Lecture 9: Numerically solving SDEs

Readings

Recommended:

- Pavliotis [2014] Section 5.2
- Hingham [2001] – a short, simple introduction to numerically solving SDEs. It is recommended that you read this *after* first trying the homework – then you can read it for clues / more background.

Optional:

- Kloeden and Platen [1999] – lots. This is the canonical textbook on numerically solving SDEs. See e.g.:
  - Ch. 5 (Itô-Taylor expansion),
  - Ch 6.3 (a version of stochastic stability),
  - Ch. 8 (a discussion of numerical methods for deterministic equations),
  - Ch. 9 (numerical discretization of SDEs, and convergence notions),
  - Ch. 10+ (a detailed discussion of specific schemes.)

Unlike the SDEs we studied last class, most SDEs don't have explicit analytical solutions. One way to gain information about them is to simulate them numerically. This gives only an approximation to the "true" solutions, so the mathematical issue is to understand how close this approximation is, and to invent schemes that approximate it more closely, given finite computational resources.

This lecture will be an introduction to the major schemes and considerations in numerical SDEs. We'll see a lot in common with numerically solving deterministic ODEs. On top of an ODE, we have stochasticity, which one might expect would make the topic of numerical methods even richer. However, as we'll see, this stochasticity will mean that fancy high-order numerical methods are usually too hard to implement, so we'll end up studying only some fairly simple numerical methods that are widely used in practice.

The bible on the topic is Kloeden and Platen [1999]. A short, basic review article (Hingham [2001]) has also been posted to the website. It is highly recommended that you read this – though it may be pedagogically better to try to homework first, and read it if/when you get stuck.

9.1 Stochastic Itô -Taylor Expansion to derive basic schemes

We'll start by considering the one-dimensional SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \tag{1}$$

with initial condition X_0, over the interval $t \in [0,T]$. A simple scheme comes from approximating the Itô integral as a sum of step functions, as in the construction of the Itô integral (Lecture 7).

Definition (Euler-Maruyama (EM) Scheme)**.** Discretize time as $0 = t_0 < t_1 < \cdots < t_N = T$ with $t_{i+1} - t_i = \Delta t$. Let Y_n be a numerical approximation to X_{t_n}, calculated recursively as

$$Y_{n+1} = Y_n + b(Y_n)\Delta t + \sigma(Y_n)\delta W_n, \tag{2}$$

where $\delta W_n \sim N(0,\Delta t)$ are i.i.d. random variables. Set $Y_0 = X_0$ if X_0 is deterministic, otherwise choose the initial condition to be sufficiently close to the desired one, e.g. $\mathbb{E}(|X_0 - Y_0^{\Delta t}|^2)^{1/2} \le C(\Delta t)^{1/2}$ for some constant C.

Remark. The random variable δW_n can be replaced by a random variable with the same mean and variance, such as a uniform random variable, or $\xi = \pm\sqrt{\Delta t}$, which can be faster than generating Gaussian random variables. This causes the solution to lose certain strong, pathwise convergence properties as $\Delta t \to 0$, but its statistical properties will converge at the same order, which is sufficient for many applications.

The Euler-Maruyama scheme is the most widely-used scheme, but also the least accurate. To derive better methods we turn to the *stochastic Itô -Taylor expansion*. It is simplest to first introduce this method in the deterministic setting, for the solution of an ODE. Suppose X_t solves

$$\frac{dX_t}{dt} = a(X_t)\,, \tag{3}$$

where $a(x)$ is a function with plenty of bounded, continuous derivatives. Write (3) in integral form as

$$X_t = X_0 + \int_0^t a(X_s)ds. \tag{4}$$

We need to approximate $a(X_s)$ over a small interval $[0,t]$. We do this using the following observation. For any continuously differentiable function f, the chain rule and (3) imply $\frac{d}{dt}f(X_t) = a(X_t)\frac{\partial f(X_t)}{\partial x}$, so

$$f(X_t) = f(X_0) + \int_0^t \mathscr{L} f(X_s)ds\,, \quad \text{where} \quad \mathscr{L} = a(x)\frac{\partial}{\partial x}\,. \tag{5}$$

We use this identity to rewrite $a(X_s)$ in (4). Choosing $f = a$ in (5) gives

$$a(X_s) = a(X_0) + \int_0^s \mathscr{L} a(X_s)ds.$$

Substituting into (4) gives

$$X_t = X_0 + a(X_0)\int_0^t ds + \int_0^t\int_0^s \mathscr{L} a(X_u)duds. \tag{6}$$

and therefore

$$X_t = X_0 + a(X_0)t + R_2\,, \qquad R_2 = \int_0^t\int_0^s \mathscr{L} a(X_u)duds.$$

If we ignore R_2 (which we can do for example if $|\mathscr{L}a(x)| \le K$, so $|R_2| \le \frac{1}{2}Kt^2 \ll a(X_0)t$ for small t), we obtain a forward Euler approximation for X_t.

To get a better approximation, we continue the expansion, this time choosing $f = \mathscr{L}a$ in (5) to approximate $\mathscr{L}a(X_u)$. This gives

$$X_t = X_0 + a(X_0)\int_0^t ds + \mathscr{L}a(X_0)\int_0^t\int_0^s duds + R_3, \qquad R_3 = \int_0^t\int_0^s\int_0^u \mathscr{L}^2 a(X_v)dvduds.$$

If we ignore R_3 we obtain another approximation for X_t. In general, we can approximate X_t up to any order by writing

$$X_t = X_0 + \sum_{m=1}^{r}\frac{t^m}{m!}\mathscr{L}^m a(X_0) + R_{r+1}, \qquad R_{r+1} = \int_0^t\cdots\int_0^{s_r} \mathscr{L}^{r+1}a(X_{s_1})ds_1\cdots ds_{r+1}. \tag{7}$$

This is the Taylor formula in integral form. It expresses X_t as a function of only X_0, t, and a remainder term.

Now let's consider the *Stochastic Itô-Taylor expansion*. Consider the SDE

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s. \tag{8}$$

For any function $f \in C^2$, Itô's formula gives

$$df(X_t) = \left(b(X_t)\frac{\partial f}{\partial x}(X_t) + \frac{1}{2}\sigma^2(X_t)\frac{\partial^2 f}{\partial x^2}(X_t) \right) dt + \sigma(X_t)\frac{\partial f}{\partial x}(X_t)dW_t.$$

Integrating gives the identity

$$f(X_t) = f(X_0) + \int_0^t \mathscr{L}_0 f(X_s)ds + \int_0^t \mathscr{L}_1 f(X_s)dW_s, \tag{9}$$

where

$$\mathscr{L}_0 = b(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}, \qquad \mathscr{L}_1 = \sigma(x)\frac{\partial}{\partial x}. \tag{10}$$

Applying (9) to $f = b$ and $f = \sigma$ in (8) gives

$$X_t = X_0 + b(X_0)\int_0^t ds + \sigma(X_0)\int_0^t dW_s + R_1, \tag{11}$$

with

$$R_1 = \int_0^t \int_0^s \mathscr{L}_0 b(X_z)dzds + \int_0^t \int_0^s \mathscr{L}_1 b(X_z)dW_z ds + \int_0^t \int_0^s \mathscr{L}_0 \sigma(X_z)dzdW_s + \int_0^t \int_0^s \mathscr{L}_1 \sigma(X_z)dW_z dW_s.$$

If we ignore the remainder term R_1 in (11) and evaluate the integrals analytically, we get the Euler-Maruyama scheme, (2)

To derive more accurate schemes we continue the expansion. The terms in R_1 have different orders in t depending on the powers of dt, dW_t. The lowest-order term in R_1 is the last one, $\int_0^t \int_0^s \mathscr{L}_1\sigma(X_z)dW_z dW_s$, which we expect to be $O(t)$; all other terms should be $O(t^{3/2})$ or higher. Apply (9) to $f = \mathscr{L}_1\sigma$ in R_1 to get

$$X_t = X_0 + b(X_0)\int_0^t ds + \sigma(X_0)\int_0^t dW_s + \mathscr{L}_1\sigma(X_0)\int_0^t \int_0^s dW_z dW_s + R_2, \tag{12}$$

where R_2 is a remainder term (see Exercise 9.1). We can compute the double integral analytically:

$$\int_0^t \int_0^s dW_z dW_s = \int_0^t W_s dW_s = \frac{1}{2}(W_t^2 - t).$$

Substituting into (12) and ignoring R_2 gives the Milstein scheme.

Definition (Milstein scheme)**.** Discretize time as $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ with $t_{i+1} - t_i = \Delta t$. Let Y_n be a numerical approximation to X_{t_n}, calculated recursively as

$$Y_{n+1} = Y_n + b(Y_n)\Delta t + \sigma(Y_n)\delta W_n + \frac{1}{2}\sigma(Y_n)\sigma'(Y_n)((\delta W_n)^2 - \Delta t), \tag{13}$$

where $\delta W_n \sim N(0, \Delta t)$ are i.i.d. random variables.

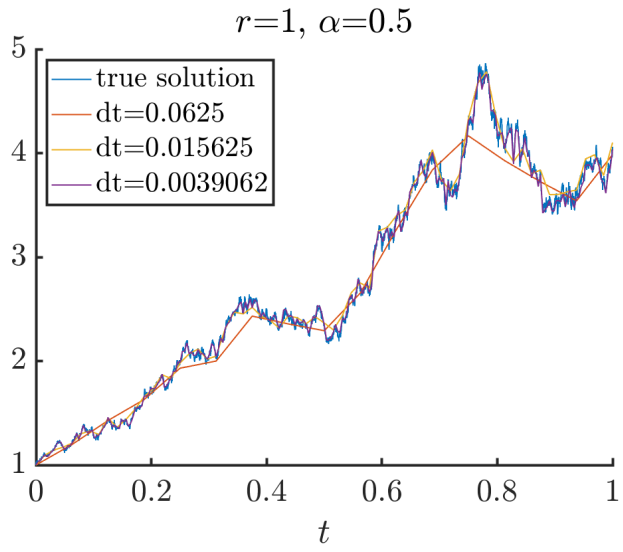


Figure 1: Numerical approximations of GBM $dX_t = rX_t dt + \alpha X_t dW_t$ using different timesteps Δt. The same realization of Brownian motion is used for each approximation.

***Exercise* 9.1.** Write down the full expression for R_2 in (12).

We could continue the Taylor expansion to derive schemes that are even more accurate. This would require calculating integrals of the form $\int_0^t dW_{s_1}^{i_1} \int_0^{s_1} dW_{s_2}^{i_2} \cdots \int_0^{s_{k-1}} dW_{s_k}^{i_k}$, where $i_j \in \{0,1\}$, and we use the notation $W_t^1 = W_t$, $W_t^0 = t$. For *scalar* equations, all the Brownian motions above are the same, and these integrals can be computed recursively – you will do this on the homework. So in principle, higher-order schemes can be derived, though they are not often used in practice, since the advantage they give is often not worth the extra programming.

For *vector* equations, deriving more accurate schemes requires calculating multiple integrals against different Brownian motions, such as

$$\int_0^t \int_0^s dW_u^{(1)} dW_s^{(2)},$$

where $W^{(1)}, W^{(2)}$ are independent Brownian motions. There are no known analytic expressions for such multiple integrals, even for the simplest case above. These multiple integrals can be approximated, for example using a Karhunen-Loeve expansion or by stochastic simulation, but often for such problems, Euler-Maruyama (or variants of it adapted to the structure of a particular problem) is often the most practical method.

9.2 Strong and weak convergence

How can we judge a scheme's quality? Commonly-used notions are

- *Consistency* – whether the mean and a notion of variance of the increment converge to those of the Itô process with time step Δt.
- *Convergence* – whether the global error over a fixed time interval $[0,T]$ converges to zero in some sense with time step Δt.
- *Stability* – whether the numerical method reproduces *qualitatively* the same long-time behaviour as the exact solution.

For each of these, there is a strong and a weak form.

- *Strong* forms deal with pathwise results.
- *Weak* forms deal with probability distributions.

In the following, let $Y^{\Delta t} = Y_0^{\Delta t}, Y_1^{\Delta t}, \ldots$ be a discrete approximation to $X = (X_t)_{t\geq 0}$, at times $0 = t_0 < t_1 < \cdots < t_N = T$, with maximum increment $\Delta t = \max_i |t_{i+1} - t_i|$.

Definition. *$Y^{\Delta t}$ converges strongly* to X at time T *with order* α if there exist constants $C > 0$, $\delta_0 > 0$, independent of Δt, such that

$$\mathbb{E}|Y_N^{\Delta t} - X_T| \leq C(\Delta t)^\alpha \qquad \forall\, \Delta t < \delta_0.$$

Definition. *$Y^{\Delta t}$ converges weakly* to X at T *with order* β with respect to a class of functions $\mathscr{C}$ if, for each $f \in \mathscr{C}$, there exist constants $C_f > 0$, $\delta_0 > 0$, independent of Δt, such that

$$|\mathbb{E}f(Y_N^{\Delta t}) - \mathbb{E}f(X_T)| \leq C_f(\Delta t)^\beta \qquad \forall\, \Delta t < \delta_0,$$

The constant C_f can depend on f but not on Δt.

Remark. The set $\mathscr{C}$ may be taken for example to be the the set of l times continuously differentiable functions whose derivatives up the the lth order have at most polynomial growth. This ensures the set contains all polynomials (Kloeden and Platen [1999], p.327).

The difference between strong and weak convergence, is that strong convergence requires the individual paths to converge as the timestep gets smaller, but for a *fixed* realization of Brownian motion. Therefore, when testing for strong convergence numerically, the same realization of Brownian motion must be used for all approximations (Figure 1). Weak convergence requires only the probability distributions to converge. Therefore a different Brownian motion can be used for each numerical approximation, or even a random process that is not Brownian motion but has increments with the same mean and variance.

Theorem. *If $\mathscr{C}$ is restricted to Lipschitz continuous functions, then $\beta \geq \alpha$ (Weak order $\geq$ Strong order).*

Proof. Suppose f has Lipschitz constant K. Then

$$\begin{aligned} |\mathbb{E}f(Y_N^{\Delta t}) - \mathbb{E}f(X_T)| &\leq \mathbb{E}|f(Y_N^{\Delta t}) - f(X_T)| \qquad \text{same as } \left|\int g - h\right| \leq \int |g-h| \\ &\leq K\mathbb{E}|Y_N^{\Delta t} - X_T|. \end{aligned}$$

Therefore if $Y^{\Delta t} \to X$ with strong order α, it also converges to X with weak order at least α. □

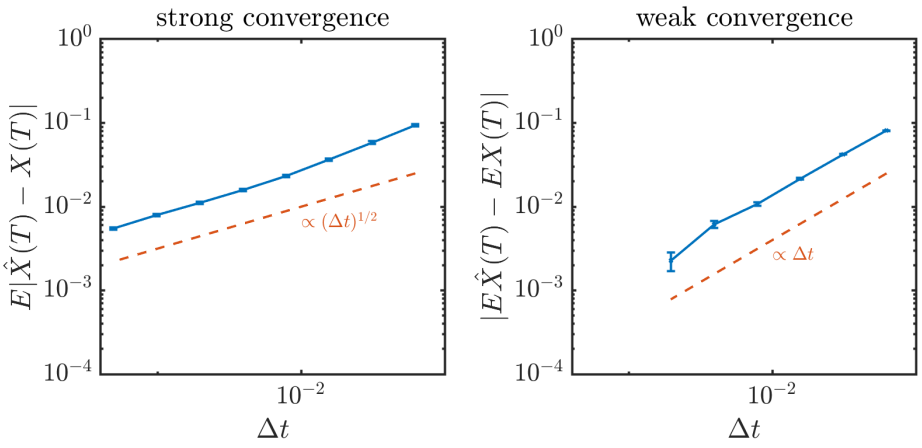


Figure 2: Demonstrating strong (left) and weak (right) convergence for a GBM $dX_t = rX_t dt + \alpha X_t dW_t$ with $r = 1, a = 0.4$, using the Euler-Maruyama discretization with different timesteps Δt. Error bars are 1 standard deviation over M independent samples, where $M = 1000$ for strong convergence, and $M = 4 \times 10^6$ for weak convergence. The total time was $T = 1$ and the initial condition was $X_0 = 1$.

Theorem. *Given an SDE whose coefficient functions and initial condition satisfy the same conditions required for existence and uniqueness of the solution to the SDE, plus additional smoothness conditions*[1]*, we have*

	Strong Order	*Weak Order*
Euler-Maruyama	*1/2*	*1*
Milstein	*1*	*1*

Remark. For additive noise (σ =const), the EM and Milstein schemes are equivalent, since $\sigma\sigma' = 0$. Therefore for additive noise EM will converge with strong and weak order 1.

We outline here a proof of strong convergence for Euler-Maruyama. Proving weak convergence requires the PDE theory of diffusion processes and we leave this until later in the course.

Outline of proof of strong convergence, Euler-Maruyama. (Kloeden and Platen [1999], Theorem 10.2.2 p.342.)

The full proof is quite involved but we outline here the major steps, referring the reader to Kloeden and Platen [1999] to fill in the details.

We list here the full set of conditions required to prove strong convergence with order $1/2$ of EM. Below,

[1]For the full list of conditions, see the following theorems in Kloeden and Platen [1999]: Theorem 10.2.2, for strong convergence of EM, Theorem 10.3.5 for strong convergence of Milstein, Theorem 10.6.3 for strong convergence of a general order γ scheme, Theorem 14.5.1 for weak convergence of a general order γ scheme.

K_1, K_2, K_3, K_4 are positive constants independent of Δt.

$$\mathbb{E}|X_0|^2 < \infty$$
$$\mathbb{E}(|X_0 - Y_0^{\Delta t}|^2)^{1/2} \le K_1(\Delta t)^{1/2}$$
$$b, \sigma \text{ are Lipschitz continuous in } x \text{ with constant } K_2$$
$$b, \sigma \text{ satisfy linear growth in } x \text{ with constant } K_3$$
$$|b(s,x) - b(t,x)| + |\sigma(s,x) - \sigma(t,x)| \le K_4(1+|x|)|s-t|^{1/2}$$

We work with a 1-dimensional process to simplify notation but the proof for a multidimensional process follows the same estimates.

We compare the solution X to (1), to an interpolated process $Y^{\Delta t}$ defined (with a slight abuse of notation) as

$$Y_t^{\Delta t} = Y_n^{\Delta t} + \int_{t_n}^t b(t_n, Y_n^{\Delta t})ds + \int_{t_n}^t \sigma(t_n, Y_n^{\Delta t})dW_s, \qquad t \in [t_n, t_{n+1}], \quad n = 0, 1, \ldots. \tag{14}$$

This interpolation is not a linear interpolation; it keeps the irregularity in the the diffusion term.

The first step is to show that

$$\mathbb{E}\left(\sup_{0\le s\le t} |X_s|^2\right) \le C_1(1+|X_0|^2), \tag{15}$$

where C_i hereafter denotes a constant independent of Δt. This can be shown by deriving an SDE for $|X_t|^2$, estimating growth bounds, and applying Doob's inequality[2] (Kloeden and Platen [1999] Theorem 4.5.4, Exercise 4.5.16). Similarly one can show that (Kloeden and Platen [1999], Lemma 10.8.1)

$$\mathbb{E}\left(\sup_{0\le s\le t} |Y_s^{\Delta t}|^2\right) \le C_2(1+|Y_0^{\Delta t}|^2). \tag{16}$$

Next we define

$$Z(t) = \mathbb{E}\left(\sup_{0\le s\le t} |X_s - Y_s^{\Delta t}|^2\right).$$

We have

$$Z(t) \le C_3\left(|X_0 - Y_0^{\Delta t}|^2 + R_t + S_t + T_t\right),$$

where, writing $n_s = \max\{n : t_n \le s\}$ for the number of discretization points less than s,

$$R_t = \mathbb{E}\left(\sup_{0\le s\le t}\left|\sum_{n=0}^{n_s}\int_{t_n}^{\min(s,t_{n+1})}(b(t_n,X_{t_n}) - b(t_n,Y_{t_n}^{\Delta t}))du + \int_{t_n}^{\min(s,t_{n+1})}(\sigma(t_n,X_{t_n}) - \sigma(t_n,Y_{t_n}^{\Delta t}))dW_u\right|^2\right)$$
$$S_t = \mathbb{E}\left(\sup_{0\le s\le t}\left|\sum_{n=0}^{n_s}\int_{t_n}^{\min(s,t_{n+1})}(b(t_n,X_u) - b(t_n,X_{t_n}))du + \int_{t_n}^{\min(s,t_{n+1})}(\sigma(t_n,X_u) - \sigma(t_n,X_{t_n}))dW_u\right|^2\right)$$
$$T_t = \mathbb{E}\left(\sup_{0\le s\le t}\left|\sum_{n=0}^{n_s}\int_{t_n}^{\min(s,t_{n+1})}(b(u,X_u) - b(t_n,X_u))du + \int_{t_n}^{\min(s,t_{n+1})}(\sigma(u,X_u) - \sigma(t_n,X_u))dW_u\right|^2\right).$$

[2]Given a martingale $M = (M_t)_{t\ge 0}$ with respect to W, $\mathbb{E}\sup_{0\le s\le t}|M_t|^p \le \frac{p}{p-1}\mathbb{E}|M_t|^p$.

Using Lipschitz continuity and Doob's inequality one can show that (Kloeden and Platen [1999], Lemma 10.8.1)

$$\begin{aligned} R_t &\le C_4 \left(\int_0^t \mathbb{E} \sup_{0\le s\le u} \left| b(t_{n_s}, X_{t_{n_s}}) - b(t_{n_s}, Y^{\Delta t}_{t_{n_s}}) \right|^2 du + \int_0^t \mathbb{E} \sup_{0\le s\le u} \left| \sigma(t_{n_s}, X_{n_s}) - \sigma(t_{n_s}, Y^{\Delta t}_{n_s}) \right|^2 dW_u \right) \\ &\le 2C_4 K_2^2 \int_0^t Z(u) du \end{aligned}$$

We also have

$$\begin{aligned} S_t &\le C_5 \left(\int_0^t \mathbb{E} \sup_{0\le s\le u} \left| X_{t_{n_s}} - X_s \right|^2 du \right) \le C_6 \left(1 + |X_0|^2\right) \Delta t, \\ T_t &\le C_7 \left(1 + |X_0|^2\right) \Delta t. \end{aligned}$$

Combining the estimates gives

$$Z(t) \le C_3 |X_0 - Y_0^{\Delta t}|^2 + C_8 (1 + |X_0|^2) \Delta t + C_9 \int_0^t Z(u) du.$$

Applying Gronwall's inequality gives

$$Z(t) \le C_{10} \left(|X_0 - Y_0^{\Delta t}|^2 + (1 + |X_0|^2) \Delta t \right)$$

from which we obtain

$$\mathbb{E}|X_T - Y_T^{\Delta t}| \le K_5 \Delta t^{1/2}.$$

This shows strong convergence of $Y^{\Delta t}$ to X with order 1/2.

□

How can you demonstrate the order of convergence numerically? Here are some possibilities.

(1) Compare to the true solution, if it is known. For example, there are analytical solutions available for the Ornstein-Uhlenbeck process, Geometric Brownian motion, and several other equations; these all involve W but you can generate exact realizations of W at a finite set of time points.

(2) Compare to a very high-resolution simulation.

(3) Look at the difference between solutions as you double the resolution. We expect

$$\frac{\|Y^{\Delta t} - Y^{\Delta t/2}\|}{\|Y^{\Delta t/2} - Y^{\Delta t/4}\|} = 2^{\alpha} + O(\Delta t),$$

where α is the order and $\|\cdot\|$ is some norm.

Options (2),(3) would show the solution converges to something, however there is no guarantee that the solution it converges to is the correct one.

To demonstrate strong convergence, one has to compare paths with different timesteps, using the *same* Brownian motion for each path. To demonstrate weak convergence, one typically picks a function, such as $f(x) = x$, and then estimates $\mathbb{E} f(Y_t^{\Delta t})$ by generating many trajectories with a fixed timestep and computing

the empirical average. See Figure 2 for an example illustrating the convergence of the Euler-Maruyama discretization of GBM.

The numerically-measured error will also contain sampling error, since you have a finite number of samples. Therefore, the number of samples must be large enough to make the sampling error smaller than the error due to discretization. It is good practice to construct error bars on each estimate, for example using the standard deviation of the estimate over independent samples.

There will also be error due to bias in the random number generator, and rounding error due to machine precision. Interestingly it is the bias in the random number generator that becomes important first, when Δt is small enough and the number of samples is large enough! See Hingham [2001].

9.3 Stochastic stability

Convergence bounds the error over time intervals $[0,T]$ using a constant $C(T)$. But typically, $C(T) \nearrow \infty$ as $T \nearrow \infty$. In many situations, such as first-passage problems, or to study long-time behaviour, we need to simulate the equation indefinitely. In such a situation, we may ask that the numerics reproduce the correct *qualitative* behaviour.

One way to do this is with the notion of asymptotic stability.[3] Typically, we pick a particular class of equations on which to study this concept. One common choice is to look at the behaviour near fixed points of linear equations.

Recall the concept of linear stability for deterministic ODEs:

- We typically study the behaviour of $\frac{dX}{dt} = \lambda X$, where λ is a complex number.
- The fixed point $X = 0$ is *asymptotically stable* if $\lim_{t\to\infty} X(t) = 0$. This happens for the equation above when $Re\{\lambda\} < 0$.
- If we discretize, we ask when the numerical scheme reproduces this same behaviour: that $Y_n \to 0$ as $n \to \infty$. This will typically be true only for some step sizes Δt. The set of values of $\lambda\Delta t$ for which the numerical scheme converges to zero forms the *domain of linear stability* of the scheme.

For SDEs, stability is commonly studied for a Geometric Brownian Motion,

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \qquad \lambda, \mu \in \mathbb{R}. \tag{17}$$

Such an equation arises from linearizing a nonlinear SDE about a fixed point, where the right-hand side of the SDE is zero. Since we are now dealing with random variables, which are infinite-dimensional objects, norms are not equivalent in general and there are different notions of "$X_t \to 0$". Two common ones are:

Definition. The solution $X_t = 0$ is *mean-square stable* (for a given pair λ, μ) if $\lim_{t\to\infty} \mathbb{E}X_t^2 = 0$ for any X_0.

Definition. The solution $X_t = 0$ is *asymptotically stable* if $P(\lim_{t\to\infty} X_t = 0) = 1$ for any X_0.

Recall in Lecture 8 we found the solution to (17) to be

$$X_t = X_0 e^{(\lambda - \frac{\mu^2}{2})t + \mu W_t}.$$

[3] Note that *numerical stability* is another concept, related to but different from that discussed here, that asks whether two nearby trajectories tend to stay together or to diverge. See Kloeden and Platen [1999], Section 9.8.

We showed in the lecture and homework that

- GBM is mean-square stable $\Leftrightarrow \lambda + \frac{1}{2}\mu^2 < 0$.
- GBM is asymptotically stable $\Leftrightarrow \lambda - \frac{1}{2}\mu^2 < 0$.

Here, for GBM, mean-square stability $\Rightarrow$ asymptotic stability, but not the reverse.

When does a numerical method give the same type of stability as the true solution? Consider Euler-Maruyama, which gives the discretization

$$Y_{n+1} = Y_n(1 + \lambda\Delta t + \mu\Delta W_n)\,.$$

Squaring gives

$$\mathbb{E}Y_{n+1}^2 = \left(|1+\lambda\Delta t|^2 + |\mu|^2\Delta t\right)\mathbb{E}Y_n^2\,.$$

Therefore mean-square stability is equivalent to the condition

$$(1+\lambda\Delta t)^2 + \mu^2\Delta t < 1\,. \tag{18}$$

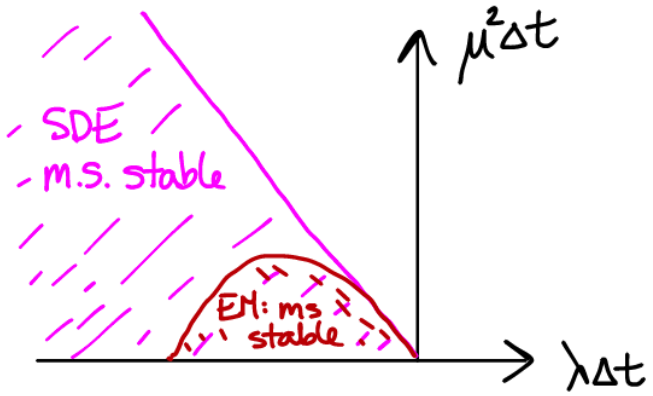
Let $y = \mu^2\Delta t$ and $x = \lambda\Delta t$. The numerical method is mean-square stable when $(1+x)^2 + y < 1$ hence

$$y < -x(2+x) \quad \Longleftrightarrow \quad \Delta t < \frac{-(\mu^2+2\lambda)}{2\lambda^2}\,.$$

The true solution is mean-square stable when

$$\mu^2 < -2\lambda \quad \Longleftrightarrow \quad y < -2x\,.$$

The region where the numerical method is mean-square stable is smaller than, and contained in, the region where the true solution is mean-square stable, as shown in the image below.



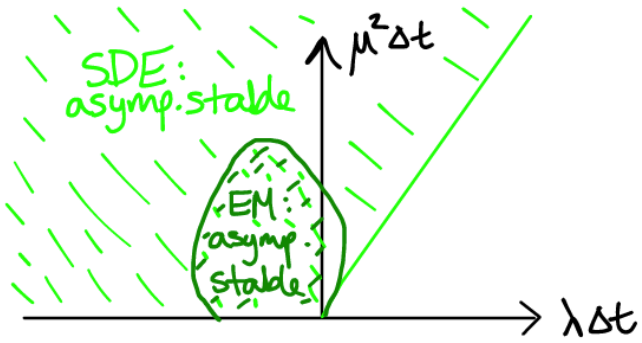
What about asymptotic stability? The true solution is asymptotically stable when $\lambda < \frac{1}{2}\mu^2$. For the EM method we write the solution as

$$Y_n = Y_0 e^{\sum_{j=0}^{n-1} A_j}, \qquad A_j = \log(1+\lambda\Delta t + \mu\Delta W_j)$$

and observe that the argument of the exponential is a sum of i.i.d. random variables. From the Strong Law of Large Numbers, we obtain that the numerical solution is asymptotically stable when

$$\mathbb{E}_Z \log|1+\Delta t\lambda+\sqrt{\Delta t}\mu Z|<0,$$

where $Z\sim N(0,1)$ is a standard normal. The regions of stability for the true solution, and numerical solution, are shown below.



9.4 Implicit methods

We now from the numerical analysis of deterministic equation that implicit methods often have better stability properties than explicit methods, allowing you to take significantly larger time steps, at the expense of solving a large nonlinear system of equations. Should such implicit methods be used for SDEs?

Let's see what would happen if we discretized (17) using a fully implicit Euler scheme [Kloeden and Platen, 1999, p.336]. We obtain

$$Y_{n+1}=Y_n+\lambda Y_{n+1}\Delta t+\mu Y_{n+1}\Delta W_n\,, \tag{19}$$

which we can solve to obtain

$$Y_{n+1}=Y_n\frac{1}{1-\lambda\Delta t-\mu\Delta W_n}\,.$$

There is a problem: the factor multiplying Y_n can approach ∞ or change sign. This is true no matter how small we choose Δt, because the probability density for ΔW has unbounded support. Indeed, $\mathbb{E}|Y_{n+1}/Y_n|=\infty$ as follows from the exercise below. Therefore, it is usually not possible to have a fully implicit scheme for an SDE.

***Exercise* 9.2.** Let $S\sim N(\mu,\sigma^2)$, for $\mu,\sigma\in\mathbb{R}$, $\sigma>0$. Show that $\mathbb{E}|S|^{-1}=\infty$.

We can still use implicit methods on the deterministic part of the equation, or in a weak approximation. For example, if one desires only a weakly convergent numerical approximation, then one can replace ΔW with a bounded random variable with the same mean and variance, such as the two-point random variable $\Delta\tilde{W}=\pm\sqrt{\Delta}$ with equal probability. One also has to modify the drift term to obtain a consistent scheme [see Kloeden and Platen, 1999, p.337]; showing that one has to be careful with implicit methods because of the

distinctions between the various kinds of integrals. A weakly convergent fully implicit Euler approximation would be

$$Y_{n+1} = Y_n + (b - \sigma\frac{\partial\sigma}{\partial x})(t_{n+1}, Y_{n+1})\Delta t + \sigma(t_{n+1}, Y_{n+1})\Delta\tilde{W}_n. \tag{20}$$

***Exercise* 9.3.** (a) Show that (19) is cannot be weakly consistent, by showing that the mean of the approximation is not consistent:

$$\mathbb{E}\left|\mathbb{E}\frac{Y^\Delta_{n+1} - Y^\Delta_n}{\Delta t} - b(t_n, Y^\Delta_n)\right|^2 \not\to 0 \qquad \text{as } \Delta t \to 0.$$

(b) Show that the mean of approximation (20) is consistent, in the sense above.

If one desires a strongly convergent numerical approximation, then one can typically only use implicit methods for the deterministic parts of the equation. For example, a semi-implicit method to solve (17) would be [Pavliotis, 2014, Section 5.2.1]

$$Y_{n+1} = Y_n + (\theta\lambda Y_{n+1} + (1-\theta)Y_n)\Delta t + \mu Y_n \Delta W_n. \tag{21}$$

This scheme is called the *stochastic theta method*.

***Exercise* 9.4.** Show that when $\theta \in [\frac{1}{2}, 1]$, the stochastic theta method in (21) is mean-square stable for all time steps Δt, provided the true solution is mean-square stable. (We say it is *A-stable*.) Derive the region of mean-square stability for $\theta \in [0, \frac{1}{2})$.

References

E, W., Li, T., and Vanden-Eijnden, E. (2014). *Applied Stochastic Analysis*. In preparation.

Hingham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546.

Kloeden, P. E. and Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*. Springer.

Pavliotis, G. A. (2014). *Stochastic Processes and Applications*. Springer.