(x_s) b(x_s) - \frac{1}{2} a(x_s)^2 dx_s \right] = 0
\]

Then

\[L_M(a) = \frac{1}{N} \sum_{i=1}^{N} \int_0^T \frac{1}{2} \bar{a} \sigma^2 g(x_s) dx_s
\]

\[-\frac{1}{2} \log \mathbb{E}[g(x_T)]
\]

\[b^c = -\frac{1}{2} \mathbb{E}[\phi(x)] \]

\[\Rightarrow \quad D_M(a) = \hat{a}^c - b
\]

\[c = A^T b \quad \in \mathbb{R}^n
\]

$\theta$ is a invertible?

$\theta \alpha \rightarrow ?$

$\Rightarrow \quad a = b$ to the left in $L_T$.
Inference and nonparametric regression, ML.

\[ f_{\theta}(x) \rightarrow Y = f(X) + \epsilon. \]

\[ \hat{f}_n = \arg\min_{f \in H_n} \sum_{i=1}^n (Y_i - f(X_i))^2 \]

\[ \implies \mathbb{E}_n(f) = \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 \]

\[ \nabla \mathbb{E}_n(f) = \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i)) \frac{X_i}{\mathbb{E}_n(f)} \]

\[ \mathbb{E}_n(f) = \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 \]

\[ \nabla \mathbb{E}_n(f) = \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i)) \frac{X_i}{\mathbb{E}_n(f)} \]

\[ \implies \hat{f}_n = \mathbb{E}_n(f) \]

Back to the context of SDE.

Inference for interacting particle system.

\[ dX_t = \frac{d}{dt} \log f(X_t) dt + \sigma dB_t. \]

Given \( \{X_t^{(i)}\}_{i=1}^n \), to estimate \( f : \mathbb{R}^d \rightarrow \mathbb{R} \)

\[ dX_t = f(X_t) dt + \sigma dB_t \]

\[ \mathbb{E}(Y_{t0} | \hat{f}) = -\int f(x) d\mu(x) + \frac{1}{2} \int [f(x)]^2 dt \quad \text{(negative log)} \]

\[ \mathbb{E}_n(\hat{f}) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_t^{(i)} | \hat{f}) \quad \rightarrow \quad \mathbb{E}_n(\hat{f}) = \frac{1}{n} \sum_{i=1}^n \left[ \int f(x) d\mu(x) - \frac{1}{2} \int [f(x)]^2 dt \right] \]

\[ = \langle \hat{f}, \mu \rangle - \frac{1}{2} \langle \hat{f}, \mu \rangle + \hat{c} \]