An Introduction to Stochastic Differential Equations

Pages 77-89

Lawrence C. Evans

presentation by: Arun Suresh

Georgia State University

Chapter 5 - Stochastic Differential Equations

Chapter 5 - Definitions and Examples

Notation

1. *Let* $\boldsymbol{W}(\cdot)$ *be an* $m-$*dimensional Wiener process and* $\boldsymbol{X}_0$ *an* $n-$*dimensional random variable (independent of* $\boldsymbol{W}(\cdot)$. *For the remainder of this chapter, we will concern ourselves with*

$$\mathcal{F}(t) = \mathcal{U}(\boldsymbol{X}_0, \boldsymbol{W}(s)(0 \leq s \leq t)) \quad (t \geq 0)$$

the $\sigma-$*algebra generated by* $\boldsymbol{X}_0$ *and* $\boldsymbol{W}(s)$ *where* $s \in [0, t]$.

2. *Say* $T > 0$ *is given and* $\boldsymbol{b} : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ *and* $\boldsymbol{B} : \mathbb{R}^n \times [0, T] \to \mathbb{R}^{n \times m}$ *are given functions. We denote the* i^{th} *component of* $\boldsymbol{b}(\boldsymbol{x}, t)$ *by* b^i *and the* $(i, j)^{th}$ *entry of* $\boldsymbol{B}(\boldsymbol{x}, t)$ *by* b^{ij}

Chapter 5 - Definitions and Examples

- Let us recall the following definitions
  - An $\mathbb{R}^{m\times n}-$valued stochastic process, say $\mathbf{G}(\mathbf{x},t) = ((G^{ij}))$ is in $\mathbb{L}^2_n(0,T)$ if $G^{ij} \in \mathbb{L}^2(0,T)$ for all $i \in \overline{1,n}$ and $j \in \overline{1,m}$.
  - An $\mathbb{R}^{n}-$valued stochastic process, say $\mathbf{F}(\mathbf{x},t) = (F^{i})$ is in $\mathbb{L}^1_n(0,T)$ if $F^{i} \in \mathbb{L}^1(0,T)$ for all $i \in \overline{1,n}$.
  - Given a probability space $(\Omega, \mathcal{F})$ and a filtration $\{\mathcal{F}_t\}$; A stochastic process $\{\mathbf{X}_t\}_{t\geq 0}$ is said to be *progressively measurable* if for every $T > 0$, $\mathbf{X}_t$ when viewed as a function on the product space $[0,T]\times\Omega$, is measurable in $\mathcal{B}([0,T]) \times \mathcal{F}_T$
    - Like mentioned earlier, the details of this definition are technical (for example, the actual definition itself concerns with $[0,T]\otimes\Omega$ and not $[0,T]\times\Omega$) and the reason for its technicality is subtle
    - One we can appreciate, but not understand just yet.
    - That being said, their existence ensures that a lot of very nice things happen
    - Chapter 6 concerns itself with *stopping process*, and in some vague sense, we need progressive measurability to make the stopping process a nice stochastic process.

Chapter 5 - Definitions and Examples

Definition

We say that $\mathbb{R}^n-$valued stochastic process $\mathbf{X}(\cdot)$ is a solution of the Itô Stochastic differential equation

$$(SDE)\begin{cases} d\mathbf{X}(t) = \mathbf{b}(\mathbf{X},t)dt + \mathbf{B}(\mathbf{X},t)d\mathbf{W} \\ \mathbf{X}(0) = \mathbf{X}_0 \end{cases}$$

if

1. $\mathbf{X}(\cdot)$ is progressively measurable with respect to $\mathcal{F}(\cdot)$
2. $\mathbf{F} := \mathbf{b}(\mathbf{X},t) \in \mathbb{L}^1_n(0,T)$
3. $\mathbf{G} := \mathbf{B}(\mathbf{X},t) \in \mathbb{L}^2_n(0,T)$ and
4. $\mathbf{X}(t) = \mathbf{X}_0 + \int_0^t \mathbf{b}(\mathbf{X}(s),s)ds + \int_0^t \mathbf{B}(\mathbf{X}(s),s)d\mathbf{W}(s)$ a.s. for all $t \in [0,T]$

Chapter 5 - Definitions and Examples

- **Remark:** Given a higher order SDE,

$$Y^{(n)} = f(t, Y, \dot{Y}, \ldots, Y^{(n-1)}) + g(t, Y, \dot{Y}, \ldots, Y^{(n-1)})\xi = f(\Delta) + g(\Delta)\xi$$

we can re-write the above SDE into a system of $n-$SDEs of first order as follows

$$\mathbf{X}(t) = \begin{bmatrix} Y(t) \\ \dot{Y}(t) \\ \vdots \\ Y^{(n-1)}(t) \end{bmatrix} = \begin{bmatrix} X^1(t) \\ X^2(t) \\ \vdots \\ X^n(t) \end{bmatrix} \tag{1}$$

This allows us to formally write

$$d\mathbf{X}(t) = \begin{bmatrix} X^2(t) \\ X^3(t) \\ \vdots \\ f(\Delta) \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ g(\Delta) \end{bmatrix} d\mathbf{W} \tag{2}$$

Example

Let $m = n = 1$ and g a given continuous function. Now, consider the SDE

$$(SDE) \quad \begin{cases} dX = fXdt + gXdW \\ X(0) = 1 \end{cases}$$

The author gives the result and verifies it for a special case (Example 1), but let us have some fun deriving the result ourselves! For this, we note that we are starting with the following SDE

$$dX(t) = Fdt + GdW$$

where $F = fX$ and $G = gX$. Now we choose an appropriate u to use in Itô's lemma. Here $u(x) = \ln(x)$ will do. Now, we know by Itô's formula that $Y(t) = u(X)$ satisfies the following SDE

$$dY(t) = \left(\frac{\partial u}{\partial t}(X,t) + \frac{\partial u}{\partial x}(X,t)F + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(X,t)G^2 \right) dt + \frac{\partial u}{\partial x}(X,t)GdW$$

Chapter 5 - Definitions and Examples

Example (Contd...)

calculating the required partials in our case, substituting them into the above SDE and renaming t to s, gives us

$$d\ln(X(s)) = \left(f - \frac{1}{2}g^2\right) ds + g dW(s)$$

integrating both sides from $0 \to t$ and noticing that $X(0) = 1; \ln(1) = 0$ yields

$$\ln(X(t)) = \int_0^t f - \frac{1}{2}g^2 ds + \int_0^t g dW(s)$$

raising both sides to the power of e provides the same result as noted by the author.

The uniqueness of this result will be discussed later in this chapter

Chapter 5 - Definitions and Examples

Example (Stock prices)

Sometimes the price of a stock can be modelled using an SDE. In particular we consider dP/P as the relative change in price (as relative changes are really what matters when comparing the performance of small and large stocks). We suppose dP/P satisfies the following SDE

$$\frac{dP}{P} = \mu dt + \sigma dW$$

where $\mu > 0$ is called the *drift* term and σ is the volatility of the stock. The initial price of the stock is assumed to be p_0. This SDE then reduces to a very simple

$$dP = \mu P dt + \sigma P dW$$

which is a special case of the previous example, so we can directly let $f = \mu$ and $g = \sigma$ to obtain

$$P(t) = p_0 e^{\sigma W(t) + (\mu - \frac{\sigma^2}{2})t}$$

Example (contd...)

In fact, we also know that since $P(t)$ is a solution to the above SDE, we know it satisfies

$$P(t) = p_0 + \int_0^t \mu P ds + \int_0^t \sigma P dW(s)$$

taking expectations on both sides, and noting $E(\int_0^t \sigma P dW(s)) = 0$ we see that

$$E(P(t)) = p_0 + \int_0^t \mu E(P(s)) ds$$

Solving the above equation (using elementary ODE techniques), we see

$$E(P(t)) = p_0 e^{\mu t} \quad \text{for } t \geq 0$$

This is exactly what we would expect!

$E\left(\int_0^t \sigma P dW(s)\right) = 0$ because of one of the main properties of the Itô integral, ref page 64 of text

A brief exploration of Numerical SDE methods (will be relevant in a moment)

- We are almost at that point where can pride ourselves with "analytically" solving SDEs in the previous examples.
- However, as reality would have it, analytical solutions are often impossible to find (and the ones we find don't really satisfy us anyway).
- So we dive into numerical methods
- Solving SDEs numerically is actually quite nice (at an elementary level at least)
- For now, we focus on one particular method: **The Euler-Maruyama method**
- Given an 1-D SDE,

$$dX = b(X,t)dt + B(X,t)dW$$

  we can discretize it by first dividing up our time interval $[0,T]$ into distinct time points

$$0 = t_0 < t_1 < t_2 < \cdots < t_k = T$$

  with $t_{i+1} - t_i = \Delta t$

The Euler Maruyama Method - an overview

- By discretizing t, we have also discretized X and W to obtain

$$X_{i+1} = X_i + b(X,t)\Delta t + B(X,t)\Delta W_i$$

- In the above discrete form, $\Delta W_i = (W(t_{i+1}) - W(t_i)) \sim \mathcal{N}(0, \Delta t)$.
- So if we take $\eta \sim \mathcal{N}(0,1)$, then we know $\sqrt{\Delta t}\eta \sim \mathcal{N}(0, \Delta t)$
- So our scheme simply becomes

$$X_{i+1} = X_i + b(X,t)\Delta t + \sqrt{\Delta t}\eta_i B(X,t)$$

  for $i \in \overline{0, k-1}$ and $X_0 = X(0)$

- **Fun fact:** Moving X_i to the other side and dividing by Δt we see that X_i is non-differentiable everywhere in general.

$$\frac{X_{i+1} - X_i}{\Delta t} = b(X,t) + B(X,t)\frac{\eta_i}{\sqrt{\Delta t}}$$

  as $\Delta t \to 0$, the LHS $\to dX/dt$, but limit doesn't exist for RHS always.

The Euler Maruyama Method - Error Analysis

- We will simply summarize the main results that detail the performance of the method
- The numerical solution given by E-M method X_i and the true solution to the SDE X satisfy

$$E(|X(t_i) - X_i|^2 | X(t_{i-1}) = X_{i-1}) = \mathcal{O}((\Delta t)^2)$$
$$E(|X(t_i) - X_i|^2 | X(0) = X_0) = \mathcal{O}(\Delta t)$$

- Each step in the E-M method has MSE $(\Delta t)^2$
- The overall error across the interval is Δt
- Small Δt improves accuracy
- Δt too small gets into territory of numerical errors.

Example (Brownian Bridge)

Given

$$dX(t) = \left[-\frac{X(t)}{1-t} \right] dt + dW(t); \quad X(0) = 0$$

Let $u(x,t) = -\frac{x}{1-t}$. Then using Itô's formula we have that

$$d\left[-\frac{X(s)}{1-s} \right] = -\frac{1}{1-s} dW(s)$$

Integrating from 0 to t on both sides, we have

$$-\frac{X(t)}{1-t} = -\int_0^t \frac{1}{1-s} dW(s)$$

Rearranging and solving for $X(t)$ we have

$$\boxed{X(t) = (1-t) \int_0^t \frac{1}{1-s} dW(s)}$$

Chapter 5 - Definitions and Examples

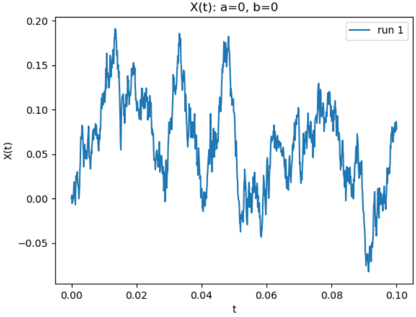


Figure: A sample Brownian bridge simulated in Python

- Utilized the Euler-Maruyama method with $dt = 0.0001$

Chapter 5 - Definitions and Examples

Example (Langevin's equation)

An improvement on existing Brownian models by accounting for friction

$$(SDE) \begin{cases} dX = -bX dt + \sigma dW \\ X(0) = X_0 \end{cases}$$

Here $X(t)$ is the velocity of the brownian particle. $b > 0$ is the coefficient of friction and σ is the diffusion coefficient. X_0 is some inital distribution. For the sake of time, we only summarize the results below

$$X(t) = e^{-bt} X_0 + \sigma \int_0^t e^{-b(t-s)} dW \quad (t \geq 0)$$

$$E(X(t)) = e^{-bt} E(X_0) \to 0$$

$$V(X(t)) = e^{-2bt} V(X_0) + \frac{\sigma^2}{2b}(1 - e^{-2bt}) \to \frac{\sigma^2}{2b}$$

From the explicit solution we see $X(t) \to \mathcal{N}(0, \frac{\sigma^2}{2b})$

Chapter 5 - Definitions and Examples

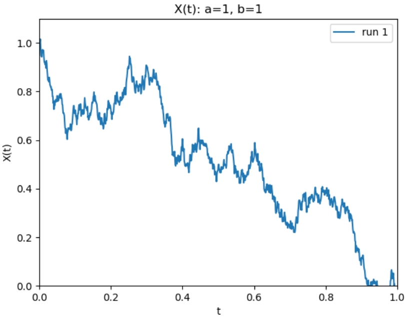


Figure: A sample path of the Langevin's equation simulated in Python

- Utilized the Euler-Maruyama method with $dt = 0.001$
- Parameters $a = X(0) = 1, b = 1, \sigma = 0.5$

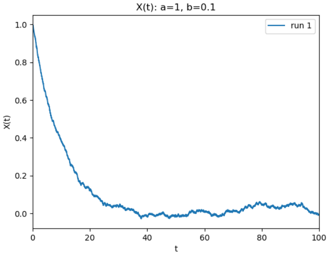


Figure: A sample path of the Langevin's equation simulated in Python

- Utilized the Euler-Maruyama method with $dt = 0.0001$
- It is visibly clear that as $t \to \infty$ $X(t) \to \mathcal{N}(0, \frac{\sigma^2}{2b})$
- Parameters $a = X(0) = 1, b = 0.1, \sigma = 0.01$
- Key-take away:
    - if σ is comparable to b, the noise factor is just as dominating as the drift factor and the process can go haywire.

Example (Ornstein-Uhlenbeck process)

A better model for Brownian motion

$$\begin{cases} \ddot{Y} = -b\dot{Y} + \sigma\xi \\ Y(0) = Y_0, \dot{Y}(0) = Y_1 \end{cases}$$

$Y(t)$ is the position of the Brownian particle at time t, Y_0, Y_1 are given Gaussian random variables. $b > 0$ is called the friction coefficient and σ is the diffusion coefficient. The above model can be repackaged into an SDE by setting $X(t) := \dot{Y}$. Then X is the solution to the the following SDE

$$(SDE) \quad \begin{cases} dX = -bXdt + \sigma dW \\ X(0) = Y_1 \end{cases}$$

This now has boiled down to the Langevin equation; to which the solution is immediate

$$X(t) = e^{-bt}Y_1 + \sigma \int_0^t e^{-b(t-s)}dW$$

Chapter 5 - Definitions and Examples

Example (contd...)

We further assume Y_1 to be a normal process, which immediately turns X into a normal process for all $t \geq 0$. Finally, we see

$$Y(t) = Y_0 + \int_0^t X ds$$

$E(Y(t))$ and $V(Y(t))$ can be calculated as

$$E(Y(t)) = E(Y_0) + \left(\frac{1 - e^{-bt}}{b}\right) E(Y_1)$$

$$V(Y(t)) = V(Y_0) + \frac{\sigma^2}{b^2} t + \frac{\sigma^2}{2b^3}(-3 + 4e^{-bt} - e^{-2bt})$$

Chapter 5 - Existence and Uniqueness of solutions

- Now that we have looked at some example computations and a copious amount of hand-waving, we would like to introduce some rigor and address questions about existence of solutions to SDEs; and uniqueness when guaranteed existence.

Example

Let us try to solve the following 1 dimensional SDE. Given $b : \mathbb{R} \to \mathbb{R}$ is C^1 and $|b'| \leq L$ and $x \in \mathbb{R}$.

$$(SDE) \begin{cases} dX = b(X)dt + dW \\ X(0) = x \end{cases}$$

This SDE now guarantees that the solution $X(t)$ (if at all it exists) satisfies

$$X(t) = x + \int_0^t b(X)ds + W(t)$$

Now, starting at $X^0(t) \equiv x$, we try successive approximation method to hopefully find a solution; and we want to show that the solution obtained in this method is nothing but $X(t)$.

Example (contd...)

So we start by defining

$$X^{n+1}(t) := x + \int_0^t b(X^n)ds + W(t)$$

and also define

$$D^n(t) := \max_{0\leq s\leq t} |X^{n+1}(s) - X^n(s)|$$

Now, given a continuous sample path of the Weiner process, we have

$$D^0(t) = \max_{0\leq s\leq t} \left| \int_0^s b(x)dr + W(s) \right| \leq C$$

for all times $0 \leq t \leq T$ where C depends on ω (as D^0 is continuous in t, and $[0, t]$ is compact - so EVT). We now want to show the following

$$D^n(t) \leq C\frac{L^n}{n!}t^n$$

Example (contd...)

For this, we notice that

$$\begin{aligned} D^n(t) &= \max_{0\le s\le t}\left|\int_0^s b(X^n(r)) - b(X^{n-1}(r))dr\right| \\ &\le L\int_0^t D^{n-1}(s)ds \\ &\le L\int_0^t C\frac{L^{n-1}s^{n-1}}{(n-1)!}ds = C\frac{L^n t^n}{n!} \end{aligned}$$

Finally, we see that for $m \ge n$ we have

$$\max_{0\le t\le T}|X^m(t) - X^n(t)| \le C\sum_{k=n}^{\infty}\frac{L^k T^k}{k!} \to 0 \text{ as } k\to\infty$$

Thus for almost all ω, $X^n(\cdot)$ converges uniformly for $0 \le t \le T$ to a limit process $X(\cdot)$ which satisfies the original equation!

Chapter 5 - Existence and Uniqueness of solutions

- We now consider solving $SDEs$ by a change of variables technique
- Say we are given the following one dimensional SDE

$$(SDE) \quad \begin{cases} dX = b(X)dt + \sigma(X)dW \\ X(0) = x \end{cases} \tag{3}$$

- We will now proceed to ignore the above SDE and try to solve the following one instead

$$(SDE) \quad \begin{cases} dY = f(Y)dt + dW \\ Y(0) = y \end{cases} \tag{4}$$

- But our anarchy isn't without reason. We will choose an f later on provided we can find a function u such that

$$X := u(Y)$$

Chapter 5 - Existence and Uniqueness of solutions

- I believe that by now, given my one solitary example, I have convinced you that it is *in principle* possible to solve equation (4)
  - All non-trivial details left as exercises to the readers.
- Let us for a moment assume there is suitable f, and therefore a corresponding u. Now for the clever bit.
- Since $X = u(Y)$, we simply use Itô's formula to say that X must satisfy the following SDE

$$dX = \left[u'f + \frac{1}{2}u''\right]dt + u'dW \tag{5}$$

- Now we compare equation (5) with (3) to conclude

$$\begin{cases} u'(Y) & = \sigma(X) = \sigma(u(Y)) \\ u'(Y)f(Y) + \frac{1}{2}u''(Y) & = b(X) = b(u(Y)) \\ u(y) & = x \end{cases} \tag{6}$$

- We first solve the following ODE system

$$(ODE) \quad \begin{cases} u'(z) = \sigma(u(z)) \\ u(y) = x \end{cases}$$

  to get u (this is not so trivial by the way. σ could very well be complicated).

- Once we get u, we use it to solve for f by noting

$$f(z) = \frac{1}{\sigma(u(z))}\left[b(u(z)) - \frac{1}{2}u''(z)\right]$$

- Finally, we use this f to solve (3) and we win.

Chapter 5 - Existence and Uniqueness of solutions

- In this process we have skipped over a mountain of crucial information and details such as:
    - What properties must u satisfy?
        - At the very least we know it must be double differentiable
    - Is this f guaranteed to exist? If so is it unique?
    - We have not yet showed uniqueness of solutions to $SDEs$ like the one in equation (3), so how can we realistically compare (5) and (3) to get (6)?
- Fortunately, the author doesn't want to talk about it either.
- Unfortunately, the paper cited by the author is pay-to-access, and we are just poor grad students trying to get by.

Chapter 5 - Existence and Uniqueness of solutions

- Now that we have honed our intuition with various examples, we are ready to grapple with a general existence and uniqueness proof.
- But before that, we start with a lemma

Lemma (Gronwall)

Let $\phi, f \in C([0,T])$ be non-negative functions. Let $C_0 \geq 0$ be some arbitrary constant. If

$$\phi(t) \leq C_0 + \int_0^t f\phi ds \quad \textit{for all } 0 \leq t \leq T$$

then

$$\phi(t) \leq C_0 e^{\int 0^t f ds}$$

Proof.

Let $\Phi(t) = C_0 + \int_0^t f\phi ds$. Then $\Phi' = f\phi \leq f\Phi$ (as $\phi \leq \Phi$ and f non-negative). So,

$$\left(e^{-\int_0^t f ds}\Phi\right)' = (\Phi' - f\Phi)e^{-\int_0^t f ds} \leq (f\phi - f\phi)e^{-\int_0^t f ds} = 0$$

Therefore

$$e^{-\int_0^t f ds}\Phi(t) \leq e^{-\int_0^0 f ds}\Phi(0) = C_0$$

and thus

$$\phi(t) \leq \Phi(t) \leq C_0 e^{\int_0^t f ds}$$

$\square$

Theorem (Existence and Uniqueness)

Suppose that $\boldsymbol{b} : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ *and* $\boldsymbol{B} : \mathbb{R}^n \times [0,T] \to \mathbb{R}^{m\times n}$ *are continuous and satisfy the following conditions.*

1. *For some constant* L *and all* $0 \leq t \leq T$, $x, \hat{x} \in \mathbb{R}^n$ *we have*

$$|\boldsymbol{b}(x,t) - \boldsymbol{b}(\hat{x},t)| \leq L|x - \hat{x}|$$
$$|\boldsymbol{B}(x,t) - \boldsymbol{B}(\hat{x},t)| \leq L|x - \hat{x}|$$

2. *For the same constant* L *(as above) and all* $0 \leq t \leq T$, $x \in \mathbb{R}^n$ *we have*

$$|\boldsymbol{b}(x,t)| \leq L(1 + |x|)$$
$$|\boldsymbol{B}(x,t)| \leq L(1 + |x|)$$

Let $\boldsymbol{X}_0$ *be any* $\mathbb{R}^n$ *valued random variable with*

3. $E(|\boldsymbol{X}_0|^2) < \infty$ *and*
4. $\boldsymbol{X}_0$ *is independent of* $\mathcal{W}^+(0)$.

Then...

Chapter 5 - Existence and Uniqueness of solutions

Theorem (contd...)

...there exists a unique solution $\boldsymbol{X} \in \mathbb{L}^2_n(0,T)$ *of the SDE*

$$\begin{cases} d\boldsymbol{X} = \boldsymbol{b}(\boldsymbol{X},t)dt + \boldsymbol{B}(\boldsymbol{X},t)d\boldsymbol{W} \quad (0 \le t \le T) \\ \boldsymbol{X}(0) = \boldsymbol{X}_0 \end{cases}$$

Remarks

1. *"Unique" means that if* $\boldsymbol{X}, \hat{\boldsymbol{X}} \in \mathbb{L}^2_n(0,T)$ *with continuous sample paths almost surely, and both solve the above SDE then*

$$P(\boldsymbol{X}(t) = \hat{\boldsymbol{X}}(t) \text{ for all } 0 \le t \le T) = 1$$

2. *Item* 1 *says that* $\boldsymbol{b}$ *and* $\boldsymbol{B}$ *are uniformly Lipschitz in the variable* x. *Hypothesis* 2 *follows directly from Hypothesis* 1. ***Personal note:*** *I am not sure if the second statement is true. In what follows, I will produce a counter-example*

Remarks

So in the following, I present my case as to why (1) *does not necessarily imply* (2). *To make the problem simple, we only consider the first part (with **b**), and we also let* $n = 1$, *which boils the implication down to the following: Given* $f : \mathbb{R} \times [0, T] \to \mathbb{R}$ *an **arbitrary** continuous function such that **for some given*** L *we have*

$$|f(x,t) - f(\hat{x},t)| \leq L|x - \hat{x}| \qquad \text{for all } x, \hat{x} \in \mathbb{R}, t \in [0,T] \tag{7}$$

*then we also satisfy **for the same L***

$$|f(x,t)| \leq L(1 + |x|) \qquad \text{for all } x \in \mathbb{R}, t \in [0,T] \tag{8}$$

Now I claim that there exists some function g *that satisfies equation (7), but not equation (8)*

Remarks

Counter-example.

Let f satisfy statement (1). By hypothesis, we are given some L such that

$$|f(x,t) - f(\hat{x},t)| \leq L|x - \hat{x}| \quad \forall x, \hat{x} \in \mathbb{R}, t \in [0,T]$$

Now consider $g(x,t) = f(x,t) + L + 1 - f(0,t)$. Notice that for the same L, g also satisfies

$$|g(x,t) - g(\hat{x},t)| \leq L|x - \hat{x}|$$

However, if equation (7) did in fact imply equation (8) this would mean $|g(0,t)| \leq L$; but by construction, we know that $|g(0,t)| = |f(0,t) + L + 1 - f(0,t)| = L + 1$, which is a contradiction! $\square$

- *So, we simply take statements (1) and (2) as necessary parts of the hypothesis for the main theorem.*

Proof.

(Uniqueness) Suppose $\mathbf{X}$ and $\hat{\mathbf{X}}$ are solutions to the above SDE. Then for all $0 \leq t \leq T$

$$\mathbf{X}(t) - \hat{\mathbf{X}}(t) = \int_0^t \mathbf{b}(\mathbf{X}, s) - \mathbf{b}(\hat{X}, s)ds + \int_0^t B(\mathbf{X}, s) - \mathbf{B}(\hat{\mathbf{X}}, s)d\mathbf{W}$$

Since $(a+b)^2 \leq 2a^2 + 2b^2$, we can estimate

$$E(|\mathbf{X}(t) - \hat{\mathbf{X}}(t)|^2) \leq 2E\left(\left|\int_0^t \mathbf{b}(\mathbf{X}, s) - \mathbf{b}(\hat{\mathbf{X}}, s)ds\right|^2\right) + 2E\left(\left|\int_0^t \mathbf{B}(\mathbf{X}, s) - \mathbf{B}(\hat{\mathbf{X}}, s)d\mathbf{W}\right|^2\right)$$

$\square$

Proof.

Notice that Cauchy Schwarz inequality gives us that for any $f : [0, t] \to \mathbb{R}^n$, we have

$$\left| \int_0^t 1 \cdot \mathbf{f} ds \right|^2 \leq t \int_0^t |\mathbf{f}|^2 ds$$

for any $t > 0$. We now use this to estimate

$$\begin{aligned} E\left(\left| \int_0^t \mathbf{b}(\mathbf{X}, s) - \mathbf{b}(\hat{\mathbf{X}}, s) ds \right|^2 \right) &\leq TE\left(\int_0^t |\mathbf{b}(\mathbf{X}, s) - \mathbf{b}(\hat{\mathbf{X}}, s)|^2 ds \right) \\ &\leq L^2 T \int_0^t E(|\mathbf{X} - \hat{\mathbf{X}}|^2) ds \end{aligned}$$

$\square$

Proof.

Furthermore, we also have

$$E\left(\left|\int_0^t \mathbf{B}(\mathbf{X},s) - \mathbf{B}(\hat{\mathbf{X}},s)d\mathbf{W}\right|^2\right) = E\left(\int_0^t |\mathbf{B}(\mathbf{X},s) - \mathbf{B}(\hat{\mathbf{X}},s)|^2 ds\right)$$
$$\leq L^2 T \int_0^t E(|\mathbf{X} - \hat{\mathbf{X}}|^2) ds$$

Therefore, for some appropriate constant C, we can say

$$E(|\mathbf{X}(t) - \hat{\mathbf{X}}(t)|^2) \leq C \int_0^t E(|\mathbf{X} - \hat{\mathbf{X}}|^2) ds$$

provided $0 \leq t \leq T$. $\square$

Proof.

Now, we set $\phi(t) := E(|\mathbf{X} - \hat{\mathbf{X}}|^2)$, which gives us that

$$\phi(t) \leq C \int_0^t \phi(s) ds; \quad \text{for all } 0 \leq t \leq T$$

Here we can apply Gronwall's lemma with $C_0 = 0$ to see that $\phi \equiv 0$. Thus $\mathbf{X}(t) = \hat{X}(t)$ a.s. for all $0 \leq t \leq T$; and so $\mathbf{X}(r) = \hat{\mathbf{X}}(r)$ for all rational $0 \leq r \leq T$ except for some set of probability 0. Since $\mathbf{X}$ and $\hat{\mathbf{X}}$ have continuous sample paths almost surely, we get

$$P(\max_{0 \leq t \leq T} |\mathbf{X}(t) - \hat{X}(t)| > 0) = 0$$

□

Proof.

(Existence) We start by defining

$$\begin{cases} \mathbf{X}^0(t) := \mathbf{X}_0 \\ \mathbf{X}^{n+1}(t) := \mathbf{X}_0 + \int_0^t \mathbf{b}(\mathbf{X}^n(s), s)ds + \int_0^t \mathbf{B}(\mathbf{X}^n(s), s)d\mathbf{W} \end{cases}$$

for $n = 0, 1, \dots$ and $0 \leq t \leq T$. As done before, we also define

$$d^n := E(|\mathbf{X}^{n+1}(t) - \mathbf{X}^n(t)|^2)$$

We once again aim to show that for some constant M depending on L, T and $\mathbf{X}_0$,

$$d^n(t) \leq \frac{(Mt)^{n+1}}{(n+1)!}$$

□

Proof.

We start with $n = 0$, to notice

$$\begin{aligned}
d^0(t) &= E(|\mathbf{X}^1(t) - \mathbf{X}^0(t)|^2) \\
&= E\left(\left|\int_0^t \mathbf{b}(\mathbf{X}_0, s)ds + \int_0^t \mathbf{B}(\mathbf{X}_0, s)d\mathbf{W}\right|^2\right) \\
&\leq 2E\left(\left|\int_0^t L(1 + |\mathbf{X}_0|)ds\right|^2\right) + 2E\left(\int_0^t L^2(1 + |\mathbf{X}_0|)^2 ds\right) \\
&\leq tM
\end{aligned}$$

for some large enough M. This proves the claim for $n = 0$. Now we assume that the claim holds for $n - 1$. □

Proof.

$$\begin{aligned}
d^n(t) &= E(|\mathbf{X}^{n+1}(t) - \mathbf{X}^n(t)|^2) \\
&= E\bigg(\bigg|\int_0^t \mathbf{b}(\mathbf{X}^n, s) - \mathbf{b}(\mathbf{X}^{n-1}, s)ds \\
&\qquad\qquad + \int_0^t \mathbf{B}(\mathbf{X}^n, s) - \mathbf{B}(\mathbf{X}^{n-1}, s)d\mathbf{W}\bigg|^2\bigg) \\
&\le 2TL^2E\bigg(\int_0^t |\mathbf{X}^n - \mathbf{X}^{n-1}|^2 ds\bigg) + 2L^2E\bigg(\int_0^t |\mathbf{X}^n - \mathbf{X}^{n-1}|^2 ds\bigg) \\
&\le 2L^2(1+T)\int_0^t \frac{M^n s^n}{n!} ds \quad (\text{ by the induction hypothesis}) \\
&\le \frac{M^{n+1}t^{n+1}}{(n+1)!}
\end{aligned}$$

The last inequality is valid only if we choose $M \ge 2L^2(1+T)$, which is what we do to prove our claim! □

Proof.

Now note

$$\max_{0\le t\le T}|\mathbf{X}^{n+1}(t)-\mathbf{X}^{n}(t)|^2 \le 2TL^2\int_0^T|\mathbf{X}^n-\mathbf{X}^{n-1}|^2ds$$
$$+2\max_{0\le t\le T}\left|\int_0^t \mathbf{B}(\mathbf{X}^n,s)-\mathbf{B}(\mathbf{X}^{n-1},s)d\mathbf{W}\right|^2$$

We now use the Martingale inequality from Chapter 2 to see

$$\begin{aligned}E(\max_{0\le t\le T}|\mathbf{X}^{n+1}(t)-\mathbf{X}^{n}(t)|^2) &\le 2TL^2\int_0^T E(|\mathbf{X}^n-\mathbf{X}^{n-1}|^2)ds\\ &\quad+8L^2\int_0^T E(|\mathbf{X}^n-\mathbf{X}^{n-1}|^2)ds\\ &\le C\frac{(MT)^n}{n!}\quad \text{(by the above claim)}\end{aligned}$$

$\square$

Proof.

We can now apply the Borel-Cantelli lemma, because

$$P\left(\max_{0\le t\le T}\left|\mathbf{X}^{n+1}(t)-\mathbf{X}^{n}(t)\right|>\frac{1}{2^n}\right)\le 2^{2n}E\left(\max_{0\le t\le T}\left|\mathbf{X}^{n+1}(t)-\mathbf{X}^{n}(t)\right|^2\right)$$
$$\le 2^{2n}\frac{C(MT)^n}{n!}$$

and

$$\sum_{n=1}^{\infty}2^{2n}\frac{(MT)^n}{n!}<\infty$$

therefore,

$$P\left(\max_{0\le t\le T}\left|\mathbf{X}^{n+1}(t)-\mathbf{X}^{n}(t)\right|>\frac{1}{2^n}\text{ i.o. }\right)=0$$

$\square$

Proof.

This means

$$\mathbf{X}^n = \mathbf{X}^0 + \sum_{j=0}^{n-1} (\mathbf{X}^{j+1} - \mathbf{X}^j)$$

converges uniformly on $[0, T]$ to a process $\mathbf{X}(\cdot)$. We pass to the limits in the definition of $\mathbf{X}^{n+1}(\cdot)$ to see that $\mathbf{X}(\cdot)$ satisfies the SDE in the theorem statement; and clearly as $n \to \infty$ we simply have

$$\mathbf{X}(t) = X_0 + \int_0^t \mathbf{b}(\mathbf{X}, s)ds + \int_0^t \mathbf{B}(\mathbf{X}, s)d\mathbf{W} \quad \text{for } 0 \leq t \leq T$$

$\square$

Proof.

We finally just have to show that this $\mathbf{X}(\cdot) \in \mathbb{L}^2_n(0,T)$. For this, we start by noticing the following

$$
\begin{aligned}
E(|\mathbf{X}^{n+1}(t)|^2) &\leq CE(\mathbf{X}_0^2) + CE\left(\left|\int_0^t \mathbf{b}(\mathbf{X}^n, s)ds\right|^2\right) \\
&\quad + CE\left(\left|\int_0^t \mathbf{B}(\mathbf{X}^n, s)d\mathbf{W}\right|^2\right) \\
&\leq C(1 + E(|\mathbf{X}_0|^2)) + C\int_0^t E(|\mathbf{X}^n|^2)ds
\end{aligned}
$$

Where C is the combination of all the constants that appear in the above simplification. Inductively, we see

$$
E(|\mathbf{X}^{n+1}(t)|^2) \leq \left[C + C^2 + \cdots + C^{n+2}\frac{t^{n+1}}{(n+1)!}\right](1 + E(|\mathbf{X}_0|^2))
$$

□

Proof.

This means,

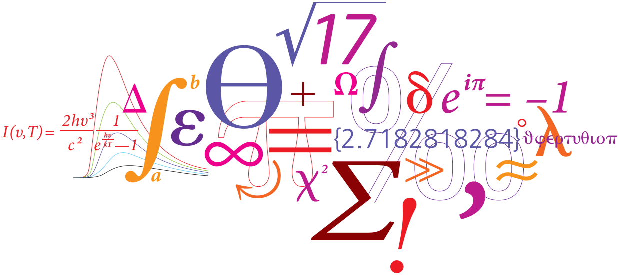
$$E(|\mathbf{X}^{n+1}(t)|^2) \leq C(1 + E(|\mathbf{X}_0|^2))e^{Ct} \qquad \text{for all } 0 \leq t \leq T$$

and letting $n \to \infty$, we see that

$$E(|\mathbf{X}(t)|^2) \leq C(1 + E(|\mathbf{X}_0|^2))e^{Ct} \qquad \text{for all } 0 \leq t \leq T$$

which shows that $\mathbf{X} \in \mathbb{L}^2_n(0, T)$, which finishes our proof! $\square$

Thank you!



$I(v,T)=\frac{2hv^3}{c^2}\frac{1}{e^{\frac{hv}{kT}}-1}$
$\sqrt{17}$
$e^{i\pi}=-1$
{2.7182818284}