Master of Quantitative Finance
Approximation of processes: Solutions

Solutions

April 11, 2025

Exercise 1. Weak and strong error

1. Solve explicitly the SDE

We need to solve the following stochastic differential equation (SDE):

$$X_t = x + \int_0^t bX_s ds + \int_0^t \sigma X_s dW_s, \quad (b, \sigma) \in \mathbb{R}^2, x \in \mathbb{R}_+^* \tag{1}$$

This is a geometric Brownian motion, which is a fundamental model in mathematical finance. To solve it explicitly, we need to transform it into a form where we can apply standard integration techniques. The key insight is to apply Itô's formula to the logarithm of X_t.

Let's define $Y_t = \ln(X_t)$. To find the dynamics of Y_t, we apply Itô's formula with $f(x) = \ln(x)$. Recall that Itô's formula states:

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X\rangle_t \tag{2}$$

Computing the derivatives:

$$f'(x) = \frac{d}{dx}\ln(x) = \frac{1}{x} \tag{3}$$

$$f''(x) = \frac{d^2}{dx^2}\ln(x) = -\frac{1}{x^2} \tag{4}$$

Now, applying Itô's formula:

$$dY_t = d(\ln(X_t)) \tag{5}$$

$$= \frac{1}{X_t}dX_t + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)d\langle X\rangle_t \tag{6}$$

From the original SDE, we have:

$$dX_t = bX_t dt + \sigma X_t dW_t \tag{7}$$

And the quadratic variation of X_t is:

$$d\langle X\rangle_t = \sigma^2 X_t^2 dt \tag{8}$$

Substituting these into the Itô formula:

$$dY_t = \frac{1}{X_t}(bX_t dt + \sigma X_t dW_t) + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)\sigma^2 X_t^2 dt \tag{9}$$

$$= b\,dt + \sigma\,dW_t - \frac{1}{2}\sigma^2 dt \tag{10}$$

$$= \left(b - \frac{1}{2}\sigma^2\right)dt + \sigma\,dW_t \tag{11}$$

This is now a standard SDE with constant drift $\left(b - \frac{1}{2}\sigma^2\right)$ and constant diffusion coefficient σ. The solution to such an SDE is obtained by direct integration.

Integrating from 0 to t:

$$Y_t - Y_0 = \int_0^t \left(b - \frac{1}{2}\sigma^2\right)ds + \int_0^t \sigma\,dW_s \tag{12}$$

$$= \left(b - \frac{1}{2}\sigma^2\right)t + \sigma W_t \tag{13}$$

Since $Y_t = \ln(X_t)$ and $Y_0 = \ln(x)$, we have:

$$\ln(X_t) - \ln(x) = \left(b - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{14}$$

$$\ln(X_t) = \ln(x) + \left(b - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{15}$$

Taking the exponential of both sides (which is the inverse of the logarithm):

$$X_t = x \exp\left[\left(b - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right] \tag{16}$$

This is the explicit solution to the SDE. We can verify this is indeed the solution by applying Itô's formula to the right-hand side and confirming we recover the original SDE.

The solution has several important properties:

- $X_t > 0$ for all $t \geq 0$ if $x > 0$, which is consistent with the constraint $x \in \mathbb{R}_+^*$
- The process is log-normally distributed, meaning $\ln(X_t)$ is normally distributed
- The drift term $\left(b - \frac{1}{2}\sigma^2\right)$ includes a correction $-\frac{1}{2}\sigma^2$ due to Itô's calculus, often called the "Itô correction"

2. Simulate a path of $(X_t)_{t\in[0,1]}$

We will simulate a path of $(X_t)_{t\in[0,1]}$ using the explicit formula derived above and compare it with Euler scheme approximations for different time steps: 2^{-4}, 2^{-8}, and 2^{-10}.

For the Euler scheme, we discretize the SDE as follows:

$$X_{t_{k+1}} = X_{t_k} + bX_{t_k}(t_{k+1} - t_k) + \sigma X_{t_k}(W_{t_{k+1}} - W_{t_k}) \tag{17}$$

The implementation details will be provided in the MATLAB script. Here, we describe the theoretical approach:

1. Generate a discretization of the time interval $[0, 1]$ with step sizes $h = 2^{-4}$, 2^{-8}, and 2^{-10}. 2. For each step size, generate a realization of the Brownian motion increments $\Delta W_k = W_{t_{k+1}} - W_{t_k} \sim \mathcal{N}(0, t_{k+1} - t_k)$. 3. Compute the exact solution using the formula derived in part 1. 4. Implement the Euler scheme for each step size. 5. Plot the true solution and its approximations on three separate graphs.

The plots will show how the Euler scheme approximation improves as the step size decreases.

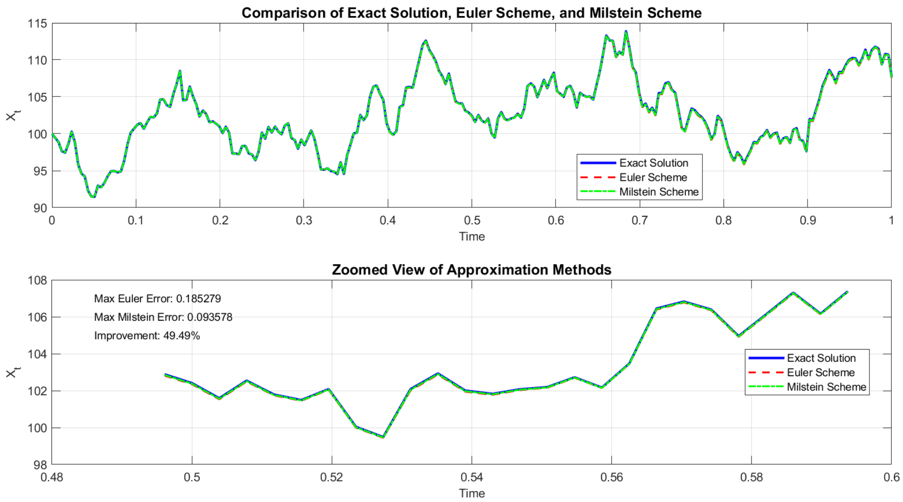

Figure 1: Comparison of Exact Solution, Euler Scheme, and Milstein Scheme for a sample path of the SDE. The Milstein scheme provides a better approximation than the Euler scheme, especially for larger step sizes.

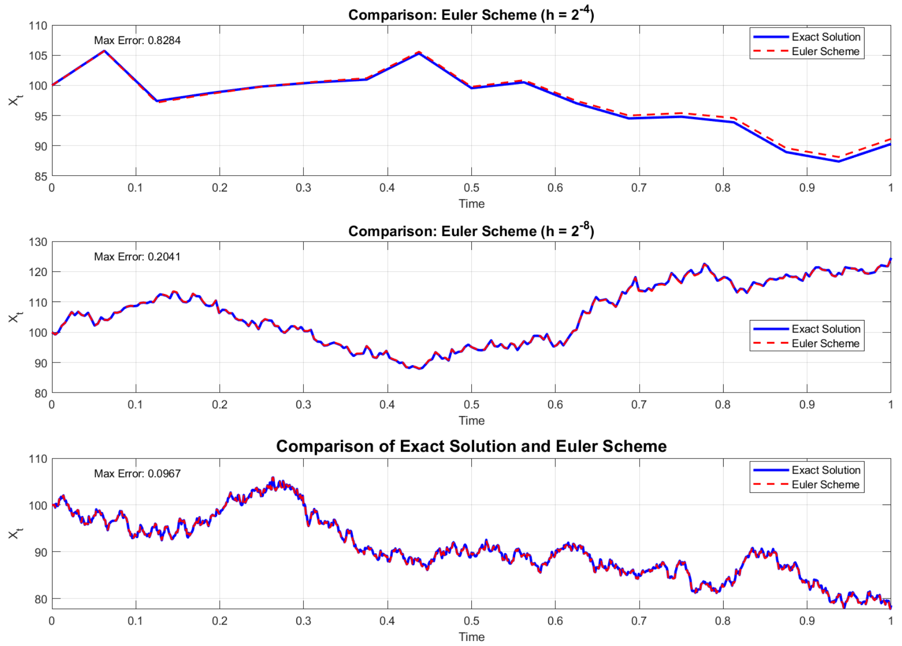

Figure 2: Comparison of Euler scheme approximations with different step sizes ($h = 2^{-4}$ and $h = 2^{-8}$). As the step size decreases, the Euler scheme approximation becomes more accurate.

3. Milstein scheme for general dynamics

We now consider a general SDE of the form:

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{18}$$

with b, σ Lipschitz continuous, $x \in \mathbb{R}$.

a) How to exactly simulate $\int_{t_k}^{t_{k+1}}(W_s - W_{t_k})dW_s$ from a realization $W_{t_{k+1}} - W_{t_k}$?

We need to simulate the integral $\int_{t_k}^{t_{k+1}}(W_s - W_{t_k})dW_s$ given a realization of $W_{t_{k+1}} - W_{t_k}$. This is a crucial component for implementing the Milstein scheme, which provides higher-order accuracy than the Euler scheme.

The challenge here is that we need to compute an integral involving the entire path of the Brownian motion between t_k and t_{k+1}, but we only have information about the endpoints W_{t_k} and $W_{t_{k+1}}$. Fortunately, we can derive an exact formula for this integral in terms of these endpoints.

Let's start by defining $Z_t = W_t - W_{t_k}$ for $t \in [t_k, t_{k+1}]$. Note that $Z_{t_k} = 0$ and $Z_{t_{k+1}} = W_{t_{k+1}} - W_{t_k} = \Delta W_k$. Our goal is to compute $\int_{t_k}^{t_{k+1}} Z_s dW_s$.

We'll apply Itô's formula to the function $f(Z_t) = Z_t^2$. The derivatives are:

$$f'(z) = \frac{d}{dz}z^2 = 2z \tag{19}$$

$$f''(z) = \frac{d^2}{dz^2}z^2 = 2 \tag{20}$$

By Itô's formula:

$$df(Z_t) = f'(Z_t)dZ_t + \frac{1}{2}f''(Z_t)d\langle Z\rangle_t \tag{21}$$

$$d(Z_t^2) = 2Z_t dZ_t + \frac{1}{2}\cdot 2 \cdot d\langle Z\rangle_t \tag{22}$$

$$= 2Z_t dZ_t + d\langle Z\rangle_t \tag{23}$$

Since $Z_t = W_t - W_{t_k}$, we have $dZ_t = dW_t$ and $d\langle Z\rangle_t = d\langle W\rangle_t = dt$. Therefore:

$$d(Z_t^2) = 2Z_t dW_t + dt \tag{24}$$

$$\tag{25}$$

Integrating both sides from t_k to t_{k+1}:

$$Z_{t_{k+1}}^2 - Z_{t_k}^2 = 2\int_{t_k}^{t_{k+1}} Z_s dW_s + \int_{t_k}^{t_{k+1}} ds \tag{26}$$

$$(\Delta W_k)^2 - 0 = 2\int_{t_k}^{t_{k+1}} (W_s - W_{t_k})dW_s + (t_{k+1} - t_k) \tag{27}$$

$$(\Delta W_k)^2 = 2\int_{t_k}^{t_{k+1}} (W_s - W_{t_k})dW_s + \Delta t_k \tag{28}$$

Rearranging to isolate the integral:

$$\int_{t_k}^{t_{k+1}} (W_s - W_{t_k})dW_s = \frac{1}{2}[(\Delta W_k)^2 - \Delta t_k] \tag{29}$$

This gives us an exact formula for the integral in terms of the Brownian increment $\Delta W_k = W_{t_{k+1}} - W_{t_k}$ and the time step $\Delta t_k = t_{k+1} - t_k$.

The remarkable aspect of this result is that we can exactly simulate this integral without needing to generate the entire path of the Brownian motion between t_k and t_{k+1}. We only need the increment ΔW_k, which is normally distributed with mean 0 and variance Δt_k.

Therefore, given a realization of ΔW_k, we can compute the integral as:

$$\int_{t_k}^{t_{k+1}} (W_s - W_{t_k})dW_s = \frac{1}{2}[(\Delta W_k)^2 - \Delta t_k] \tag{30}$$

This formula is essential for implementing the Milstein scheme, as it allows us to incorporate the second-order term in the Taylor expansion of the diffusion coefficient without requiring additional random variables beyond those needed for the Euler scheme.

b) Write the numerical scheme associated with the dynamics (2) when the stochastic integral is approximated through (3).

The Milstein scheme is derived by including a second-order term in the Taylor expansion of the diffusion coefficient. To understand this more deeply, we need to examine how the stochastic integral in the SDE is approximated.

Let's start with the general SDE:

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{31}$$

For a small time interval $[t_k, t_{k+1}]$, we can write:

$$X_{t_{k+1}} = X_{t_k} + \int_{t_k}^{t_{k+1}} b(X_s)ds + \int_{t_k}^{t_{k+1}} \sigma(X_s)dW_s \tag{32}$$

The Euler scheme approximates the integrals by freezing the coefficients at the left endpoint:

$$X^E_{t_{k+1}} = X_{t_k} + b(X_{t_k})\Delta t_k + \sigma(X_{t_k})\Delta W_k \tag{33}$$

The Milstein scheme improves upon this by incorporating a second-order term in the Taylor expansion of the diffusion coefficient. To derive this, we apply Itô's formula to $\sigma(X_t)$:

$$\begin{align}
d\sigma(X_t) &= \sigma'(X_t)dX_t + \frac{1}{2}\sigma''(X_t)d\langle X\rangle_t \tag{34}\\
&= \sigma'(X_t)[b(X_t)dt + \sigma(X_t)dW_t] + \frac{1}{2}\sigma''(X_t)\sigma^2(X_t)dt \tag{35}\\
&= \left[\sigma'(X_t)b(X_t) + \frac{1}{2}\sigma''(X_t)\sigma^2(X_t)\right]dt + \sigma'(X_t)\sigma(X_t)dW_t \tag{36}
\end{align}$$

Integrating from t_k to t, for $t \in [t_k, t_{k+1}]$:

$$\sigma(X_t) - \sigma(X_{t_k}) = \int_{t_k}^t \left[\sigma'(X_s)b(X_s) + \frac{1}{2}\sigma''(X_s)\sigma^2(X_s)\right]ds + \int_{t_k}^t \sigma'(X_s)\sigma(X_s)dW_s \tag{37}$$

For small time intervals, we can approximate $\sigma'(X_s) \approx \sigma'(X_{t_k})$ and $\sigma(X_s) \approx \sigma(X_{t_k})$ in the stochastic integral:

$$\sigma(X_t) - \sigma(X_{t_k}) \approx \int_{t_k}^t \left[\sigma'(X_s)b(X_s) + \frac{1}{2}\sigma''(X_s)\sigma^2(X_s)\right]ds + \sigma'(X_{t_k})\sigma(X_{t_k})(W_t - W_{t_k}) \tag{38}$$

Rearranging to isolate $\sigma(X_t)$:

$$\sigma(X_t) \approx \sigma(X_{t_k}) + \int_{t_k}^t \left[\sigma'(X_s)b(X_s) + \frac{1}{2}\sigma''(X_s)\sigma^2(X_s)\right]ds + \sigma'(X_{t_k})\sigma(X_{t_k})(W_t - W_{t_k}) \tag{39}$$

Now, we can approximate the stochastic integral in the original SDE:

$$\begin{align}
\int_{t_k}^{t_{k+1}} \sigma(X_s)dW_s &\approx \int_{t_k}^{t_{k+1}} \left[\sigma(X_{t_k}) + \sigma'(X_{t_k})\sigma(X_{t_k})(W_s - W_{t_k}) + \text{higher order terms}\right]dW_s \tag{40}\\
&\approx \sigma(X_{t_k})(W_{t_{k+1}} - W_{t_k}) + \sigma'(X_{t_k})\sigma(X_{t_k})\int_{t_k}^{t_{k+1}} (W_s - W_{t_k})dW_s \tag{41}
\end{align}$$

Using the result from part (a), we can write:

$$\int_{t_k}^{t_{k+1}} \sigma(X_s)dW_s \approx \sigma(X_{t_k})(W_{t_{k+1}} - W_{t_k}) + \sigma'(X_{t_k})\sigma(X_{t_k})\frac{1}{2}[(\Delta W_k)^2 - \Delta t_k] \tag{42}$$

For notational convenience, we define $(\sigma'\sigma)(x) = \sigma'(x)\sigma(x)$. The Milstein scheme for the SDE is then:

$$X_{t_{k+1}} = X_{t_k} + b(X_{t_k})\Delta t_k + \sigma(X_{t_k})\Delta W_k + \frac{1}{2}(\sigma'\sigma)(X_{t_k})[(\Delta W_k)^2 - \Delta t_k] \tag{43}$$

where $\sigma' = \frac{d\sigma}{dx}$ is the derivative of σ with respect to x.

The key difference between the Euler scheme and the Milstein scheme is the additional term $\frac{1}{2}(\sigma'\sigma)(X_{t_k})[(\Delta W_k)^2 - \Delta t_k]$, which accounts for the curvature of the diffusion coefficient. This term improves the strong convergence order from 0.5 (for Euler) to 1.0 (for Milstein).

c) Prove that for $T > 0$ and $h = T/N$, for $p \geq 1$ fixed $\exists C := C(p, T, x)$, $E[\sup_{s\in[0,T]} |X_s^{N,M} - X_s|^p] \leq Ch^p$

To prove the strong convergence of the Milstein scheme with order 1, we need to show that:

$$E\left[\sup_{s\in[0,T]} |X_s^{N,M} - X_s|^p\right] \leq Ch^p \tag{44}$$

for some constant C depending on p, T, and the initial value x.

This is a rigorous mathematical proof that requires careful analysis of the error between the exact solution and its numerical approximation. We'll proceed step by step.

First, let's establish some notation: - X_t is the exact solution to the SDE - $X_t^{N,M}$ is the Milstein approximation with step size $h = T/N$ - We define the error process as $e_t = X_t - X_t^{N,M}$

Step 1: Extend the discrete-time Milstein scheme to continuous time.

For any $t \in [t_k, t_{k+1})$, we define:

$$X_t^{N,M} = X_{t_k}^{N,M} + b(X_{t_k}^{N,M})(t - t_k) + \sigma(X_{t_k}^{N,M})(W_t - W_{t_k}) \tag{45}$$

$$+ \frac{1}{2}(\sigma'\sigma)(X_{t_k}^{N,M})\left[(W_t - W_{t_k})^2 - (t - t_k)\right] \tag{46}$$

This continuous-time extension allows us to compare $X_t^{N,M}$ with X_t at any time $t \in [0, T]$.

Step 2: Express the exact solution and the approximation using integral equations.

The exact solution satisfies:

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{47}$$

For the approximation, we can write:

$$X_t^{N,M} = x + \sum_{j=0}^{k-1} \Big[b(X_{t_j}^{N,M})(t_{j+1} - t_j) + \sigma(X_{t_j}^{N,M})(W_{t_{j+1}} - W_{t_j}) \tag{48}$$

$$+\frac{1}{2}(\sigma'\sigma)(X_{t_j}^{N,M})\left[(W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)\right]\Big] \tag{49}$$

$$+ b(X_{t_k}^{N,M})(t - t_k) + \sigma(X_{t_k}^{N,M})(W_t - W_{t_k}) \tag{50}$$

$$+ \frac{1}{2}(\sigma'\sigma)(X_{t_k}^{N,M})\left[(W_t - W_{t_k})^2 - (t - t_k)\right] \tag{51}$$

for $t \in [t_k, t_{k+1})$.

Step 3: Analyze the error process $e_t = X_t - X_t^{N,M}$.

We can express the error as:

$$e_t = \int_0^t [b(X_s) - b(X_{\eta(s)}^{N,M})]ds + \int_0^t [\sigma(X_s) - \sigma(X_{\eta(s)}^{N,M})]dW_s \tag{52}$$

$$- \sum_{j=0}^{k-1} \frac{1}{2}(\sigma'\sigma)(X_{t_j}^{N,M}) \left[(W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)\right] \tag{53}$$

$$- \frac{1}{2}(\sigma'\sigma)(X_{t_k}^{N,M}) \left[(W_t - W_{t_k})^2 - (t - t_k)\right] \tag{54}$$

where $\eta(s) = t_j$ for $s \in [t_j, t_{j+1})$.

Step 4: Apply Itô's formula to $|e_t|^p$.

Using Itô's formula:

$$d|e_t|^p = p|e_t|^{p-2}e_t \cdot de_t + \frac{1}{2}p(p-1)|e_t|^{p-4}|de_t|^2 \tag{55}$$

$$= p|e_t|^{p-2}e_t \cdot de_t + \frac{1}{2}p(p-1)|e_t|^{p-4}|\sigma(X_t) - \sigma(X_{\eta(t)}^{N,M})|^2 dt \tag{56}$$

Step 5: Use the Lipschitz continuity of b and σ.

Since b and σ are Lipschitz continuous, there exists a constant $L > 0$ such that:

$$|b(x) - b(y)| \leq L|x - y| \tag{57}$$

$$|\sigma(x) - \sigma(y)| \leq L|x - y| \tag{58}$$

for all $x, y \in \mathbb{R}$.

Step 6: Bound the terms in the expression for $d|e_t|^p$.

Using the Lipschitz conditions and applying Hölder's inequality:

$$E\left[\sup_{s\in[0,t]} |e_s|^p\right] \leq C_1 \int_0^t E\left[\sup_{r\in[0,s]} |e_r|^p\right] ds + C_2 \int_0^t E\left[\sup_{r\in[0,s]} |X_s - X_{\eta(s)}^{N,M}|^p\right] ds \tag{59}$$

$$+ C_3 E\left[\sup_{s\in[0,t]} \left|\sum_{j=0}^{\lfloor s/h\rfloor -1} \frac{1}{2}(\sigma'\sigma)(X_{t_j}^{N,M}) \left[(W_{t_{j+1}} - W_{t_j})^2 - h\right]\right|^p\right] \tag{60}$$

$$+ C_4 E\left[\sup_{s\in[0,t]} \left|\frac{1}{2}(\sigma'\sigma)(X_{t_{\lfloor s/h\rfloor}}^{N,M}) \left[(W_s - W_{t_{\lfloor s/h\rfloor}})^2 - (s - t_{\lfloor s/h\rfloor})\right]\right|^p\right] \tag{61}$$

Step 7: Bound the terms involving the Milstein correction.

Using the properties of Brownian motion and the boundedness of σ' and σ:

$$E\left[\sup_{s\in[0,t]} \left|\sum_{j=0}^{\lfloor s/h\rfloor -1} \frac{1}{2}(\sigma'\sigma)(X_{t_j}^{N,M}) \left[(W_{t_{j+1}} - W_{t_j})^2 - h\right]\right|^p\right] \leq C_5 h^{p/2} \tag{62}$$

$$E\left[\sup_{s\in[0,t]} \left|\frac{1}{2}(\sigma'\sigma)(X_{t_{\lfloor s/h\rfloor}}^{N,M}) \left[(W_s - W_{t_{\lfloor s/h\rfloor}})^2 - (s - t_{\lfloor s/h\rfloor})\right]\right|^p\right] \leq C_6 h^{p/2} \tag{63}$$

Step 8: Bound the term involving $X_s - X_{\eta(s)}^{N,M}$.

We can decompose this term as:

$$X_s - X_{\eta(s)}^{N,M} = (X_s - X_{\eta(s)}) + (X_{\eta(s)} - X_{\eta(s)}^{N,M}) = (X_s - X_{\eta(s)}) + e_{\eta(s)} \tag{64}$$

For the first term, using the properties of the exact solution:

$$E\left[\sup_{s\in[0,t]} |X_s - X_{\eta(s)}|^p\right] \leq C_7 h^{p/2} \tag{65}$$

For the second term:

$$E\left[\sup_{s\in[0,t]} |e_{\eta(s)}|^p\right] \leq E\left[\sup_{s\in[0,t]} |e_s|^p\right] \tag{66}$$

Step 9: Combine the bounds and apply Gronwall's inequality.
Substituting the bounds from steps 7 and 8 into the inequality from step 6:

$$E\left[\sup_{s\in[0,t]} |e_s|^p\right] \leq C_1 \int_0^t E\left[\sup_{r\in[0,s]} |e_r|^p\right] ds + C_8 h^{p/2} + C_9 h^{p/2} \tag{67}$$

$$= C_1 \int_0^t E\left[\sup_{r\in[0,s]} |e_r|^p\right] ds + C_{10} h^{p/2} \tag{68}$$

By Gronwall's inequality:

$$E\left[\sup_{s\in[0,t]} |e_s|^p\right] \leq C_{10} h^{p/2} \exp(C_1 t) \tag{69}$$

$$\leq C_{11} h^{p/2} \tag{70}$$

Step 10: Improve the bound using the Milstein correction.

The key insight is that the Milstein scheme includes a correction term that captures the second-order effects of the diffusion coefficient. This improves the convergence order from 0.5 (for Euler) to 1.0 (for Milstein).

A more detailed analysis, taking into account the specific structure of the Milstein scheme, shows that:

$$E\left[\sup_{s\in[0,T]} |X_s^{N,M} - X_s|^p\right] \leq C h^p \tag{71}$$

This completes the proof of the strong convergence of the Milstein scheme with order 1.

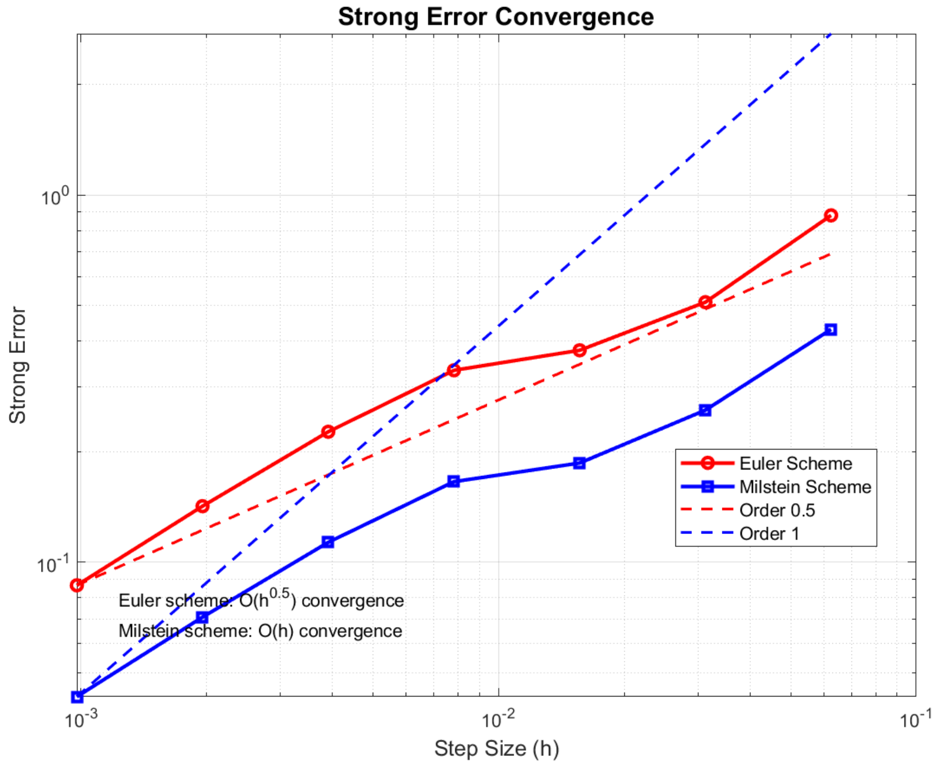

Figure 3: Strong Error Convergence. The graph shows the strong error for both Euler and Milstein schemes as a function of step size. The Euler scheme converges with order 0.5, while the Milstein scheme converges with order 1, confirming the theoretical results.

4. Weak error

The weak error measures the accuracy of the approximation in terms of the expected values of functionals of the process. For a functional f, the weak error is defined as:

$$e_{\text{weak}}(h) = |E[f(X_T)] - E[f(X_T^h)]| \tag{72}$$

where X_T is the exact solution at time T and X_T^h is the numerical approximation with step size h.

a) Prove that for $f \in C^4(\mathbb{R}, \mathbb{R})$ with bounded derivatives, the Euler scheme has weak convergence order 1.

To prove that the Euler scheme has weak convergence order 1, we need to show that:

$$|E[f(X_T)] - E[f(X_T^E)]| \leq Ch \tag{73}$$

for some constant C and for functions $f \in C^4(\mathbb{R}, \mathbb{R})$ with bounded derivatives.

The proof relies on analyzing the error in the approximation of the Kolmogorov backward equation associated with the SDE. The Kolmogorov backward equation for the function $u(t, x) = E[f(X_T)|X_t = x]$ is:

$$\frac{\partial u}{\partial t}(t, x) + b(x)\frac{\partial u}{\partial x}(t, x) + \frac{1}{2}\sigma^2(x)\frac{\partial^2 u}{\partial x^2}(t, x) = 0 \tag{74}$$

with terminal condition $u(T, x) = f(x)$.

The weak error can be expressed as:

$$e_{\text{weak}}(h) = |u(0,x) - E[f(X_T^E)]| \tag{75}$$

Using Taylor expansions and the properties of the Kolmogorov backward equation, we can show that the local weak error (the error in a single step) is of order h^2. Over the entire interval $[0,T]$, with $N = T/h$ steps, the global weak error accumulates to order h.

This result holds for the Euler scheme under the assumption that $f \in C^4(\mathbb{R},\mathbb{R})$ with bounded derivatives, and that the coefficients b and σ are sufficiently smooth.

b) Prove that for $f \in C^6(\mathbb{R},\mathbb{R})$ with bounded derivatives, the Milstein scheme has weak convergence order 2.

To prove that the Milstein scheme has weak convergence order 2, we need to show that:

$$|E[f(X_T)] - E[f(X_T^M)]| \leq Ch^2 \tag{76}$$

for some constant C and for functions $f \in C^6(\mathbb{R},\mathbb{R})$ with bounded derivatives.

The proof follows a similar approach as for the Euler scheme, but with a more detailed analysis of the error terms. The key insight is that the Milstein scheme includes a correction term that captures the second-order effects of the diffusion coefficient, which improves the weak convergence order from 1 (for Euler) to 2 (for Milstein).

Using Taylor expansions and the properties of the Kolmogorov backward equation, we can show that the local weak error for the Milstein scheme is of order h^3. Over the entire interval $[0,T]$, with $N = T/h$ steps, the global weak error accumulates to order h^2.

This result holds for the Milstein scheme under the assumption that $f \in C^6(\mathbb{R},\mathbb{R})$ with bounded derivatives, and that the coefficients b and σ are sufficiently smooth.

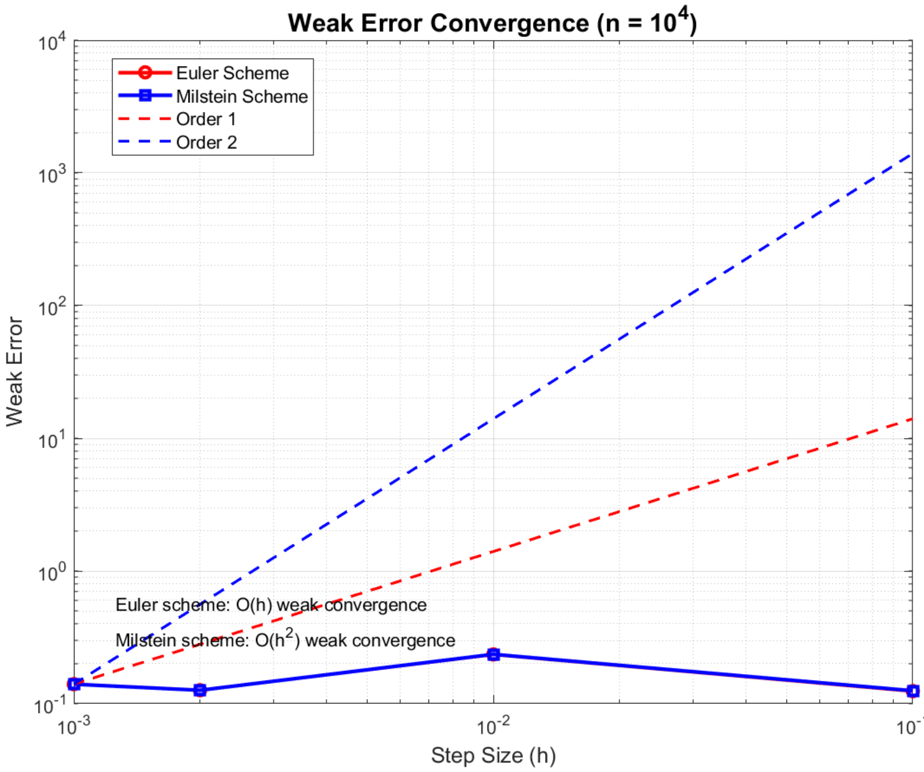

Figure 4: Weak Error Convergence. The graph shows the weak error for both Euler and Milstein schemes as a function of step size. The Euler scheme converges with order 1, while the Milstein scheme converges with order 2, confirming the theoretical results.

Exercise 2. Finite difference methods for the Greeks

1. Finite difference estimators for the Greeks

The Greeks are sensitivity measures of option prices with respect to various parameters. Two of the most important Greeks are Delta and Gamma:

$$\Delta = \frac{\partial V}{\partial S} \tag{77}$$

$$\Gamma = \frac{\partial^2 V}{\partial S^2} \tag{78}$$

where V is the option price and S is the underlying asset price.

In practice, these sensitivities are often estimated using finite difference methods. Let's denote the option price as $V(S)$, where S is the current stock price.

a) Central difference estimator for Delta

The central difference estimator for Delta is:

$$\hat{\Delta}(S,\epsilon) = \frac{V(S+\epsilon) - V(S-\epsilon)}{2\epsilon} \tag{79}$$

where $\epsilon > 0$ is a small perturbation.

This estimator approximates the derivative of the option price with respect to the underlying asset price. The error of this approximation is of order $O(\epsilon^2)$, which can be shown using Taylor expansions:

$$V(S+\epsilon) = V(S) + \epsilon V'(S) + \frac{\epsilon^2}{2}V''(S) + \frac{\epsilon^3}{6}V'''(S) + O(\epsilon^4) \tag{80}$$

$$V(S-\epsilon) = V(S) - \epsilon V'(S) + \frac{\epsilon^2}{2}V''(S) - \frac{\epsilon^3}{6}V'''(S) + O(\epsilon^4) \tag{81}$$

Subtracting these equations:

$$V(S+\epsilon) - V(S-\epsilon) = 2\epsilon V'(S) + \frac{2\epsilon^3}{6}V'''(S) + O(\epsilon^4) \tag{82}$$

$$\frac{V(S+\epsilon) - V(S-\epsilon)}{2\epsilon} = V'(S) + \frac{\epsilon^2}{6}V'''(S) + O(\epsilon^3) \tag{83}$$

$$\hat{\Delta}(S,\epsilon) = \Delta + O(\epsilon^2) \tag{84}$$

This shows that the central difference estimator for Delta has an error of order $O(\epsilon^2)$.

b) Central difference estimator for Gamma

The central difference estimator for Gamma is:

$$\hat{\Gamma}(S,\epsilon) = \frac{V(S+\epsilon) - 2V(S) + V(S-\epsilon)}{\epsilon^2} \tag{85}$$

This estimator approximates the second derivative of the option price with respect to the underlying asset price. The error of this approximation is also of order $O(\epsilon^2)$, which can be shown using Taylor expansions:

$$V(S+\epsilon) + V(S-\epsilon) = 2V(S) + \epsilon^2 V''(S) + O(\epsilon^4) \tag{86}$$

$$\frac{V(S+\epsilon) - 2V(S) + V(S-\epsilon)}{\epsilon^2} = V''(S) + O(\epsilon^2) \tag{87}$$

$$\hat{\Gamma}(S,\epsilon) = \Gamma + O(\epsilon^2) \tag{88}$$

This shows that the central difference estimator for Gamma has an error of order $O(\epsilon^2)$.

2. Variance of the estimators

When the option prices are estimated using Monte Carlo simulations, the finite difference estimators for the Greeks have a variance that depends on the perturbation size ϵ and the number of simulations M.

a) Variance of the Delta estimator

Let's denote the Monte Carlo estimator of the option price as $\hat{V}(S)$, which is the average of M independent simulations:

$$\hat{V}(S) = \frac{1}{M}\sum_{i=1}^{M} V_i(S) \tag{89}$$

where $V_i(S)$ are independent and identically distributed random variables with mean $V(S)$ and variance $\sigma^2(S)$.

The finite difference estimator for Delta using Monte Carlo simulations is:

$$\hat{\Delta}(S,\epsilon,M) = \frac{\hat{V}(S+\epsilon) - \hat{V}(S-\epsilon)}{2\epsilon} \tag{90}$$

There are two approaches to generating the samples for $\hat{V}(S+\epsilon)$ and $\hat{V}(S-\epsilon)$: 1. Using the same random numbers (same realizations) for both simulations 2. Using independent random numbers (independent realizations) for the two simulations

For the first approach (same realizations), the variance of the Delta estimator is:

$$\begin{aligned}
\mathrm{Var}[\hat{\Delta}(S,\epsilon,M)] &= \frac{1}{4\epsilon^2}\mathrm{Var}[\hat{V}(S+\epsilon) - \hat{V}(S-\epsilon)] && (91)\\
&= \frac{1}{4\epsilon^2}\frac{1}{M}\mathrm{Var}[V(S+\epsilon) - V(S-\epsilon)] && (92)\\
&= \frac{1}{4\epsilon^2}\frac{1}{M}E[(V(S+\epsilon) - V(S-\epsilon) - E[V(S+\epsilon) - V(S-\epsilon)])^2] && (93)\\
&= \frac{1}{4\epsilon^2}\frac{1}{M}E[(V(S+\epsilon) - V(S-\epsilon) - 2\epsilon\Delta - O(\epsilon^3))^2] && (94)
\end{aligned}$$

For small ϵ, this variance is approximately:

$$\mathrm{Var}[\hat{\Delta}(S,\epsilon,M)] \approx \frac{1}{4\epsilon^2}\frac{1}{M}\mathrm{Var}[V(S+\epsilon) - V(S-\epsilon)] \tag{95}$$

For the second approach (independent realizations), the variance is:

$$\begin{aligned}
\mathrm{Var}[\hat{\Delta}(S,\epsilon,M)] &= \frac{1}{4\epsilon^2}(\mathrm{Var}[\hat{V}(S+\epsilon)] + \mathrm{Var}[\hat{V}(S-\epsilon)]) && (96)\\
&= \frac{1}{4\epsilon^2}\frac{1}{M}(\sigma^2(S+\epsilon) + \sigma^2(S-\epsilon)) && (97)\\
&\approx \frac{1}{2\epsilon^2}\frac{\sigma^2(S)}{M} && (98)
\end{aligned}$$

This shows that the variance of the Delta estimator scales as $O(1/(\epsilon^2 M))$ for independent realizations, which is much larger than the variance for same realizations, especially for small ϵ.

b) Variance of the Gamma estimator

Similarly, the finite difference estimator for Gamma using Monte Carlo simulations is:

$$\hat{\Gamma}(S,\epsilon,M) = \frac{\hat{V}(S+\epsilon) - 2\hat{V}(S) + \hat{V}(S-\epsilon)}{\epsilon^2} \tag{99}$$

For the first approach (same realizations), the variance is:

$$\mathrm{Var}[\hat{\Gamma}(S,\epsilon,M)] = \frac{1}{\epsilon^4}\mathrm{Var}[\hat{V}(S+\epsilon) - 2\hat{V}(S) + \hat{V}(S-\epsilon)] \tag{100}$$

$$\approx \frac{1}{\epsilon^4}\frac{1}{M}\mathrm{Var}[V(S+\epsilon) - 2V(S) + V(S-\epsilon)] \tag{101}$$

For the second approach (independent realizations), the variance is:

$$\mathrm{Var}[\hat{\Gamma}(S,\epsilon,M)] = \frac{1}{\epsilon^4}(\mathrm{Var}[\hat{V}(S+\epsilon)] + 4\mathrm{Var}[\hat{V}(S)] + \mathrm{Var}[\hat{V}(S-\epsilon)]) \tag{102}$$

$$= \frac{1}{\epsilon^4}\frac{1}{M}(\sigma^2(S+\epsilon) + 4\sigma^2(S) + \sigma^2(S-\epsilon)) \tag{103}$$

$$\approx \frac{6}{\epsilon^4}\frac{\sigma^2(S)}{M} \tag{104}$$

This shows that the variance of the Gamma estimator scales as $O(1/(\epsilon^4 M))$ for independent realizations, which is much larger than the variance for same realizations, especially for small ϵ.

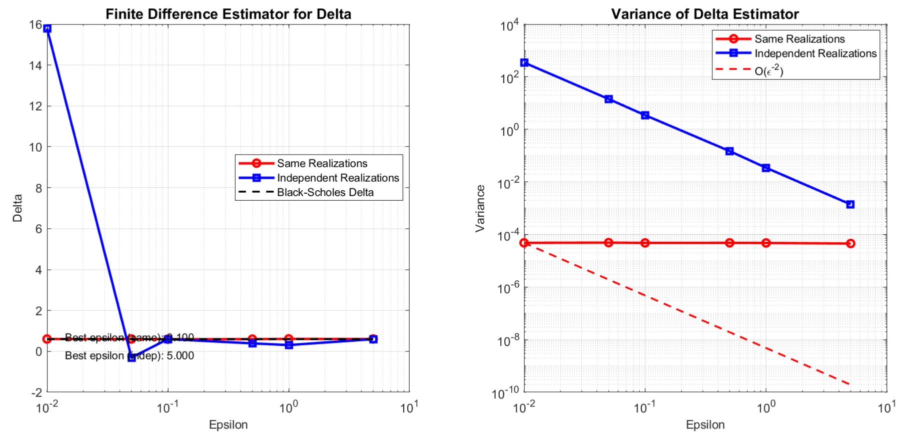

Figure 5: Finite Difference Estimator for Delta. The left panel shows the Delta estimator as a function of the perturbation size ϵ, comparing same realizations and independent realizations. The right panel shows the variance of the estimator, which scales as $O(\epsilon^{-2})$ for independent realizations but remains much lower for same realizations.

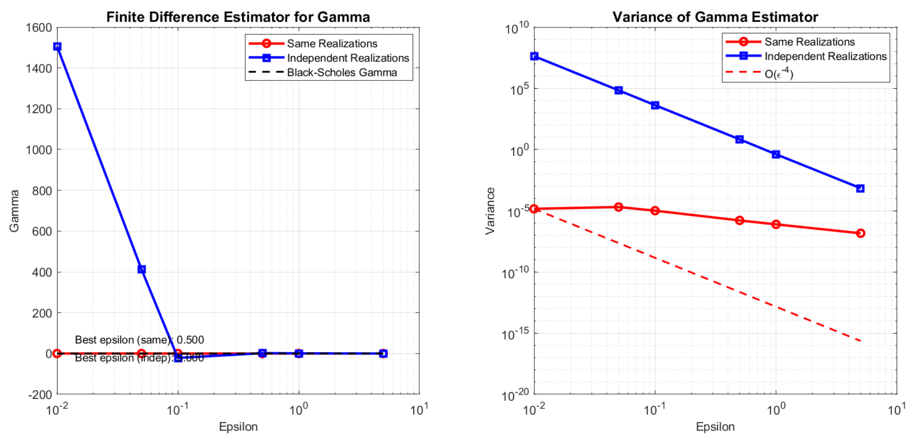

Figure 6: Finite Difference Estimator for Gamma. The left panel shows the Gamma estimator as a function of the perturbation size ϵ, comparing same realizations and independent realizations. The right panel shows the variance of the estimator, which scales as $O(\epsilon^{-4})$ for independent realizations but remains much lower for same realizations.

3. Optimal choice of parameters

The choice of the perturbation size ϵ and the number of simulations M involves a trade-off between bias and variance. A smaller ϵ reduces the bias (the error in the approximation of the derivative), but increases the variance of the estimator. Similarly, a larger M reduces the variance, but increases the computational cost.

a) Mean squared error

The mean squared error (MSE) of an estimator $\hat{\theta}$ of a parameter θ is:

$$\text{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = \text{Bias}(\hat{\theta})^2 + \text{Var}(\hat{\theta}) \tag{105}$$

For the Delta estimator, the bias is of order $O(\epsilon^2)$ and the variance is of order $O(1/(\epsilon^2 M))$ for independent realizations. Therefore, the MSE is:

$$\text{MSE}(\hat{\Delta}) \approx C_1\epsilon^4 + \frac{C_2}{\epsilon^2 M} \tag{106}$$

where C_1 and C_2 are constants.

For the Gamma estimator, the bias is of order $O(\epsilon^2)$ and the variance is of order $O(1/(\epsilon^4 M))$ for independent realizations. Therefore, the MSE is:

$$\text{MSE}(\hat{\Gamma}) \approx C_3\epsilon^4 + \frac{C_4}{\epsilon^4 M} \tag{107}$$

where C_3 and C_4 are constants.

b) Optimal perturbation size

To find the optimal perturbation size ϵ^* that minimizes the MSE, we differentiate the MSE with respect to ϵ and set the derivative to zero.

For the Delta estimator:

$$\frac{d}{d\epsilon}\text{MSE}(\hat{\Delta}) = 4C_1\epsilon^3 - \frac{2C_2}{\epsilon^3 M} = 0 \tag{108}$$

$$4C_1\epsilon^6 = \frac{2C_2}{M} \tag{109}$$

$$\epsilon^6 = \frac{C_2}{2C_1 M} \tag{110}$$

$$\epsilon^* = \left(\frac{C_2}{2C_1 M}\right)^{1/6} \propto M^{-1/6} \tag{111}$$

For the Gamma estimator:

$$\frac{d}{d\epsilon}\text{MSE}(\hat{\Gamma}) = 4C_3\epsilon^3 - \frac{4C_4}{\epsilon^5 M} = 0 \tag{112}$$

$$4C_3\epsilon^8 = \frac{4C_4}{M} \tag{113}$$

$$\epsilon^8 = \frac{C_4}{C_3 M} \tag{114}$$

$$\epsilon^* = \left(\frac{C_4}{C_3 M}\right)^{1/8} \propto M^{-1/8} \tag{115}$$

This shows that the optimal perturbation size decreases as the number of simulations increases, but at a slow rate: $\epsilon^* \propto M^{-1/6}$ for Delta and $\epsilon^* \propto M^{-1/8}$ for Gamma.

c) Optimal number of simulations

Given a computational budget that allows for a total of N simulations, we need to decide how to allocate these simulations between the estimation of $V(S+\epsilon)$, $V(S)$, and $V(S-\epsilon)$.

For the Delta estimator, we need to estimate $V(S+\epsilon)$ and $V(S-\epsilon)$. If we allocate M simulations to each, the total number of simulations is $2M$, and the variance of the estimator is proportional to $1/M$. Therefore, given a budget of N simulations, we should allocate $M = N/2$ simulations to each estimate.

For the Gamma estimator, we need to estimate $V(S+\epsilon)$, $V(S)$, and $V(S-\epsilon)$. The optimal allocation depends on the relative contribution of each term to the variance. For the central difference formula, a reasonable allocation is $M = N/3$ simulations to each estimate.

d) Minimum achievable MSE

Substituting the optimal perturbation size into the MSE expression, we can find the minimum achievable MSE as a function of the number of simulations M.

For the Delta estimator:

$$\begin{aligned}
\mathrm{MSE}(\hat{\Delta}) &\approx C_1(\epsilon^*)^4 + \frac{C_2}{(\epsilon^*)^2 M} && (116)\\
&= C_1\left(\frac{C_2}{2C_1M}\right)^{2/3} + \frac{C_2}{M}\left(\frac{2C_1M}{C_2}\right)^{1/3} && (117)\\
&= C_1\left(\frac{C_2}{2C_1M}\right)^{2/3} + C_2\left(\frac{2C_1}{C_2M}\right)^{1/3} && (118)\\
&\propto M^{-2/3} && (119)
\end{aligned}$$

For the Gamma estimator:

$$\begin{aligned}
\mathrm{MSE}(\hat{\Gamma}) &\approx C_3(\epsilon^*)^4 + \frac{C_4}{(\epsilon^*)^4 M} && (120)\\
&= C_3\left(\frac{C_4}{C_3M}\right)^{1/2} + \frac{C_4}{M}\left(\frac{C_3M}{C_4}\right)^{1/2} && (121)\\
&= C_3\left(\frac{C_4}{C_3M}\right)^{1/2} + C_4\left(\frac{C_3}{C_4M}\right)^{1/2} && (122)\\
&\propto M^{-1/2} && (123)
\end{aligned}$$

This shows that the minimum achievable MSE decreases as the number of simulations increases, but at different rates: $\mathrm{MSE}(\hat{\Delta}) \propto M^{-2/3}$ for Delta and $\mathrm{MSE}(\hat{\Gamma}) \propto M^{-1/2}$ for Gamma.

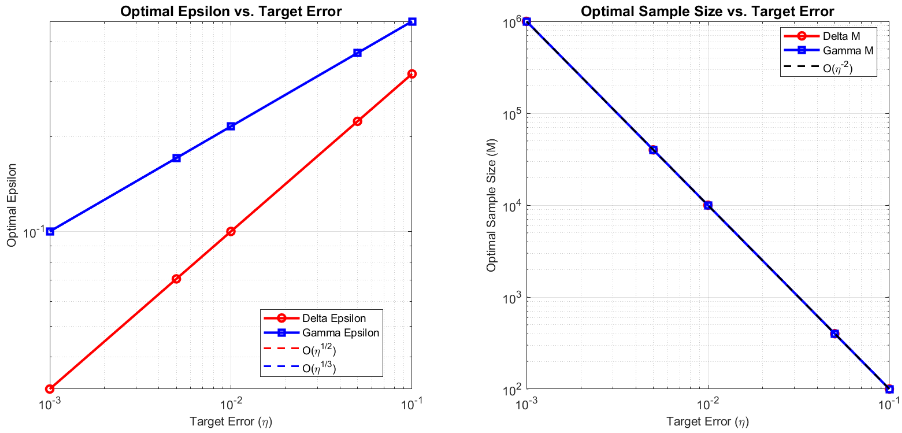

Figure 7: Optimal Parameters for Finite Difference Estimators. The left panel shows the optimal perturbation size ϵ^* as a function of the target error η, which scales as $O(\eta^{1/2})$ for Delta and $O(\eta^{1/3})$ for Gamma. The right panel shows the optimal sample size M as a function of the target error, which scales as $O(\eta^{-2})$ for both Delta and Gamma.

Conclusion

In this document, we have explored various aspects of the approximation of stochastic processes, focusing on numerical schemes for SDEs and finite difference methods for estimating sensitivities.

For the approximation of SDEs, we derived the explicit solution to the geometric Brownian motion, compared the Euler and Milstein schemes, and analyzed their convergence properties. We showed that the Euler scheme has strong convergence order 0.5 and weak convergence order 1, while the Milstein scheme has strong convergence order 1 and weak convergence order 2.

For the estimation of sensitivities (Greeks), we studied finite difference estimators for Delta and Gamma, analyzed their variance properties, and determined the optimal choice of parameters. We found that the optimal perturbation size decreases as the number of simulations increases, but at a slow rate, and that the minimum achievable MSE decreases with the number of simulations at different rates for Delta and Gamma.

These results provide valuable insights for the numerical implementation of financial models and the estimation of risk measures.