

Stochastic Differential Equations

Jun Wei Tan*

Julius-Maximilians-Universität Würzburg

(Dated: January 28, 2025)

My presentation will be on stochastic differential equations. The first question is: What is a stochastic differential equation?

(Write definition)

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

$$\frac{dX_t}{dt} = \mu(t, X_t) + \sigma(t, X_t)W_t, \quad (1)$$

where W_t is white noise.

White noise can be understood in the intuitive way — as a perturbation whose value at all time t is uncorrelated.

Physically, we can intuitively think of white noise as the derivative of Brownian motion

$$W_t = \frac{dB_t}{dt}.$$

However, as we know from the previous presentations, B_t is almost certainly nowhere differentiable. Hence, we cannot take the derivative. Instead, we are to interpret this in a *distributional* sense: (Write definition)

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

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s$$

This should be reminiscent of the integral form of an ordinary differential equation. Let us look at some solutions.

The standard method to solve a stochastic differential equation is the Itô formula , which

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

Theorem 3 (The 1-Dimensional Itô Formula). Suppose X_t is an Itô process. Let $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$. Then

$$Y_t = g(t, X_t)$$

is also an Itô process, and

$$dY_t = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \cdot (dX_t)^2$$

where $(dX_t)^2$ is computed according to the rules $dt \cdot dt = dt \cdot dB_t = 0$ and $dB_t \cdot dB_t = dt$.

where an Itô process is defined as follows

Definition 4 (Itô Process). An Itô process is a stochastic process of the form

$$X_t = X_a + \int_a^t \mu(s) ds + \int_a^t \sigma(s) dB(s)$$

Using this, we can solve some stochastic differential equations by guessing the solutions.

Let us look at an example:

Example 5. A model of a population is given by the stochastic differential equation

$$\frac{dN_t}{dt} = rN_t + \alpha W_t N_t.$$

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

This is the well known model for a population growth, where the growth of the population is proportional to the population itself. However, here, we introduce a perturbation that could be due to environmental factors. The size of this perturbation is controlled by the parameter α .

~~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:

$$dN_t = rN_t dt + \alpha N_t dB_t,$$

or

$$\frac{dN_t}{N_t} = r dt + \alpha dB_t.$$

Inspired by the solution in the deterministic case, we guess $g(t, x) = \ln x$ in Itô's formula, and let $Y_t = g(t, N_t)$. Then, we have

$$dY_t = \frac{dN_t}{N_t} + \frac{1}{2}\alpha^2 dt.$$

Substituting, we have

$$dY_t = \left(r - \frac{1}{2}\alpha^2\right) dt + \alpha dB_t$$

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 [?].

As in the theory of deterministic differential equations, one of the central results is that of existence and uniqueness of solutions. As an ordinary differential equation is a special case of a stochastic differential equation when we use the dirac delta distribution, or formally, the dirac measure, we begin by recalling the basic theorem for an ordinary differential equation, also known as the Picard-Lindelöf theorem: If

$$\dot{x} = f(t, x) \text{ (write ODE)}$$

is a differential equation, and f is Lipschitz continuous with respect to x , with Lipschitz constant independent of time, then we have a unique solution of the ODE.

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

Theorem 6 (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. This extends to the stochastic case:

Theorem 7. Let $\sigma(t, x)$ and $f(t, x)$ be measurable functions on $[a, b] \times \mathbb{R}$ satisfying the Lipschitz condition in x , and \mathcal{F} a filtration such that the Brownian motion is adapted to it. Suppose ξ is an \mathcal{F}_a -measurable random variable satisfying $\mathbb{E}[\xi^2] < \infty$. Then the stochastic differential equation

$$\frac{dX_t}{dt} = \mu(t, X_t) + \sigma(t, X_t)W_t$$

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.

$$Z_t = \int_a^t (\sigma(s, X_s) - \sigma(s, Y_s)) dB(s) + \int_a^t (\mu(s, X_s) - \mu(s, Y_s)) ds.$$

2. We obtain an integral inequality

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

This is a deterministic differential equation in the squared expectation value, and we have seen this in the theory of deterministic differential equations too. The standard method is to apply what is called Gronwall's inequality, which shows that

$$E(Z_t^2) = 0 \quad \forall t$$

3. At this point, it is important to comment about the importance of the condition of continuity in the theorem, which we have not used yet. What we have shown is that for all time t , we have a set

$$\forall t, \exists \Omega_t \text{ s.t. } \mathbb{P}(\Omega_t) = 1, Z_t = 0 \text{ on } \Omega_t$$

This does not suffice — we need a slightly different condition, that there exists a set Ω , with probability one such that Z_t is 0 on Ω for *all* time!

$$\exists \Omega, \mathbb{P}(\Omega) = 1, Z_t = 0 \text{ on } \Omega \forall t$$

This follows by the definition of sample path continuity, that

$$t \mapsto Z_t(\omega) \quad \forall \omega$$

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

The existence theorem is as follows:

Theorem 8. The stochastic differential equation (1) has a unique solution with initial condition ξ , where ξ^2 has finite expectation, σ and μ 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

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

Then, the proof proceeds in two steps. First, we show that this sequence converges. The estimate we get will be of the form

$$\mathbb{E}|X_{n+1}(t) - X_n(t)| \leq 2B^2(T+1) \int_0^t \mathbb{E}|X_s^{(n)} - X_s^{(n-1)}|^2 ds.$$

This is reminiscent of a geometric sequence, which converges to 0. Using this, we can see that we have uniform convergence on compact intervals. Then, by the Lipschitz continuity, we have $\mu(X^{(n)}) \rightarrow \mu(X)$ and $\sigma(X^{(n)}) \rightarrow \sigma(X)$.

Then, by the dominated convergence theorem, we can take the limit into the integral and show that the limit function X is a solution of the SDE. If it does, and converges in a dominated manner, it satisfies the integral equation by the dominated convergence theorem. \square

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 $x(t)$. Draw graph 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. Perhaps it is best to define the Markov property for discrete time differential equations first. A discrete time system is said to have the Markov property if for all time, we have

$$\mathbb{P}(X_{t+1} \leq x | X_t = x_t, \dots, x_0 = x_0) = \mathbb{P}(X_{t+1} \leq x | X_t = x_t),$$

or that the distribution only depends on the timestep just before where we are — alternatively, where we are going next only depends on where we are now, and not where we were.

In a continuous time process, there is no immediate predecessor, and thus we require the alternative definition

Definition 9. A stochastic process X_t , with $a \leq t \leq b$, is said to have the *Markov property* if for all sequences $a < t_1 < \dots < t_n < t < b$ and corresponding x_1, \dots, x_n , we have

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

As an example, all processes with independent increments have the Markov property. The theorem we seek is thus

Theorem 10. The solution to (1) is a Markov process.

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 $\frac{d}{dt}\varphi(t - t_0) = \varphi'(t - t_0)$. Draw graph

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

Definition 11. A stochastic process X is called stationary if the moments are time translation invariant:

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

for all n, τ and t_1, \dots, t_n .

Thus, we have our final theorem

Theorem 12. Suppose that $\mu(x)$ and $\sigma(x)$ are functions satisfying the Lipschitz condition. Then the solution to

$$\frac{dX_t}{dt} = \mu(X_t) + \sigma(X_t)W_t, \quad (2)$$

is a stationary Markov process.

As an example of this, we solve the Langevin equation. The Langevin equation can be understood to be the brownian motion with an extra damping force proportional to the velocity.

Example 13. The Langevin equation is the SDE given by

$$dX_t = \mu X_t dt + \sigma dB_t.$$

It has solutions

$$X_t = e^{\mu t} X_0 + \int_0^t \sigma e^{\mu(t-s)} dB_s$$

Proof. We multiply by the “integrating factor” $e^{-\mu t}$ and consider

$$Y_t = e^{-\mu t} X_t.$$

By Itô’s formula, we have

$$\begin{aligned} dY_t &= -\mu e^{-\mu t} X_t dt + e^{-\mu t} dX_t \\ &= -\mu e^{-\mu t} X_t dt + e^{-\mu t} (\mu X_t dt + \sigma dB_t) \\ &= e^{-\mu t} \sigma dB_t \end{aligned}$$

implying that

$$X_t = e^{\mu t} X_0 + \int_0^t \sigma e^{\mu(t-s)} dB_s. \quad \square$$

Finally, we note that some stochastic processes can be described through density functions. Where such a density function is available, it satisfies the *Fokker-Planck Equation*

Theorem 14. The probability density of the solution to Eq. (1) $p(x, t)$ satisfies the

equation

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x}[\mu(x, t)p(x, t)] + \frac{\partial^2}{\partial x^2}[D(x, t)p(x, t)]$$

where $D(x, t) = \frac{\sigma^2(X_t, t)}{2}$ is the diffusion coefficient.