1 Stochastic Taylor Expansion

In this lecture, we discuss the stochastic version of the Taylor expansion to understand how stochastic integration methods are designed. In addition, we illustrate why the Euler method is strongly convergent with order 1/2 and is weakly convergent with order 1.

1.1 Ito-Taylor Expansion

First, let us recall how we can obtain a Taylor expansion from an integral representation for the deterministic case. Consider the autonomous ODE,

$$\frac{d}{dt}X(t) = a[X(t)].$$

Let f be a function of $X(t)$, then the evolution of the function f is governed by

$$\frac{d}{dt}f[X(t)] = a[X(t)]\frac{\partial}{\partial X}f[X(t)] \tag{1}$$

via the chain rule.

Integral Equation: By defining a linear operator,

$$\mathcal{L} \equiv a(X)\frac{\partial}{\partial X},$$

we can rewrite Eq. (1) in terms of the integral equation:

$$f[X(t)] = f[X(t_0)] + \int_{t_0}^{t} \mathcal{L}f[X(\tau_1)]\,d\tau_1 \tag{2}$$

where t_0 is the initial time.

Iterations: Iterating Eq. (2) leads to

$$\begin{aligned}
f[X(t)] &= f[X(t_0)] + \int_{t_0}^{t} \mathcal{L}\left[f[X(t_0)] + \int_{t_0}^{\tau_1} \mathcal{L}f[X(\tau_2)]\,d\tau_2\right]d\tau_1 \\
&= f[X(t_0)] + \mathcal{L}f[X(t_0)]\int_{t_0}^{t} d\tau_1 + \int_{t_0}^{t}\int_{t_0}^{\tau_1} \mathcal{L}^2 f[X(\tau_2)]\,d\tau_2 d\tau_1 \\
&= f[X(t_0)] + \mathcal{L}f[X(t_0)](t-t_0) + \int_{t_0}^{t}\int_{t_0}^{\tau_1} \mathcal{L}^2 f[X(\tau_2)]\,d\tau_2 d\tau_1
\end{aligned}$$

If we iterate once more using Eq. (2) for $f(X(\tau_2))$ in the integrand of the double integral,

$$\begin{aligned}\mathcal{L}^2 f[X(\tau_2)] &= \mathcal{L}^2\left[f[X(t_0)] + \int_{t_0}^{\tau_2} \mathcal{L} f[X(\tau_3)]\,d\tau_3\right] \\ &= \mathcal{L}^2 f[X(t_0)] + \underbrace{\mathcal{L}^2 \int_{t_0}^{\tau_2} \mathcal{L} f[X(\tau_3)]\,d\tau_3}_{O(\mathcal{L}^3)}\end{aligned}$$

then the contribution of the first term above, which is a constant, to the double integral is

$$\begin{aligned}\int_{t_0}^{t}\int_{t_0}^{\tau_1} \mathcal{L}^2 f[X(t_0)]\,d\tau_2 d\tau_1 &= \mathcal{L}^2 f[X(t_0)] \int_{t_0}^{t}\int_{t_0}^{\tau_1} d\tau_2 d\tau_1 \\ &= \frac{1}{2}\mathcal{L}^2 f[X(t_0)](t-t_0)^2\end{aligned}$$

Putting all these together, we have

$$f[X(t)] = f[X(t_0)] + \mathcal{L} f[X(t_0)](t-t_0) + \frac{1}{2}\mathcal{L}^2 f[X(t_0)](t-t_0)^2 + O\left(\mathcal{L}^3\right)$$

which is precisely the Taylor expansion we are familiar with.

1.1.1 Ito-Taylor Expansion

Now we discuss a similar expansion for the stochastic differential equation:

$$dX(t) = a[X(t)]\,dt + b[X(t)]\,dW(t)\,. \tag{3}$$

Again, for simplicity, we consider the autonomous case, i.e., $a = a[X(t)]$, $b = b[X(t)]$ and they do not depend on time explicitly.

The Ito lemma leads to

$$df[X(t)] = \left\{a\frac{\partial}{\partial X} f[X(t)] + \frac{1}{2}b^2[X(t)]\frac{\partial^2}{\partial X^2} f[X(t)]\right\} dt + b[X(t)]\frac{\partial}{\partial X} f[X(t)]\,dW(t) \tag{4}$$

Defining

$$\begin{aligned}\mathcal{L}^0 &\equiv a\frac{\partial}{\partial X} + \frac{1}{2}b^2[X]\frac{\partial^2}{\partial X^2} \\ \mathcal{L}^1 &\equiv b[X]\frac{\partial}{\partial X}\end{aligned}$$

then Eq. (4) becomes

$$df[X(t)] = \mathcal{L}^0 f[X(t)]\,dt + \mathcal{L}^1 f[X(t)]\,dW(t)$$

i.e.,

$$f[X(t)] = f[X(t_0)] + \int_{t_0}^{t} \mathcal{L}^0 f[X(s)]\,ds + \int_{t_0}^{t} \mathcal{L}^1 f[X(s)]\,dW(s) \tag{5}$$

In equation (5), if we choose the following $f(x)$:

1. Choose $f(x) = x$, then Eq. (5) becomes

$$X(t) = X(t_0) + \int_{t_0}^{t} a[X(s)]\,ds + \int_{t_0}^{t} b[X(s)]\,dW(s) \tag{6}$$

2. Choose $f(x) = a(x)$, then Eq. (5) becomes

$$a[X(t)] = a[X(t_0)] + \int_{t_0}^{t} \mathcal{L}^0 a[X(s)]\,ds + \int_{t_0}^{t} \mathcal{L}^1 a[X(s)]\,dW(s) \tag{7}$$

3. Similarly, choose $f(x) = b(x)$, then Eq. (5) becomes

$$b[X(t)] = b[X(t_0)] + \int_{t_0}^{t} \mathcal{L}^0 b[X(s)]\,ds + \int_{t_0}^{t} \mathcal{L}^1 b[X(s)]\,dW(s) \tag{8}$$

Substituting Eqs. (7) and (8) into Eq. (6) leads to

$$\begin{aligned} X(t) &= X(t_0) + \int_{t_0}^{t} \left\{ a[X(t_0)] + \int_{t_0}^{s_1} \mathcal{L}^0 a[X(s_2)]\,ds_2 + \int_{t_0}^{s_1} \mathcal{L}^1 a[X(s_2)]\,dW(s_2) \right\} ds_1 \\ &\quad + \int_{t_0}^{t} \left\{ b[X(t_0)] + \int_{t_0}^{s_1} \mathcal{L}^0 b[X(s_2)]\,ds_2 + \int_{t_0}^{s_1} \mathcal{L}^1 b[X(s_2)]\,dW(s_2) \right\} dW(s_1) \end{aligned} \tag{9}$$

Note that

$$\begin{aligned} \mathcal{L}^0 a &= a\frac{\partial}{\partial X}a + \frac{1}{2}b^2[X]\frac{\partial^2}{\partial X^2}a \equiv aa' + \frac{1}{2}b^2 a'' \\ \mathcal{L}^0 b &= ab' + \frac{1}{2}b^2 b'' \\ \mathcal{L}^1 a &= b\frac{\partial}{\partial X}a = ba' \\ \mathcal{L}^1 b &= bb' \end{aligned}$$

Separating the constant terms out in the integrand in Eq. (9) from the remaining terms, which are the double integral terms:

$$\begin{aligned} R &\equiv \int_{t_0}^{t}\int_{t_0}^{s_1} \mathcal{L}^0 a[X(s_2)]\,ds_2 ds_1 + \int_{t_0}^{t}\int_{t_0}^{s_1} \mathcal{L}^1 a[X(s_2)]\,dW(s_2)\,ds_1 \\ &\quad + \int_{t_0}^{t}\int_{t_0}^{s_1} \mathcal{L}^0 b[X(s_2)]\,ds_2 dW(s_1) + \int_{t_0}^{t}\int_{t_0}^{s_1} \mathcal{L}^1 b[X(s_2)]\,dW(s_2)\,dW(s_1), \end{aligned}$$

leads to

$$X(t) = X(t_0) + a[X(t_0)]\int_{t_0}^{t} ds_1 + b[X(t_0)]\int_{t_0}^{t} dW(s_1) + R \tag{10}$$

Note that the essence of the method is to use the substitution repeatedly to obtain constant integrands in higher and higher order terms. For example, the last term in the remainder, R, (which is of the lowest order in Δt in R if $\Delta t \equiv t - t_0$ is small) is

$$\begin{aligned}
& \int_{t_0}^{t}\int_{t_0}^{s_1} \mathcal{L}^1 b\left[X\left(s_2\right)\right] dW\left(s_2\right) dW\left(s_1\right) \\
= & \int_{t_0}^{t}\int_{t_0}^{s_1} \left\{\mathcal{L}^1 b\left[X\left(t_0\right)\right] + \int_{t_0}^{s_2} \mathcal{L}^0\mathcal{L}^1 b\left[X\left(s_3\right)\right] ds_3 + \int_{t_0}^{s_2} \mathcal{L}^1\mathcal{L}^1 b\left[X\left(s_3\right)\right] dW\left(s_3\right)\right\} dW\left(s_2\right) dW\left(s_1\right)
\end{aligned}$$

by selecting $f = \mathcal{L}^1 b$ in Eq. (5). The first term in the last line of the above equation is

$$b\left[X\left(t_0\right)\right] b'\left[X\left(t_0\right)\right] \int_{t_0}^{t}\int_{t_0}^{s_1} dW\left(s_2\right) dW\left(s_1\right)$$

by noting $\mathcal{L}^1 b = bb'$. Therefore, Eq. (10) becomes

$$\begin{aligned}
X\left(t\right) &= X\left(t_0\right) + a\left[X\left(t_0\right)\right] \int_{t_0}^{t} ds_1 + b\left[X\left(t_0\right)\right] \int_{t_0}^{t} dW\left(s_1\right) \\
&\quad + b\left[X\left(t_0\right)\right] b'\left[X\left(t_0\right)\right] \int_{t_0}^{t}\int_{t_0}^{s_1} dW\left(s_2\right) dW\left(s_1\right) + \tilde{R} \qquad (11)
\end{aligned}$$

where $\tilde{R}$ is a new remainder. Eq. (11) is the Ito-Taylor expansion for the process (3).

Note that the double integral in Eq. (11) is evaluated to be

$$\int_{t_0}^{t}\int_{t_0}^{s_1} dW\left(s_2\right) dW\left(s_1\right) = \frac{1}{2}\left[W\left(t\right) - W\left(t_0\right)\right]^2 - \frac{1}{2}\left(t - t_0\right),$$

which can be shown as follows

$$\begin{aligned}
\int_{t_0}^{t}\int_{t_0}^{s_1} dW\left(s_2\right) dW\left(s_1\right) &= \int_{t_0}^{t} \left[W\left(s_1\right) - W\left(t_0\right)\right] dW\left(s_1\right) \\
&= \int_{t_0}^{t} W\left(s_1\right) dW\left(s_1\right) - \int_{t_0}^{t} W\left(t_0\right) dW\left(s_1\right) \\
&= \frac{1}{2}\int_{t_0}^{t} \left[dW^2\left(s_1\right) - dt\right] - \int_{t_0}^{t} W\left(t_0\right) dW\left(s_1\right) \\
&\qquad \left(\because dW^2 = 2W dW + dt \quad \text{(Ito Lemma)}\right) \\
&= \frac{1}{2}\left[W^2\left(t\right) - W^2\left(t_0\right) - \left(t - t_0\right)\right] - W\left(t_0\right)\left[W\left(t\right) - W\left(t_0\right)\right] \\
&= \frac{1}{2}\left[W\left(t\right) - W\left(t_0\right)\right]^2 - \frac{1}{2}\left(t - t_0\right) \qquad (12)
\end{aligned}$$

1.1.2 Numerical Integration Schemes

Once we have the Ito-Taylor expansion, we can construct numerical integration schemes for the process (3). For the interval $[t_i, t_{i+1}]$, by choosing

$$\begin{aligned} t_0 &= t_i, \quad t = t_{i+1}, \\ \Delta t &= t_{i+1} - t_i \\ \Delta W_i &= W(t_{i+1}) - W(t_i), \end{aligned}$$

combining Eq. (11) and Eq. (12) yields

$$X(t_{i+1}) = X(t_i) + a[X(t_i)]\Delta t + b[X(t_i)]\Delta W_i + \frac{1}{2}b[X(t_i)]b'[X(t_i)]\left[(\Delta W_i)^2 - \Delta t\right] + \tilde{R}. \tag{13}$$

1. Keeping the first three terms in Eq. (13) gives us the explicit Euler method:

$$\hat{X}_{i+1} = \hat{X}_i + a\left[\hat{X}_i\right]\Delta t + b\left[\hat{X}_i\right]\Delta W_i$$

2. Keeping all terms of $O(\Delta t)$ gives us the Milstein method:

$$\hat{X}_{i+1} = \hat{X}_i + a\left[\hat{X}_i\right]\Delta t + b\left[\hat{X}_i\right]\Delta W_i + \frac{1}{2}b\left[\hat{X}_i\right]b'\left[\hat{X}_i\right]\left[(\Delta W_i)^2 - \Delta t\right] \tag{14}$$

Note that the Milstein needs to compute the derivative b' for the last term.

Runge-Kutta Methods Note that the Milstein method requires to compute the derivative of b. Some times it is computationally costly to compute derivatives. We can construct Runge-Kutta schemes to avoid this, as for the deterministic case.

For $\Delta X = a\Delta t + b\Delta W$, we have

$$b(X + \Delta X) - b(X) = b'(X)\Delta X + O\left((\Delta X)^2\right),$$

therefore,

$$\begin{aligned} b(X + \Delta X) - b(X) &= b'(X)[a(X)\Delta t + b(X)\Delta W] + O(\Delta t) \qquad \left(\because (\Delta X)^2 \sim \Delta t\right) \\ &= b'(X)b(X)\Delta W + O(\Delta t) \end{aligned}$$

$$\begin{aligned} &\therefore \\ b'(X_i)b(X_i) &\sim \frac{1}{\sqrt{\Delta t}}[b(X_i + a(X_i)\Delta t + b(X_i)\Delta W_i) - b(X_i)] + O(\Delta t)^{1/2} \\ &\sim \frac{1}{\sqrt{\Delta t}}\left[b\left(X_i + a(X_i)\Delta t + b(X_i)\sqrt{\Delta t}\right) - b(X_i)\right] + O(\Delta t)^{1/2} \end{aligned}$$

Thus, we have the following RK method

$$\begin{cases} \tilde{X}_i = X_i + a(X_i)\Delta t + b(X_i)\sqrt{\Delta t} \\ X_{i+1} = X_i + a(X_i)\Delta t + b(X_i)\Delta W_i + \dfrac{1}{2\sqrt{\Delta t}}\left[b\left(\tilde{X}_i\right) - b(X_i)\right]\left[(\Delta W_i)^2 - \Delta t\right] \end{cases}$$

which has order-one strong convergence. (Question: how would you demonstrate this convergence of order one numerically?)

Higher order RK schemes can be constructed similarly.
Clearly, if $b =$constant, the RK method above reduces to the Euler scheme.

1.2 Weak and Strong Convergence for Euler Scheme

Now we turn to getting some intuitive feeling why the Euler scheme has strong order 1/2 and weak order 1. For simplicity, we analyze the convergence for the geometric Brownian motion

$$dS(t) = \mu S(t)\,dt + \sigma S(t)\,dW(t)$$

where μ and σ are constant. Recall that the exact solution for this process is

$$S(t) = S(0)\,e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)}. \tag{15}$$

The Euler scheme for this process is

$$\begin{aligned} \hat{S}(t_{i+1}) &= S(t_i) + \mu\hat{S}(t_i)\Delta t + \sigma\hat{S}(t_i)\Delta W(t_i) \\ \textit{i.e.,}\quad \hat{S}(t_{i+1}) &= \hat{S}(t_i)\left[1 + \mu\Delta t + \sigma\Delta W(t_i)\right] \end{aligned}$$

where $\Delta W(t_i) = \sqrt{\Delta t}Z_i$, $Z_i \sim \mathcal{N}(0,1)$. Then, obviously, the numerical solution at $t = t_k$ is

$$\hat{S}(t_k) = S(0)\prod_{i=0}^{k-1}\left[1 + \mu\Delta t + \sigma\Delta W(t_i)\right]$$

1.2.1 The Error for Strong Convergence

For $t_0 = 0, t_1 = \frac{T}{n}, t_2 = \frac{2T}{n}, \cdots, t_n = T$, by definition, the error in strong convergence is

$$\begin{aligned} &\mathbb{E}\left|\hat{S}(T) - S(T)\right| \\ =\ & S(0)\,\mathbb{E}\left|\prod_{i=0}^{n-1}\left[1 + \mu\Delta t + \sigma\Delta W(t_i)\right] - e^{\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W(T)}\right| \end{aligned} \tag{16}$$

where $\hat{S}(T)$ is numerically evaluated using the Euler method and $S(T)$ is the exact solution (Eq.(15)).

To $\mathcal{O}(\Delta t)^2$, Taylor expansion leads to

$$\begin{aligned}
e^{\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)} &= 1+\left[\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)\right] \\
&\quad +\frac{1}{2}\left[\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)\right]^2 \\
&\quad +\frac{1}{6}\left[\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)\right]^3 \\
&\quad +\cdots \\
&= 1+\left[\left(\mu-\underline{\frac{1}{2}\sigma^2}\right)\Delta t+\sigma W(t_i)\right] \\
&\quad +\left(\mu-\frac{1}{2}\sigma^2\right)\sigma\Delta t\Delta W(t_i)+\underline{\frac{1}{2}\sigma^2\left[\Delta W(t_i)\right]^2} \qquad \left(\text{N.B. } (\Delta W)^2=\Delta t\right) \\
&\quad +\frac{1}{6}\sigma^3\left[\Delta W(t_i)\right]^3 \quad +\mathcal{O}(\Delta t)^2 \\
&= 1+\mu\Delta t+\sigma W(t_i)+\mu'\sigma\Delta t\Delta W(t_i)+\frac{1}{6}\sigma^3\left[\Delta W(t_i)\right]^3 \quad +\mathcal{O}(\Delta t)^2
\end{aligned}$$

where $\mu' \equiv \left(\mu-\frac{1}{2}\sigma^2\right)$. Therefore,

$$\begin{aligned}
&1+\mu\Delta t+\sigma W(t_i) \\
&= e^{\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)}-\mu'\sigma\Delta t\Delta W(t_i)-\frac{1}{6}\sigma^3\left[\Delta W(t_i)\right]^3-\mathcal{O}(\Delta t)^2
\end{aligned}$$

$\therefore$

$$\prod_{i=0}^{n-1}\left[1+\mu\Delta t+\sigma W(t_i)\right]=\prod_{i=0}^{n-1}\left[e^{\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma W(t_i)}-\mu'\sigma\Delta t\Delta W(t_i)-\frac{1}{6}\sigma^3\left[\Delta W(t_i)\right]^3\right]$$

$$\therefore \prod_{i=0}^{n-1}\left[1+\mu\Delta t+\sigma W(t_i)\right]=e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}+n\mathcal{O}(\Delta t\Delta W)+n\mathcal{O}(\Delta W)^3+n\mathcal{O}(\Delta t)^2 \quad (17)$$

Hence, the error (16) is

$$\begin{aligned}
&\mathbb{E}\left|\hat{S}(T)-S(T)\right| \\
&=\mathbb{E}\left|n\mathcal{O}(\Delta t\Delta W)+n\mathcal{O}(\Delta W)^3+n\mathcal{O}(\Delta t)^2\right| \\
&=\mathbb{E}\left|\frac{T}{\Delta t}\mathcal{O}(\Delta t\Delta W)+\frac{T}{\Delta t}\mathcal{O}(\Delta W)^3+\frac{T}{\Delta t}\mathcal{O}(\Delta t)^2\right| \qquad \left(\because\ n=\frac{T}{\Delta t}\right) \\
&=T\mathbb{E}\left|\underbrace{\frac{1}{\Delta t}\mathcal{O}(\Delta t\Delta W)}_{\mathcal{O}(\Delta t)^{1/2}}+\underbrace{\frac{1}{\Delta t}\mathcal{O}(\Delta W)^3}_{\mathcal{O}(\Delta t)^{1/2}}+\underbrace{\frac{1}{\Delta t}\mathcal{O}(\Delta t)^2}_{\mathcal{O}(\Delta t)}\right| \\
&=\mathcal{O}(\Delta t)^{1/2}
\end{aligned}$$

This demonstrates that the Euler scheme has strong convergence of order 1.

1.2.2 The Error for Weak Convergence

For weak convergence, we need to specify the function f in the error estimate:

$$\left|\mathbb{E}\left[f\left(\hat{S}(T)\right)\right]-\mathbb{E}\left[f(S(T))\right]\right|\le\beta\Delta t^q$$

For the convergence of expectation, i.e, the first moment, we choose

$$f(x)=x$$

For the geometric Brownian motion, we thus have the "weak" error

$$\begin{aligned}
&\left|\mathbb{E}\left[\hat{S}(T)\right]-\mathbb{E}\left[S(T)\right]\right| \\
&=S(0)\left|\mathbb{E}\left[\prod_{i=0}^{n-1}\left[1+\mu\Delta t+\sigma W(t_i)\right]\right]-\mathbb{E}\left[e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}\right]\right| \\
&=S(0)\left|\mathbb{E}\left[e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}+n\mathcal{O}(\Delta t\Delta W)+n\mathcal{O}(\Delta W)^3+n\mathcal{O}(\Delta t)^2\right]-\mathbb{E}\left[e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}\right]\right| \\
&\qquad\left(\because\prod_{i=0}^{n-1}\left[1+\mu\Delta t+\sigma W(t_i)\right]\right. \\
&\qquad\qquad\left.=e^{\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}+n\mathcal{O}(\Delta t\Delta W)+n\mathcal{O}(\Delta W)^3+n\mathcal{O}(\Delta t)^2\right) \\
&=S(0)\left|\mathbb{E}\left[n\mathcal{O}(\Delta t\Delta W)+n\mathcal{O}(\Delta W)^3+n\mathcal{O}(\Delta t)^2\right]\right|
\end{aligned}$$

Note that

$$\mathbb{E}\left[\mathcal{O}\left(\Delta t \Delta W\right)\right]=0, \quad \mathbb{E}\left[\mathcal{O}\left(\Delta W\right)^3\right]=0$$

this leads to the order of weak convergence for Euler scheme:

$$\begin{aligned}\left|\mathbb{E}\left[\hat{S}\left(T\right)\right]-\mathbb{E}\left[S\left(T\right)\right]\right| &= S\left(0\right)\left|\mathbb{E}\left[n\mathcal{O}\left(\Delta t\right)^2\right]\right| \\ &= S\left(0\right)n\mathcal{O}\left(\Delta t\right)^2 \\ &= \frac{T}{\Delta t}\mathcal{O}\left(\Delta t\right)^2 \\ &= \mathcal{O}\left(\Delta t\right)\end{aligned}$$

which has order one.

The main difference between the strong and weak convergence lies in whether the odd powered terms, such as ΔW, ΔW^3, can be eliminated:

$$\begin{aligned}\text{For strong convergence:} \quad & \mathbb{E}\left[\left|\mathcal{O}\left(\Delta W\right)+h.o.t.\right|\right]=\mathbb{E}\left|\mathcal{O}\left(\Delta t\right)^{1/2}+h.o.t.\right| \\ \text{For weak convergence:} \quad & \left|\mathbb{E}\left[\mathcal{O}\left(\Delta W\right)+h.o.t.\right]\right|=\left|0+h.o.t.\right|\end{aligned}$$

where $h.o.t.$ is "higher order terms".

Finally, we try to get a little more intuition of why the Euler scheme has order one of strong convergence for the process

$$dS=\mu dt+\sigma dW$$

when σ is *deterministic*. In this case, the Euler scheme is

$$\hat{S}\left(t_{i+1}\right)=\hat{S}\left(t_i\right)+\mu\left(t_i\right)\Delta t+\sigma\left(t_i\right)\Delta W\left(t_i\right)$$

therefore at $t_n=T$,

$$\hat{S}\left(T\right)=S\left(0\right)+\sum_{i=0}^{n-1}\mu\left(t_i\right)\Delta t+\sum_{i=0}^{n-1}\sigma\left(t_i\right)\Delta W\left(t_i\right).$$

Since the exact solution is

$$S\left(T\right)=S\left(0\right)+\int_0^T\mu\left(t\right)dt+\int_0^T\sigma\left(t\right)dW\left(t\right),$$

we have

$$\begin{aligned}\hat{S}\left(T\right)-S\left(T\right) &= \sum_{i=0}^{n-1}\mu\left(t_i\right)\Delta t-\int_0^T\mu\left(t\right)dt \\ &\quad +\underbrace{\sum_{i=0}^{n-1}\sigma\left(t_i\right)\Delta W\left(t_i\right)}_{\equiv A}-\underbrace{\int_0^T\sigma\left(t\right)dW\left(t\right)}_{\equiv B}\end{aligned}$$

the first line of which clearly has the $\mathcal{O}(\Delta t)$ error. Now we analyze the error for the second line in the above equation. Define

$$\begin{aligned} A &\equiv \sum_{i=0}^{n-1} \sigma(t_i)\,\Delta W(t_i), \\ B &\equiv \int_0^T \sigma(t)\,dW(t) \end{aligned}$$

Note that

1. A is a Gaussian (since a linear superposition of Gaussian random variables is a Gaussian random variable) with the expectation

$$\mathbb{E}A = \mathbb{E}\left(\sum_{i=0}^{n-1} \sigma(t_i)\,\Delta W(t_i)\right) = 0$$

   and variance

$$\mathrm{Var}A = \sum_{i=0}^{n-1} \sigma^2(t_i)\,\Delta t. \tag{18}$$

   Therefore, the Gaussian random variable A can be expressed as

$$A = \frac{W(T)}{\sqrt{T}} \sqrt{\sum_{i=0}^{n-1} \sigma^2(t_i)\,\Delta t}$$

   since the right-hand side of the equation above is a Gaussian random variable with mean zero and variance the same as Eq. (18).

2. B is also a Gaussian random variable with mean and variance:

$$\begin{aligned} \mathbb{E}B &= \mathbb{E}\left(\int_0^T \sigma(t)\,dW(t)\right) = 0 \\ \mathrm{Var}B &= \int_0^T \sigma^2(t)\,dt \end{aligned}$$

   therefore, B can be expressed as

$$B = \frac{W(T)}{\sqrt{T}} \sqrt{\int_0^T \sigma^2(t)\,dt}.$$

Hence,

$$A - B = \frac{W(T)}{\sqrt{T}} \left[\sqrt{\sum_{i=0}^{n-1} \sigma^2(t_i)\,\Delta t} - \sqrt{\int_0^T \sigma^2(t)\,dt}\right].$$

Since

$$\sum_{i=0}^{n-1} \sigma^2(t_i)\,\Delta t - \int_0^T \sigma^2(t)\,dt = \mathcal{O}(\Delta t),$$

$$\sum_{i=0}^{n-1} \sigma^2(t_i)\,\Delta t = \int_0^T \sigma^2(t)\,dt + \mathcal{O}(\Delta t).$$

$\therefore$ By defining $\Sigma \equiv \int_0^T \sigma^2(t)\,dt$, we have

$$\begin{aligned} A - B &= \frac{W(T)}{\sqrt{T}}\left[\sqrt{\Sigma + \mathcal{O}(\Delta t)} - \sqrt{\Sigma}\right] \\ &= \frac{W(T)}{\sqrt{T}}\left[\sqrt{\Sigma}\sqrt{1 + \frac{\mathcal{O}(\Delta t)}{\Sigma}} - \sqrt{\Sigma}\right] \\ &= \frac{W(T)}{\sqrt{T}}\left[\sqrt{\Sigma}\left(1 + \frac{1}{2}\frac{\mathcal{O}(\Delta t)}{\Sigma} + \mathcal{O}(\Delta t)^2\right) - \sqrt{\Sigma}\right] \\ &= \frac{W(T)}{\sqrt{T}}\frac{1}{2}\frac{\mathcal{O}(\Delta t)}{\sqrt{\Sigma}} \end{aligned}$$

Therefore,

$$\begin{aligned} \hat{S}(T) - S(T) &= \mathcal{O}(\Delta t) + \frac{1}{2}\frac{W(T)}{\sqrt{T}}\frac{\mathcal{O}(\Delta t)}{\sqrt{\Sigma}} \\ &= \mathcal{O}(\Delta t) \end{aligned}$$

i.e., the Euler scheme has strong convergence of order one for deterministic $\sigma(t)$.