Statistical Inference for Diffusion Processes

Radu Herbei

The Ohio State University
The Mathematical Biosciences Institute

July 7, 2015

Motivation

Let $x(t)$ be the *state of a system* at time $t \geq 0$. Assume that the time evolution of $x(\cdot)$ can be described via

$$\begin{cases} \frac{d}{dt}x(t) = b(x(t)), & \text{for } t > 0 \\ x(0) = x_0 \end{cases} \tag{1}$$

where $b(\cdot)$ is a given, smooth function. Under conditions which will not be discussed here, the problem above can be *solved*, i.e., one can find a function $x(t)$ satisfying (1). This function is necessarily smooth and its graph may take the following form.

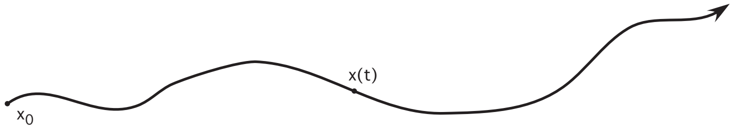


Figure: Trajectory of a solution $x(\cdot)$.

In many cases, one can obtain measurements of the variable x (at many time points). When plotted against time, trajectories behave as follows:

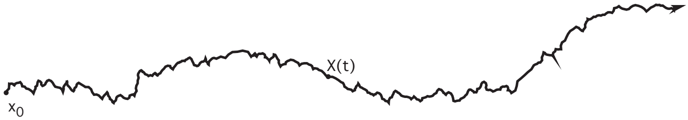


Figure: Trajectory of a "measured" solution $X(\cdot)$.

Note that

- We are plotting observations X, not the variable x;
- There are many dissimilarities between the two graphs;
- There are many similarities between the two graphs;

- Our goal is to **understand** how X changes in time, accounting for various sources of uncertainty: measurement error, approximate dynamics, etc.
- Why ? Ultimately we would like to predict the value of the system at a a future time point, or a spatial location of interest. For the time being, we will ignore the fact that X may contain measurement error – this can be dealt with later.
- Clearly x and X are different and one cannot use (1) to describe how X behaves in time.
- On the other hand one can observe that the evolution of X is very similar to that of x, which indicates that

  $$\frac{d}{dt}X(t) = b(X(t))$$

  is a "good" place to start in describing how X changes in time.

- The little wiggles that appear in the graph of X can be thought of as "noise" - something that we cannot explain, but something that doesn't seem to change the overall dynamics.
- This suggests the following modification

$$\begin{cases} \frac{d}{dt}X(t) = b(X(t)) + \text{"noise"}, & \text{for } t > 0 \\ X(0) = X_0 \end{cases}$$

Questions:

- define "noise" in a rigorous way; define what it means for $X(\cdot)$ to solve the system above;
- discuss uniqueness, asymptotic behavior, dependence upon X_0, $b(\cdot)$, etc

These questions are addressed by the classical SDE theory. In many cases, $b(\cdot)$ is also **unknown**. This raises some additional questions:

- estimate b (parametric, non-parametric, Bayes, etc.);
- if "noise" involves parameters, estimate those too;
- what statistical properties do all the estimators have ? (consistency, asymptotics);
- computational issues

Outline (part I)

- Primer on stochastic processes
- Brownian Motion
- Stochastic integrals
- Itô proceses, stochastic differential equations, Itô formula
- Solutions of diffusion processes
- Girsanov formula

References

- [L&S] Statistics of Stochastic Processes I and II by R.S. Lipster, A.N. Shiryaev and B. Aries, Springer, 2000;
- [K&S] Brownian Motion and Stochastic Calculus by I. Karatzas and S. Shreve, Springer, 1991;

Probability spaces, random variables

- Let $(\Omega, \mathcal{B}, \mathrm{P})$ be a probability space.
  - $\Omega \neq \emptyset$ is the sample space;
  - $\mathcal{B} \subseteq 2^{\Omega}$ is a σ-field (its elements are called events);
  - $\mathrm{P} : \mathcal{B} \to [0, 1]$ is a probability measure.
- A Borel-measurable map $X : \Omega \to \mathbb{R}^k$ is called a random vector (or variable, if $k = 1$). In general, a Borel-measurable map $X : \Omega \to \mathbb{D}$ is called a random element (of $\mathbb{D}$). Here $\mathbb{D}$ is a generic metric space.
- The law of X or, the distribution of X is the probability measure $\mathrm{P}X^{-1} : \mathcal{B}(\mathbb{D}) \to [0, 1]$

$$
\begin{aligned}
\mathbb{P}X^{-1}(B) &= \mathrm{P}(X^{-1}(B)) \\
&= \mathrm{P}(\{\omega \in \Omega \ : \ X(\omega) \in B\}) \quad \forall \ \ B \in \mathcal{B}(\mathbb{D})
\end{aligned}
$$

Stochastic Processes

View 1: A collection of random variables $\{X_t,\ t \in \mathcal{T}\}$.
Typically $\mathcal{T} = [0, \infty)$ or $\mathcal{T} = [0, T]$.

$$X_t : \Omega \to \mathbb{R} \qquad t \in \mathcal{T}$$

For each $\omega \in \Omega$, the map

$$t \mapsto X_t(\omega) \qquad t \in \mathcal{T}$$

is called a sample path.

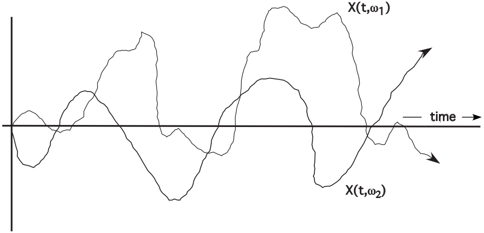


Figure: Two sample paths of a stochastic process.

View 2: A map

$$X : \mathcal{T} \times \Omega \to \mathbb{R} \qquad (t, \omega) \mapsto X(t, \omega) \equiv X_t(\omega)$$

View 3: A map

$$X : \Omega \to \mathbb{R}^{\mathcal{T}} \qquad \omega \mapsto X_\omega \quad \text{where} \quad X_\omega : \mathcal{T} \to \mathbb{R}$$

Unless otherwise specified, we will assume that $\mathcal{T} = [0, T]$.

- A family of σ-fields $(\mathcal{F}_t)$, $t \in \mathcal{T}$ such that $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ if $t_1 < t_2$ is called a filtration.
- A σ-field $\mathcal{F}_t$ is viewed as "information". Thus, $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ can be interpreted as "information accumulates in time".
- The process $(X_t), \ t \in \mathcal{T}$ is adapted to the filtration $\mathcal{F}_t$ if

$$X_t \in \mathcal{F}_t / \mathcal{B}(\mathbb{R})$$

- The process $(X_t), \ t \in \mathcal{T}$ is measurable if the map

$$(t, \omega) \mapsto X(t, \omega) \quad t \in [0, T] \ \omega \in \Omega$$

  is measurable wrt the product σ-field $\mathcal{B}([0, T]) \times \mathcal{B}$.
- The process (X_t) is progressively measurable if , for each $t \in [0, T]$ the map

$$(s, \omega) \mapsto X(s, \omega) \quad s \in [0, t] \ \omega \in \Omega$$

  is measurable wrt the product σ-field $\mathcal{B}([0, t]) \times \mathcal{B}_t$.

Some classes of stochastic processes

Stationary processes.
The process $X^T = \{X_t,\ t \in \mathcal{T}\}$ is called stationary in a narrow sense if

$$P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) = P(X_{t_1+\delta} \in A_1, \ldots X_{t_n+\delta} \in A_n)$$

The process $X^T = \{X_t,\ t \in \mathcal{T}\}$ is called stationary in a wide sense if

$$E(X_t) < \infty \quad E(X_t) = E(X_{t+\delta}) \quad E(X_s X_t) = E(X_{s+\delta} X_{t+\delta})$$

The process X^T has independent increments if, for any $t_1 < t_1 < \cdots < t_n$, the increments

$$X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots, X(t_n) - X(t_{n-1})$$

are independent.

Markov processes

The stochastic process X^T is called Markov wrt the filtration $(\mathcal{F}_t)$ if

$$\mathrm{P}(A \cap B \,|\, X_t) = \mathrm{P}(A \,|\, X_t)\mathrm{P}(B \,|\, X_t)$$

for any $t \in \mathcal{T}, \quad A \in \mathcal{F}_t, \quad B \in \mathcal{F}_{[t,\infty)} \equiv \sigma(X_s, s \geq t)$.

Theorem(1.12, L&S)
The process X_t is Markov iff for each measurable function $f(x)$ with $\sup_x |f(x)| < \infty$ and any $0 \leq t_1 \leq \ldots, \leq t_n \leq t$,

$$\mathrm{E}(f(X_t) \mid X_{t_1}, \ldots, X_{t_n}) = \mathrm{E}(f(X_t) \mid X_{t_n})$$

Stochastic processes with independent increments are an important subclass of Markov processes.

Martingales

The stochastic process $(X_t),\ t \in \mathcal{T}$ is called a martingale with respect to the filtration $(\mathcal{F}_t)$ if $E(X_t) < \infty,\ t \in \mathcal{T}$ and

$$\mathrm{E}(X_t \,|\, \mathcal{F}_s) = X_s \quad \text{a.s.} \quad t \geq s.$$

Exercise Let $Y_1, Y_2, \ldots$ be such that
$(Y_1, Y_2, \ldots, Y_n) \sim p_n(y_1, \ldots, y_n)$ wrt λ.
Let $q_n(y_1, \ldots, y_n)$ be an *alternative* pdf (wrt λ). Then

$$X_n = \frac{q_n(Y_1, \ldots, Y_n)}{p_n(Y_1, \ldots, Y_n)}$$

is a martingale wrt $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$.

Brownian Motion (BM)

- discovered by Robert Brown (1828);
- first quantitative work on BM due to Bachelier (1900) – in the context of stock price fluctuations;
- Einstein (1905) derived the transition density for BM from molecular-kinetic theory of heat;
- Wiener (1923,1924) – first rigorous treatment of BM; first proof of existence;
- P. Lévy (1939, 1948) – most profound work (construction by interpolation, first passage times, more).

Definition of a BM

A real-valued continuous time stochastic process $\mathbb{W}^T = \{\mathbb{W}_t, t \geq 0\}$ is called a Brownian motion if

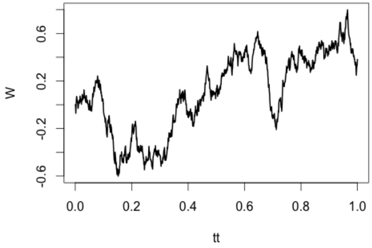
- $\mathbb{W}_0 = 0$ a.s.;
- $\mathbb{W}^T$ has stationary and independent increments;
- If $s < t$, $\mathbb{W}_t - \mathbb{W}_s$ is a Gaussian variate with

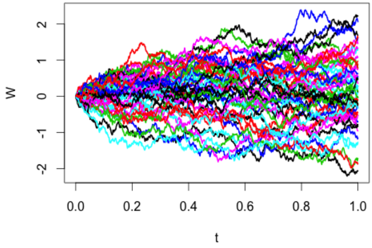
$$\mathrm{E}(\mathbb{W}_t - \mathbb{W}_s) = 0 \quad \mathrm{Var}(\mathbb{W}_t - \mathbb{W}_s) = \sigma^2(t - s)$$

- For almost all $\omega \in \Omega$, the sample path $t \mapsto \mathbb{W}_t(\omega)$ is a continuous function of $t \geq 0$

If $\sigma = 1$ the process $(\mathbb{W}_t)$ is called a standard BM.

Simulation





Properties

Let $\mathbb{W}^T$ be a SBM.

- The natural filtration generated by a BM process is

$$\mathcal{F}_t = \sigma\big(\mathbb{W}_s,\ 0 \le s \le t\big)$$

- $\mathrm{E}(\mathbb{W}_t) = 0$, $\mathrm{Var}(\mathbb{W}_t) = t$
- SBM is a martingale wrt $(\mathcal{F}_t)$
- Independent increments $\Rightarrow$ Markov process.

Exercise Let $t_1 < t_2 < \cdots < t_n$. Derive the joint distribution of $(\mathbb{W}(t_1), \mathbb{W}(t_2), \ldots, \mathbb{W}(t_n))$.

Existence

Constructive method.

Let $\eta_1, \eta_2, \ldots$ be iid $N(0,1)$ variates and $\phi_1(t), \phi_2(t), \ldots$ be an arbitrary complete orthonormal sequence in $L_2[0, T]$. Define

$$\Phi_j(t) = \int_0^t \phi_j(s)\, ds \quad j = 1, 2, \ldots$$

Theorem. The series

$$\mathbb{W}_t = \sum_{j=1}^{\infty} \eta_j \Phi_j(t)$$

converges P-a.s. and defines a Brownian motion process on $[0, T]$.

Brownian motion as a limit of a random walk

Let $X_n = \pm 1$ with probability $1/2$ and consider the partial sum

$$S_n = X_1 + X_2 + \cdots + X_n \ .$$

Then, as $n \to \infty$,

$$\mathrm{P}\Big(\frac{S_{[nt]}}{\sqrt{n}} < x\Big) \to \mathrm{P}(\mathbb{W}_t < x)$$

(discussion)

Strong Markov property

Let τ be a Markov time wrt $\mathcal{F}_t$, assume that $P(\tau \leq T) = 1$.
Fix s such that $\mathrm{P}(s + \tau \leq T) = 1$.

$$\mathrm{E}(f(\mathbb{W}_{\tau+s}) \,|\, \mathcal{F}_\tau) = \mathrm{E}(f(\mathbb{W}_{\tau+s}) \,|\, \mathbb{W}_\tau)$$

This is equivalent to saying that

$$\widetilde{\mathbb{W}}_t = \mathbb{W}_{\tau+t} - \mathbb{W}_\tau$$

is a SBM, independent of $\mathcal{F}_\tau$.

Reflection principle

Let $\mathbb{W}^T$ be a SBM and τ a Markov time. The process

$$\mathbb{W}^*(t) = \begin{cases} \mathbb{W}_t & \text{if } t \leq \tau \\ \mathbb{W}_\tau - (\mathbb{W}_t - \mathbb{W}_\tau) & \text{if } t \geq \tau \end{cases}$$

is a SBM.

Let $\tau = \inf\{t \geq 0, \ \mathbb{W}_t \geq x\}$ where $x > 0$, and let

$$M_t = \sup_{0 \leq s \leq t} \mathbb{W}_s$$

Then,

$$\mathrm{P}(M_t \geq x) = \mathrm{P}(\tau \leq t) = 2\mathrm{P}(\mathbb{W}_t \geq x)$$

Stochastic Integral

Let $(\Omega, \mathcal{B}, \mathrm{P})$ be a prob. space, $\mathbb{W}^T$ be a SBM.
The quadratic variation (on $[0, T]$) is defined as

$$[\mathbb{W}_T, \mathbb{W}_T] = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} |\mathbb{W}_{t_{i+1}} - \mathbb{W}_{t_i}|^2$$

where $\Pi = (0 = t_0 < t_1 < \cdots < t_n = T)$ is a partition of $[0, T]$.

Lemma. The quadratic variation of a Brownian motion is

$$[\mathbb{W}_T, \mathbb{W}_T] = T \quad \text{a.s.}$$

Differential forms (stochastic calculus)

Recall that $[\mathbb{W}_T, \mathbb{W}_T] = T \quad \text{a.s.}$. In short, we write that

$$d\mathbb{W}_t \; d\mathbb{W}_t = dt$$

It can also be shown that

$$dt \; d\mathbb{W}_t = 0 \quad \text{and } dt \; dt = 0$$

Higher order variations are all equal to zero.

Stochastic integrals

Let X^T be a stochastic process (random function). Define

$$\mathcal{M}_T = \Big\{ X^T - \text{ prog. meas. } : \mathsf{P}\Big(\int_0^T X^2(t,\omega)dt < \infty \Big) = 1 \Big\}$$

This is the class of all progressively measurable functions which are square integrable a.s. Also, define

$$\mathcal{M}_T^2 = \Big\{ X^T \in \mathcal{M}_T \ : \ \mathsf{E}\Big(\int_0^T X^2(t,\omega)dt \Big) < \infty \Big\}$$

Consider $h \in \mathcal{M}_T^2$ and $\mathbb{W}^T$ – Brownian motion. We aim to define the Itô integral

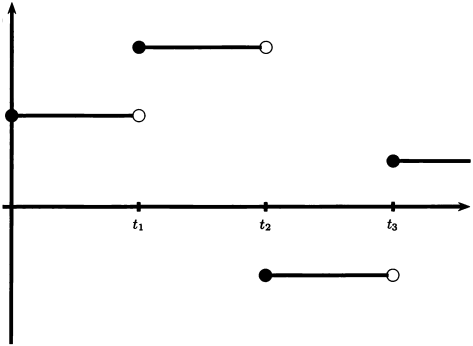
$$I_T(h) = \int_0^T h(s,\omega) d\mathbb{W}_s$$

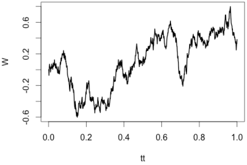
Case 1: h is a simple function.

$$h : [0, T] \times \Omega \to \mathbb{R} \qquad (t, \omega) \mapsto h(t, \omega)$$

Assume that there exists $0 = t_0 < t_1 < \cdots < t_n = T$ such that

$$h(t) = h_i \qquad \text{if} \qquad t \in [t_i, t_{i+1})$$





The Itô integral $I_T(h)$ is defined as

$$\begin{aligned} I_T(h) &= \int_0^T h(t,\omega)\, d\mathbb{W}_t \\ &= h_0(\mathbb{W}_{t_1} - \mathbb{W}_{t_0}) + h_1(\mathbb{W}_{t_2} - \mathbb{W}_{t_1}) + \cdots + h_{n-1}(\mathbb{W}_{t_n} - \mathbb{W}_{t_{n-1}}) \\ &= \sum_{i=0}^{n-1} h_i(\mathbb{W}_{t_{i+1}} - \mathbb{W}_{t_i}) \end{aligned}$$

Properties of the Itô integral

- $I_T(h)$ is a martingale. That is,

$$E(I_T(h)) = 0, \quad E\big(I_T(h) \mid \mathcal{F}_t\big) = I_t(h), \quad , t < T \ ,$$

where $\mathcal{F}_t = \sigma(\mathbb{W}_u, \ 0 \le u \le t)$.

- For any simple functions $h, g \in \mathcal{M}_T^2$,

$$E\Big(I_T(h) \cdot I_T(g)\Big) = E\Big(\int_0^T h(t,\omega)g(t,\omega)dt\Big)$$

thus,

$$E\Big(I_T(h)^2\Big) = E\Big(\int_0^T h(t,\omega)^2 dt\Big)$$

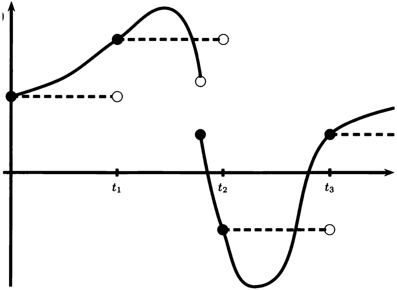
- The quadratic variation is

$$[I_T(h), I_T(h)] = \int_0^T h(t,\omega)^2 dt$$

Case 2: General h.

Lemma There exists a sequence of simple random functions h_n such that

$$\int_0^T |h_n(t,\omega) - h(t,\omega)|^2 dt \xrightarrow{\mathrm{P}} 0 \quad \text{as } n \to \infty$$



The stochastic integral $I_T(h)$ is defined as the limit

$$\int_0^T h_n(t,\omega)d\mathbb{W}_t \xrightarrow{\text{P}} I_T(h) = \int_0^T h(t,\omega)d\mathbb{W}_t \quad \text{as } n \to \infty$$

Important observation: The Ito integral is defined as a limit of a Riemann-Stieltjes sum, where the intermediate points are taken to be the lower limits of the partition intervals.

$$\sum_{i=0}^{n-1} h(t_i)\big(\mathbb{W}_(t_{i+1}) - \mathbb{W}(t_i)\big) \xrightarrow{\text{P}} I_T(h) = \int_0^T h_t \, d\mathbb{W}_t$$

Properties

- As a function of t, the paths $I_t(h)$ are continuous;
- for each t, $I_t(h)$ is measurable wrt $\mathcal{F}_t$;
- $\alpha I_t(h) + \beta I_t(g) = I_t(\alpha h + \beta g)$
- $I_t(h)$ is a martingale;
- $$E(I_t^2(h)) = E \int_0^t h^2(s)\ ds$$
- $$[I, I](t) = \int_0^t h^2(s)\ ds$$

Differential form

$$I_t(h) = \int_0^t h(u,\omega) d\mathbb{W}_u \quad \Leftrightarrow \quad dI_t(h) = h(t,\omega)\ d\mathbb{W}_t$$

Exercise

Show that:

$$\int_0^T \mathbb{W}_t d\mathbb{W}_t = \frac{1}{2}\mathbb{W}_T^2 - \frac{1}{2}T$$

Itô processes

Let $h \in \mathcal{M}_T^2$ and g be such that $P\Big(\int_0^T |g(t,\omega)|dt < \infty\Big) = 1$.
The stochastic process

$$X_t = X_0 + \int_0^t g(s,\omega)ds + \int_0^t h(s,\omega)d\mathbb{W}_s$$

is called an Itô process.

In differential form,

$$dX_t = g(t,\omega)dt + h(t,\omega)d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0,T]$$

Examples

- The Ornstein-Uhlenbeck (OU) process is defined as

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

where $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_3 > 0$.

- The Geometric Brownian Motion (GBM) process is defined as

$$dX_t = \theta_1 X_t dt + \theta_2 X_t d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

where $\theta_1 \in \mathbb{R}$, $\theta_2 > 0$

- The Cox-Ingersoll-Ross (CIR) process is defined as

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

where $\theta_1, \theta_2 \in \mathbb{R}$, $\theta_3 > 0$.

Stochastic integral wrt an Itô process

As before, let $f \in \mathcal{M}_T^2$ and X^T be an Itô process defined via the SDE

$$dX_t = g(t,\omega)dt + h(t,\omega)d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

The Itô integral wrt X^T is defined as

$$\int_0^T f(t,\omega)dX_t = \int_0^T f(t,\omega)g(t,\omega)dt + \int_0^T f(t,\omega)h(t,\omega)d\mathbb{W}_t$$

Here we assume that all the above integrals are well defined.

Itô formula

The class of Itô processes is closed with respect to smooth transformations, in the following sense. Let X^T be an Itô process defined by

$$dX_t = g(t,\omega)dt + h(t,\omega)d\mathbb{W}_t \quad X(0) = X_0, t \in [0, T]$$

Also let $G(t,x)$ be a "smooth" function: the derivatives G_t, G_x, G_{xx} exist and are continuous. Then the stochastic process $Y_t = G(t, X_t)$ is an It^o process with the stochastic differential

$$\begin{aligned} dY_t &= \Big[G_t(t,X_t) + G_x(t,X_t)g(t,\omega) + \frac{1}{2}G_{xx}(t,X_t)h(t,\omega)^2\Big]dt + \\ &\quad + \Big[G_x(t,X_t)h(t,\omega)\Big]d\mathbb{W}_t \end{aligned}$$

or,

$$dY_t = \Big[G_t(t,X_t) + \frac{1}{2}G_{xx}(t,X_t)h(t,\omega)^2\Big]dt + \Big[G_x(t,X_t)\Big]dX_t$$

Application

Consider the OU process

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

and the transformation $Y(t) = X(t)e^{\theta_2 t}$.

Application

Consider the OU process

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 d\mathbb{W}_t \quad X(0) = X_0, \;\; t \in [0, T]$$

and the transformation $Y(t) = X(t)e^{\theta_2 t}$.

The solution is

$$X_t = X_0 e^{-\theta_2 t} + \frac{\theta_1}{\theta_2}\left(1 - e^{-\theta_2 t}\right) + \theta_3 \int_0^t e^{-\theta_2 (t-s)} d\mathbb{W}_s$$

Application

Consider the Geometric Brownian Motion (GBM)

$$dX_t = \theta_1 X_t dt + \theta_2 X_t d\mathbb{W}_t \quad X(0) = X_0, \;\; t \in [0, T]$$

and the transformation $Y(t) = \log(X(t))$.

Application

Consider the Geometric Brownian Motion (GBM)

$$dX_t = \theta_1 X_t dt + \theta_2 X_t d\mathbb{W}_t \quad X(0) = X_0,\ t \in [0, T]$$

and the transformation $Y(t) = \log(X(t))$.

The solution is

$$X_t = X_0 e^{(\theta_1 - (1/2)\theta_2^2)t + \theta_2 \mathbb{W}_t}$$

Another (important) application of Itô formula

$$dX_t = S(X_t)dt + \sigma(X_t)d\mathbb{W}_t \quad X(0) = X_0, \quad 0 \le t \le T$$

If $h(\cdot)$ is a smooth function, show that

$$\int_0^T h(X_s)\ dX_s = \int_{X_0}^{X_T} h(s)ds - \frac{1}{2}\int_0^T h'(X_s)\sigma(X_s)^2\ ds$$

Solution: Define $G(x) = \int_0^x h(s)ds$ and use Itô formula for $Y(t) = G(X(t))$.

Diffusion processes

A homogeneous diffusion process is a particular case of an Itô process, defined as the solution of the stochastic differential equation (SDE)

$$dX_t = S(X_t)dt + \sigma(X_t)d\mathbb{W}_t \quad X(0) = X_0, \quad 0 \leq t \leq T$$

Notation: $\{X_t, 0 \leq t \leq T\} \equiv X^T$.

The functions $S(\cdot)$ and $\sigma(\cdot)$ are called the *drift* and *diffusion* coefficients, respectively. In integral form, the process X^T is represented as

$$X_t = X_0 + \int_0^t S(X_u)du + \int_0^t \sigma(X_u)d\mathbb{W}_u, \quad 0 \leq t \leq T$$

Strong solution of an SDE

$$dX_t = S(X_t)dt + \sigma(X_t)d\mathbb{W}_t \ , \quad X(0) = X_0, \quad 0 \le t \le T$$

The SDE above has a strong solution $\{X_t, t \in [0, T]\}$ on $(\Omega, \mathcal{F}, \mathrm{P})$ wrt the Wiener process $\{\mathbb{W}_t, t \in [0, T]$ and initial condition X_0 if:

- $\{X_t, t \in [0, T]\}$ is adapted to $\mathcal{F}_t = \sigma(X_0, \mathbb{W}_u, u \in [0, t])$;
- $P(X(0) = X_0) = 1$;
- $\{X_t, t \in [0, T]\}$ has continuous sample paths;
- $$\mathrm{P}\Big\{ \int_0^T \big[S(X_t) + \sigma(X_t)^2\big] dt < \infty \Big\} = 1$$
- $$X_t = X_0 + \int_0^t S(X_u)du + \int_0^t \sigma(X_u)d\mathbb{W}_u$$

  holds a.s. for each $0 \le t \le T$.

The crucial requirement of this definition is captured in the highlighted condition; it corresponds to our intuitive understanding of X_t as the output of a dynamical system described by $[S(\cdot), \sigma(\cdot)]$, whose input is $\mathbb{W}^T$ and X_0. The **principle of causality** for dynamical systems requires that the output X_t at time t depend only on X_0 and the input $\{\mathbb{W}_u, 0 \leq u \leq t\}$.

$(\mathcal{GL})$ Globally Lipschitz condition

$$|S(x) - S(y)| + |\sigma(x) - \sigma(y)| \leq L|x - y| \quad \forall \;\; x, y \in \mathbb{R}^k$$

This implies a linear growth condition

$$|S(x) + \sigma(x)| \leq \tilde{L}(1 + |x|)$$

Theorem. Let the condition $\mathcal{GL}$ be fulfilled and $P(|X_0| < \infty) = 1$. Then the SDE above has a unique strong solution.

Proof can be found in L&S, Theorem 4.6.

($\mathcal{LL}$) Locally Lipschitz condition. For any $N < \infty$ and $|x|, |y| < N$, there exists a constant $L_N > 0$ such that

$$|S(x) - S(y)| + |\sigma(x) - \sigma(y)| \leq L_N |x - y|$$

and

$$2xS(x) + \sigma(x)^2 \leq B(1 + x^2) \ .$$

Theorem Let the condition $\mathcal{LL}$ be fulfilled and $P(|X_0| < \infty) = 1$. Then the SDE above has a unique strong solution.

Proof can be found in K&S.

Weak Solution of an SDE

$$dX_t = S(X_t)dt + \sigma(X_t)d\mathbb{W}_t \ , \quad X(0) = X_0, \quad 0 \le t \le T$$

A weak solution to the SDE above is a triplet $(\Omega, \mathcal{F}, \mathrm{P})$, $\{\mathcal{F}_t, 0 \le t \le T\}$, $(X^T, \mathbb{W}^T)$ where

- $(\Omega, \mathcal{F}, \mathrm{P})$ is a probability space and $\{\mathcal{F}_t, 0 \le t \le T\}$ is a filtration of $\mathcal{F}$;
- $X^T = \{X_t, 0 \le t \le T\}$ is a continuous, adapted process;
- $\mathbb{W}^T = \{\mathbb{W}_t, 0 \le t \le T\}$ is a Brownian motion;
- $$\mathrm{P}\Big\{ \int_0^T \big[S(X_t) + \sigma(X_t)^2\big] dt < \infty \Big\} = 1$$
- $$X_t = X_0 + \int_0^t S(X_t)dt + \int_0^t \sigma(X_t)d\mathbb{W}_t$$

  holds a.s. for each $0 \le t \le T$.

Uniqueness of solutions for SDEs

Suppose $(\Omega, \mathcal{F}, \mathrm{P})$, $\{\mathcal{F}_t, 0 \le t \le T\}$, $(X^T, \mathbb{W}^T)$ and $(\Omega, \mathcal{F}, \mathrm{P})$, $\{\tilde{\mathcal{F}}_t, 0 \le t \le T\}$, $(\tilde{X}^T, \mathbb{W}^T)$ are two weak solutions with common initial value $\mathrm{P}(X_0 = \tilde{X}_0) = 1$. We say that *pathwise uniqueness holds* if

$$\mathrm{P}(X_t = \tilde{X}_t \ \forall\ 0 \le t \le T) = 1$$

Suppose $(\Omega, \mathcal{F}, \mathrm{P})$, $\{\mathcal{F}_t, 0 \le t \le T\}$, $(X^T, \mathbb{W}^T)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathrm{P}})$, $\{\tilde{\mathcal{F}}_t, 0 \le t \le T\}$, $(\tilde{X}^T, \tilde{\mathbb{W}}^T)$ are two weak solutions with the same initial distribution $\mathcal{L}(X_0) = \mathcal{L}(\tilde{X}_0)$. We say that *uniqueness in the sense of probability law holds* if the two processes X^T and $\tilde{X}^T$ have the same law.

($\mathcal{ES}$) The function $S(\cdot)$ is locally bounded, the function $\sigma^2(\cdot)$ is continuous and for some $A > 0$

$$xS(x) + \sigma(x)^2 \leq A(1 + x^2)$$

Theorem. Suppose that condition $\mathcal{ES}$ is fulfilled, then the SDE has a unique (in law) weak solution.

Proof can be found in K&S.

Absolute continuity of measures on $\mathcal{C}([0, T])$

Preamble

- Assume that $X = 3.5$ is one observation, generated from one of two possible probability distributions P_1 and P_2 (over $\mathbb{R}$)
- Q: Is X likely to be a draw from P_1 or P_2 ?

Absolute continuity of measures on $\mathcal{C}([0, T])$

Preamble

- Assume that $X = 3.5$ is one observation, generated from one of two possible probability distributions P_1 and P_2 (over $\mathbb{R}$)
- Q: Is X likely to be a draw from P_1 or P_2 ?
- A: Calculate the likelihood ratio

$$\frac{dP_1}{dP_2}(X) = \frac{dP_1}{dP_2}(3.5)$$

- If this quantity is "large", X is likely a draw from P_1
- Otherwise, X is likely a draw from P_2

Absolute continuity of measures on $\mathcal{C}([0, T])$

Similar ideas can be applied to stochastic processes.
(Reference: L&S, Chapter 7)

- $(\Omega, \mathcal{F}, \mathrm{P})$ - prob. space, $(\mathcal{F}_t)$ - filtration, $(\mathbb{W}_t)$ - SBM
- $\mathcal{C}([0, T])$ = space of continuous functions on $[0, T]$
- Let X_t be a homogeneous Itô process

$$\begin{aligned} dX_t &= \beta(X_t)dt + d\mathbb{W}_t \quad X_0 = 0 \\ d\mathbb{W}_t &= \qquad\qquad\quad d\mathbb{W}_t \quad \mathbb{W}_0 = 0 \end{aligned}$$

- let μ_X, $\mu_{\mathbb{W}}$ be the probability measures on $\mathcal{C}_T([0, T])$ induced by $X_t, \mathbb{W}_t$.
- Task : define

$$\frac{d\mu_X}{d\mu_{\mathbb{W}}}$$

Does it exist ?

Theorem

Under some conditions, $\mu_X \sim \mu_{\mathbb{W}}$ and

$$\frac{d\mu_{\mathbb{W}}}{d\mu_X}(X) = \exp\left\{-\int_0^T \beta(X_t)\ dX_t + \frac{1}{2}\int_0^T \beta(X_t)^2\ dt\right\}$$

(Girsanov formula)

Example

$$dX_t = -\theta dt + d\mathbb{W}_t \quad 0 \leq t \leq 1 \quad X_0 = 0$$

Assume $\theta > 0$.

Girsanov formula, more general case

$$dX_t = A(X_t)dt \; + \; b(X_t)d\mathbb{W}_t$$
$$dY_t = a(Y_t)dt \; + \; b(Y_t)d\mathbb{W}_t$$

Assume that $X_0 = Y_0$. Under <u>some conditions</u>,

$$\frac{d\mu_Y}{d\mu_X}(X) = \exp\left\{-\int_0^T \frac{A(X_t) - a(X_t)}{b(X_t)^2}\, dX_t + \frac{1}{2}\int_0^T \frac{A(X_t)^2 - a(X_t)^2}{b(X_t)^2}\, dt\right\}$$