Stochastic Differential Equations

Jun Wei Tan[*]

Julius-Maximilians-Universität Würzburg

(Dated: January 22, 2025)

Definition 1. A stochastic differential equation is a (formal) equation of the form

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = b(t, X_t) + \sigma(t, X_t)W_t, \tag{1}$$

where W_t is white noise.

This equation is to be interpreted as follows:

Definition 2. We say that the stochastic process X_t is a *solution* of the SDE (1) if

$$X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s$$

Example 3. A model of a population is given by the stochastic differential equation

$$\frac{\mathrm{d}N_t}{\mathrm{d}t} = rN_t + \alpha W_t N_t.$$

Here, r, α are constants and W_t is white noise.

This is the well known model for a population, except that we have allowed r to vary by a white noise term. The solution in the nonstochastic limit is given by a simple exponential. To solve this, we first rewrite the SDE in standard form:

$$\mathrm{d}N_t = rN_t\,\mathrm{d}t + \alpha N_t\,\mathrm{d}B_t\,,$$

or

$$\frac{\mathrm{d}N_t}{N_t} = r\,\mathrm{d}t + \alpha\,\mathrm{d}B_t\,.$$

Inspired by the solution in the deterministic case, we guess $g(t, x) = \ln x$ in Ito's formula,

[*] jun-wei.tan@stud-mail.uni-wuerzburg.de

and let $Y_t = g(t, N_t)$. Then, we have

$$\mathrm{d}Y_t = \frac{\mathrm{d}N_t}{N_t} + \frac{1}{2}\alpha^2 \,\mathrm{d}t\,.$$

Substituting, we have

$$\mathrm{d}Y_t = \left(r - \frac{1}{2}\alpha^2\right) \mathrm{d}t + \alpha B_t$$

which integrates easily to yield

$$N_t = N_0 \exp\left(\left(r - \frac{1}{2}\alpha^2\right) t + \alpha B_t\right).$$

Clearly, for $\alpha = 0$, this reduces to the well known exponential solution. This process is known as geometric Brownian motion, which, for example, is a solution to the Black-Scholes equation [1].

Now, we turn to the questions of existence and uniqueness. Recall the basic theorem for existence and uniqueness of a deterministic differential equation

Theorem 4 (Picard-Lindelöf). Let $\dot{x} = f(t, x)$ be a differential equation with f defined on a rectangle $[a, b] \times \mathbb{R}^n$. If f is Lipschitz continuous in x, with Lipschitz constant independent of time, and continuous in time, then the differential equation has a unique global solution on $[a, b]$.

Note that for uniqueness we do not need the continuity in time; the Lipschitz condition alone suffices.

Definition 5 (Linear Growth). A function $g(t, x)$ on $[a, b] \times \mathbb{R}$ is said to satisfy the linear growth condition in x if there exists $K > 0$ such that

$$|g(t, x)| \leq K(1 + |x|)$$

for all $t \in [a, b]$ and $x \in \mathbb{R}$.

Because of the inequality

$$1 + x^2 \leq (1 + x)^2 \leq 2(1 + x^2),$$

the earlier definition is equivalent to the following definition:

Proposition 6. A function $g(t, x)$ satisfies the linear growth condition if there exists $K > 0$ such that

$$|g(t, x)|^2 \leq K(1 + x^2)$$

The second definition will be more useful in the following proofs. Note that it is indeed stronger

Proposition 7. Globally Lipschitz continuous functions satisfy the linear growth condition.

Proof. If $f(x)$ is Lipschitz, we have

$$|f(x) - f(0)| \leq K|x|,$$

or

$$|f(x)| \leq |f(0)| + K|x|$$

by the reverse triangle inequality. □

Conversely, linear growth does not imply Lipschitz continuity, or any kind of continuity

Example 8. Let

$$f(x) = \begin{cases} x & x \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, f grows linearly, but is not Lipschitz continuous, or even continuous at any point.

We can prove uniqueness with just Lipschitz continuity, as we do in the deterministic case.

Theorem 9. Let $\sigma(t, x)$ and $f(t, x)$ be measurable functions on $[a, b] \times \mathbb{R}$ satisfying the Lipschitz condition in x. Suppose ξ is an $\mathcal{F}_a$-measurable random variable satisfying $\mathbb{E}[\xi^2] < \infty$. Then the stochastic differential equation

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = f(t, X_t) + \sigma(t, X_t)W_t$$

has at most one continuous solution on $[a, b]$.

Proof. We will not go through the proof in detail. Instead, we will talk about the main steps. Assume we have two solutions X_t and Y_t. We seek to estimate $Z_t = X_t - Y_t$

1. We estimate the expectation value $\mathbb{E}(Z_t^2)$, using the Lipschitz condition.

2. We obtain an integral inequality

$$E(Z_t^2) \leq 2K^2(1 + b - a) \int_a^t \mathbb{E}(Z_s^2)\,\mathrm{d}s\,,$$

   which, by the theory of classical differential equations, implies that Z_t is 0 almost surely for all t.

3. Then, we extend the solution to show that Z_t is 0 almost surely, using sample path continuity. □

The existence theorem is as follows:

Theorem 10. The stochastic differential equation (1) has a unique solution with initial condition ξ, where ξ^2 has finite expectation, σ and b are Lipschitz in x, with Lipschitz constant independent of t, and continuous in t.

Proof. The proof follows by Picard Iteration much as it does in the deterministic case. We define $X_0 = \xi$ and

$$X_t^{(n+1)} = X_0 + \int_0^t b(s, X_s^{(n)})\,\mathrm{d}s + \int_0^t \sigma(s, X_s^{(n)})\,\mathrm{d}B_s\,.$$

Then, the proof proceeds in two steps. First, we show that this sequence converges, If it does, and converges in a dominated manner, it satisfies the integral equation by the dominated convergence theorem. □

I. THE MARKOV PROPERTY

It is a known property of initial value problems that the future solution is not dependent on the past. In particular, we can imagine that we have some initial state $x(0)$ and let it evolve a time t to $x(t)$. Then we can let it evolve further. Alternatively, we can consider an initial value problem that has the value $x(t)$ at time t. We expect that these two solutions are identical. In a stochastic differential equation, this "'memory" property is known as the Markov property. Before we define it, let us recall a few definitions:

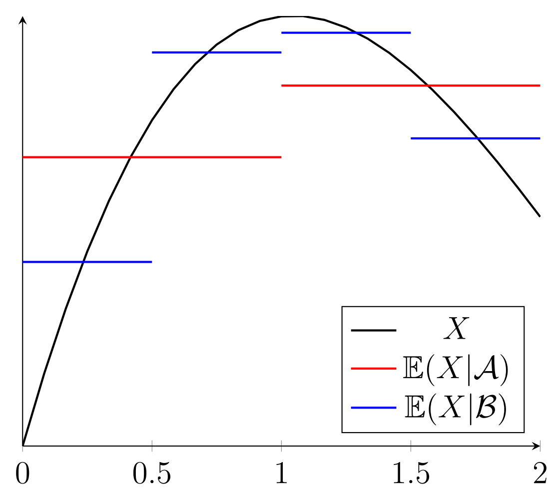
Definition 11 (Conditional Expectation)**.** Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and X is a real-valued random variable on that space with finite expectation. Let $\mathcal{H} \subseteq \mathcal{F}$ be a sub σ-algebra of $\mathcal{F}$. Then the conditional expectation of X relative to $\mathcal{H}$ is a $\mathcal{H}$-$\mathcal{B}$ measurable function $\mathbb{E}(X|\mathcal{H})$ such that

$$\int_H \mathbb{E}(X|\mathcal{H})\,\mathrm{d}\mathbb{P} = \int_H X\,\mathrm{d}\mathbb{P}$$

for all $H \in \mathcal{H}$.

Examples are as follows

Example 12.



Where $Y_1, \dots, Y_n$ is a collection of random variables, we use $\mathbb{E}(X|Y_1, \dots, Y_n)$ to denote $\mathbb{E}(X|\sigma(Y_1, \dots, Y_n))$, the conditional expectation on the σ-algebra generated by the Y_is.

Using this, we can define the conditional probability of an event,

Definition 13. The conditional probability of an event A, conditional on a σ-algebra $\mathcal{A}$, $\mathbb{P}(A|\mathcal{A})$ is given by $\mathbb{E}(1_A|\mathcal{A})$.

Now, we may define the Markov property

Definition 14. A stochastic process X_t, with $a \leq t \leq b$, is said to have the *Markov property* if for all sequences $a < t_1 < \cdots < t_n < t < b$ and corresponding $x_1, \ldots, x_n$, we have

$$\mathbb{P}(X_t \leq x | X_{t_1} = x_1, \ldots, X_{t_n} = x_n) = \mathbb{P}(X_t \leq x | X_{t_n} = x_n).$$

As an example, all processes with independent increments have the Markov property. Before we move on, let us prove a useful lemma

Lemma 15. Suppose a stochastic process X_t defined on $[a, b]$ is adapted to a filtration $\mathcal{F}_t$. Suppose also that

$$\mathbb{P}(X_t \leq x | \mathcal{F}_s) = \mathbb{P}(X_t \leq x | X_s).$$

Then X_t is a Markov process.

Finally, we move to characterise a Markov process by its marginal distributions. Recall that a discrete-time Markov process is characterised by its transition matrix — that is, the probability to transition to a certain state given its current state.

We define a Markov process on $[a, b]$ and let $a = t_1 < t_2$. We have, by Bayes's Theorem

$$\mathbb{P}(X_{t_1} \leq c_1 \cap X_{t_2} \leq c_2) = \mathbb{P}(X_{t_2} \leq c_2 | X_{t_1} \leq c_1)\mathbb{P}(X_{t_1} \leq c_1)$$

The final property that is of interest to us is time translation invariance. For a deterministic differential equation $\dot{x} = f(x)$, we know that the solution exhibits time translation invariance. In the deterministic case, this is quite easy to see, and follows from the fact that $\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t - t_0) = \varphi'(t - t_0)$.

In this case, the relevant property is called the *stationary markov property*

Definition 16. A stochastic process X is called stationary if the moments are time translation invariant:

$$\langle X_{t_1+\tau} X_{t_2+\tau} \ldots X_{t_n+\tau} \rangle = \langle X_{t_1} X_{t_2} \ldots X_{t_n} \rangle$$

for all n, τ and $t_1, \ldots, t_n$.

Thus, we have our final theorem

Theorem 17. Suppose that $b(x)$ and $\sigma(x)$ are functions satisfying the Lipschitz condition. Then the solution to Eq. (1) is a stationary Markov process.

[1] T. Sauer, *Numerical Solution of Stochastic Differential Equations in Finance*, pp. 529–550. 07 2012.