Stochastic Differential Equations

Apurva Nakade

2025-02-17

Table of contents

In this module, we will look at the Euler-Maruyama method for solving stochastic differential equations. Our focus will be on understanding numberical stability and convergence of the method.

Recall that the most basic differential equation, which is at the foundation of theory of ordinary differential equations, is the first order ordinary differential equation (ODE) of the form

$$\frac{dy}{dt} = \lambda y \tag{1}$$

where λ is a constant. The solution to this ODE is given by

$$y(t) = y(0)e^{\lambda t}. \tag{2}$$

We want to modify the ODE (Equation 1) to include noise

$$\frac{dy}{dt} = \lambda y + \text{ noise.}$$

To make sense of this, we'll make two changes:

1. We'll replace the function $y(t)$ with a stochastic process $X(t)$.
2. We'll model the noise as a Wiener process $W(t)$.

Stochastic Processes

A stochastic process is a time dependent random variable. More precisely, it is a function of the form

$$X(t, \omega) : \mathbb{R} \times \Omega \to \mathbb{R}$$

where Ω is the sample space. For example, if Ω is the set of gas molecules in a room, then $X(t, i)$ could be the position of the i-th molecule at time t. For at fixed t, $X(t, \cdot)$ is a random variable. For a fixed ω, $X(\cdot, \omega)$ is a function of time. This function, $X(\cdot, \omega)$ is called a **sample path** of the stochastic process.

Wiener Process

A **Wiener process**, $W(t)$, also known as Brownian motion, is a stochastic process with the following properties:

1. $W(0) = 0$.
2. For $0 \leq s < t$, the increment $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$.
3. The increments are independent.
4. The process is continuous but nowhere differentiable.

We can derive the properties of the Wiener process by taking the continuous limit of a random walk. Note that for all time t, the mean of $W(t)$ is 0. However the variance grows with time. Imagine a box of gas particles all starting at the origin without any initial velocity or external forces. As there is no external force or initial velocity, the center of mass of the gas particles will remain at the origin. However, the gas particles will spread out over time. The Wiener process models this spreading out of gas particles.

Sample paths of Wiener process

Distribution at time 1   Distribution at time 4
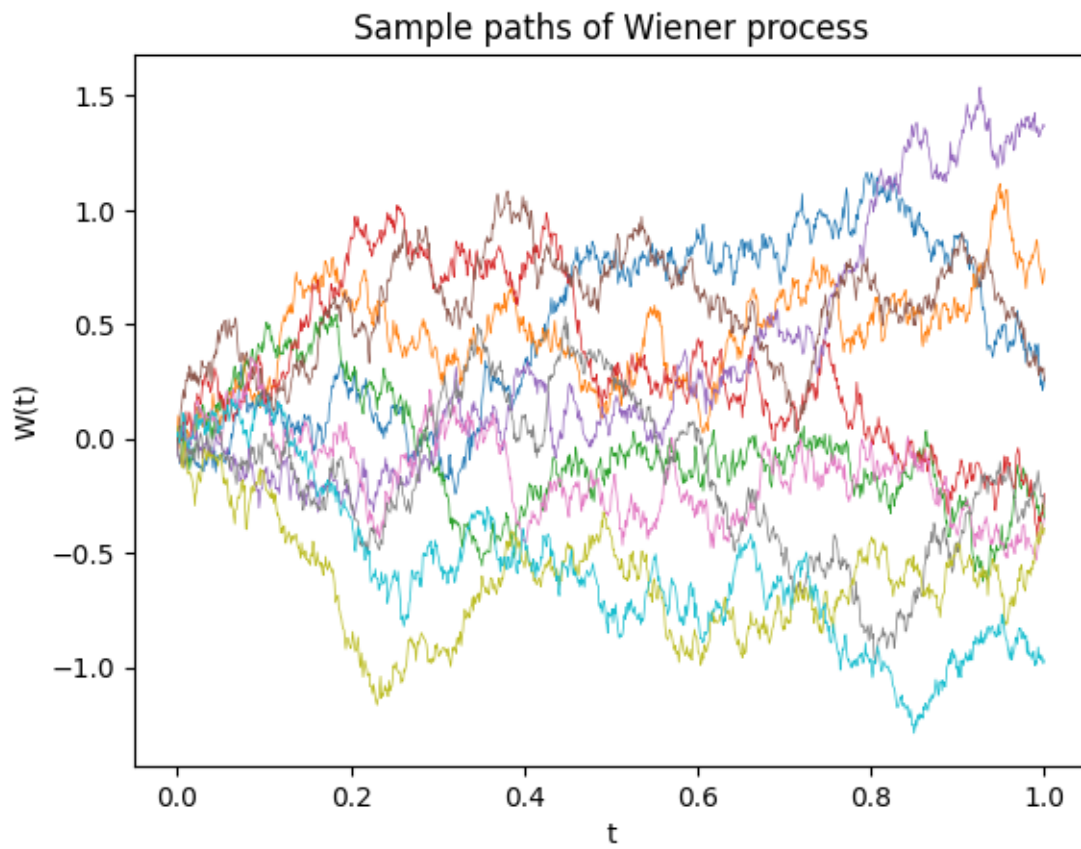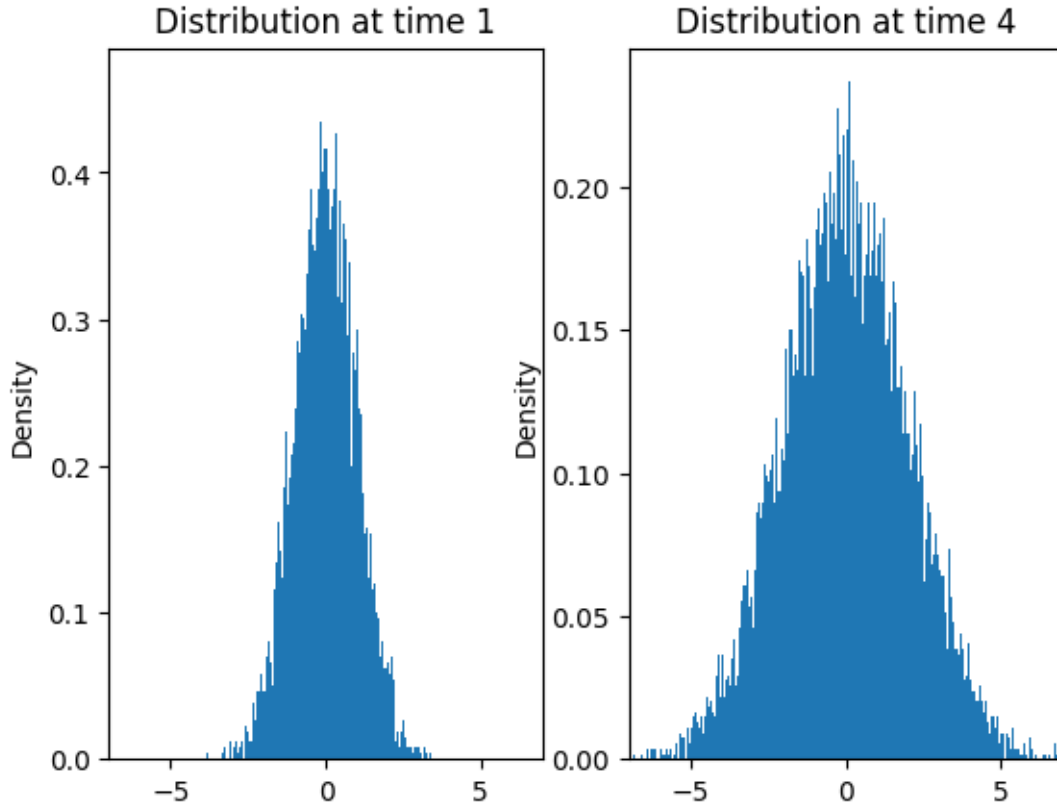
Stochastic Differential Equations

A **stochastic differential equation** (SDE) is an equation of the form

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

where a and b are functions of $X(t)$ and t. The term $a(X(t), t)dt$ is the deterministic part of the equation and $b(X(t), t)dW(t)$ is the stochastic part. The solution to an SDE is a stochastic process $X(t)$. $W(t)$ is the Wiener process.

We write the equation in this form because the Wiener process is not differentiable. We interpret the equation as saying

$$X(t) - X(0) = \int_0^t a(X(s), s)ds + \int_0^t b(X(s), s)dW(s).$$

Ito vs Stratonovich Interpretation

We can try to define the above integrals in the usual sense using Riemann sums. The first integral can be written as the limit

$$\int_0^t a(X(s),s)ds = \lim_{n\to\infty} \sum_{i=0}^{n-1} a(X(t_i),t_i)\Delta t_i$$

where t_i is some point in the interval $[t_i, t_{i+1}]$ and $\Delta t = t/n$. Even though function $a(X(s),s)$ is a stochastic process, it is mathematically still just a function. The integral can be defined in the usual sense.

However, the second integral is more problematic. We can try to define it as

$$\int_0^t b(X(s),s)dW(s) = \lim_{n\to\infty} \sum_{i=0}^{n-1} b(X(t_i),t_i)(W(t_{i+1}) - W(t_i))$$

where t_i is some point in the interval $[t_i, t_{i+1}]$, $\Delta t = t/n$, and $W(t_{i+1}) - W(t_i)$ is normally distributed with mean 0 and variance Δt. The problem is that this integral does not converge in the usual sense. Where the sum converges depends on which point in the interval $[t_i, t_{i+1}]$ we choose to evaluate the integrand. This happens because the Wiener process is not differentiable. There are two commonly used interpretations of the integral:

1. **Ito Interpretation**: In this interpretation, we use the left end point of the interval to evaluate the integrand. This is the most common interpretation.
2. **Stratonovich Interpretation**: In this interpretation, we use the midpoint of the interval to evaluate the integrand.

It is possible to convert between the two interpretations using the Ito formula. We will assume the Ito interpretation in this module.

Geometric Brownian Motion

A simple example of a stochastic differential equation is the **geometric Brownian motion**. This is a model for the evolution of stock prices. The equation is

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$$

where μ is the drift and σ is the volatility. This is one of the few SDEs for which we can find an exact solution. The solution to the Ito version of the equation is

$$X(t) = X(0)e^{(\mu - \sigma^2/2)t + \sigma W(t)}. \tag{3}$$

This is a log-normal distribution whose mean is given by $X(0)e^{\mu t}$ and variance is given by $X(0)^2 e^{2\mu t}(e^{\sigma^2 t} - 1)$. Note that when $\sigma = 0$, this reduces to a standard ODE. Unlike (Equation 2), the solution to the SDE has a quadratic term $\sigma^2/2$ in the exponent. This term appears due to Ito's lemma, which is the stochastic analog of the chain rule.

20 Sample paths of geometric Brownian motion
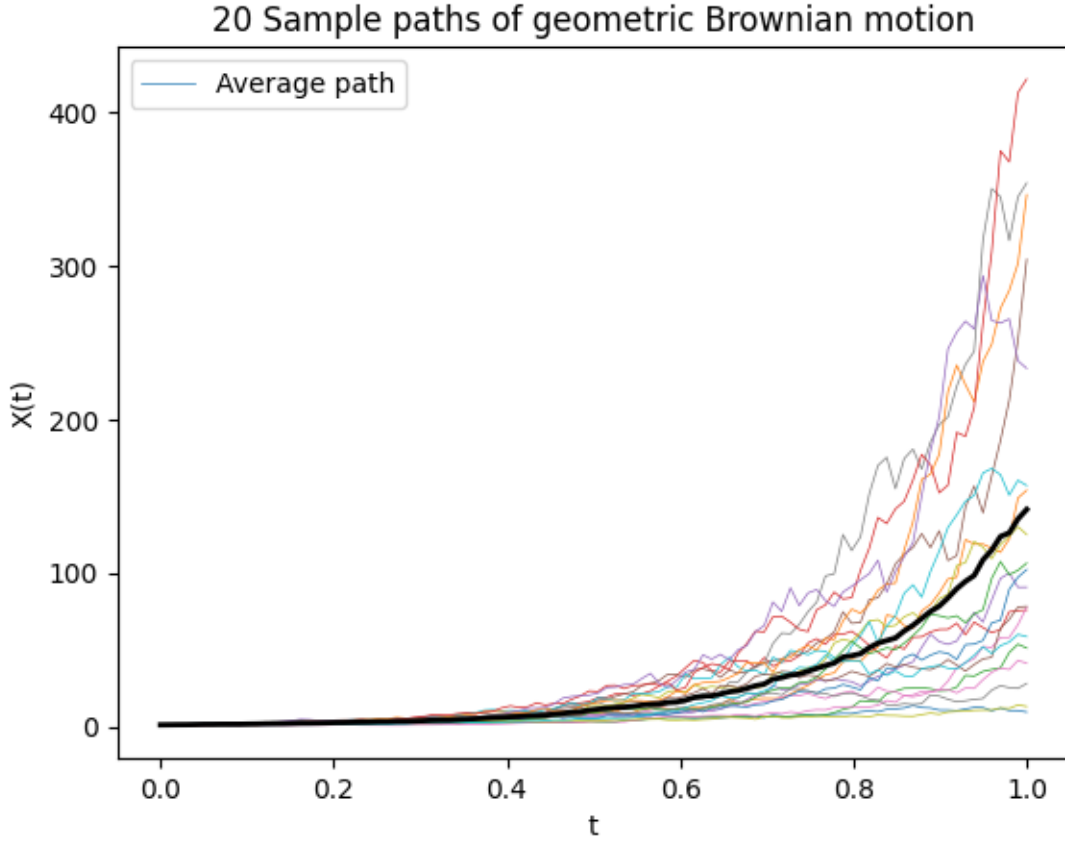
Euler-Maruyama Method

Most SDEs do not have analytical solutions. We need to solve them numerically. Similar to ODEs, some simple regularity conditions on the coefficients a and b imply that the SDE has a unique solution. Most common SDEs satisfy these conditions. However, unlike ODEs, the solution to an SDE is a stochastic process. We can't just evaluate the solution at a few points to get an approximate solution. We need to generate several sample paths of the stochastic process to get an approximate solution.

The simplest method for solving SDEs is the **Euler-Maruyama method**. This is a stochastic analog of the Euler method for ODEs. The Euler-Maruyama method is a recursive method. Given the value of the stochastic process X_n at time t_n, we can find the value of the process at time $t_{n+1} = t_n + \Delta t$ using the formula

$$X_{n+1} = X_n + a(X_n, t_n)\Delta t + b(X_n, t_n)\Delta W_n$$

where $\Delta t = t_{n+1} - t_n$ and $\Delta W_n = W(t_{n+1}) - W(t_n) \sim \mathcal{N}(0, \Delta t)$. This results in a simple algorithm for solving SDEs:

1. Initialize X_0.
2. For $n = 0, 1, 2, ..., N - 1$, do

    1. Generate a random number $\Delta W_n \sim \mathcal{N}(0, \Delta t)$.
    2. Compute $X_{n+1} = X_n + a(X_n, t_n)\Delta t + b(X_n, t_n)\Delta W_n$.

3. Return $X_0, X_1, ..., X_N$.

Convergence of the Euler-Maruyama Method

The EM method provides a numerical approximation to the solution of the SDE. We want to understand how good this approximation is.

Fix a time interval $[0, T]$. Divide the interval $[0, T]$ into N subintervals of length $\Delta t = T/N$. Let $t_i = i\Delta t$ so that $t_0 = 0$ and $t_N = T$. Let $X(t)$ be the analytical solution to the SDE at time t with initial condition $X(0)$. We first think of the EM method as generating a sequence of random variables $X_0, X_1, ..., X_N$ defined recursively by

$$\begin{aligned} X_0 &= X(0), \\ X_{i+1} &= X_i + a(X_i, t_i)\Delta t + b(X_i, t_i)\Delta W_i. \end{aligned}$$

where $\Delta W_i \sim \mathcal{N}(0, \Delta t)$.

Then we are interested in the question of how good the sequence $X_0, X_1, ..., X_N$ is as an approximation to the solution $X(t_0), X(t_1), ..., X(t_N)$. If the EM method is a good approximation method, we should get

$$X_i \to X(t_i) \text{ as } \Delta t \to 0.$$

However, as X_i and $X(t_i)$ are random variables, we need to be more precise about what we mean by convergence. There are two notions of convergence that we are interested in: weak convergence and strong convergence.

Weak Convergence

With the setup as above, we say that the EM method converges weakly to the solution over the interval $[0, T]$ if

$$\mathbb{E}[X_i] \to \mathbb{E}[X(t_i)] \text{ as } \Delta t \to 0.$$

Since we are interested in using the EM method to approximate the solution to the SDE, we need more than just convergence in expectation. We need to know how fast the method converges. We say that the EM method converges weakly to the solution over the interval $[0, T]$ with order p if

$$|\mathbb{E}[X_i] - \mathbb{E}[X(t_i)]| \le C\Delta t^p$$

For some constant C that does not depend on Δt. Note that C is allowed to depend on T and $X(0)$. We often write this as

$$|\mathbb{E}[X_i] - \mathbb{E}[X(t_i)]| = O(\Delta t^p)$$

to emphasize the rate of convergence.

Strong Convergence

Strong convergence is the convergence of the sample paths. We say that the EM method converges strongly to the solution over the interval $[0, T]$ if

$$X_i \xrightarrow{a.s.} X(t_i) \text{ as } \Delta t \to 0.$$

Recall that convergence almost surely means that the probability of the event that the sequence X_i does not converge to $X(t_i)$ is 0. By Markov's inequality it is enough to say that the expected value of the distance between X_i and $X(t_i)$ goes to 0.

$$\mathbb{E}[|X_i - X(t_i)|] \to 0 \text{ as } \Delta t \to 0.$$

We say that the EM method converges strongly to the solution over the interval $[0, T]$ with order p if

$$\mathbb{E}[|X_i - X(t_i)|] = O(\Delta t^p)$$

In the homework assignment, you will derive the rates of convergence for the EM method for the geometric Brownian motion *empirically.* However, it is necessary to derive these rates theoretically as for arbitrary SDEs, it is not possible to write an exact analytical solution. The rates of convergence tell us how close the analytical solution is to the numerical solution.

Also, you'll only estimate p at the end of the interval i.e. $t = T$. This is technically not correct. We should find a p that works for each time step t_i. However, this is computationally expensive and the assumption is that error grows with time so that the error at the end of the interval is the largest.

Final remarks

The EM method has a slower rate of strong convergence compared to weak convergence. This is because even though the EM method is a natural extension of the Euler method for ODEs, there are second order terms in the Ito formula (chain rule for stochastic processes) that contribute to the error. This is fixed in the Milstein method, which is a second order method for solving SDEs. However the Milstein method is more computationally expensive compared to the EM method and has the same weak rate of convergence as the EM method.