(1) Brownian Motion

Definition and Existence of Brownian Motion

Definition 4.2.1: Definition of Brownian motion

A stochastic process $\{B_t\}_{t\geq 0}$ on the probability triple $(\Omega, \mathscr{F}, \mathbb{P})$ is a Brownian motion if:

1. $B_0 = 0$.
2. The increments are independent, hence for $0 \leq t_0 \leq t_1 \leq t_2 \leq t_3$ then $B_{t_1} - B_{t_0}$ and $B_{t_2} - B_{t_1}$ are independent.
3. The increments are Gaussian, i.e. $0 \leq s \leq t$ then $B_t - B_s \sim N(0, t-s)$.
4. For almost all realizations $\omega \in \Omega$, the sample path $t \mapsto B_t(\omega)$ is continuous.

How can we be sure this exists?

Theorem 4.2.1: Brownian Motion Exists

There exists a probability triple $(\Omega, \mathscr{F}, \mathbb{P})$ and a stochastic process $\{B_t\}_{t\geq 0}$ which together satisfy all four conditions in def. 4.2.1.

Key Properties: Total Variation and Quadratic Variation

Unbounded Total Variation

Partition an interval $[S, T]$ into $\#\Delta$ sub-intervals, and let $|\Delta|$ denote the largest. The discretized total variation is then

$$V_\Delta = \sum_{i=1}^{\#\Delta} \left| B_{t_i} - B_{t_{i-1}} \right| .$$

As $V = \limsup_{|\Delta| \to 0} V_n$, the partition becomes finer, and $V = \infty$ w.p. 1.

Bounded Quadratic Variation

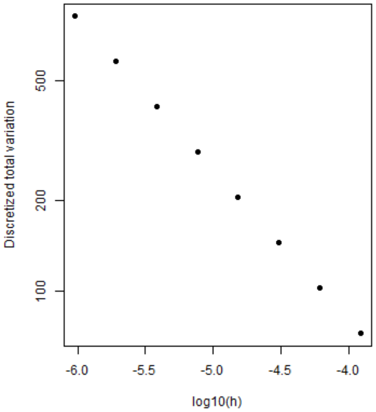
In contrast, the quadratic variation of Brownian motion, given as

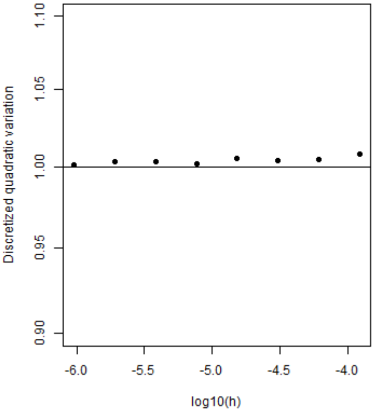
$$[B]_t = \lim_{|\Delta| \to 0} \sum_{i=1}^{\#\Delta} \left| B_{t_i} - B_{t_{i-1}} \right|^2 ,$$

on the interval $[0, t]$, is equal to the time. That is, $[B]_t = t$.

Note, for a differentiable function f, the total variation on $\int_0^1 |f'(t)| \, \mathrm{d}t$ is finite, and the quadratic variation on is zero. Hence, Brownian motion is not differentiable *a.s.*

Key Properties: Total Variation and Quadratic Variation





Key Properties: Reflection Principle

Define the maximum process

$$S_t = \max\{B_s \mid 0 \leq s \leq t\}.$$

By Theorem 4.3.2, we have

$$\mathbb{P}(S_t \geq x) = 2\mathbb{P}(B_t \geq x) = 2\Phi(-x/\sqrt{t}).$$

Let $x > 0$ be an arbitrary point and τ the hitting time of x. Then

$$\mathbb{P}(\tau \leq t) = \mathbb{P}(S_t \geq x) = 2\Phi(-x/\sqrt{t}).$$

For $x < 0$, we can use symmetry. Letting $t \to \infty$, we have $\mathbb{P}(\tau < \infty) = 1$, but it can also be shown that $\mathbf{E}\tau = \infty$. Continuing in the same direction, it can be shown that Brownian motion is null recurrent.

Filtrations and Martingales

Filtration

Recall, we use σ-algebra of events to model 'static' information. For stochastic processes, we want to model changes in information in time. We define introduce a family of σ-algebras in time, i.e.

$$\{\mathscr{F}_t : t \in \mathbb{R}\}.$$

We consider the accumulation of information without any loss. The family is increasing in the sense that

$$\mathscr{F}_s \subset \mathscr{F}_t \text{ whenever } t > s.$$

Martingales, def. 4.5.1

Brownian motion is a part of a class of stochastic processes called *martingales*.

Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$, a stochastic process $\{M_t\}_{t\geq 0}$ must satisfy the following to be a martingale:

1. M_t is adapted to the filtration $\mathscr{F}_t$.
2. For all times $t \geq 0$, $E|M_t| < \infty$.
3. $E\{M_t \mid \mathscr{F}_s\} = M_s$ when $t \geq s \geq 0$.

Brownian Motion is a Martingale and defintion of a Markov Process

Let $\{\mathscr{F}_t\}_{t\geq 0}$ be the filtration generated by $\{B_t\}_{t\geq 0}$. Condition 1 and 2 are obvious. Condition 3 follows from the independence of increments. Let $0 \leq s < t$, then

$$\mathbb{E}[B_t \mid \mathscr{F}_s] = B_s + \mathbb{E}[B_t - B_s \mid \mathscr{F}_s] = B_s.$$

Definition 9.1.1: Markov process

A process $\{X_t \in \mathbf{R}^n : t \geq 0\}$ is said to be a Markov process w.r.t. the filtration $\{\mathcal{F}_t\}$ if:

1. $\{X_t\}$ is adapted to $\{\mathcal{F}_t\}$, and
2. for any bounded and Borel-measurable test function $h : \mathbf{R}^n \mapsto \mathbf{R}$, and any $t \geq s \geq 0$, it holds almost surely that
$$\mathbf{E}\left\{h\left(X_t\right) \mid \mathcal{F}_s\right\} = \mathbf{E}\left\{h\left(X_t\right) \mid X_s\right\}.$$
i.e. X_s is sufficient statistics of $\{\mathcal{F}_s\}$

Brownian Motion as Driving Engine of SDE

Consider a non-linear system perturbed by a driving noise signal $\{\xi_t : t \geq 0\}$

$$\frac{dX_t}{dt} = f(X_t) + g(X_t)\xi_t,$$

we formally define ξ as the "derivative"of Brownian motion hence, $\xi = \frac{\mathrm{d}B_t}{\mathrm{d}t}$.

The integral form gives the a way for us to specify what we mean.

$$X_t = X_0 + \int_0^t f(X_s)\,\mathrm{d}s + \underbrace{\int_0^t g(X_s)\,\mathrm{d}B_s}_{\text{Itô integral}}.$$

where we define

$$\lim_{|\Delta|\to 0} \sum_{i=0}^{n-1} g(X_{t_i}, t_i)(B_{t_{i+1}} - B_{t_i}) = \int_0^t g(X_s, s)\,\mathrm{d}B_s.$$

Backward and Forward Kolmogorov for Brownian Motion

We have previously seen that Brownian motion is a stochastic process satisfying the SDE $\mathrm{d}X_t = \mathrm{d}B_t$. Hence, it is an Ito process with drift $f(X_t) = 0$ and diffusion $g(X_t) = 1$.

Applying the expression for the backward Kolmogorov equation, we find

$$-\dot{\psi} = L\psi = \psi' u + D\psi'' = \frac{1}{2}\psi'',$$

where $D = \frac{1}{2}g^2$ and $u = f - D'$. Similarly for the forward Kolmogorov equation, we find

$$\dot{\varphi} = L^*\varphi = -\left(u\varphi + D\varphi'\right)' = \frac{1}{2}\varphi''.$$

In this case, $L^* = L$.

Recall Fick's second law for diffusion, and the conservation law. Thus, the connection between diffusion and Brownian motion is clear.

Brownian Motion is a Markov Process

Let $0 \leq s \leq t$. The distribution of a random variable X is uniquely determined by its moment generating function $M(k) = \mathbb{E}\mathrm{e}^{kX}$, where $k \in \mathbb{R}$.

Let $\{\mathscr{F}_t\}_{t\geq 0}$ be the filtration generated by $\{B_t\}_{t\geq 0}$. Then

$$\begin{aligned}
\mathbb{E}[\mathrm{e}^{kB_t} \mid \mathscr{F}_s] &= \mathrm{e}^{kB_s}\mathbb{E}[e^{k(B_t - B_s)} \mid \mathscr{F}_s] \\
&= \mathrm{e}^{kB_s}\mathbb{E}[\mathrm{e}^{k(B_t - B_s)}] \\
&= \mathrm{e}^{kB_s}\mathbb{E}[\mathrm{e}^{k(B_t - B_s)} \mid B_s] \\
&= \mathbb{E}[\mathrm{e}^{kB_t} \mid B_s],
\end{aligned}$$

which shows that $\mathbb{E}[B_t \mid \mathscr{F}_s] = \mathbb{E}[B_t \mid B_s]$.

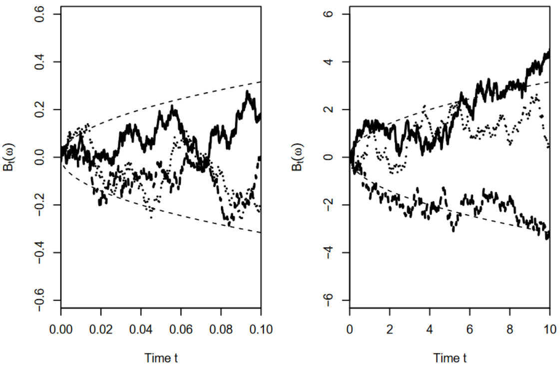
Key Properties: Self-similarity

If we rescale time for a Brownian motion $\{B_t\}_{t\geq 0}$, we can also rescale motion by α to recover the same process. That is,

$$\alpha^{-1}B_{\alpha^2 t}$$

for any $\alpha > 0$. In particular, this implies that moments of Brownian motion also scale with time

$$\mathbb{E}|B_t|^p = \mathbb{E}|\sqrt{t}B_1|^p = t^{p/2}\mathbb{E}|B_1|^p.$$



Statistically Indistinguishable

(2) Linear Systems

Multivariate Linear System

Narrow-sense multivariate linear system

We consider a stochastic process $\{X_t\}_{t\geq 0}$, $X_t \in \mathbb{R}^n$, which satisfies the Itô stochastic differential equation

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + G\,\mathrm{d}B_t,$$

with initial condition $X_0 = x$. The solution is then given by

$$X_t = \mathrm{e}^{At}x + G\int_0^t \mathrm{e}^{A(t-s)}\,\mathrm{d}B_s.$$

Proof.

Define $Y_t = h(X_t, t) = \mathrm{e}^{-At}X_t$. By Itô's Lemma, we get

$$\begin{aligned}\mathrm{d}Y_t &= \frac{\partial h}{\partial t}\,\mathrm{d}t + \frac{\partial h}{\partial x}\,\mathrm{d}X_t + \frac{1}{2}\,\mathrm{d}X_t^\top\frac{\partial^2 h}{\partial x^2}\mathrm{d}X_t\\ &= -A\mathrm{e}^{-At}X_t\,\mathrm{d}t + \mathrm{e}^{-At}AX_t\,\mathrm{d}t + \mathrm{e}^{-At}G\,\mathrm{d}B_t.\end{aligned}$$

Since $A\mathrm{e}^{-At} = \mathrm{e}^{-At}A$, we can multiply Y_t by e^{At} from the left and get the result. $\square$

Mean and Variance

Differential equation governing the mean

Define $\bar{x}(t) = \mathbb{E}X_t$. By Fubini's Theorem, we get $\bar{x}(t) = \mathrm{e}^{At}x$, so

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{x}(t) = A\bar{x}(t),$$

with initial condition $\bar{x}(0) = x$.

Differential equation governing the variance

Let $\rho_X(t,s)$ denote the covariance function of $\{X_t\}_{t\geq 0}$ and define the variance $\Sigma(t) = \rho_X(t,t)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma(t) = \underbrace{A\Sigma(t) + \Sigma(t)A^\top}_{\textit{sys respond to noise}} + \underbrace{GG^\top}_{\textit{new noise to sys}}$$

with initial condition $\Sigma(0) = 0$. This is the Lyaponov equation.

Autocovariance Function

Autocovariance function

For a narrow sense linear system, we can write the autocovariance function as:

$$\rho_X(s,t) = \int_0^s \mathrm{e}^{A(s-v)} G G^\top \mathrm{e}^{A^\top (t-v)} \, \mathrm{d}v = \Sigma(s) \cdot e^{A^\top (t-v)}$$

At stationary, we find Σ the solution to the *algebraic* Lyapunov equation:

$$A\Sigma + \Sigma A^\top + GG^\top = 0. \tag{0.1}$$

and write the autocovariance function as:

$$\rho(h) = \Sigma \exp\left(A^\top h\right) \text{ for } h \geq 0. \tag{0.2}$$

Variance Spectrum at Stationarity

Variance Spectrum

Let $\{X : t \geq 0\}$ be a stationary process with an $\mathcal{L}_2$ autocorrelation function ρ_X, then let S_X denote the variance spectrum defined as the Fourier transform of ρ_X:

$$S_X(\omega) = \int_{-\infty}^{+\infty} \rho_X(t) \exp(-i\omega t) \mathrm{d}t$$

Consider the inverse Fourier transform of S_X and evaluate at $t = 0$,

$$\mathbf{V} X_t = \rho_X(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_X(\omega) \mathrm{d}\omega,$$

which justifies the name.

Example with Mass-Spring-Damper System

Consider the standard-form system for $X_t = (Q_t, V_t)^\top$ given by

$$\mathrm{d}X_t = AX_t \, \mathrm{d}t + G \, \mathrm{d}B_t,$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ \sigma/m \end{bmatrix}.$$

Hence, the system is being driven by a white noise force $\{F_t\}_{t \geq 0}$ with intensity $S_F(\omega) = \sigma^2$.

We solve the Lyapunov equation by setting up the matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P = A \otimes I + I \otimes A, \quad Q = GG^\top.$$

We then solve $-PM = \mathrm{Vec}(Q)$. The elements of M yield the stationary variance matrix, In particular, $\Sigma = \mathrm{Vec}^{-1}_{(2\times 2)}(M)$.

Example with Mass-Spring-Damper System

We find that

$$\Sigma = \begin{bmatrix} \frac{\sigma^2}{2ck} & 0 \\ 0 & \frac{\sigma^2}{2cm} \end{bmatrix}.$$

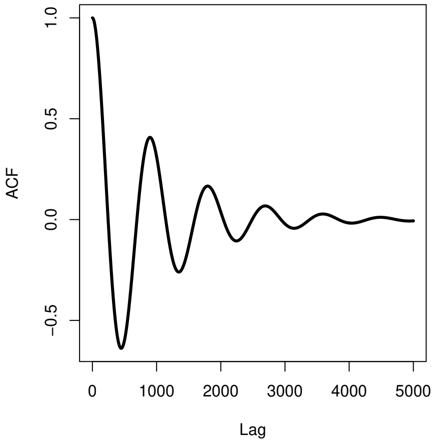
and as it is a narrow-sense system, the autocovariance at stationarity is

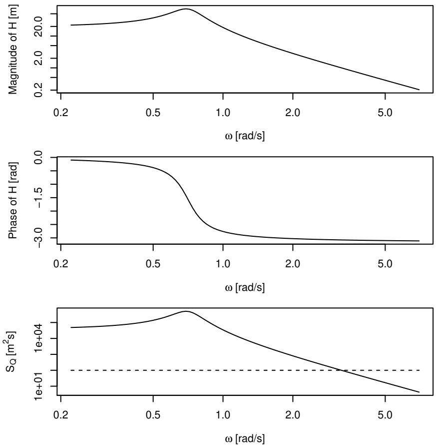
$$\begin{aligned} \rho_X(h) &= \mathbb{E} X_t X_{t+h}^\top \\ &= \Sigma \, \mathrm{e}^{A^\top h}. \end{aligned}$$

Furthermore, we calculate the frequency response

$$H(\omega) = (i\omega I - A)^{-1} G = \begin{bmatrix} \frac{\sigma}{ic\omega - m\omega^2 + k} \\ \frac{i\omega\sigma}{ic\omega - m\omega^2 + k} \end{bmatrix}.$$

Example with Mass-Spring-Damper System





- Parameters: $m = 1$, $k = 0.5$, $c = 0.2$, $\sigma = 10$.
- Spectra: $H_Q(\omega) = H^{(1)}(\omega)$, $S_Q(\omega) = \sigma^2 |H_Q(\omega)|^2$.

(3) The Itô Integral

Objectives and Specification of The Itô-Integral

Consider a non-linear system perturbed by a driving noise signal $\{\xi_t : t \geq 0\}$

$$\frac{dX_t}{dt} = f(X_t) + g(X_t)\xi_t,$$

we formally define ξ as the "derivative"of Brownian motion hence, $\xi = \frac{\mathrm{d}B_t}{\mathrm{d}t}$.

The integral form gives the a way for us to specify what we mean.

$$X_t = X_0 + \int_0^t f(X_s)\,\mathrm{d}s + \underbrace{\int_0^t g(X_s)\,\mathrm{d}B_s}_{\text{Itô integral}}.$$

We know from classical calculus to do $\int_0^t f(X_s)\,\mathrm{d}s$, but we need to formulate a clear understanding of the Itô integral.

Help From an Old Friend

Consider applying the Euler-Maruyama scheme directly

$$X_{t+h} = X_t + f(X_t, t)h + g(X_t, t)(B_{t+h} - B_t).$$

Introduce a partition of time, $0 = t_0 < t_1 < \ldots < t_n = t$, and apply Euler-Maruyama scheme for each timepoint t_i:

$$X_t = X_0 + \sum_{i=0}^{n-1} f(X_{t_i}, t_i)(t_{i+1} - t_i) + \sum_{i=0}^{n-1} g(X_{t_i}, t_i)(B_{t_{i+1}} - B_{t_i}),$$

where we know $(B_{t_{i+1}} - B_{t_i}) \sim \mathcal{N}(0, t_{i+1} - t_i)$. Now we let the time steps become very fine, i.e., $|\Delta| = \max\{t_i - t_{i-1} : i = 1, \ldots, n\}$ and let $|\Delta| \to 0$, then

$$\lim_{|\Delta| \to 0} \sum_{i=0}^{n-1} g(X_{t_i}, t_i)(B_{t_{i+1}} - B_{t_i}) = \int_0^t g(X_s, s)\, \mathrm{d}B_s.$$

Definition of the Itô integral

Definition 6.3.1 (Itô Integral: $\mathcal{L}^2$ version)

Let $0 \leq S \leq T$ and let $\{G_t : S \leq t \leq T\}$ be a real-valued stochastic process, which has left-continuous sample paths, which is adapted to $\{\mathcal{F}_\ell\}$, and for which $\int_S^T \mathbf{E} |G_t|^2 \, dt < \infty$. Then the limit

$$I = \lim_{|\Delta| \to 0} \sum_{i=1}^{\#\Delta} G_{l_{i-1}} \left(B_{l_i} - B_{t_{i-1}} \right) \text{ (limit in mean square)} \tag{0.3}$$

exists. Here $\Delta = \{S = t_0 < t_1 < t_2 < \cdots < t_n = T\}$ is a partition of $[S, T]$. We say that $\{G_t\}$ is $\mathcal{L}_2$ Itô integrable, and that I is the Itô integral:

$$I = \int_S^T G_t \mathrm{d} B_t \tag{0.4}$$

Explicitly, we evaluate the integrand at the LHS.

Properties of The Itô Integral

Theorem 6.3.2

Let $0 \le S \le T \le U$; let $\{F_t : 0 \le t \le U\}$ and $\{G_t : 0 \le t \le U\}$ be $\mathcal{L}_2$ Itô integrable on $[0, U]$ with respect to $\{B_t\}$. Then the following holds:

- *Additivity*: $\int_S^U G_t \, \mathrm{d}B_t = \int_S^T G_t \, \mathrm{d}B_t + \int_T^U G_t \, \mathrm{d}B_t$.
- *Linearity*: $\int_S^T aF_t + bG_t \, \mathrm{d}B_t = a \int_S^T F_t \, \mathrm{d}B_t + b \int_S^T G_t \, \mathrm{d}B_t$ when $a, b \in \mathbb{R}$.
- *Measurability*: $\int_S^T G_t \, \mathrm{d}B_t$ is $\mathscr{F}_T$-measurable.
- *Continuity*: $\{I_t\}_{t \ge 0}$ is continuous in the mean square and almost surely.
- *The martingale property*: $\{I_t\}_{0 \le t \le U}$, where $I_t = \int_0^t G_s \, \mathrm{d}B_s$, is a martingale w.r.t. $\{\mathscr{F}_t\}_{t \ge 0}$.
- *The Itō isometry*: $\mathbb{E} \left| \int_S^T G_t \, \mathrm{d}B_t \right|^2 = \int_S^T \mathbb{E} \left| G_t \right|^2 \, \mathrm{d}t$.

Implications of Defining Properties

Recall $\{X_t\}$ and consider this a $\mathcal{L}_2$ Itô process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$:

$$X_t = X_0 + \int_0^t F_s ds + \int_0^t G_s dB_s \tag{0.5}$$

Implication of the Markov Property

The expected future value is always, $\mathbf{E}\left\{\int_0^t G_t dB_t \mid \mathcal{F}_S\right\} = 0$, hence the mean of X_t is:

$$\mathbf{E}X_t = \mathbf{E}X_0 + \int_0^t \mathbf{E}F_s ds \tag{0.6}$$

Implications of the Itô isometry

We go from stochastic integral to Riemann integral which crazy: $\mathbb{E}\left|\int_S^T G_t \,\mathrm{d}B_t\right|^2 = \int_S^T \mathbb{E}\left|G_t\right|^2 \,\mathrm{d}t$. We can e.g. consider conditional variance of X_{t+h} given $\mathcal{F}_t$ useful for filtering:

$$\mathbf{E}\{X_{t+h} \mid \mathcal{F}_t\} = X_t + F_t \cdot h + o(h), \quad \mathbf{V}\{X_{t+h} \mid \mathcal{F}_t\} \approx |G_t|^2 \cdot h + o(h) \tag{0.7}$$

Integration of Brownian Motion

Let I_t^L denote the Itô integral, then we can analytically find integral of the Browanian:

$$\begin{aligned} I_t^L &= \sum_{i=1}^{n} B_{t_{i-1}} \cdot (B_{t_i} - B_{t_{i-1}}) \\ &= \frac{1}{2}\sum_{i=1}^{n} (B_{t_i} + B_{t_{i-1}}) \cdot (B_{t_i} - B_{t_{i-1}}) - \frac{1}{2}\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}}) \cdot (B_{t_i} - B_{t_{i-1}}) \\ &= \frac{1}{2}\sum_{i=1}^{n} \left(B_{t_i}^2 - B_{t_{i-1}}^2\right) - \frac{1}{2}\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2 \\ &= \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2 \end{aligned} \tag{0.8}$$

Now let the mesh approach 0 i.e. $|\Delta| \to 0$:

$$\int_0^t B_s dB_s = \lim_{|\Delta| \to 0} I_t^L = \frac{1}{2}B_t^2 - \frac{1}{2}t. \tag{0.9}$$

The Time of Evaluation of the Intensity Matters!

We consider three different ways:

$$I_t^L = \sum_{i=1}^{n} B_{t_{i-1}} \cdot (B_{t_i} - B_{t_{i-1}}) \quad \text{and} \quad I_t^R = \sum_{i=1}^{n} B_{t_i} \cdot (B_{t_i} - B_{t_{i-1}})$$
$$I_t^S = \sum_{i=1}^{n} \frac{1}{2} (B_{t_{i-1}} + B_{t_i}) \cdot (B_{t_i} - B_{t_{i-1}}) = 1/2 \left(I_t^L + I_t^R\right) \tag{0.10}$$

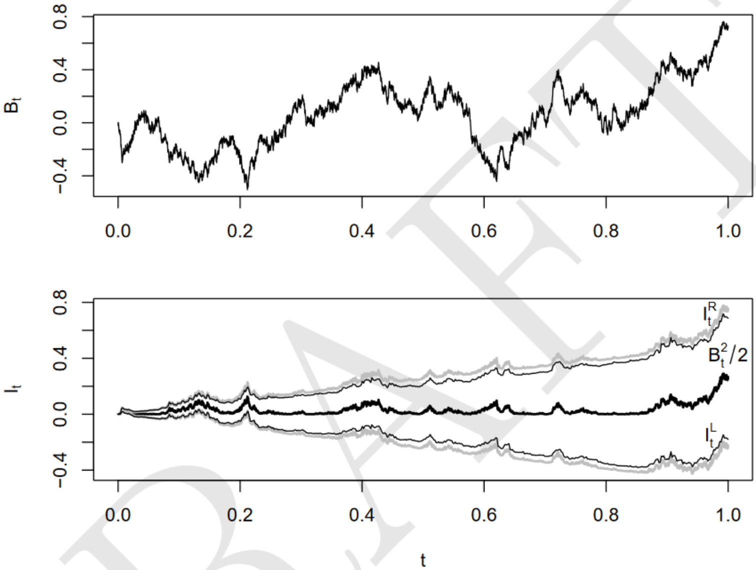
If we use the same argument as for I_t^L and let $n \to \infty$, we find $I_t^R = \frac{1}{2}B_t^2 + \frac{1}{2}t$. Additionally,

$$I_t^S = \frac{1}{2}\left(I_t^L + I_t^R\right) = \frac{1}{2}B_t^2. \tag{0.11}$$

I_t^S is special and well studied along the Itô integral and is denoted the Stratonovich integral:

$$\int_0^t B_s \circ \mathrm{d}B_s = \lim_{|\Delta| \to 0} \sum_{i=1}^{\#\Delta} \frac{1}{2}(B_{t_{i-1}} + B_{t_i})(B_{t_i} - B_{t_{i-1}}).$$

Numerical Simulation



(4) Itô’s Lemma

Itô's Lemma

Theorem 7.3.1 (Itô's lemma)

Let $\{X_t : t \geq 0\}$ be an Itô process as in definition 6.6.1, taking values in $\mathbf{R}^n$ and given by

$$\mathrm{d}X_t = F_t \mathrm{d}t + G_t \mathrm{d}B_t$$

where $\{B_t : t \geq 0\}$ is d-dimensional Brownian motion. Let $h : \mathbf{R}^n \times \mathbf{R} \mapsto \mathbf{R}$ be differentiable w.r.t. time t and twice differentiable w.r.t. x, with continuous derivatives and define $Y_t = h(X_t, t)$. Then $\{Y_t\}$ is an Itô process given by

$$\begin{aligned} \mathrm{d}Y_t &= \dot{h}\mathrm{d}t + \nabla h \mathrm{d}X_t + \frac{1}{2}\mathrm{d}X_t^\top \mathbf{H}h \mathrm{d}X_t \\ &= \dot{h}\mathrm{d}t + \left(\nabla h F_t + \frac{1}{2} \operatorname{tr} G_t^\top \mathbf{H}h G_t \right) \mathrm{d}t + \nabla h G_t \mathrm{d}B_t \end{aligned}$$

The New Guy:

$$\frac{1}{2}\mathrm{d}X_t^\top \mathbf{H}h \mathrm{d}X_t = \frac{1}{2}\mathrm{d}B_t^\top G_t^\top \mathbf{H}h G_t \mathrm{d}B_t = \frac{1}{2} \operatorname{tr} G_t^\top \mathbf{H}h G_t \mathrm{d}t$$

Simple Analytical Solution

Consider the Itô integral of Brownian Motion $\int_0^t B_s \mathrm{d}B_s = \left(B_t^2 - t\right)/2$. Identify the components

$$X_t = B_t \quad Y_t = h\left(X_t, t\right) \quad h(x) = \frac{x^2 - t}{2}$$

now we can find directly $\dot{h} = -1/2$, $h' = x$, and $h'' = 1$ and apply Itôs lemma:

$$\mathrm{d}Y_t = \dot{h}dt + h'\mathrm{d}X_t + \frac{1}{2}h''\left(\mathrm{d}X_t\right)^2 = -\frac{\mathrm{d}t}{2} + B_t\mathrm{d}B_t + \frac{\mathrm{d}t}{2} = B_t\mathrm{d}B_t$$

as we have $B_0 = 0$ in turn $Y_0 = 0$ hence we can easily find $Y_t = \int_0^t B_s \mathrm{d}B_s$ and by definition of h, we just found:

$$Y_s = \int_0^t B_s \mathrm{d}B_s = \frac{B_t^2 - t}{2}.$$

Multivariate Narrow-Sense Process

Consider the multivariate SDE given by

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + G\,\mathrm{d}B_t, \quad X_0 = x \in \mathbb{R}^n.$$

$\{B_t\}_{t\geq 0}$ is n-dimensional Brownian motion. Take $Y_t = h(X_t, t) = \mathrm{e}^{-At}X_t$ and identify

$$\dot{h} = -A\mathrm{e}^{At}, \qquad \nabla h = \mathrm{e}^{-At}, \qquad Hh = 0, \tag{0.12}$$

apply Itô's lemma to find

$$\mathrm{d}Y_t \quad = -A\mathrm{e}^{At}X_t\,\mathrm{d}t + \left(A\mathrm{e}^{At}X_t + 0\right)\mathrm{d}t + \mathrm{e}^{-At}G\,\mathrm{d}B_t \quad = \mathrm{e}^{-At}G\,\mathrm{d}B_t,$$

with the initial condition $Y_0 = h(x,0) = x$, we have $Y_t = x + \int_0^t \mathrm{e}^{-As}G\,\mathrm{d}B_s$ and hence the integral form of X_t

$$X_t = \mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}G\,\mathrm{d}B_s$$

The Lamperti Transformation

This is a transformation of coordinates such that the noise becomes *additive*.

Consider the scalar SDE

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + g(X_t)\,\mathrm{d}B_t.$$

The Lamperti transformed process $\{Y_t\}_{t\geq 0}$ is defined by $Y_t = h(X_t)$, where

$$h(x) = \int^x \frac{1}{g(y)}\,\mathrm{d}y, \quad h'(x) = \frac{1}{g(x)}, \quad h''(x) = -\frac{g'(x)}{g^2(x)},$$

Using Itô's Lemma we get

$$\begin{aligned}
\mathrm{d}Y_t &= h'(X_t)\,\mathrm{d}X_t + \frac{1}{2}h''(X_t)g^2(X_t)\,\mathrm{d}t \\
&= \left[\frac{f(h^{-1}(Y_t))}{g(h^{-1}(Y_t))} - \frac{1}{2}g'(h^{-1}(Y_t))\right]\mathrm{d}t + \mathrm{d}B_t.
\end{aligned}$$

We see that the noise term enters with constant diffusivity. Useful for simulation.

Lamperti Transformed Geometric Brownian Motion

We consider the geometric Brownian motion process

$$\mathrm{d}X_t = rX_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}B_t.$$

Using the transformation

$$h(x) = \int^x \frac{1}{\sigma y}\,\mathrm{d}y = \sigma^{-1}\log x,$$

the transformed process $\{Y_t\}_{t\geq 0}$ is given by $Y_t = h(X_t) = \sigma^{-1}\log X_t$, and additionally $h^{-1}(Z_t) = \exp(Z_t\sigma)$ which is governed by the SDE

$$\mathrm{d}Y_t = \left(\frac{rX_t}{\sigma X_t} - \frac{1}{2}\sigma\right)\mathrm{d}t + \mathrm{d}B_t = \left(\frac{r}{\sigma} - \frac{1}{2}\sigma\right)\mathrm{d}t + \mathrm{d}B_t.$$

Hence, we went from multiplicative noise to additive noise.

The Scale Function

The scale function can be used in the hunt of finding the martingale. It removes the drift term for scalar SDEs. Consider the scalar SDE

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + g(X_t)\,\mathrm{d}B_t.$$

Define $Y_t = s(X_t)$ where $s : \mathbb{R} \to \mathbb{R}$ such that the Y_t is an Itô process:

$$\mathrm{d}Y_t = \left[f\left(X_t\right) s'\left(X_t\right) + \frac{1}{2} g^2\left(X_t\right) s''\left(X_t\right)\right] \mathrm{d}t + s'\left(X_t\right) g\left(X_t\right) \mathrm{d}B_t$$

We want to find s such that the drift disappears i.e. $f + s' + \frac{1}{2} g^2 s'' = 0$:

$$\phi(x) = \exp\left(-\int_{x_0}^{x} \frac{2f(y)}{g^2(y)} \mathrm{d}y\right) \quad s(x) = \int_{x_1}^{x} \phi(y) \mathrm{d}y \tag{0.13}$$

Example Scale Function

Consider the Geometric Brownian Motion X_t given by the Itô SDE

$$\mathrm{d}X_t = \mu X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}B_t.$$

The scale function is

$$\begin{aligned} s(x) &= \int^x \exp\left(\int^y -\frac{2\mu}{\sigma^2 z}\mathrm{d}z\right)\mathrm{d}y \\ &= \int^x \exp\left(-\frac{2\mu}{\sigma^2}\log y\right) \\ &= \int^x y^{-\frac{2\mu}{\sigma^2}} \\ &= \begin{cases} \frac{1}{\nu}x^\nu & \text{when} \quad \nu := 1 - 2\mu/\sigma^2 \\ \log x & \text{when} \quad \nu = 0 \end{cases} \end{aligned}$$

(5) Existence and Uniqueness

Non-existence

In the following example, we consider an explosion. This means that there exists a stopping time τ such that X_t is defined in $[0, \tau[$, but $X_t \to \infty$ as $t \to \tau$.

Explosion of an SDE

Consider the Stratonovich SDE

$$\mathrm{d}X_t = (1 + X_t^2) \circ \mathrm{d}B_t.$$

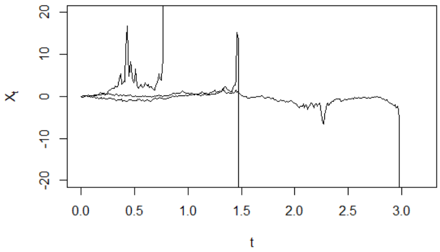
In the localized situation, it follows from the chain rule of Stratonovich calculus that

$$X_t = \tan B_t.$$

This solution is well-defined until the stopping time

$$\tau = \inf\left\{t \geq 0 : |B_t| \geq \frac{\pi}{2}\right\}.$$

Explosions at τ



Non-uniqueness

Non-uniqueness often happens when there exits singularities. In this case, it happens at the initial condition

Consider the Stratonovich SDE

$$\mathrm{d}X_t = 3|X_t|^{2/3} \circ \mathrm{d}B_t, \quad X_0 = 0.$$

Applying the chain rule of Stratonovich calculus, we may end up with the solution

$$X_t = \begin{cases} 0 & \text{for } 0 \leq t \leq T, \\ (B_t - B_T)^3 & \text{for } t > T, \end{cases}$$

for some deterministic $T \geq 0$. Indeed, we can choose T however we like, so this is certainly not a unique solution.

So what happened? The function $g(x) = 3|x|^{2/3}$ is not differentiable at $x = 0$.

Sufficient Conditions for Uniqueness

Next: When can we be certain that our solution is unique?

Definition 8.2.1: Lipschitz continuity

Let f be a function from one normed space $\mathbf{X}$ to another $\mathbf{Y}$. We say that f is globally Lipschitz continuous if there exists a constant $K > 0$ such that

$$|f(x_1) - f(x_2)| \leq K |x_1 - x_2|$$

holds for any $x_1, x_2 \in \mathbf{X}$. We say that f is locally Lipschitz continuous if for each $x \in \mathbf{X}$ there exists a neighborhood A of x such that the restriction of f to A is Lipschitz continuous.

Theorem 8.1.1: Uniqueness

Let $T > 0$ and assume that (f, g) are locally Lipschitz continuous for $0 \leq t \leq T$. Then the solution $\{X_t\}_{0 \leq t \leq T}$ to

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + g(X_t, t)\,\mathrm{d}B_t, \quad X_0 = x,$$

is unique *if such a solution exits*.

Sufficient Conditions for Existence

"linear bounds in dynamics gives exponential bounds on solutions"

Theorem 8.3.1:

Let the Itô process $\{X_t : 0 \leq t \leq T\}$ satisfy the initial value problem (8.1), (8.2) for $t \in [0, T]$ where $T > 0$. If (f, g) satisfy the bound

$$x^\top f(x,t) \leq C \cdot \left(1 + |x|^2\right), \quad |g(x,t)|^2 \leq C \cdot \left(1 + |x|^2\right)$$

for $C > 0$, all $x \in \mathbf{R}^n$, and all $t \in [0, T]$, then

$$\mathbf{E}\,|X_t|^2 \leq \left(x_0^2 + 3Ct\right) e^{3Ct}.$$

In particular, $\mathbf{E}\,|X_t|^2$ is finite and bounded on $[0, T]$.

Theorem 8.3.2:

If we have IVP, f, g local Lipschitz condition for uniqueness and linear growth bounds, then there exits a unique solution $\{X_t\}_{0 \leq t \leq T}$.

Examples of Existence and Uniqueness

We will now consider a few examples.

Bounded derivatives

$$\mathrm{d}X_t = X_t\,\mathrm{d}t + X_t\,\mathrm{d}B_t, \quad X_0 = 0.$$

Note, a C^1 function is necessarily locally Lipschitz. Furthermore, it is globally Lipschitz if and only if the derivative is bounded.

Locally Lipschitz and superlinear growth

$$\mathrm{d}X_t = -X_t^3\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 = 0.$$

Note, f is not globally Lipschitz since $f'(x) = -3x^2$ is unbounded. However, $xf(x) \leq 0$ for all x.

Simulation of Geometric Brownian

In the following, we will consider the Geometric Brownian Motion:

$$\mathrm{d}X_t = rX_t\mathrm{d}t + \sigma X_t\mathrm{d}B_t, \quad X_0 = x \tag{0.14}$$

with analytical solution $X_t = x\exp\left(\left(r - 1/2\sigma^2\right)2 + \sigma B_t\right)$. We will consider 3 different simulation methods that all stat at $X_0 = x$

1. The EM:

$$Z_{t+h} = Z_t + rZ_th + \sigma Z_t\left(B_{t+h} - B_t\right), \quad Z_0 = x \tag{0.15}$$

2. The Mil'shtein:

$$X_{t+h} = X_t + rX_th + \sigma X_t\left(B_{t+h} - B_t\right) + \frac{\sigma^2 X_t}{2}\left[\left(B_{t+h} - B_t\right)^2 - h\right] \tag{0.16}$$

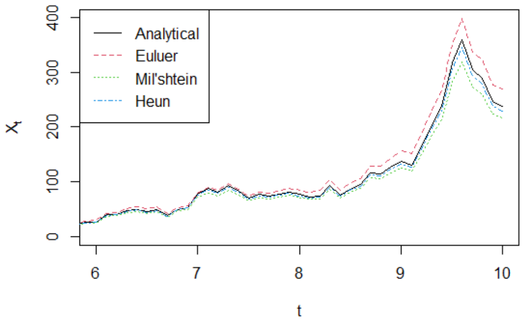
3. The Heun where we introduce

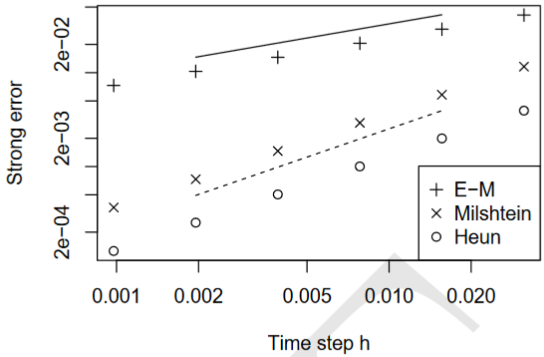
$$\bar{f} = \frac{1}{2}\left(f\left(X_t, t\right) + f\left(Z_{t+h}, t+h\right)\right), \quad \bar{g} = \frac{1}{2}\left(g\left(X_t, t\right) + g\left(Z_{t+h}, t+h\right)\right)$$

and do the update using:

$$X_{t+h} = X_t + \bar{f}h + \bar{g}\left(B_{t+h} - B_t\right) \tag{0.17}$$

Simulation of Geometric Brownian





- One Sample Path Using each scheme

Strong Order: $\mathbb{E}\left[\|x - X\|\right] \le Ch^{\gamma}$.

- EM: 0.5
- Mil'shtein: 1
- Heun: 1

(6) Transition Probabilities

Markov Processes

Definition 9.1.1: Markov process

A process $\{X_t \in \mathbf{R}^n : t \geq 0\}$ is said to be a Markov process w.r.t. the filtration $\{\mathcal{F}_t\}$ if:

1. $\{X_t\}$ is adapted to $\{\mathcal{F}_t\}$, and
2. for any bounded and Borel-measurable test function $h : \mathbf{R}^n \mapsto \mathbf{R}$, and any $t \geq s \geq 0$, it holds almost surely that

$$\mathbf{E}\left\{h\left(X_t\right) \mid \mathcal{F}_s\right\} = \mathbf{E}\left\{h\left(X_t\right) \mid X_s\right\}.$$

i.e. X_s is sufficient statistics of $\{\mathcal{F}_s\}$

Theorem 9.2.1: Solution to SDE is a Markov process

Let $\{X_t \in \mathbf{R}^n\}_{t\geq 0}$ be the unique solution to the SDE on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$

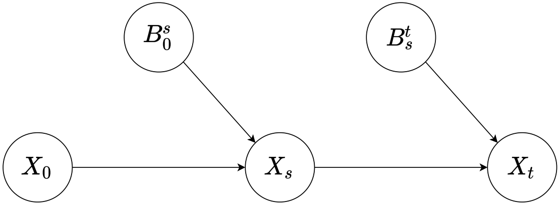
$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + g(X_t, t)\,\mathrm{d}B_t,$$

with $X_0 = x$. Then the process $\{X_t\}_{t\geq 0}$ is a Markov process with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ and its own filtration.

Diffusions are Markov Processes

Graphical 'proof'

Let B_s^t be the shorthand of $\{B_u - B_s\}_{s \le u \le t}$. Then, for $0 \le s \le t$, consider



Conditional on X_s, X_t is independent of X_0 and the Brownian Motion up to time s. From the property of Brownian motion, $\{B_u - B_s\}_{s \le u \le t}$ is independent of $\mathscr{F}_s$. Hence, the figure applies.

Transition Probabilities

Probelm Formulation

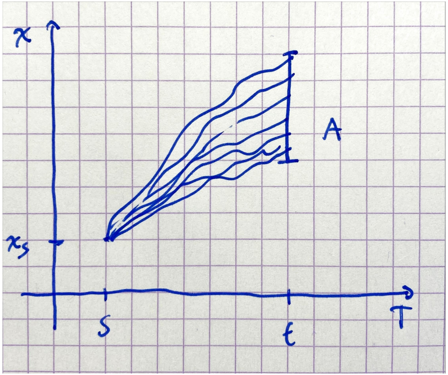
Assume X_t admits a transition density $p(s, x, t, y,$ then we can imagine a question:

$$\mathbf{P}^{X_s=x}\left(X_t \in A\right) = \int_A p(s, x, t, y)\mathrm{d}y$$

Note that the finite dimensional distribution can be found from the transition density:

$$\prod_{i=1}^{n} p\left(t_{i-1}, x_{t_{i-1}}, t_i, x_{t_i}\right)$$

Graphical Interpretation



Forward and Backward Kolmogorov Equations

Assume that X_t admits a *transition density* $p(s \mapsto t, x \mapsto y)$. If we fix (s, x), we get

$$\varphi(t, y) = p(s \mapsto t, x \mapsto y),$$

which is not a function of y.

Theorem 9.5.1: Forward Kolmogorov equation

Under the same assumptions as for the Backward Kolmogorov, we get for fixed (x, s) that

$$\begin{aligned} \dot{\varphi} &= L^* \varphi, \\ &= -(f\varphi)' + \left(\frac{1}{2} g^2 \varphi\right)'' \\ &= -\nabla \left(u\varphi - D\nabla\varphi\right) \end{aligned}$$

where L^* is the adjoint operator of L. By repetitive use of the Divergence Theorem, we find explicitly that $L^*\varphi = -\nabla \cdot (f\varphi) + \nabla \cdot \nabla(D\varphi)$.

Forward and Backward Kolmogorov Equations

Assume that X_t admits a *transition density* $p(s \mapsto t, x \mapsto y)$. If we fix (t, y), we get the likelihood

$$\psi(s, x) = p(s \mapsto t, x \mapsto y),$$

being a function of a parameter.

Theorem 9.4.1: Backward Kolmogorov equation

For a bounded C^2-function $h : \mathbb{R}^n \to \mathbb{R}$, let $k(X_s, s) = \mathbb{E}^{X_s = x} h(X_t)$. Then for fixed (t, y), we get

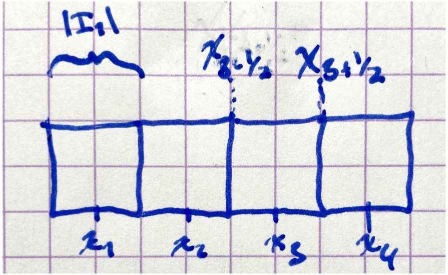
$$\begin{aligned} -\dot{\psi} &= L\psi \\ &= \psi' f + \frac{1}{2} g g^\top \psi'' \\ &= \nabla + \nabla\left(D \nabla \psi\right) \end{aligned}$$

where $L\psi = \nabla\psi \cdot f + \mathrm{tr}\left(\frac{1}{2} g g^\top \mathrm{H}\psi\right)$ is the familiar terms from Itô's Lemma with H beeing the Hessian.

Numerical Implementation

Split the domain into N grid cells. Let φ satisfy the forward Kolmogorov equation. Define the cell probability

$$\bar{\varphi}_i \approx \int_{x_{i-1/2}}^{x_{i+1/2}} \varphi(x)\,\mathrm{d}x$$



The diffusive part of the flux between cell I_i and I_{i+1} is approximated as

$$J_D(x_{i+1/2}) = -D(x_{i+1/2})\varphi'(x_{i+1/2}) \approx -D(x_{i+1/2})\frac{\bar{\varphi}_{i+1}/|I_{i+1}| - \bar{\varphi}_i/|I_i|}{x_{i+1} - x_i},$$

while the advective part of the flux between cell I_i and I_{i+1} is approximated as

$$J_A(x_{i+1/2}) = u(x_{i+1/2})\varphi(x_{i+1/2}) \approx \begin{cases} u(x_{i+1/2})\bar{\varphi}_i/|I_i|, & \text{if} u(x_{i+1/2}) > 0, \\ u(x_{i+1/2})\bar{\varphi}_{i+1}/|I_{i+1}|, & \text{if} u(x_{i+1/2}) < 0. \end{cases}$$

Numerical Implementation

Using these approximations, we can discretize the generator into two matrices. Consider the matrices going from i to $(i+1)$:

$$G^D_{i(i+1)} = \frac{D\left(x_{i+1/2}\right)}{|I_i|\,(x_{i+1} - x_i)} \qquad G^A_{i(i+1)} = \begin{cases} u\left(x_{i+1/2}\right)/\,|I_i| & \text{when } u\left(x_{i+1/2}\right) > 0 \\ 0 & \text{else} \end{cases}$$

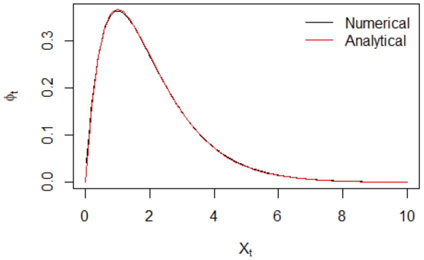
$G = G^A + G^D$ is a tridiagonal matrix with $G_{ii} = -G_{i(i+1)} - G_{i(i-1)}$.

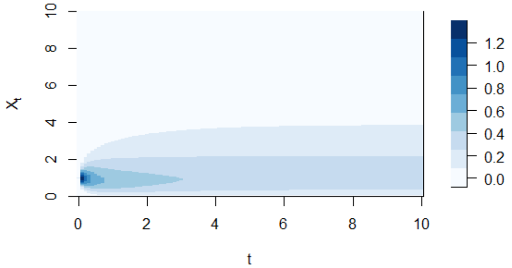
The discretized forward Kolmogorov equation becomes

$$\dot{\bar{\varphi}} = \bar{\varphi} G,$$

which has the solution $\varphi_t = \varphi_0 \mathrm{e}^{Gt}$.

Numerical Example: The CIR process





We consider the CIR process

$$dX_t = \lambda(\xi - X_t)\,dt + \gamma\sqrt{X_t}\,dB_t,$$

where λ, ξ and γ are positive parameters. On the left we have the steady-state p.d.f. and on the right the φ_t.

(7) State Estimation

State Estimation Problem

State Estimation Problem

Let $\{X_t : t \geq 0\}$ be a diffusive process the admit the Itô equation:

$$\mathrm{d}X_t = f\left(X_t, t\right)\mathrm{d}t + g\left(X_t, t\right)\mathrm{d}B_t$$

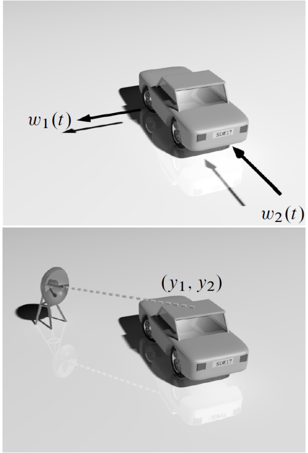
with the measurement equation:

$$Y_i = c\left(X_i\right) + s\left(X_i\right)\xi_i$$

At a time t_i, we essentially want to determine the following probability density:

$$\rho\left(x(t)|y_{1:k}\right) \qquad t \in [t_k, t_{k+1})$$

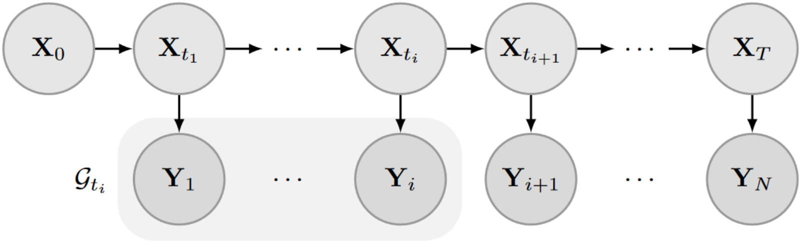
Graphical Interpretation



Särkkä, Simo and Solin, Arno, 2019, *Applied Stochastic Differential Equations*

Recursive filtering

Consider the graphical representation of the Hidden Markov Model:



Three things to note:

1. $\{X_t\}_{t\geq 0}$ is continuous but sub-sampled at times t_i, $\{X_{t_i}\}_{t_i\in\mathbb{N}}$
2. $\{Y_i\}_{i\in\mathbb{N}}$ are measurements used to infer $\{X_{t_i}\}_{t_i\in\mathbb{N}}$
3. $\{\mathscr{G}_t\}_{t\geq 0}$ is the filtration generated by the measurements i.e. $\mathscr{G}_t = \sigma(\{Y_i\}_{t_i\leq t})$.

The Algorithm

At a time t_i, we essentially want to determine the following probability density:

$$\rho\left(x(t)|y_{1:k}\right) \qquad t \in [t_k, t_{k+1})$$

We split this problem into two. Let $\varphi_i(x)$ as the p.d.f of X_{t_i} given $\mathcal{G}_{t_{i-1}}$, and $\psi_i(x)$ as the p.d.f of X_{t_i} given $\mathcal{G}_{t_i}$.

$$\begin{aligned} &\varphi_i(x) \text{ is the predicted distribution, } \mathcal{G}_{t_{i-1}} - \text{measurable} \\ &\psi_i(x) \text{ is the estimated distribution, } \mathcal{G}_{t_i} - \text{measurable} \end{aligned}$$

Recursive filtering

1. Start at t_0 with $\psi_0(\cdot)$. Set $i = 0$.
2. *Time update* - Advance time to t_{i+1} and set $\varphi_{i+1}(x)$ equal to the solution $\rho(x, t_{i+1})$.
3. *Data update* - Compute $\psi_{i+1}(\cdot)$ from Bayes' rule.
4. Advance i and go to the second step.

The Updates

Time Update

Let $\rho(x,t)$ be the p.d.f of X_t conditioned on $\mathcal{G}_{t_i}$ and evaluated at x. Then, for $t \geq t_i$, ρ is governed by

$$\dot{\rho} = -\nabla \cdot (u\rho - D\nabla\rho), \qquad \rho(x,t_i) = \psi_i(x),$$

Specifically, we can find $\varphi_{i+1}(x) = \rho(x,t_{i+1})$ and for a time-invariant SDE, we find:

$$\phi_{i+1} = e^{L^*(t_{i+1}-t_i)}\psi_i$$

The Data Update

Consider an observation $y_i = Y_i(\omega)$, then we calculate the likelihood given our internal state:

$$l_i(x) = f_{Y_i|X_{t_i}}(y_i \mid x)$$

In the time update we determined $f_{X_{t_i}} = \varphi_{t_i}$. Using Bayes rule, we can now find $f_{X_{t_i}|Y_i} = \psi_i$:

$$\psi_i = f_{X_{t_i}} = \frac{f_{X_{t_i}} f_{Y_i|X_{t_i}}}{f_Y} = \frac{\varphi_i \ell_i(x)}{f_{Y_i}}$$

The normalization constant, we can find by marginalizing x with

$$f_{Y_i}(y_i) = \int_x f_{X_{t_i}} f_{Y_i|X_{t_i}} = c_i$$

Likelihood Estimation and Normalizing Constant

Assume our data (f, g) depends on some unknown parameter θ, then we find the likelihood function

$$\begin{aligned}\Lambda(\theta) &= f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N;\theta) \\ &= f_{Y_1}(y_1;\theta) f_{Y_2,\ldots,Y_N|Y_1}(y_1,\ldots,y_N;\theta) \\ &= \prod_{i=1}^{N} f_{Y_i|Y_1,\ldots,Y_N}(y_1,\ldots,y_N;\theta) \\ &= \prod_{i=1}^{N} c_i(\theta)\end{aligned}$$

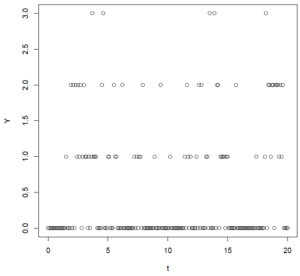
Maximising the likelihood function corresponds to tuning the predictive filter to predict the next measurement optimally

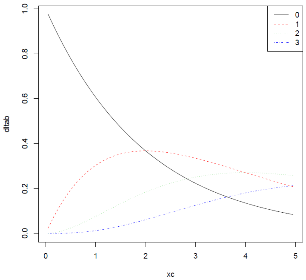
Example

Let X_t be the measure of bacteria cells pr volume described by CIR:

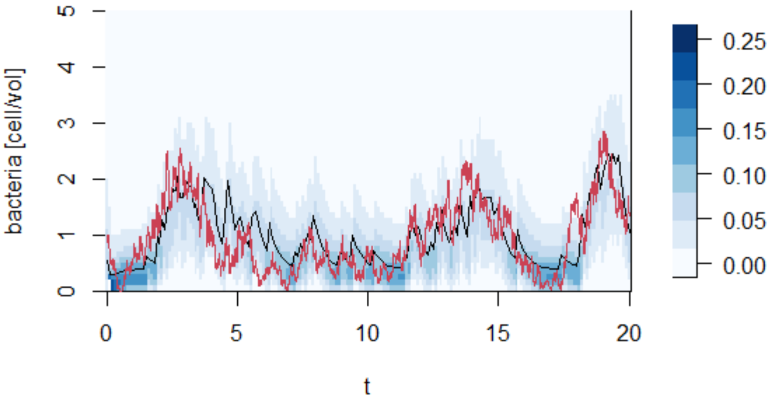
$$\mathrm{d}X_t = \lambda \cdot (\xi - X_t)\,\mathrm{d}t + \gamma\sqrt{X_t}\mathrm{d}B_t \quad \text{with } \lambda = \xi = \gamma = 1.$$

$$Y_t \mid X_t \sim \mathrm{Poisson}\,(\mathrm{mean} = v \cdot X_t)$$

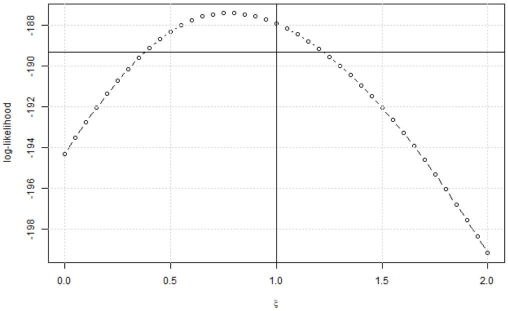




Predicted State



Likelihood of growth parameter



The log-likelihood estimate of the mean-reversion parameter is $\hat{\theta} = 0.75$ and the 95% CI is $[0.40; 1.20]$, based on $\{\theta \mid L(\theta) \geq L(\hat{\theta}) - \chi_1^2(0.95)\}$. The real value is 1.

Kalman Filter

We briefly introduce another technique in the case of linear stochastic differential equations with linear observations. Here, the conditional differential equations of the state given the observation is also Gaussian.

Time update

Advance in time from t_i to t_{i+1} by solving the equations

$$\begin{aligned} \frac{d}{dt}\mu_{t|t_i} &= A\mu_{t|t_i} + u_t, \\ \frac{d}{dt}\Sigma_{t|t_i} &= A\Sigma_{t|t_i} + \Sigma_{t|t_i}A^\top + GG^\top, \end{aligned}$$

with $\mu_{t|t_i} = \mathbb{E}\{X_{t_i}|\mathscr{G}_{t_i}\}$ and $\Sigma_{t_i|t_i} = \mathbb{V}\{X_{t_i}|\mathscr{G}_{t_i}\}$.

Data update

Compute first the Kalman gain

$$K_{i+1} = \Sigma_{t|t_i}C^\top(C\Sigma_{t|t_i}C^\top + DD^\top)^{-1},$$

and from here

$$\begin{aligned} \mu_{t_{i+1}|t_{i+1}} &= \mu_{t_{i+1}|t_i} + K_{i+1}(y_{i+1} - C\mu_{t_{i+1}|t_i}) \\ \Sigma_{t_{i+1}|t_{i+1}} &= \Sigma_{t_{i+1}|t_i} - K_{i+1}\ C\ \Sigma_{t_{i+1}|t_i}. \end{aligned}$$

(8) Applications of Dynkin's Lemma

Dynkin's Lemma

Theorem 11.2.1: Dynkin's Lemma

Let $h \in C_0^2(\mathbb{R}^n)$ and let τ be a stopping time such that $\mathbb{E}^x \tau < \infty$. Then

$$\mathbb{E}^x h(X_\tau) = h(x) + \mathbb{E}^x \int_0^\tau (Lh)(X_s)\,\mathrm{d}s.$$

Two important notions for this machinery to work $h \in C_0^2(\mathbb{R}^n)$ and $\mathbb{E}^x \tau < \infty$

Regular Diffusions on Bounded Domain (Theorem 11.2.2)

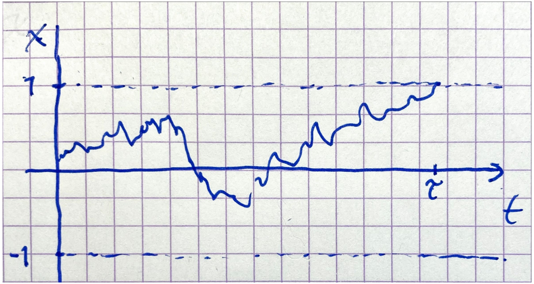
Let the diffusion $\{X_t : t \geq 0\}$ be regular in the sense that there exists a $d > 0$ such that

$$\frac{1}{2} g(x) g^\top(x) > dI$$

for all $x \in \mathbf{\Omega}$. As before, let $\tau = \inf\{t \geq 0 : X_t \notin \mathbf{\Omega}\}$ be the time of first exit. Then $\mathbf{E}^x \tau < \infty$ for all $x \in \mathbf{\Omega}$.

Applications and Different Versions

Questions	Formula $x \in \Omega$	Boundary $x \in \delta\Omega$	Expectations
Exit Time	$Lh(x) + 1 = 0$	$h(x) = 0$	$\mathbb{E}^x \tau$
Point of Exit	$Lh(x) = 0$	$h(x) = c(x)$	$\mathbb{E}^x c(X_\tau)$
Rewards	$Lh(x) + r(x) = 0$	$h(x) = c(x)$	$h(x) = \mathbf{E}^x \left[c(X_\tau) + \int_0^\tau r(X_t)\, dt\right]$
Discounting	$\dot{h} + Lh(x) - \mu h = 0$	$h(x, T) = k(x)$	$h(x,t) = \mathbf{E}^{X_t = x}\left\{e^{-\int_t^T \mu(X_s) ds} k(X_T)\right\}$



Example with O-U Process

The Example

Let $\{X_t : t \geq 0\}$ be an Itô diffusion with the SDE

$$\mathrm{d}X_t = -X_t \mathrm{d}t + \sigma \mathrm{d}B_t, \quad X_0 = x$$

$X_t \in (-1,1)$ Let $\tau = \inf\{t : |X_t| \geq 1\}$ with $\sigma = 1$

Find $\mathbb{P}^x(X_\tau = 1)$

We need $h(x) = c(x)$ on $x \in \delta\Omega$ and know $h(x) = \mathbb{E}^x c(x) = \mathbb{P}^x(X_\tau = 1)$ if $c(x) = \mathbb{1}_{X_\tau = 1}$ hence set $h(-1) = 0$, $h(1) = 1$ on the boundary

$$Lh(x) = -xh' + \frac{1}{2}\sigma^2 h'' = 0 \quad \text{on } x \in (-1,1)$$

Solve numerically to find $\mathbb{P}^x(X_\tau = 1)$.

Find $\mathbb{E}^x\tau$

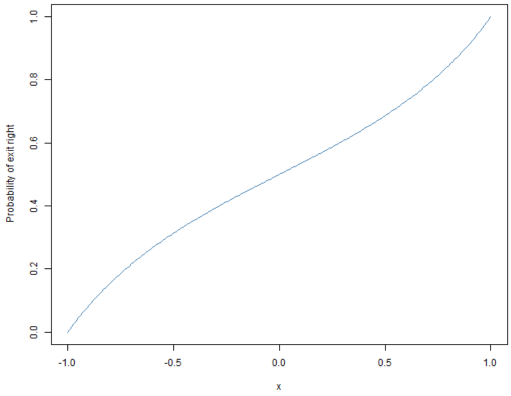
We consult our table and find $h(x) = 0$ on $x \in \delta\Omega$ and know $h(x) = \mathbb{E}^x\tau$ hence set $h(-1) = 0$, $h(1) = 0$ on the boundary and solve:

$$Lh(x)+1 = -xh'+\frac{1}{2}\sigma^2 h''+1 = 0 \quad \text{on } x \in (-1,1)$$

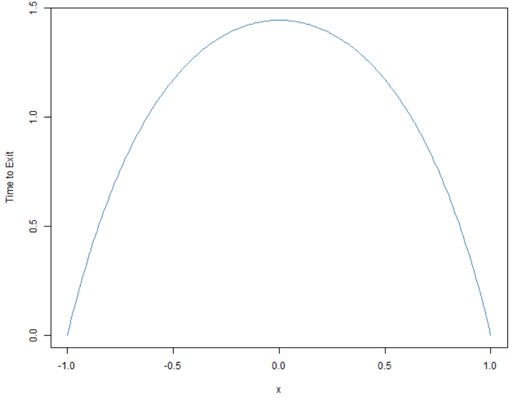
This BVP, we solve numerically to find expected time of exit.

Example with O-U Process

Find $\mathbb{P}^x(X_\tau = 1)$



Find $\mathbb{E}^x\tau$



Scale Function as Solution in Sclar Case

Consider the more general equation $\{X_t\}_{t\geq 0}$ be governed by $\mathrm{d}X_t = f\,\mathrm{d}t + g\,\mathrm{d}B_t$ on a domain (a,b) and we want to find $\mathbb{P}^x(X_\tau = b)$ hence solve:

$$h'f + \frac{1}{2}g^2h'' = 0 \ \text{ for } \ x \in (a,b), \quad h(a) = 0, \quad h(b) = 1,$$

which has the full solution

$$h(x) = c_1 s(x) + c_2.$$

Introduce the scale function s and set $\varphi = s'$ to obtain:

$$\varphi(x) = \exp\left(\int^x \frac{-2f(y)}{g^2(y)}\,\mathrm{d}y\right), \qquad s(x) = \int^x \varphi(y)\,\mathrm{d}y.$$

for an arbitrary lower bound. Insert boundary conditions to obtain

$$\mathbb{P}^x(X_\tau = b) = h(x) = \frac{s(x) - s(a)}{s(b) - s(a)}.$$

Singular Boundary Point

Consider the Bessel Process X_t given by the Itô SDE

$$\mathrm{d}X_t = \mu\,\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}B_t.$$

The point $x = 0$ is singular since $g(0) = 0$. The scale function is

$$\begin{aligned} s(x) &= \int^x \exp\left(\int^y -\frac{2\mu}{\sigma^2 z}\mathrm{d}z\right)\mathrm{d}y \\ &= \frac{1}{\nu}x^\nu \text{ when } \nu := 1 - 2\mu/\sigma^2 \neq 0 \end{aligned}$$

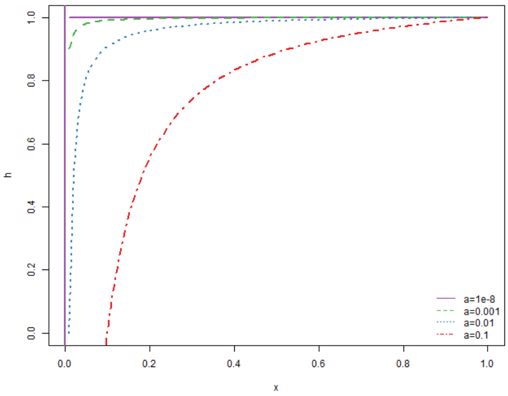
hence the probability of exit at b before a is:

$$h(x) = \mathbf{P}^x\left\{X_\tau = b\right\} = \frac{x^\nu - a^\nu}{b^\nu - a^\nu} \tag{0.18}$$

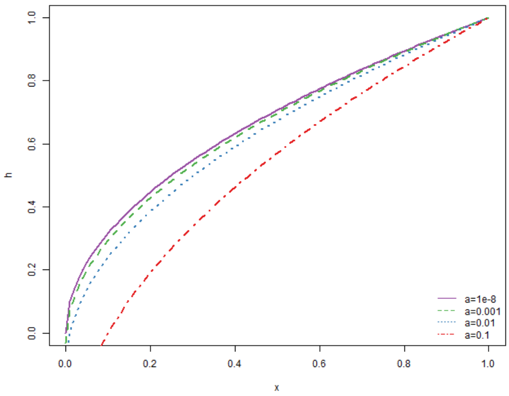
We will consider two cases with different parameters.

Bessel Process, $\mathrm{d}X_t = \mu\,\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}B_t.$, i.e. $a \to 0$ (can we sleep at night?)

h with $\mu = 1$ and $\sigma^2 = 1$



h with $\mu = 0.25$ and $\sigma^2 = 1$



(9) Stochastic Stability Theory

Notions of Stability

3 Notions of Stability for Engineers

1. Stochastic Stability (the Lyapunov's)
   - Stochastic Lyapunov Exponent
   - Stochastic Lyapunov Function
2. Bounded Mean
3. Bounded Variance of Perturbations

Sensitivity of Initial Condition

Sensitivity of the Initial Condition

Let $\{X_t : t \geq 0\}$ be an Itô diffusion with the SDE:

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + g(X_t, t)\,\mathrm{d}B_t, \quad X_0 = x,$$

For each initial condition $x \in \mathbb{R}^n$, define the state transition map Φ_t govern by:

$$\mathrm{d}\Phi_t(x) = f(\Phi_t(x))\,\mathrm{d}t + g(\Phi_t(x))\,\mathrm{d}B_t, \quad \Phi_0(x) = x.$$

Then we ask $\Phi_t(x + \delta x) \approx \Phi_t(x) + S_t(x)\delta x$ and define $\mathrm{d}S_t(x) = \frac{\partial \Phi_t(x)}{\partial x}$, $S_t \in \mathbb{R}^{n \times n}$ found by:

$$dS_t(x) = \nabla f\left(\Phi_t(x), t\right) \cdot S_t(x) dt + \sum_{i=1}^{m} \nabla g_i\left(\Phi_t(x), t\right) \cdot S_t(x) dB_t^{(i)}, \qquad \Phi_0(x) = x.$$

Stochastic Lyapunov Exponent

To measure the magnitude of $S_t(x)$, take the operator norm $\sup\{\|S_t(x)\tilde{x}\| \,|\, \tilde{x} \in \mathbb{R}^n, \|\tilde{x}\| = 1\}$. This is the largest singular value of $S_t(x)$, denoted by $\bar{\sigma}(S_t(x))$. We then consider the average growth rate of $\bar{\sigma}(S_t(x))$ on the time interval $[0, t]$

$$\lambda_t = \frac{1}{t} \log \bar{\sigma}(S_t(x)),$$

which we name the *stochastic finite-time Lyapunov exponent*. Going to the limit, we get

$$\bar{\lambda} = \limsup_{t\to\infty} \lambda_t, \quad \text{a.s.}$$

Definition 12.3.2: Stochastic stability from Lyapunov exponent

- $\{X_t\}_{t\geq 0}$ is stable if $\bar{\lambda} < 0$ a.s.
- $\{X_t\}_{t\geq 0}$ is unstable if $\bar{\lambda} > 0$ a.s.
- $\{X_t\}_{t\geq 0}$ is marginally stable if $\bar{\lambda} = 0$ a.s.

Example using Geometric Brownian

Sensitivity Equation

Consider the stochastic process $\{X_t : t \geq 0\}$ with sensitivity $\{S_t : t \geq 0\}$

$$\begin{aligned} \mathrm{d}X_t &= rX_t\mathrm{d}t + \sigma X_t \mathrm{d}B_t, \qquad X_0 &= x \\ &= rS_t\mathrm{d}t + \sigma S_t \mathrm{d}B_t, \qquad S_0 = 1 \end{aligned} \tag{0.19}$$

This is also just a GBM which has solution $S_t = \mathrm{e}^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma B_t}$

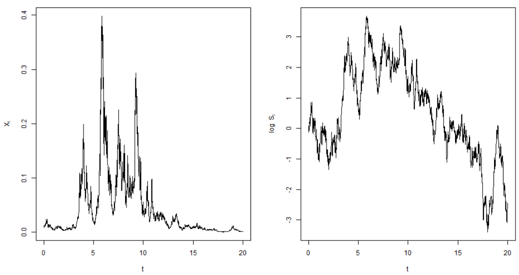
Stochastic Lyapunov Exponent

$$\lambda_t = \frac{1}{t}\log S_t = r - \frac{1}{2}\sigma^2 + \frac{1}{t}\sigma B_t.$$

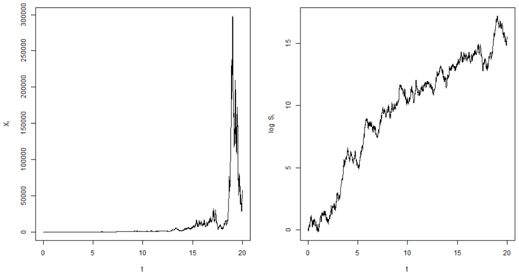
Since $\{B_t\}_{t \geq 0}$ scales with the square root of time, we get $\bar{\lambda} = r - \sigma^2/2$ w.p.1. We know that the only equilibrium point is $x^* = 0$ as it is where $f(0) = g(0) = 0$. If X_t is a population it means that if $r > \frac{\sigma^2}{2}$, then the population will die due to strong fluctuations.

Stability with Different Parameters

$x_0 = 0.01$, $r = 0.1$, $\sigma = 1.2$ i.e. $r < \frac{1}{2}\sigma^2$



$x_0 = 0.01$, $r = 1$, $\sigma = 1.2$ i.e. $r > \frac{1}{2}\sigma^2$



Stochastic Lyapunov Functions

Theorem 12.7.1: Stability assessment using stochastic Lyapunov functions

Let x^* be an equilibrium point of our SDE, meaning that $f(x^*) = 0$ and $g(x^*) = 0$. Assume that there exists a function $V(x)$ defined on a domain D containing x^* such that

1. V is C^2 on $D \setminus \{x^*\}$.
2. There exist continuous, strictly increasing functions a and b with $a(0) = b(0) = 0$ such that $a(|x - x^*|) \leq V(x) \leq b(|x - x^*|)$
3. $LV(x) \leq 0$ for $x \in D \setminus \{x^*\}$.

Then

$$\lim_{x \to x^*} \mathbb{P}^{X_0 = x}\{\sup |X_t - x^*| > \varepsilon\} = 0.$$

The third assumption outlines the supermartingale property. The conclusion states that the probability of leaving the permitted region $\{\xi \mid |\xi - x^*| \leq \varepsilon\}$ is zero.

Application of Stochastic Lyapunov Functions

Consider $\{X_t : t \geq 0\}$ and a the following Lyapunov function $V(x)$:

$$\mathrm{d}X_t = rX_t\mathrm{d}t + \sigma X_t \mathrm{d}B_t, \qquad X_0 = x$$
$$V(x) = |x|^p$$

To analyse the equilibrium at $X_t = 0$, we check criteria

1. if $p > 0$, then V is C^2 on $D \setminus \{x^*\}$ ✓
2. we can bound $a(|x - x^*|) \leq V(x) \leq b(|x - x^*|)$✓
3. need check $LV(x) \leq 0$ for $x \in D \setminus \{x^*\}$?

$$\begin{aligned} LV(x) &= px^{p-1}f(x) + \frac{1}{2}p(p-1)x^{p-2}g^2(x) \\ &= p\left(pr + \sigma^2(p-1)\right)|x|^2 \end{aligned}$$

This is negative if $p\left(pr + \sigma^2(p-1)\right) \leq 0$ and with $p > 0$, then $0 < 1 - 2\frac{r}{\sigma^2}$. We get to the same conclusion that it is stable for

$$r < \frac{\sigma^2}{2}$$

Mean Square Stability

Theorem 12.8.1: Stability in the mean square

Assume that there eixsts a Lyapunov function V such that

$$k_1|x - x^*| \leq V(x) \leq k_2|x - x^*|^2 \quad \text{and} \quad LV(x) \leq -k_3|x - x^*|^2,$$

for all x, where k_1, k_2 and k_3 are positive constants. Then the equilibrium x^* is exponentially stable in the mean square.

Lyapunov function GBM

With $V(x) = \frac{1}{2}x^2$ we get $LV(x) = \left(r + \frac{1}{2}\sigma\right)x^2$ and we are bounded above by

$$r + \sigma/2 < 0$$

Could we have seen this?

We know the moment of the geometric Brownian:

$$\mathbf{E}X_t = xe^{rt}$$

$$\mathbf{E}X_t^2 = x^2 e^{\left(2r+\sigma^2\right)t}$$

And we see directly the same result $2r + \sigma^2 \leq 0$

Notions of Stability

3 Notions of Stability for Geometric Brownian Motion

1. Stochastic Stability (the Lyapunov's)
   - Stochastic Lyapunov exponent, $r < \frac{1}{2}\sigma^2$
   - Stochastic Lyapunov function $r < \frac{1}{2}\sigma^2$
2. Bounded Mean, L_1, $r < 0$
3. Bounded Variance of Perturbations, L_2, $2r + \sigma^2 \leq 0$

(10) Optimal Control

Optimal Control for Stochastic Differential Equations

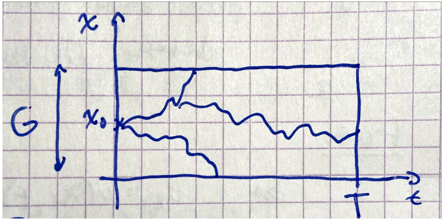
Control Problem

Let $\{X_t : t \geq 0\}$ be a diffusive process determined by the Itô equation:

$$\mathrm{d}X_t = f(X_t, U_t)\,\mathrm{d}t + g(X_t, U_t)\,\mathrm{d}B_t, X_0 = x \in \mathbb{R}^n,$$

with $\tau = \min\{T, \inf\{t \in [0,T] : X_t \notin G\}\}$ and a control signal $\{U_t\}_{t \in \mathbf{T}}$.

Graphical Interpretation



Our Approach to Optimal Control

To maintain the Markov property, we consider *state feedback* controls $U_t = \mu(X_t, t)$ for some function $\mu : \mathbf{X} \times \mathbf{T} \to \mathbf{U}$ such that existence and uniqueness holds for $\{X_t\}_{t \in \mathbf{T}}$. We want to find:

$$\sup_{\mu} J(x, \mu, s) = \sup_{\mu} \mathbb{E}^{\mu, X_s = x}\left[k(X_\tau, \tau) + \int_s^\tau h(X_t, U_t, t)\,\mathrm{d}t\right],$$

Hamilton-Jacobi-Bellman Equation

Theorem 13.3.1

Let the domain G be open and bounded, let $V : G \times [0,T] \mapsto \mathbf{R}$ be $C^{2,1}$ and satisfy the Hamilton-Jacobi-Bellman equation

$$\frac{\partial V}{\partial t} + \sup_{u \in \mathbf{U}} [L^u V + h] = 0$$

on $G \times [0,T]$, along with boundary and terminal conditions

$$V = k \text{ on } \partial(G \times [0,T])$$

Let $\mu^* : G \times [0,T] \mapsto \mathbf{U}$ be such that

$$\sup_{u \in \mathbf{U}} [L^u V + h] = L^{\mu^*} V + h$$

on $G \times [0,T]$, and assume that with this μ^*, the closed-loop system (13.6) satisfies the conditions in theorem 8.3.2. Then, for all $x \in \mathbf{G}$ and all $s \in [0,T]$, we have that μ^* is the optimal strategy.

$$V(x,s) = \sup_{\mu} J(x,\mu,s) = J\left(x,\mu^*,s\right)$$

Steady State Control: Fisheries Management System

Consider again a closed loop system for a fish population $\{X_t : t \geq 0\}$ govern by the following SDE and total profit of harvesting defined by:

$$\mathrm{d}X_t = X_t(1 - X_t) - U_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}B_t, \quad X_0 = x,$$

$$J = \int_0^T \sqrt{U_t}\mathrm{d}t, \quad , \quad \text{i.e.} \quad h = \sqrt{U_t}$$

where $\{U_t\}$ is the catch rate. For a stationary control problem, we seek a time-invariant control strategy $U_t = \mu(X_t)$ such that the system admits a stationary solution $\{X_t\}_{t\in\mathbf{T}}$ within the time horizon $\mathbf{T}$.

We define L^u and now write the Hamilton-Jacob-Belmann equation becomes:

$$(L^u V)(x) = \frac{\partial V(x)}{\partial x} f(x,u) + \frac{1}{2}\,\mathrm{tr}\left[g^\top(x,u)\frac{\partial^2 V}{\partial x^2} g(x,u)\right].$$

$$\frac{\partial V}{\partial t} + \sup_{u\in\mathbf{U}}\left[L^u V + h\right] = \dot{V} + \sup_{u\geq 0}\underbrace{\left[V' x(1-x) - V' u + \frac{1}{2}\sigma^2 V'' + \sqrt{u}\right]}_{\textit{optim w.r.t } u} = 0$$

Finding the Optimal Strategy

Simple Optimization

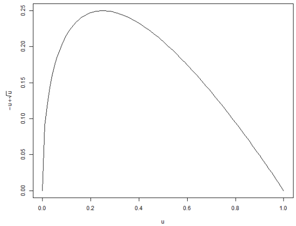
We collect terms only with u i.e.

$$\sup_{u \geq 0} \left[-V'u + \sqrt{u} \right]$$

This is a concave function if we assume $V' > 0$, hence we can easily find

$$-V' + 1/2u^* = 0 \qquad \Longrightarrow \qquad u^* = \frac{1}{4\left(V'\right)^2}$$

Visuals with $V' = 1$



Steady State Control: Fisheries Management System

We now have the HJB for the optimal control $u^* = \frac{1}{4(V')^2}$:

$$\dot{V} + V'x(1-x) + \frac{1}{4V'} + \frac{1}{2}\sigma^2 x^2 V'' = 0.$$

Inspiration from above solution of the form $V(x,t) = V_0(x) - \gamma t$ with $V_0(x) = \frac{1}{2}\log x + b$:

$$\dot{V} = -\gamma, \quad V' = \frac{1}{2x}, \quad V'' = -\frac{1}{2x^2}$$

$$-\gamma + \frac{1}{2}(1-x) + \frac{1}{2}x - \frac{1}{4}\sigma^2 = 0 \qquad \Longrightarrow \qquad \gamma = \frac{1}{2} - \frac{1}{4}\sigma^2$$

and we can easily find the optimal control as:

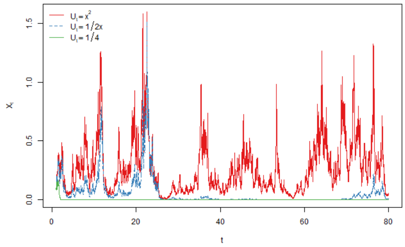
$$u = \mu^*(x) = \frac{1}{4\left(V'\right)^2} = x^2 \tag{0.20}$$

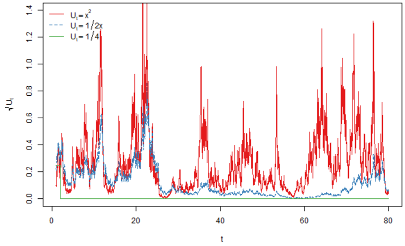
Steady State Control: Different Strategies one realization

Strategies Considered

In the following, we will consider the optimal strategy and some suboptimal strategies:

$$u_t^* = x^2$$
$$u_t = \frac{1}{2}x$$
$$u_t = \frac{1}{4}$$





Different Strategies Longer Term, Average over 300 Simulations