SDE Exam: Anton Ruby Larsen

DTU Compute
Institut for Matematik og Computer Science

(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 $(\Omega, \mathcal{F}, \mathbb{P})$ is a Brownian motion if:

1. $B_0 = 0$.
2. The increments are independent, i.e. $0 \leq s \leq t \leq u \leq v \Rightarrow B_t - B_s$ and $B_v - B_u$ are independent.
3. The increments are Gaussian, i.e. $0 \leq s \leq t \Rightarrow B_t - B_s \sim N(0, t-s)$.
4. For almost all realizations $\omega \in \Omega$, the sample path $t \mapsto B_t(\omega)$ is continuous.

Remark that the definition of Brownian motion did not tell whether such a process exists. Fortunately, we have the following result.

Theorem 4.2.1: Brownian motion exists

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

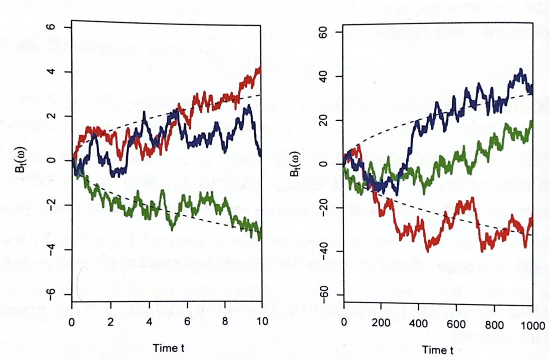
Self-similarity

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

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

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

$$\mathbb{E}|B_t|^p = (\sqrt{t})^p \mathbb{E}|B_1|^p.$$



Total Variation and Quadratic Variation

Unbounded total variation [Def. 4.3.1]

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

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

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

Bounded quadratic variation [Def. 4.3.2]

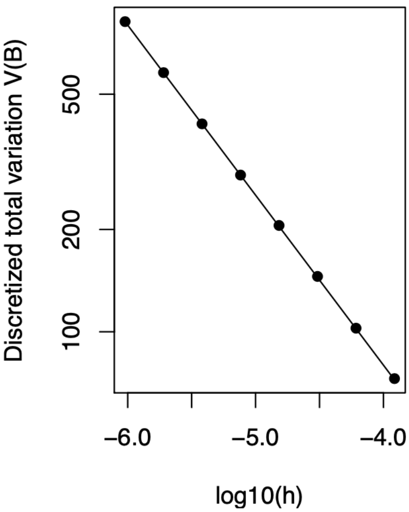
The quadratic variation of Brownian motion is given as

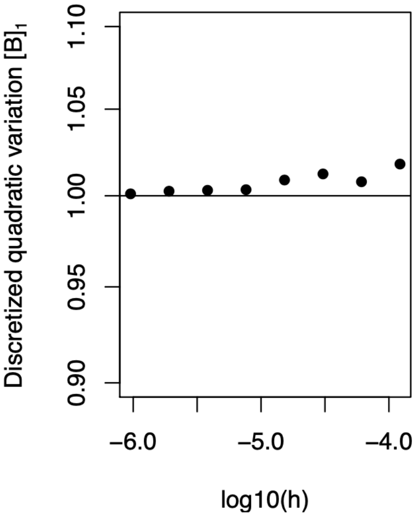
$$[B]_t = \lim_{|\Delta|\to 0} \sum_{i=1}^{\#\Delta} \left| B_{t_i} - B_{t_{i-1}} \right|^2 .$$

Cf. Thm. 4.3.1, the quadratic variation on the interval $[0, t]$, equals t. That is, $[B]_t = t$.

To appreciate these results, notice that for a differentiable function, the total variation on $[0, 1]$, is finite, and equals $\int_0^1 |f'(t)|\, dt$, while the quadratic variation is 0. We can therefore conclude that the sample paths of Brownian are nowhere differentiable w.p. 1.

Total Variation and Quadratic Variation





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, so that we obtain a family of σ-algebras in time, i.e.

$$\{\mathcal{F}_t\}_{t\in\mathbb{R}}.$$

Of special interest is accumulation of information without any loss. For this purpose, we define filtrations as

$$\mathcal{F}_s \subset \mathcal{F}_t \text{ when } t > s.$$

Martingales [Def. 4.5.1]

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

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

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

Brownian Motion is a Martingale

Let $\{\mathcal{F}_t\}_{t\geq 0}$ 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

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

We can similarly show that $\{B_t^2 - t\}_{t\geq 0}$ is a martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$. Condition 3 follows from the calculation

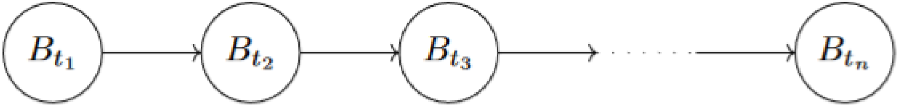
$$\begin{aligned}
\mathbb{E}[B_t^2 - t \mid \mathcal{F}_s] &= \mathbb{E}[(B_s + B_t - B_s)^2 \mid \mathcal{F}_s] - t \\
&= B_s^2 + 2B_s\mathbb{E}[B_t - B_s \mid \mathcal{F}_s] + \mathbb{E}[(B_t - B_s)^2 \mid \mathcal{F}_s] - t \\
&= B_s^2 + (t - s) - t \\
&= B_s^2 - s.
\end{aligned}$$

Brownian Motion is a Markov Process

Markov Process [Def 9.1.1]

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



Brownian Motion as a SDE

Brownian motion can be written the following SDE

$$\mathrm{d}X_t = \mathrm{d}B_t$$

, where we by use of the Itô integral can solve the SDE.

$$X_t = \int \mathrm{d}X_t = \int \mathrm{d}B_t = \lim_{|\Delta| \to 0} \sum_{i=0}^{\#\Delta} (B_{t_i} - B_{t_{i-1}}) = B_t$$

Hence, we see that Brownian motion is an Itô process with drift $f(X_t) = 0$ and diffusion $g(X_t) = 1$.

Backward and Forward Kolmogorov for Brownian Motion

We can use the SDE form of the Brownain motion to apply the backward Kolmogorov equation. We find

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

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

$$\dot{\phi} = L^*\phi = -(f\phi)' + (D\phi)'' = \frac{1}{2}\phi''.$$

In this case, $L^* = L$, but this will almost never be the case. In both cases, the probability density functions, ϕ and ψ, solve the classic heat equation.

We recall Fick's second law for diffusion with constant diffusivity, which has the same form as both the forward and backward. Hence, we see that Brownain motion is the path of a single particle out of many, only subject to diffusion, while Fick's second law describes the expansion of all particles.

Appendix

Maximum, Hitting time and null recurrent

Maximum of the Brownian

Define the maximum process

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

By Theorem 4.3.2, we have

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

Hitting time of the Brownian

Let $x > 0$ be an arbitrary point and τ the hitting time, i.e. $\tau = \inf\{t > 0 : B_t = x\}$. Then cf. thm 4.3.3 we have

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

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 $\mathbf{E}\tau = \infty$. Hence, it can be shown that Brownian motion is null recurrent, i.e. it always hits any given point on the real line, and always returns to the origin again, but the expected time until it does so is infinite.

Total Variation and Quadratic Variation

Note, for $X \sim \mathcal{N}(0, \sigma^2)$, we have

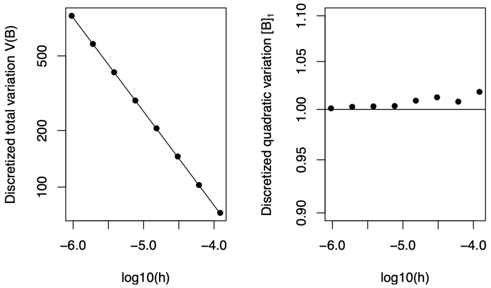
$$\mathbb{E}|X| = \int_{-\infty}^{\infty} |x| \frac{1}{\sigma} \phi\left(\frac{x}{\sigma}\right) \, \mathrm{d}x = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \int_0^{\infty} x \mathrm{e}^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \, \mathrm{d}x$$
$$= \sqrt{2\sigma^2/\pi}.$$

Hence, for Brownian increments of Δt, we get

$$\mathbb{E}|\Delta B| = \sqrt{2\Delta t/\pi}.$$

Assume a regular grid of n subintervals of length $\Delta t = 1/n$ on the time interval, $[0, 1]$, then

$$\mathbb{E}V_\Delta = \mathbb{E}\sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}| = n\sqrt{2/(n\pi)} = \sqrt{2n/\pi}.$$



Brownian Motion is a Markov Process

Markov Process [Def 9.1.1]

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

To show Brownian motion in Markov we need to show that $\mathbb{E}\left[h\left(B_t\right) \mid \mathcal{F}_s\right] = \mathbb{E}\left[h\left(B_t\right) \mid B_s\right]$. We write

$$E\left[h\left(B_t\right) \mid \mathcal{F}_s\right] = E\left[h\left(B_t - B_s + B_s\right) \mid \mathcal{F}_s\right] = E\left[g\left(B_t - B_s, B_s\right) \mid \mathcal{F}_s\right],$$

where $g(x, y) = h(x + y)$. We now know that B_s is $\mathcal{F}_s$-measurable and that $B_t - B_s$ is independent of $\mathcal{F}_s$. Hence we know that

$$E\left[g\left(B_t - B_s, B_s\right) \mid \mathcal{F}_s\right] = E\left[g\left(B_t - B_s, B_s\right) \mid B_s\right] = E\left[h\left(B_t\right) \mid B_s\right]$$

(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

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

Multivariate Linear System, cont'd

Proof.

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

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

We can now solve the system with initial condition $Y_0 = h(0, X_0) = x$.

$$Y_t = x + \int_0^t e^{-As} G dB_s$$

And back-transform by multiplying with e^{At}.

$$X_t = e^{At} Y_t = e^{At} x + \int_0^t e^{A(t-s)} G dB_s.$$

□

Mean and Variance

Differential equation governing the mean

Define $\bar{x}(t) = \mathbb{E}X_t$. By Fubini's theorem and knowing an Itô integral is martingale, we get $\bar{x}(t) = \exp(At)x$, 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

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma(t) = A\Sigma(t) + \Sigma(t)A^\top + GG^\top,$$

with initial condition $\Sigma(0) = 0$. This is the Lyaponov equation and given as equation 5.21.

Autocovariance Function

Autocovariance function

For a narrow multivariat linear system the autocovariance is given by,

$$\rho_X(s,t) = \int_0^s \exp(A(s-v))GG^\top \exp\big(A^\top(t-v)\big)\,\mathrm{d}v = \Sigma(s)\cdot\exp\big(A^\top(t-s)\big).$$

If the system is stationary we have

$$\rho_X(h) = \Sigma\exp\big(A^\top h\big) = \Sigma\exp\big(A^\top\big)^h.$$

Variance Spectrum

Variance spectrum

From the solution $\{X_t\}_{t\geq 0}$ and the superposition principle, we get the frequency response

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\exp\{-i\omega t\}dt = \int_{0}^{\infty} \exp(At)G\exp\{-i\omega t\}dt = (i\omega I - A)^{-1}G.$$

If all eigenvalues of A have negative real part, there exists a wide-sense stationary solutions $\{X_t\}_{t\geq 0}$ for which the variance spectrum is given by,

$$S_X(\omega) = \int_{-\infty}^{+\infty} \rho_X(h)\exp(-i\omega h)dh = H(-\omega)S_U(\omega)H^\top(\omega).$$

Example with Mass-Spring-Damper System, 1

Consider the standard-form system for $X_t = (Q_t, V_t)^\top$ where Q is position and V is velocity.

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

where

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

The system is being driven by a white noise force $\{U_t\}_{t\geq 0}$ with intensity $S_U(\omega) = \sigma^2$.

We solve the Lyapunov equation gouverning the variance, by setting up the matrices

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

We then solve the linear system, $-PM = \mathrm{Vec}(C)$. 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, 2

We find that

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

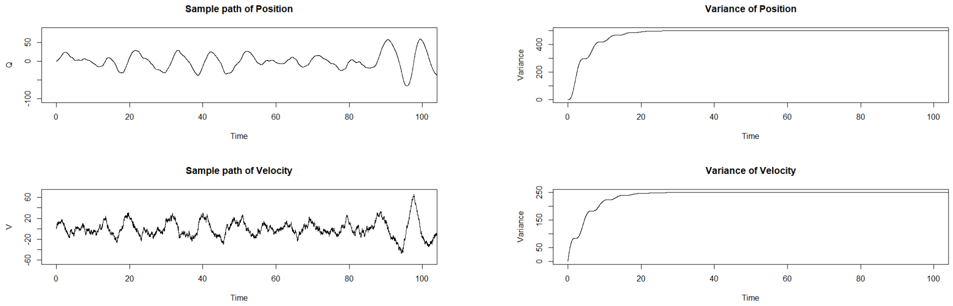
and the autocovariance can then be found from

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

Furthermore, we calculate the frequency response

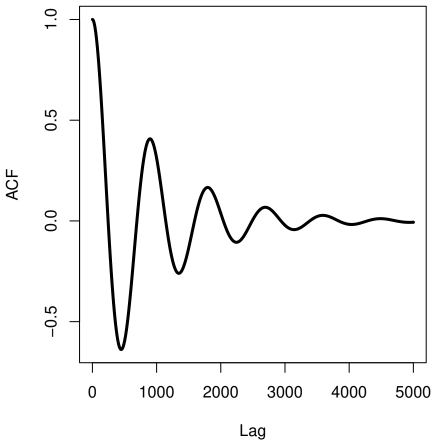
$$H(\omega) = (i\omega I - A)^{-1} G = \begin{bmatrix} \dfrac{\sigma}{ic\omega - m\omega^2 + k} \\ \dfrac{i\omega\sigma}{ic\omega - m\omega^2 + k} \end{bmatrix}.$$

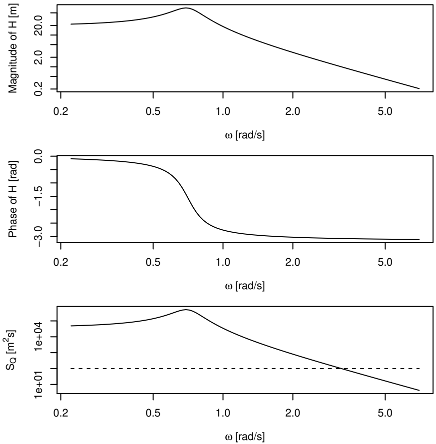
Example with Mass-Spring-Damper System, 3



- Parameters: $m = 1$, $k = 0.5$, $c = 0.2$, $\sigma = 10$.

Example with Mass-Spring-Damper System, 4





- 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

Specifying The Itô-Integral

Consider the following non-linear perturbed differential equation

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

where ξ is white noise, which formally is the "derivative" of Brownian motion, i.e., $\xi = \frac{\mathrm{d}B_t}{\mathrm{d}t}$.

We give the equation a precise meaning by integrating

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

The drift dependent integral is a classic Riemann integral while the intensity dependent integral we call an Itô integral which requires more work.

Defining The Itô integral

Thm 6.3.1: The 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 continous sample paths, which are adapted to $\{\mathcal{F}_t\}_{S \leq t \leq T}$, and for which $\int_S^T \mathbb{E}|G_t|^2 dt < \infty$.
Then the limit

$$I = \lim_{|\Delta| \to 0} \sum_{i=1}^{\#\Delta} G_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}),$$

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:

$$I = \int_S^T G_t dB_t$$

Hence by this theorem we have defined the meaning of the Itô SDE seen on the previous slide.

Numerical Approximation of The Itô Integral

The definition of the Itô integral is closely related to the Euler-Maruyama scheme

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

We now partition the time interval $[0, t]$ into n subintervals $0 = t_0 < t_1 < \ldots, < t_n = t)$. When we then apply the Euler-Maruyama scheme to each subinterval we get

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

When the time discretization becomes finer, we see that the diffusion term becomes the Itô integral

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

Properties of The Itô Integral

- *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 $\mathcal{F}_T$-measurable.
- *Continuity*: $\{I_t\}_{t\geq 0}$ is continuous in the mean square and almost surely.
- *The martingale property*: $\{I_t\}_{0\leq t\leq U}$, where $I_t = \int_0^t G_s \,\mathrm{d}B_s$, is a martingale w.r.t. $\{\mathcal{F}_t\}_{t\geq 0}$. (If the not $\mathcal{L}_2$ Itô integrable then only local martingale)
- *The Itō isometry*: $\mathbb{E}\left|\int_S^T G_t \,\mathrm{d}B_t\right|^2 = \int_S^T \mathbb{E}\left|G_t\right|^2 \,\mathrm{d}t$.

Integral of Brownian Motion

We want to solve the Itô integral $\int_0^t B_s \, \mathrm{d}B_s$. We hence calculate

$$\begin{aligned}
\int_0^t B_s \, \mathrm{d}B_s &= \lim_{n\to\infty} \sum_{i=1}^n B_{t_{i-1}} \cdot \left(B_{t_i} - B_{t_{i-1}} \right) \\
&= \lim_{n\to\infty} \frac{1}{2} \sum_{i=1}^n \left(B_{t_i}^2 - B_{t_{i-1}}^2 \right) - \frac{1}{2} \sum_{i=1}^n \left(B_{t_i} - B_{t_{i-1}} \right)^2 \\
&= \frac{1}{2} B_t^2 - \frac{1}{2} \lim_{n\to\infty} \sum_{i=1}^n \left(B_{t_i} - B_{t_{i-1}} \right)^2
\end{aligned}$$

We recognize $\lim_{n\to\infty} \sum_{i=1}^n \left(B_{t_i} - B_{t_{i-1}} \right)^2$ as the quadratic variation and hence it equals t. So we get

$$\int_0^t B_s \, \mathrm{d}B_s = \frac{1}{2} B_t^2 - \frac{1}{2} t$$

Integral of Brownian Motion

To see how Itô integrals are different from Riemann integrals we discritize the integral of Brownian Motion by the left and right end-point as well as the mid-point.

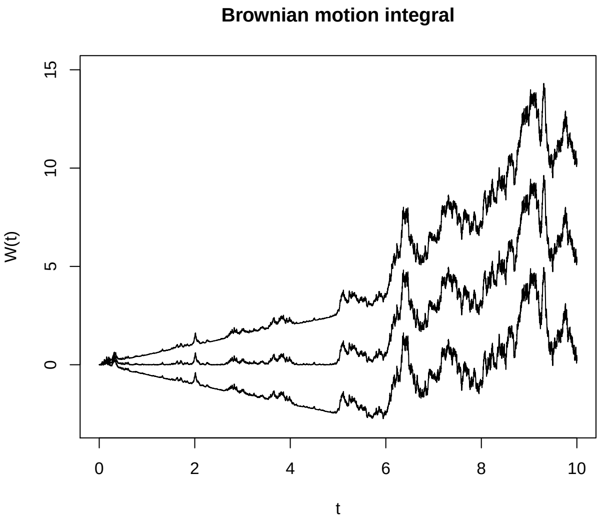
$$I_t^L = \sum_{i=1}^{n} B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) \quad , \quad I_t^R = \sum_{i=1}^{n} B_{t_i}(B_{t_i} - B_{t_{i-1}})$$

$$\text{and} \quad I_t^M = \sum_{i=1}^{n} \frac{1}{2}(B_{t_{i-1}} + B_{t_i})(B_{t_i} - B_{t_{i-1}}).$$

We approximate the integral by picking a finite n and then simulate it using the same approach as for the Euler-Maruyama method.

Integral of Brownian Motion

We have that in the limit $\lim_{n\to\infty} I^L = \frac{1}{2}B_t^2 - \frac{1}{2}t$, $\lim_{n\to\infty} I^R = \frac{1}{2}B_t^2 + \frac{1}{2}t$ and $\lim_{n\to\infty} I^M = \frac{1}{2}B_t^2$ which agrees with the numerical simulation.



(4) Itô's Lemma

Itô's Lemma - The Stochastic Chain Rule

Itô Process ($\mathcal{L}_2$ version) [Def 6.5.1]

Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$, we say that a process $\{X_t : t \geq 0\}$ given by

$$X_t = X_0 + \int_0^t F_s ds + \int_0^t G_s dB_s$$

is an $\mathcal{L}_2$ Itô process, provided that the initial condition X_0 is $\mathcal{F}_0$-measurable (for example, deterministic), and that $\{F_t\}$ and $\{G_t\}$ are adapted and have left-continuous sample paths and locally integrable variance, i.e. $\int_S^T \mathbb{E}|G_t|^2 dt < \infty$.

Itô's Lemma - The Stochastic Chain Rule

Thm 7.3.1: Itô's Lemma

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

$$dX_t = F_t dt + G_t dB_t$$

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

$$\begin{aligned} dY_t &= \dot{h} dt + \nabla h dX_t + \frac{1}{2} dX_t^\top \mathbf{H} h dX_t \\ &= \dot{h} dt + \left(\nabla h F_t + \frac{1}{2} \operatorname{tr} G_t^\top \mathbf{H} h G_t \right) dt + \nabla h G_t dB_t . \end{aligned}$$

Itô's Lemma - The Stochastic Chain Rule

We get from the first form to the second form of Itô's lemma by use of cross-variation and quadratic variation of Brownian motion. First cross-variation is given by

$$\langle X, Y\rangle_t = \lim_{|\Delta|\to 0} \sum_{i=1}^{\#\Delta} \left(X_{t_i} - X_{t_{i-1}}\right)\left(Y_{t_i} - Y_{t_{i-1}}\right)$$

And we have the following properties

- Let X_t be real-valued stochastic processes. Then $\langle X, X\rangle_t = [X]_t$.
- Let T be deterministic and B_t be Brownian motion. Then $\langle T, B\rangle_t = 0$.
- We can write $dL_t dM_t = d\langle L, M\rangle_t$.

We can now conclude that $dBdB = d[B_t] = t$, $dtdt = d[t]_t = 0$ and $dBdt = d\langle T, B\rangle_t = 0$

The Lamperti Transformation

This is a transformation of coordinates such that the noise becomes *additive*. It holds generally, but here we only consider a state dependent 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.

Norm of Multivariate Process

Consider the SDE

$$\mathrm{d}X_t = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} X_t \,\mathrm{d}t + \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} \mathrm{d}B_t, \quad X_0 = x.$$

Here, $X_t \in \mathbb{R}^2$ and $\{B_t\}_{t \geq 0}$ is two-dimensional Brownian motion. Take $S_t = \|X_t\|^2$. Using Itô's Lemma, and $\dot{h} = 0$, $\nabla h = 2X$, and $\nabla^2 h = 2$, we find

$$\mathrm{d}S_t = (2X_t^\top A X_t + \operatorname{tr} G^\top G)\,\mathrm{d}t + 2X_t^\top G \,\mathrm{d}B_t.$$

Multivariate Narrow-Sense Process

Consider the multivariate SDE given by

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

Here, $X_t \in \mathbb{R}^n$ and $\{B_t\}_{t\geq 0}$ is n-dimensional Brownian motion. Take $Y_t = h(X_t, t) = \mathrm{e}^{At}X_t$. Then the process $\{Y_t\}_{t\geq 0}$ is governed by the SDE

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

which is equivalent with

$$Y_t = x + \int_0^t \mathrm{e}^{As}G\,\mathrm{d}B_s,$$

since $Y_0 = h(X_0, 0) = x$.

Finding the Martingale

Dynkins Lemma: Let $h \in C_0^2(\mathbb{R}^n \to \mathbb{R})$, and $\mathbb{E}^x[\tau] < \infty$. Then

$$\mathbb{E}^x[h(X_\tau)] = h(x) + \mathbb{E}^x[\int_0^\tau Lh(X_s)ds]$$

Proof: Define $Y_t = h\left(X_t\right)$ and use Itô's lemma:

$$\mathrm{d}Y_t = \left(\frac{\partial h}{\partial x} f\left(X_t\right) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2} g^2\left(X_t\right)\right)\mathrm{d}t + \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t = Lh\left(X_t\right)\mathrm{d}t + \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t$$

And then integrate this from $t = 0$ to the provided stopping time $t = \tau$:

$$Y_\tau = Y_0 + \int_0^\tau Lh\left(X_t\right)\mathrm{d}t + \int_0^\tau \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t$$

All Itō integrals are martingales, so the latter integral's conditional expectation is zero for the "initial" condition X_0. The rest of the above equality is also a martingale:

$$0 = \mathbb{E}\left[Y_\tau - Y_0 - \int_0^\tau Lh\left(X_t\right)\mathrm{d}t \mid X_0\right]$$

Isolating this equation for $\mathbb{E}\left[Y_\tau \mid X_0\right]$ then gives Dynkin's formula.

(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

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

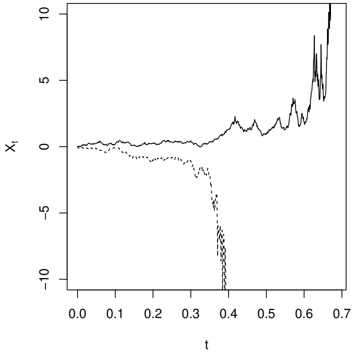
In the localized situation, it follows from the chain rule of Stratonovich calculus and $Y_t = h(x) = \arctan(x)$ that

$$X_t = \tan B_t.$$

This solution is well-defined until the stopping time

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

Explosions at τ



Non-uniqueness

In the case of non-uniqueness, singularities are a common reason. For simplicity, we have chosen an example where the singularity occurs 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 and $Y_t = h(x) = x^{1/3} + B_T$, we can show that

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

for some deterministic $T \geq 0$. Hence by picking T arbitarily we will obtain infinite solutions.

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

Sufficient Conditions for Existence

We have now seen an examples of non-existence and one of non-uniqueness. First, we will ensure existence of solutions.

Thm 8.3.1: Existence

Let the Itô process $\{X_t : 0 \leq t \leq T\}$ satisfy the initial value problem,

$$dX_t = f\left(X_t, t\right) dt + g\left(X_t, t\right) dB_t, \quad X_0 = x$$

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

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

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

Sufficient Conditions for Existence

Proof.

Define $S_t = |X_t|^2$. By Itô's Lemma, we have

$$\mathrm{d}S_t = 2X_t^\top \mathrm{d}X_t + |\mathrm{d}X_t|^2 = 2X_t^\top f(X_t, t)\,\mathrm{d}t + \mathrm{tr}[g^\top(X_t, t)g(X_t, t)]\,\mathrm{d}t + 2X_t^\top g(X_t, t)\,\mathrm{d}B_t.$$

Taking expectations using Proposition 6.5.1, we get

$$\begin{aligned}\mathbb{E}S_t &= S_0 + \mathbb{E}\int_0^t 2X_s^\top f(X_s, s) + \mathrm{tr}[g^\top(X_s, s)g(X_s, s)]\,\mathrm{d}s \\ &\le S_0 + \int_0^t 3C(1 + \mathbb{E}[S_s])ds.\end{aligned}$$

We recognize $v(t) = \mathbb{E}[S_t]$, $b(t) = S_0 + 3Ct$ and $a(t) = 3C$. We can hence use the Grönwall-Bellman inequality and get

$$\mathbb{E}S_t \le (S_0 + 3Ct)e^{3Ct}.$$

So "Linear bounds in the dynamics give exponential bounds in the solution". $\square$

Sufficient Conditions for Uniqueness

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

Thm 8.2.1: Uniqueness

Let $T > 0$ and assume that for any $R > 0$ there exist $K_f > 0, K_g > 0$ such that

$$|f(x,t) - f(y,t)| \leq K_f \cdot |x - y|, \quad |g(x,t) - g(y,t)| \leq K_g \cdot |x - y|$$

whenever $0 \leq t \leq T$ and $|x|, |y| < R$. Then there can exist at most one Itô process $\{X_t : 0 \leq t \leq T\}$ which satisfies the initial value problem (8.1), (8.2). That is, if $\{X_t : t \geq 0\}$ and $\{Y_t : t \geq 0\}$ are two Itô processes which both satisfy the initial value problem, then $X_t = Y_t$ for all t, almost surely.

The proof is left out, but it follows the same structure as for existence but also includes the Cauchy-Schwartz inequality.

Examples of Existence and Uniqueness

We will now consider a few examples.

Bounded derivatives

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

Note, all wide-sense linear SDEs are necessarily locally Lipschitz, and hence has a unique solution.

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 superlinear, however, $xf(x) \leq 0$ for all x. This is because the superlinearity is stabilizing, i.e. directed towards the origin.

Euler-Maruyama and Milstein

Euler-Maruyama:

$$X_{t+h}^{(h)} = X_t^{(h)} + f(X_t, t) X_t^{(h)} h + g(X_t, t) X_t^{(h)} \left(B_{t+h} - B_t\right), \quad X_0^{(h)} = x.$$

Strong and weak order

$$\mathbb{E}[||x - X||] \leq Ch^{\gamma} \quad |\mathbb{E}[x] - \mathbb{E}[X]| \leq Ch^{\delta}$$

Convergence of the Euler-Maruyama scheme:

$$\text{Strong order: } \gamma = 0.5, \text{ Weak order: } \delta = 1$$

Largest error:

$$gg' \int_0^h B_s dB_s = gg' (\left(B_h^2 - h\right) / 2)$$

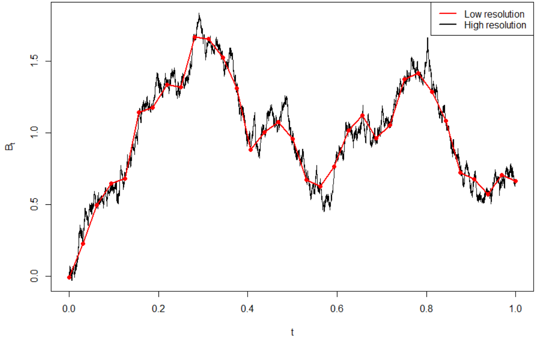
Milstein:

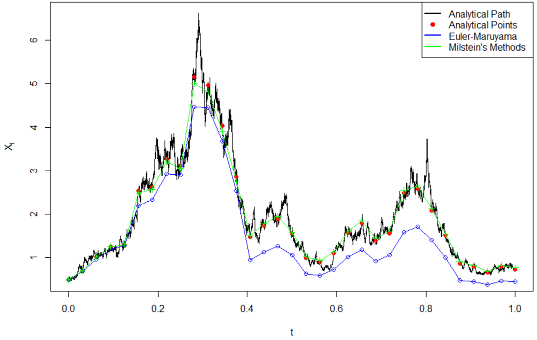
$$X_{t+h} = X_t + fh + \sum_{k=1}^{m} g_k \Delta B^{(k)} + \sum_{k,l=1}^{m} \nabla g_k g_l \int_t^{t+h} B_s^{(l)} dB_s^{(k)}$$

Milstein for scalar noise:

$$X_{t+h} = X_t + f\left(X_t\right) h + g\left(X_t, t\right) \Delta B + \frac{1}{2} g'\left(X_t\right) g\left(X_t\right) \left[(\Delta B)^2 - h\right]$$

Examples of Milstein and Euler-Maruyama





- $dX_t = \frac{1}{2} X_t \mathrm{d}t + \frac{3}{2} X_t \mathrm{d}B_t$
- $h = 2^{-3}$

(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 $L \geq s \geq 0$, then a.s.

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

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 stochastic differential equation

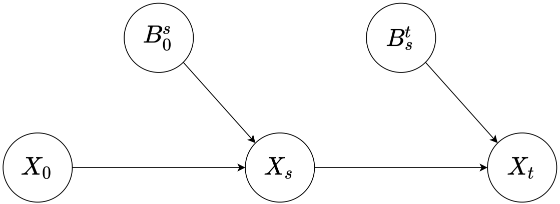
$$dX_t = f\left(X_t, t\right) dt + g\left(X_t, t\right) dB_t$$

where $\{B_t : t \geq 0\}$ is Brownian motion with respect to a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}_t\})$, the initial condition X_0 is $\mathcal{F}_0$-measurable, and f and g satisfy existence and uniqueness. Then the process $\{X_t : t \geq 0\}$ is Markov with respect to $\{\mathcal{F}_t\}$ as well as with respect to its own filtration.

Markov Processes

Proof.

Let B_s^t denote $\{B_u - B_s\}_{s \leq u \leq t}$. Then, for $0 \leq s \leq t$, consider



X_t can be computed from X_s and the noise between time s and t. From the property of independent increments of Brownian motion, the figure applies. $\square$

Forward and Backward Kolmogorov Equations

If we fix (s, x), we get the density

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

Theorem 9.5.1: Forward Kolmogorov equation

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

$$\begin{aligned}\dot{\phi} &= L^* \phi \\ &= -\nabla \cdot (f\phi) + \nabla \cdot \nabla(D\phi) \\ &= -\nabla(u\phi - D\nabla\phi)\end{aligned}$$

where again $D = \frac{1}{2} g g^\top$ and $u = f - \nabla D$.

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

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

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 if the transmission density, $p(s \mapsto t, x \mapsto y)$, exists, the backward Kolmogorov equation with fixed terminal condition (t, y), is given by

$$\begin{aligned} -\dot{\psi} &= L\psi \\ &= \nabla\psi \cdot f + \mathrm{tr}(D\mathrm{H}\psi) \\ &= \nabla\psi u + \nabla(D\nabla\psi), \end{aligned}$$

where $D = \frac{1}{2} g g^\top$, $u = f - \nabla D$ and H is the Hessian.

Adjoint Operators

Below we apply integration of parts under compact support: $\int u \mathrm{d}v = [uv]_{-\infty}^{\infty} - \int v \mathrm{d}u = -\int v \mathrm{d}u$

$$
\begin{aligned}
\langle v, L^* u \rangle &= \int v(x) L^*(u(x)) \mathrm{d}x \\
&= \int v(x) \left(-\frac{\partial}{\partial x}(f(x)u(x)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g^2(x)u(x)\right) \mathrm{d}x \\
&= -\int v(x) \frac{\partial}{\partial x}\left(f(x)u(x)\right) \mathrm{d}x + \frac{1}{2}\int v(x)\frac{\partial^2}{\partial x^2}\left(g^2(x)u(x)\right) \mathrm{d}x \\
&= \int \frac{\partial v(x)}{\partial x} f(x)u(x)\mathrm{d}x + \frac{1}{2}\int \frac{\partial^2 v(x)}{\partial x^2} g^2(x)u(x)\mathrm{d}x \\
&= \int \left(\frac{\partial v(x)}{\partial x} f(x) + \frac{1}{2}\int \frac{\partial^2 v(x)}{\partial x^2} g^2(x) \right) u(x)\mathrm{d}x \\
&= \int L(v(x))u(x)\mathrm{d}x = \langle Lv, u \rangle
\end{aligned}
$$

Numerical Implementation

The transition probabilities can only be found analytically in special cases so we want to approximate numerically. We discritize space while keeping time continuous which results in

$$\dot{\phi} = L^* \phi \quad \mapsto \quad \dot{\bar{\phi}} = \bar{\phi} G \quad \wedge \quad -\dot{\psi} = L\psi \quad \mapsto \quad -\dot{\bar{\psi}} = G\bar{\psi}.$$

where $\bar{\phi}, \bar{\psi} \in \mathbb{R}^N$ and $G \in \mathbb{R}^{N \times N}$ is the generator matrix. Hence we reduced the SDE to a continuous-time Markov chain.
To obtain this discritization we apply a finite volume method.

1. We truncate the real axis, $\mathbb{R} \ \mapsto \ [a, b]$ and partition into N grid cells $\{I_i : i = 1, \dots, N\}$
2. We let x_i be the center point in each grid cell and use $x_{i-1/2}$ and $x_{i+1/2}$ as boundary points in each grid cell. Thus $x_i = \left(x_{i-1/2} + x_{i+1/2}\right)/2$ and $|I_i| = x_{i+1/2} - x_{i-1/2}$.

Numerical Implementation

❸ We build the interior of the generator G

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

$$G_{(i+1)i} = \begin{cases} \frac{D\left(x_{i+1/2}\right)}{|I_{i+1}|(x_{i+1}-x_i)} - u\left(x_{1+1/2}\right) / |I_{i+1}| & \text{when } u\left(x_{i+1/2}\right) < 0, \\ \frac{D\left(x_{i+1/2}\right)}{|I_{i+1}|(x_{i+1}-x_i)} & \text{else.} \end{cases}$$

$$G_{ii} = -G_{i(i+1)} - G_{i(i-1)}$$

Numerical Implementation

4. To finish the generator G we must decide on a boundary condition. This could for example be reflective boundaries, $G_{1(1)} = -G_{1(2)}$ and $G_{N(N)} = -G_{N-1(N)}$
5. We can now solve the system of N ODEs. If we have a time-invariant system the solution to the forward equation is

$$\phi_t = \phi_0 \exp(Gt)$$

Another task could be to find the stationary solution which is given by the two equations $\phi G = 0$ and $\phi \mathbf{e} = 1$ which can be written as a system of equations as

$$[\phi \delta] \begin{bmatrix} G & \mathbf{e} \\ \mathbf{e}^T & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & 1 \end{bmatrix}$$

Numerical Implementation

As an example, consider the CIR process

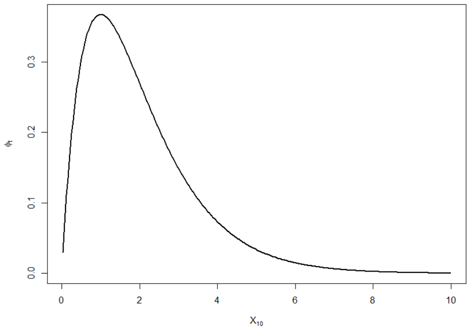
$$\mathrm{d}X_t = \lambda(\xi - X_t)\,\mathrm{d}t + \gamma\sqrt{X_t}\,\mathrm{d}B_t,$$

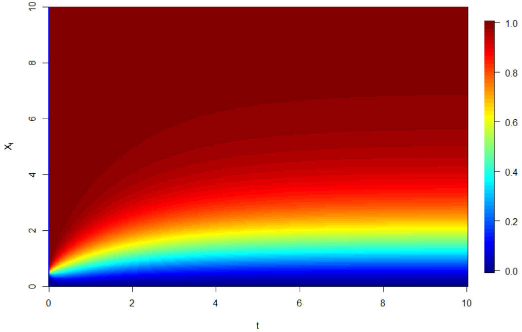
where λ, ξ and γ are positive parameters.
In our particular example we have

$$\lambda = 1/2, \quad \xi = 2 \text{ and } \gamma = 1$$

and we discritize space in $N = 300$ grid cells spanning the interval $x \in [0, 10]$.

Numerical Implementation



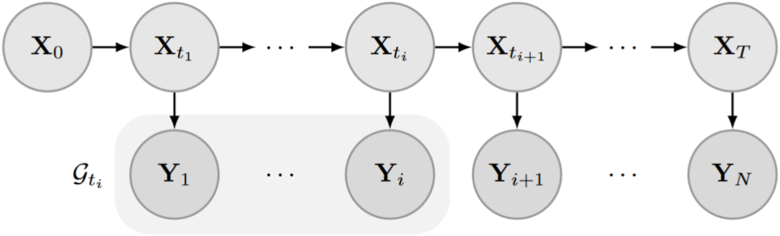


- Left plot: Steady state p.d.f
- Right plot: Time development of c.d.f

(7) State Estimation

Recursive filtering

The idea behind recursive filtering is to view the problem as a Hidden Markov Model.



Note, the continuous process $\{X_t\}_{t\geq 0}$ is sub-sampled at time t_i to constitute a discrete-time Markov Process $\{X_{t_i}\}_{t_i\in\mathbb{N}}$, which we aim to infer indirectly from the measurements $\{Y_i\}_{i\in\mathbb{N}}$. We note that $\{\mathcal{G}_t\}_{t\geq 0}$ is the filtration generated by $\{Y_i\}_{i\in\mathbb{N}}$, i.e.

$$\mathcal{G}_t = \sigma(\{Y_i\}_{t_i\leq t}).$$

The Setup

Our state is given by some SDE

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

which we only partially observe through measurements,

$$y_i = c(X_{t_i}) + s(X_{t_i})\varepsilon.$$

Hence we can construct a conditional distribution of the state,

$$f_{X_{t_i}|Y_i}\left(x \mid y_i\right) = \frac{1}{c} f_{X_{t_i}}(x) f_{Y_i|X_{t_i}}\left(y_i \mid x\right),$$

where $f_{X_{t_i}}$ is given through our SDE and $f_{Y_i|X_{t_i}}\left(y_i \mid x\right)$ is the state likelihood. If we had assumed $\varepsilon \sim \mathcal{N}(0, 1)$ and independent of the Brownian motion then the state likelihood would have the form

$$l_i(x) = \frac{1}{\sqrt{2\pi} s(x)} \exp\left(-\frac{1}{2}\frac{\left(y_i - c(x)\right)^2}{s(x)^2}\right)$$

The Algorithm

We define two distributions: $\phi_i \in \sigma(\{Y_i\}_{t_0 \le s \le t_{i-1}})$ and $\psi_i \in \sigma(\{Y_i\}_{t_0 \le s \le t_i})$
And we define two likelihoods: $\mu_i \in \sigma(\{Y_i\}_{t_{i-1} \le s \le t_N})$ and $\lambda_i \in \sigma(\{Y_i\}_{t_i \le s \le t_N})$

Forward in time

1. Init with t_0, $\psi_0(\cdot)$ and set $i = 0$.
2. *Time Update(Find ϕ_{i+1}):* We solve the Forward Kolmogorov equation

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

   and set $\phi_{i+1} = \rho(x, t_{i+1})$
3. *Data Update(Find ψ_{i+1}):* Solve Bayes

$$\psi_{i+1}(x) = \frac{1}{\int \phi_{i+1}(x) l_{i+1}(x)\, \mathrm{d}x} \phi_{i+1}(x) l_{i+1}(x),$$

Backwards in time

1. Init with t_N, $\lambda_N(x) = l_N(x)$ and set $i = N$.
2. *Time Update(Find μ_{i-1}):* We solve the Backward Kolmogorov equation

$$-\dot{h} = u \cdot \nabla h + \nabla \cdot (D\nabla h), \quad h(x, t_i) = \lambda_i(x),$$

   and set $\mu_{i-1} = h(x, t_{i-1})$
3. *Data Update(Find λ_{i-1}):* Likelihood update

$$\lambda_{i-1}(x) = \mu_{i-1}(x) l_{i-1}(x)$$

Posterior Distribution

Having completed the recursion, we find the posterior distribution

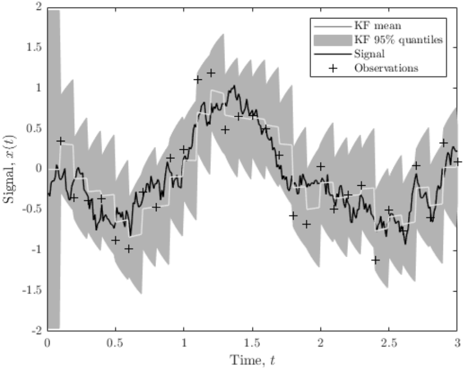
$$\pi_i(x) = \frac{1}{k_i}\psi_i(x)\mu_i(x)$$

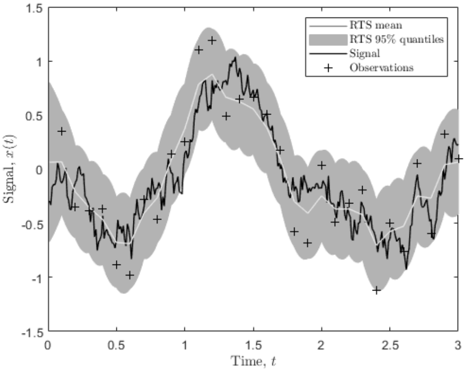
which is the conditional distribution of X_{t_i} given all measurements, past, present, and future. Here k_i is a normalization constant ensuring that π_i integrates to 1 and is given by

$$k_i = \int_X \psi_i(x)\mu_i(x)dx$$

We can use the posterior distribution, $\pi_i(x)$, to sample "typical tracks"to make Monte Carlo estimates of statistics that are otherwise difficult to compute.

Recursive filtering, example





- State: $\mathrm{d}X = -1/2 X_t \mathrm{d}t + \mathrm{d}B_t$
- Observations: $Y_t = X_t + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, 1/10)$

Likelihood Estimation

We can apply filters to estimate unknown parameters.

Assume our SDE model depends on some unknown parameter θ. Now we can find the likelihood function

$$\begin{aligned}
\Lambda(\theta) &= f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N;\theta) \\
&= f_{Y_1}(y_1;\theta)\, f_{Y_2|Y_1}(y_1,y_2;\theta)\, f_{Y_3,\ldots,Y_N|Y_1,Y_2}(y_1,\ldots,y_N;\theta) \\
&= \prod_{i=1}^{N} f_{Y_i|Y_1,\ldots,Y_{i-1}}(y_1,\ldots,y_i;\theta) \\
&= \prod_{i=0}^{N-1} \int \phi_{i+1}(x) l_{i+1}(x)\,\mathrm{d}x = \prod_{i=0}^{N-1} c_{i+1}(x)
\end{aligned}$$

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

Example

We consider the Cox-Ingersoll-Ross model

$$dX_t = \lambda \cdot (\xi - X_t)\, dt + \gamma \sqrt{X_t} dB_t \quad \text{with } \lambda = \xi = \gamma = 1.$$

We assume that ξ is *unknown*, while λ and γ are *known*. The measurements are count data following a Poisson distribution with conditional mean $\mathbb{E}[Y_t \mid X_t] = v \cdot X_t$. This results in the following state likelihood

$$\mathbb{P}\left\{Y_{\iota_i} = y \mid X_{\iota_i} = x\right\} = e^{-xv} \frac{(vx)^y}{y!}$$

where $v = 0.5$.
We have 201 observations from a simulation of the Cox-Ingersoll-Ross model with $\xi = 1$. We will try to estimate the parameter $\xi = 1$ from these observations.

Example

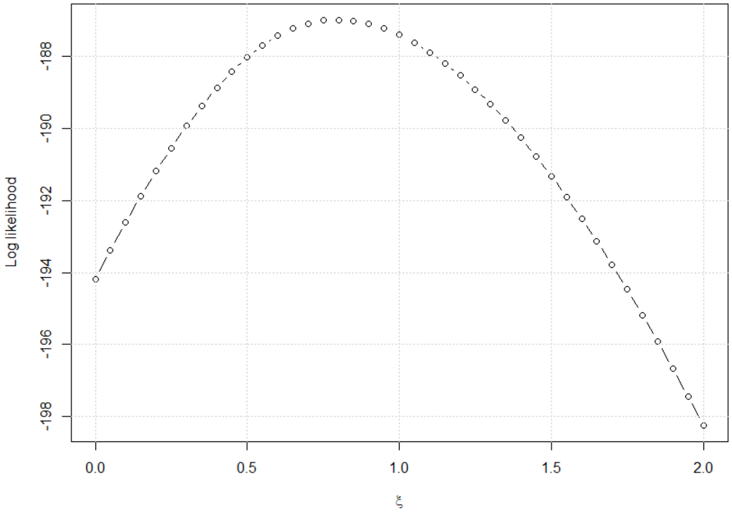
We can calculate the log likelihood by performing a forward pass.

Here we can not reduce to algebraic equations so we approximate the SDE with a continuous-time Markov chain.

```r
## Filter
hmmfilter <- function(G)
{
    phi <- Matrix(array(0,c(length(obs$t),length(xc))))
    psi <- phi

    ## Initialize with the stationary distribution
    mu <- StationaryDistribution(G)
    mu <- mu/sum(mu)

    ## Compute transition probabilities
    P <- expm(G*tsample)
    const <- numeric(length(obs$t))

    phi[1,] <- mu

    ## Include the first data update
    psi[1,] <- phi[1,] * ltab[,1]
    const[1] <- sum(psi[,1] )
    psi[,1] <- psi[,1] / const[1]

    ## Main time loop over the remaining time steps
    for(i in 2:length(obs$t))
    {
        phi[i,] = psi[i-1,] %*% P     # Time update
        psi[i,] = phi[i,] * ltab[,i] # Data update
        const[i] <- sum(psi[i,])      # Normalization
        psi[i,] = psi[i,] / const[i]
    }

    return(list(c=const,phi=as,matrix(phi),psi=as.matrix(psi),loglik=sum(log(const)))
}

## Run the filter with default values for parameters
est <- hmmfilter(G)
```

Example

The log-likelihood estimate is found to be $\xi = 0.8$.



Appendix

Kalman Filter

In the case of linear stochastic differential equations with linear observations things simplifies greatly. The foward pass is given by the Kalman filter and the backward pass is given by the Rauch-Tung-Striebel smoother. Here we have given the Kalman filter.

Time update

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

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

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

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

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

(8) Applications of Dynkin's Lemma

Dynkin's Lemma

Theorem 11.1.1: Dynkin's Lemma

Let $h \in C_0^2(\mathbb{R}^n \to \mathbb{R})$, and $\mathbb{E}^x[\tau] < \infty$. Then

$$\mathbb{E}^x[h(X_\tau)] = h(x) + \mathbb{E}^x[\int_0^\tau Lh(X_s)ds]$$

Dynkin's Lemma

Proof.

Define $Y_t = h\left(X_t\right)$ and use Itô's lemma:

$$\mathrm{d}Y_t = \left(\frac{\partial h}{\partial x} f\left(X_t\right) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2} g^2\left(X_t\right)\right)\mathrm{d}t + \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t = Lh\left(X_t\right)\mathrm{d}t + \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t$$

And then integrate this from $t = 0$ to the provided stopping time $t = \tau$:

$$Y_\tau = Y_0 + \int_0^\tau Lh\left(X_t\right)\mathrm{d}t + \int_0^\tau \frac{\partial h}{\partial x} g\left(X_t\right)\mathrm{d}B_t$$

All Itō integrals are martingales, so the latter integral's conditional expectation is zero for the "initial"condition X_0. The rest of the above equality is also a martingale:

$$0 = \mathbb{E}\left[Y_\tau - Y_0 - \int_0^\tau Lh\left(X_t\right)\mathrm{d}t \mid X_0\right]$$

Isolating this equation for $\mathbb{E}\left[Y_\tau \mid X_0\right]$ then gives Dynkin's formula. □

Applications of Dynkin's Lemma

Applications

We assume that $h \in C_0^2(\mathbb{R}^n \to \mathbb{R})$, the diffusion is regular and $\mathbb{E}^x[\tau] < \infty, \quad \forall x \in \Omega$

Thm	Usage	Interior condition	Boundary condition	Then
11.2.1	Time to exit	$Lh(x) + 1 = 0$	$h(x_\tau) = 0$	$h(x) = \mathbb{E}^x[\tau]$
11.4.1	Point of exit	$Lh(x) = 0$	$h(x_\tau) = c(x)$	$h(x) = \mathbb{E}^x[c(x_\tau)]$
11.6.1	Total reward until exit	$Lh(x) + r(x) = 0$	$h(x_\tau) = c(x)$	$h(x) = \mathbb{E}^x[c(x_\tau) + \int_0^\tau r(x_t)\mathrm{d}t]$

Theorem 11.2.2: Regular diffusion

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 $x \in \mathbf{\Omega}$.

Point of Exit and Time to Exit

We consider the Ornstein-Uhlenbeck process $\mathrm{d}X_t = -X_t\mathrm{d}t + \sqrt{2}\mathrm{d}B_t$ on the domain $]-l, l[\subset \mathbb{R}$. We define the time of exit as $\tau = \inf\{t \geq 0 \mid X_t \notin]a, b[\}$ and aim to find time to exit, $\mathbb{E}^x[\tau]$ and point of exit $\mathbb{P}^x[X_\tau = l]$.

First we focus on $\mathbb{P}^x[X_\tau = l]$. We use thm 12.4.1 and get the following problem

$$-xh' + h'' = 0 \quad \text{for} \quad x \in]-l, l[, \quad h(-l) = 0, \quad h(l) = 1,$$

This has the full solution $h(x) = c_1 s(x) + c_2$ where $s(x)$ is the scale function which is given by

$$s(x) = \int^x \phi(y)\,\mathrm{d}y,$$

where $\phi(x) = \exp\left(\int^x \frac{-2f(y)}{g^2(y)}\,\mathrm{d}y\right)$. We insert boundary conditions and get

$$\mathbb{P}^x(X_\tau = b) = h(x) = \frac{s(x) - s(-1)}{s(l) - s(-1)} = \frac{\operatorname{erfi}\left(\frac{x\sqrt{2}}{2}\right) + \operatorname{erfi}\left(\frac{l\sqrt{2}}{2}\right)}{2\operatorname{erfi}\left(\frac{l\sqrt{2}}{2}\right)}.$$

Point of Exit and Time to Exit

Next we focus on the time to exit, $\mathbb{E}^x[\tau]$. We use thm 12.2.1 and get the following problem

$$1 - xh' + h'' = 0 \text{ for } x \in]-l, l[, \quad h(-l) = 0, \quad h(l) = 0,$$

To solve this equation, we first reduce it to a first order equation. Define $k = h'$, then

$$k'(x) - xk(x) + 1 = 0$$

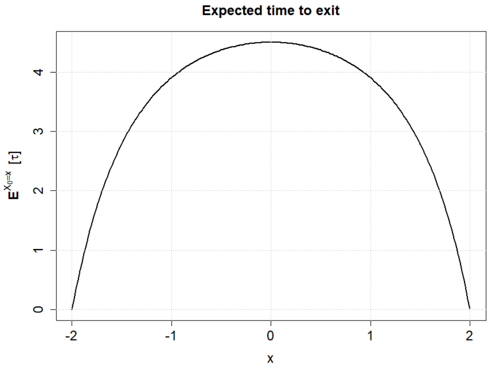
and using symmetry $k(0) = 0$, we find

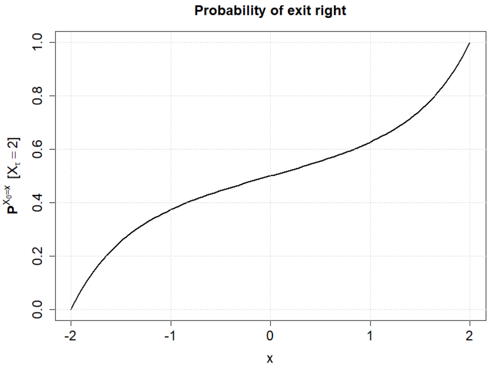
$$k(x) = -\int_0^x e^{x^2/2 - y^2/2} dy = -\sqrt{2\pi} e^{x^2/2} \left(\Phi(x) - \frac{1}{2} \right)$$

where $\Phi(x)$ is the c.d.f. of a standard Gaussian random variable. We therefore get

$$h(x) = h(0) - \sqrt{2\pi} \int_0^x e^{y^2/2} \left(\Phi(y) - \frac{1}{2} \right) dy.$$

Point of Exit and Time to Exit





- Left: Time to exit, $\mathbb{E}^x[\tau]$.
- Right: Probability of exit to the right, $\mathbb{P}^x[X_\tau = l]$.

Singular Boundary Point

We consider the Bessel process given by

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

On the domain $\Omega = (a, b)$ with $0 < a < b$. We aim to find the probability of exiting at b, $\mathbb{P}^x[X_\tau = b]$. As earlier we again use thm 12.4.1 and the scale function to solve for h.

$$s(x) = \int^x \exp\left(\int^y -\frac{2\mu}{\sigma^2 z} dz\right) dy = \frac{1}{\nu} x^\nu,$$

when $\nu := 1 - 2\mu/\sigma^2 \neq 0$. So the probability of exit at b is

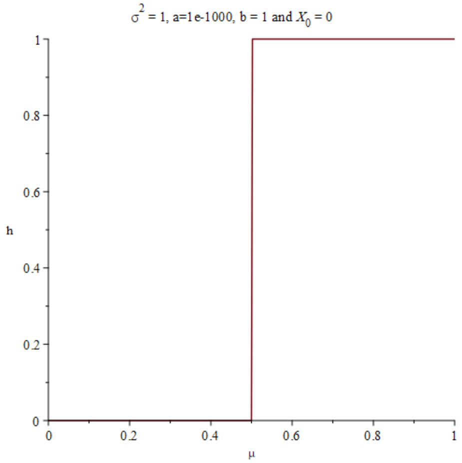
$$h(x) = \mathbf{P}^x \left\{X_\tau = b\right\} = \frac{x^\nu - a^\nu}{b^\nu - a^\nu}.$$

Singular Boundary Point

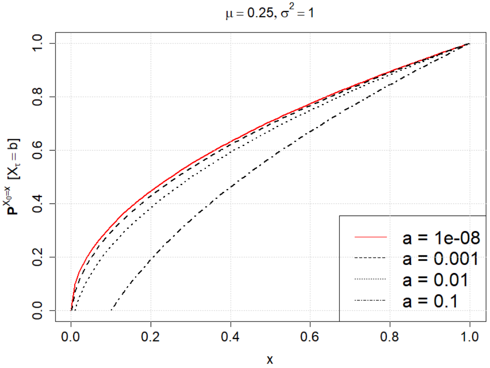
We now want to investigate if it is possible to reach the singularity at $x = 0$. We fix $\sigma^2 = 1$ and plot the value of h for $x = 0$. We see when $\mu > 1/2$ the drift is stronger than the diffusion. Hence we exit to the right.

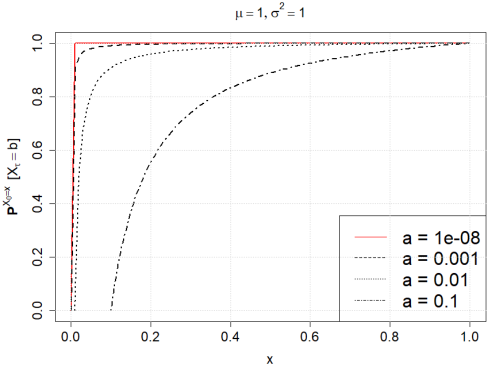
When $\mu < 1/2$ we will exit to the left if we start at $x = 0$.

Hence we conclude that if $\mu > \frac{1}{2}\sigma^2$ then we will never be able to reach the singularity at $x = 0$.



Singular Boundary Point





- Left: We see that we have a probability of $1 - \sqrt{x}$ of exiting to the left because $\nu = 1/2$.
- Right: We see that we have a probability of 0 of exiting to the left because $\nu = -1/2$.

(9) Stochastic Stability Theory

Stability of Stochastic Differential Equations

As always, we consider a general non-linear SDE

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

where $X_t \in \mathbb{R}^n$ and $B_t \in \mathbb{R}^m$. We propose the question:

- ***How sensitive are trajectories to perturbations in the initial condition?***

To clarify the dependency on the initial condition, we introduce the state transition map $\Phi_t(x) \in \mathbb{R}^n$

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

Assuming differentiability, we define the sensitivity $S_t(x) \in \mathbb{R}^{n\times n}$ by

$$S_t(x) = \frac{\partial \Phi_t(x)}{\partial x},$$

which satisfy the sensitivity equation

$$\mathrm{d}S_t(x) = \frac{\partial f}{\partial x}(\Phi_t(x))S_t(x)\,\mathrm{d}t + \sum_{i=1}^{m} \frac{\partial g_i}{\partial x}(\Phi_t(x))S_t(x)\,\mathrm{d}B_t^{(i)}, \quad S_0(x) = I.$$

Stochastic Lyapunov Exponent

To summarize $S_t(x)$ in one number, take the operator norm $\sup\{\|S_t(x)\tilde{x}\| \,|\, \tilde{x} \in \mathbb{R}^n, \|\tilde{x}\| = 1\}$. This is exactly 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.

Application of Stochastic Lyapunov Exponent

Consider the geometric Brownian motion

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

We have that $f'(0) = r$ and $g'(0) = \sigma$ which results in the following sensitivity equation

$$\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}B_t.$$

We see that the sensitivity equation of geometric Brownian motion is also a geometric Brownian motion. Hence the solution is well known

$$S_t = \mathrm{e}^{\left(r-\frac{1}{2}\sigma^2\right)t+\sigma B_t}.$$

We calculate the finite-time Lyapunov exponent

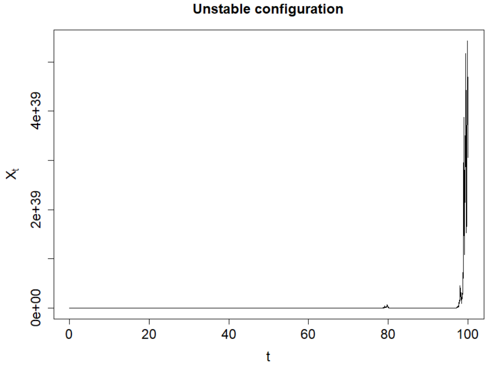
$$\lambda_t = \frac{1}{t}\log(\bar{\sigma}(S_t)) = \frac{1}{t}\log\left(\mathrm{e}^{\left(r-\frac{1}{2}\sigma^2\right)t+\sigma B_t}\right) = r - \frac{1}{2}\sigma^2 + \frac{1}{t}\sigma B_t.$$

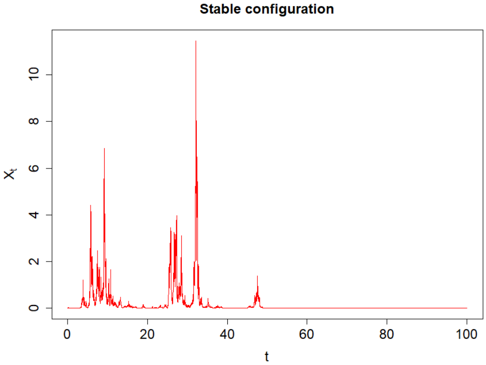
Since $\{B_t\}_{t\geq 0}$ scales with $\sqrt{t}$, we get $\bar{\lambda} = \lim_{t\to\infty}\sup r - \sigma^2/2 + \frac{1}{t}\sigma B_t = r - \sigma^2/2.$

Application of Stochastic Lyapunov Exponent

We just calculated that to ensure stable sample paths for the geometric Brownian motion we must have $r < \sigma^2/2$.

- Stable case: $r = 1.2$ and $\sigma = 2$, so $\bar{\lambda} = -0.8$.
- Unstable case: $r = 2.8$ and $\sigma = 2$, so $\bar{\lambda} = 0.8$.





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

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

Then

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

The second assumptions makes the conclusion useful due to the continuity and the third assumption ensures $V(x)$ is a supermartingale. Hence, cf. thm. 12.7.2, giving the stochastic Lyapunov function the same properties, on average, as the deterministic counter part.

Application of Stochastic Lyapunov Functions

We again consider the geometric Brownian motion

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

We try the Lyapunov function $V(x) = |x|^p$ for $p > 0$. It satisfy condition 1 and 2 in thm. 12.7.1 trivially. We check condition 3.

$$\begin{aligned} LV(x) &= f(x)V'(x) + DV''(x) \\ &= rX_t p|X_t|^{p-1} + \frac{1}{2}\sigma^2 X_t^2 p(p-1)|X_t|^{p-2} \\ &= p|X_t|^p(r + (p-1)\sigma^2/2) \end{aligned}$$

We see that for $LV(x) \leq 0$ it must hold that $r + \sigma^2(p-1)/2 > 0$ or $p \leq 1 - 2r/\sigma^2$. Because $p > 0$ such p only exists if $1 - 2r/\sigma^2 > 0$. We see that by a bit of algebra $1 - 2r/\sigma^2 > 0 \Leftrightarrow r < \sigma^2/2$ as we concluded by use of Lyapunov exponents.

Other types of Stability

Until now we have only considered sample path stability as in the deterministic theory. This is though not the only notion of stability we have in the stochastic theory. Every mode of a process can be stable or unstable independent of the stability of the sample path.

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.

We again consider geometric Brownian motion. We now choose $V(x) = x^2/2$. We calculate $LV(x) = (r + \sigma^2/2)x^2$ hence for $(r + \sigma^2/2)x^2 \leq -k_3 x^2$ it must hold that $r + \sigma^2/2 < 0$.

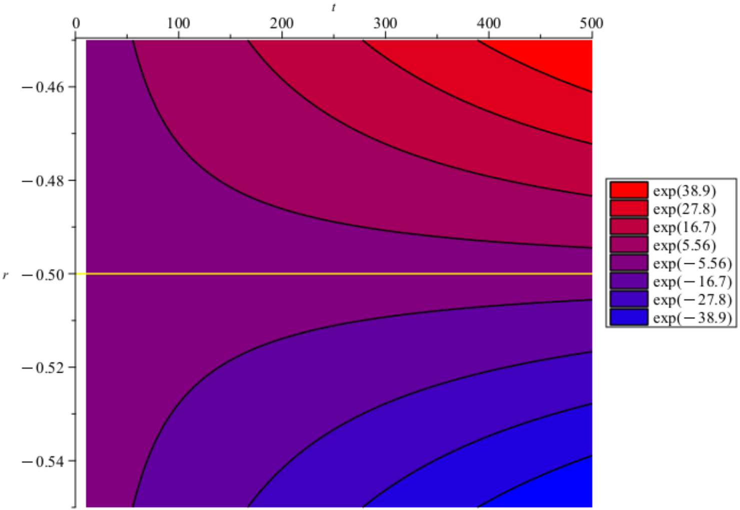
Mean square stability of GBM

We illustrate a GBM with $\sigma = 1$ meaning for the process to be mean square stable

$$r < -\sigma^2/2 = -1/2$$

. We see that the variance explodes for

$$r > -1/2$$



Appendix

Stationarity and Stability

Lastly we will discuss connections between stationarity and stability of a solution.

In 1 Dimension

We assume stationrity and hence, cf. prop. 9.9.1, we also have detailed balance, i.e. no flux ($u\phi - D\phi' = 0$). At a stationary point, x^*, we have that $\phi'(x^*) = 0$ and the advection, u, must vanish $u(x^*) = 0$. Hence the forward Kolmogorov equation at stationarity must be $u'(x^*)\phi(x^*) = D(x^*)\phi''(x^*)$. Thus we know that

1. x^* is a local minimum for the stationary distribution ($\phi''(x) < 0$), iff is a stable equilibrium for the advection ($u'(x) < 0$).
2. x^* is a local maximum for the stationary distribution ($\phi''(x) > 0$), iff is a unstable equilibrium for the advection ($u'(x) > 0$).

In +2 Dimension

In more than 1 dimension we have the same result if detailed balance hold. But if the only thing we know is that the process is stationary we cannot claim anything anymore because if the flux is only curl we can still have a stationary distribution but not satisfy detailed balance.

(10) Optimal Control

Optimal Control for Stochastic Differential Equations

We now consider a domain where we introduce control. This gives us the controlled diffusion,

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

where $t \in \mathbf{T} = [0, T]$ and $\{U_t\}_{t\in\mathbf{T}}$ is the control signal, which, for each $t \geq 0$, is chosen from a compact set $\mathbf{U}$ of permissible controls. 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 $\{X_t\}_{t\in\mathbf{T}}$.

Termination happens when the state exits a region $\Omega \subset \mathbf{X}$ or if the time reaches T. Hence, we de define the stopping time

$$\tau = \min\{T, \inf\{t \in [0, T] \mid X_t \notin \Omega\}.$$

Optimal Control for Stochastic Differential Equations

For a given control strategy $\{U_t = \mu(X_t, t)\}_{t \in \mathbf{T}}$ and a given initial condition $X_s = x$, we want to maximize the performance objective

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

To this end, we include the lower bound $s = t_0$ as a parameter, but to get the total reward, we should take $s = 0$.

Furthermore, for a fixed control $u \in \mathbf{U}$, we define the generator L^u as

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

Hamilton-Jacobi-Bellman Equation

Theorem 13.3.1: The Verification Theorem

Let $V : \Omega \times \mathbf{T} \to \mathbb{R}$ be $C^{2,1}$ and satisfy the Hamilton-Jacobi-Bellman equation

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

on the interior $\Omega^o \times \mathbf{T}^o$, along with the boundary and terminal condition $V = k \; on \; \partial(\Omega \times \mathbf{T})$.

Let $\mu^* : \Omega \times \mathbf{T} \to \mathbf{U}$ be such that

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

on $\Omega^o \times \mathbf{T}^o$, and assume that μ^* satisfy existence and uniqueness. Then, for all $x \in \Omega$ and all $s \in \mathbf{T}$

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

Steady State Control: Fisheries Management System

We now consider the fish population given by

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

where U_t is the possibility to harvest from the population. We are given that our total profit is given by $J = \int_0^T \sqrt{U_t} dt$. We want to varify that a steady-state solution of the HJB equation is $V(x,t) = V_0(x) - \gamma t$ with $V_0(x) = \frac{1}{2}\log x$, $\gamma = \frac{1}{2}\left(1 - \frac{1}{2}\sigma^2\right)$, with the optimal control $\mu^*(x) = x^2$.

We use the Verification theorem and setup Hamilton-Jacobi-Bellman equation

$$\begin{aligned} 0 &= \nabla_t V + \sup_{u \geq 0}\left[f\nabla_x V + D\nabla^2_{xx}V + h\right] \\ &= \nabla_t V + \sup_{u \geq 0}\left[(x(1-x) - u)\nabla_x V + \sigma^2 x^2/2\nabla^2_{xx}V + \sqrt{u}\right] \end{aligned}$$

Steady State Control: Fisheries Management System

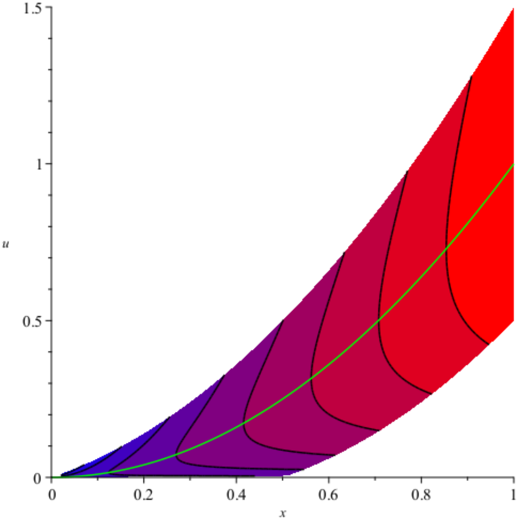
We see that only two terms depend on u.

$$-u\nabla_x V + \sqrt{u}$$

We see in the figure the function is concave. Hence we can optimize and be assured that the found maximum is global.

$$\nabla_u(-u\nabla_x V + \sqrt{u}) = 0 \Rightarrow$$

$$u = \frac{1}{4(\nabla_x V)^2}$$



Steady State Control: Fisheries Management System

We reinsert

$$\nabla_t V + (x(1-x) - \frac{1}{4(\nabla_x V)^2})\nabla_x V + \sigma^2 x^2/2\nabla^2_{xx} V + \sqrt{\frac{1}{4(\nabla_x V)^2}} = 0$$

We have that $\nabla_t V = -\gamma$, $\nabla_x V = \frac{1}{2x}$ and $\nabla^2_{xx} V = -\frac{1}{2x^2}$, so together with $\gamma = \frac{1}{2}\left(1 - \frac{1}{2}\sigma^2\right)$,

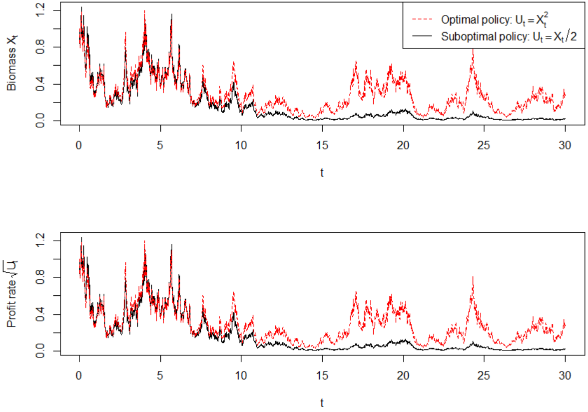
$$-\gamma + \frac{1}{2}(1-x) + \frac{1}{2}x - \frac{1}{4}\sigma^2 =$$
$$\frac{1}{2}x + \frac{1}{2}\left(1 - \frac{1}{2}\sigma^2\right) - \frac{1}{2}x - \frac{1}{2}\left(1 - \frac{1}{2}\sigma^2\right) = 0$$

Hence the policy $\mu = x^2$ is optimal.

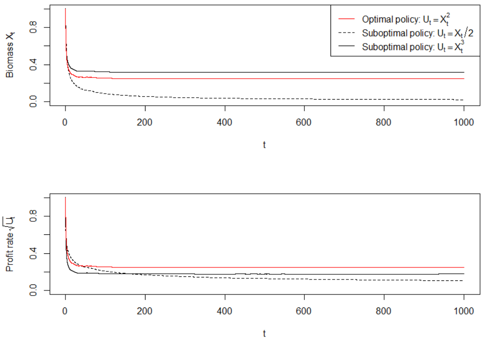
$$u = \frac{1}{4(\nabla_x V)^2} = \frac{1}{4}(2x)^2 = x^2$$

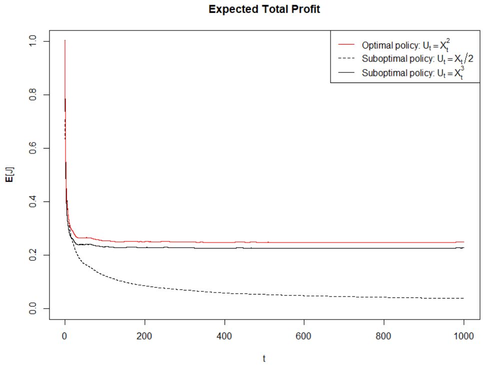
Steady State Control: Fisheries Management System

Using the optimal policy we see that the population recovers from hits, compared to the suboptimal policy, $\mu = x/2$.



Steady State Control: Fisheries Management System





- Left: We see that the policy x^3 harvests the population too little while $x/2$ harvest too much
- Right: We see that in the long run the policy $x/2$ makes the population go extinct and hence earns 0 in average.