Advanced computational methods-Lecture 7

1 Continued from last week

If we include the dt^2, we get

$$X^{n+1} = X^n + b(X^n)\Delta t + \sigma(X^n)\Delta W_n + \frac{1}{2}\sigma(X^n)\sigma'(X^n)(\Delta W_n^2 - \Delta t)$$
$$+(L^1)^2\sigma(X^n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{s}(W(s_1)-W(t_n))dW_{s_1}dW_s + L^1 b(X^n)\int_{t_n}^{t_{n+1}}(W(s)-W(t_n))ds$$
$$+ L^0\sigma(X^n)\int_{t_n}^{t_{n+1}}(s-t_n)dW(s) + \frac{1}{2}L^0 b(X^n)\Delta t^2.$$

Let

$$\rho := X_{t,x}(t+h) - \bar{X}_{t,x}(t+h).$$

We recall

$$|\mathbb{E}\rho| \le K\sqrt{1+|x|^2}h^{p_1}, \quad \sqrt{\mathbb{E}|\rho|^2} \le K\sqrt{1+|x|^2}h^{p_2},$$

For the scheme

$$p_1 = 3, p_2 = 2.$$

These two numbers can be read directly from the truncation errors. The mean error is only determined by the integration on time (no dW should be involved) while the p_2 indices is determined by the lemma (counting dW as one half and ds as 1). The order is 3/2.

Exercise: Compute p_1 and p_2 for $X^{n+1} = X^n + \sigma(X^n)\Delta W_n$. Does this scheme converge?

1.1 Modeling the integrals

The question remains: how do we model the above stochastic integrals? Let us start with the simpler case: the noise is constant.

$$X^{n+1} = X^n + b(X^n)\Delta t + \sigma(X^n)\Delta W_n$$
$$+ L^1 b(X^n)\int_{t_n}^{t_{n+1}}(W(s)-W(t_n))ds + \frac{1}{2}L^0 b(X^n)\Delta t^2.$$

Then, we need to simulate

$$\Delta Z := \int_{t_n}^{t_{n+1}}(W(s)-W(t_n))ds.$$

If the noise depends on time t or in general multiplicative, we can have

$$\int_{t_n}^{t_{n+1}}(s-t_n)dW(s)=h\Delta W-\int_{t_n}^{t_{n+1}}(W(\theta)-W(t_n))d\theta.$$

This stochastic integration by parts can be derived by apply Itô's formula to $(t-t_n)(W(t)-W(t_n))$.

For this random variable, it can be shown

$$\mathbb{E}\Delta W\Delta Z=\frac{1}{2}h^2.$$

This means ΔW is independent of $\zeta-\frac{1}{2}h\Delta W$ (two normal variables are independent if they are uncorrelated). Moreover,

$$\mathbb{E}(\zeta-\frac{1}{2}h\Delta W)^2=\frac{1}{12}h^3.$$

Hence, ΔZ is with mean 0, variance $\frac{1}{3}h^3$ such that

$$\mathbb{E}(\Delta Z\Delta W)=\frac{1}{2}h^2.$$

A way to sample such a variable is

$$\Delta Z=\frac{1}{2}h(\Delta W+\Delta V/\sqrt{3}),$$

where ΔV is i.i.d with ΔW.

The above implementation was due to Platen and Wagner. In the full scheme, it is given by

$$\begin{aligned}X^{n+1}&=X^n+bk+\sigma\Delta W_n+\frac{1}{2}\sigma\sigma'(\Delta W^2-k)\\&\quad+b'\sigma\Delta Z_n+\frac{1}{2}(bb'+\frac{1}{2}\sigma^2b'')k^2+(b\sigma'+\frac{1}{2}\sigma^2\sigma'')(\Delta W_nk-\Delta Z_n)\\&\quad+\frac{1}{2}\sigma(\sigma\sigma''+(\sigma')^2)(\frac{1}{3}\Delta W_n^2-k)\Delta W_n.\end{aligned}$$

A fact is that only using the increments of Brownian motion ΔW_n, the strong order is at most 1.

Remark 1. *Compared with the high order strong schemes, high order weak schemes are more easily constructed and are much simpler. In fact, a part of the above strong* 1.5 *scheme gives the weak 2 scheme:*

$$\begin{aligned}X^{n+1}&=X^n+bk+\sigma\Delta W_n+\frac{1}{2}\sigma\sigma'(\Delta W^2-k)\\&\quad+b'\sigma\Delta Z_n+\frac{1}{2}(bb'+\frac{1}{2}\sigma^2b'')k^2+(b\sigma'+\frac{1}{2}\sigma^2\sigma'')(\Delta W_nk-\Delta Z_n).\end{aligned}$$

Hence, in practice, people usually only care about high order weak schemes.

2 Implicit schemes

2.1 Drift implicit schemes

For deterministic cases, the implicit schemes usually have better stability conditions. However, for SDE, this is harder.

Often, one can make the drift part implicit while the stochastic part is still explicit. For example,

$$X^{n+1} = X^n + \alpha b(X^n)k + (1-\alpha)b(X^{n+1})k + \sigma(X^n)\Delta W_n.$$

The strong convergence order of such schemes is 1/2.

One may also consider using Itô-Taylor expansion trying to involve more terms. However, there may be some issues when doing so. Let us briefly investigate these issues here.

We again take $1D$ as the example. The general dimension is similar. The Itô's formula reads

$$\begin{aligned} f(X_t) = f(X_{t_0}) + \int_{t_0}^t \left(b(X_s)\frac{\partial}{\partial x} f(X_s) + \frac{1}{2}\sigma^2(X_s)\frac{\partial^2}{\partial x^2} f(X_s) \right) ds + \int_{t_0}^t \sigma(X_s)\frac{\partial f(X_s)}{\partial x} dW_s \\ = f(X_{t_0}) + \int_{t_0}^t L^0 f(X_s) ds + \int_{t_0}^t L^1 f(X_s) dW_s, \end{aligned}$$

where L^0 is the generator of the scheme.

Instead of subtituting the same expansions into the functions in the integrand, we persue an implicit way. This Itô's integral tells us that

$$f(X_{t_1}) = f(X_t) - \int_{t_1}^t L^0 f(X_s) ds - \int_{t_1}^t L^1 f(X_s) dW_s.$$

Using this to replace the terms in the integrand can yield implicit schemes. For the $L^0 f$ term, we then have

$$\begin{aligned} f(X_t) = f(X_{t_0}) + \int_{t_0}^t \Bigg[L^0 f(X_t) - \int_s^t L^0 f(X_{s_1}) ds_1 \\ - \int_s^t L^1 f(X_{s_1}) dW_{s_1} \Bigg] ds + \int_{t_0}^t L^1 f(X_s) dW_s, \end{aligned}$$

This expansion could yield some the implicit scheme above for $\alpha = 0$.

However, one might be attempted to do the same trick for the Itô integral term, which is, however, unacceptable. The reason is that this expansion makes the integrand involve the futur, thus not adapted to the underlying filtration. This is not desired for Itô's integrals. This is the first issue for doing such expansion.

Another way is to use the forward substitution to make this explicit first and then make it implicit:

$$L^1 f(X_{t_0})\Delta W + \int_{t_0}^{t}\int_{t_0}^{s_1} L^0(L^1 f)(X_{s_2})\, ds_2 dW_{s_1} + \int_{t_0}^{t}\int_{t_0}^{s_1} (L^1)^2 f(X_{s_2}) dW_{s_2} dW_{s_1}.$$

Then, for the first term, apply this backward expansion:

$$\left[L^1 f(X_t) - \int_{t_0}^{t} L^0 L^1 f(X_s) ds - \int_{t_0}^{t} (L^1)^2 f(X_s) dW_s\right]\Delta W \\ + \int_{t_0}^{t}\int_{t_0}^{s_1} (L^1)^2 f(X_{s_2}) dW_{s_2} dW_{s_1}.$$

This then gives some implicit scheme as

$$X^{n+1} = X^n + b(X^{n+1})\Delta t - \sigma\sigma'(X^n)h + \sigma(X^{n+1})\Delta W.$$

The term "$-\sigma\sigma'(X^n)h$" comes from $-(L^1)^2 f(\Delta W)^2$, which is clearly not small.

This, however, is also undesired. The reason is that one must solve X^{n+1}. For example, if it is the geometric Brownian motion,

$$X^{n+1} = \frac{X^n - \sigma\sigma'(X^n)h}{1 - \lambda\Delta t - \sigma\Delta W}$$

Since ΔW is unbounded and it takes normal distribution, the moments of this numerical solution is often infinity. Hence, there can be no convergence. This is the second issue. One will see that later, we will take some truncation to make it fully implicit.

To conclude, using the Taylor expansion to obtain higher implicit schemes is chanllenging, especially for the stochastic part.

2.2 Balanced schemes

Let us try to introduce the implicity in the stochastic term. Consider the geometric Brownian motion with diffusion only:

$$dX = \sigma X dW.$$

The Euler-Maruyama scheme is

$$X^{n+1} = X^n + \sigma X^n \Delta W.$$

Consider making the second implicit.

$$X^{n+1} = X^n - \sigma^2 X^n h + \sigma X^{n+1} \Delta W.$$

The term $-\sigma^2 X_n h$ is here because we have Itô's integral, using the terminal value for the integral must be balanced by some correction terms. This implicit method, as have mentioned, will not work since the moments of X^{n+1} is often infinity.

One possible way to deal with this is to make the part that does not depend on ΔW implicit. For example, in the Milstein scheme:

$$X^{n+1} = X^n + \sigma X^n \Delta W + \frac{1}{2}\sigma^2 X^n (\Delta W^2 - h).$$

One can make the term for h implicit:

$$X^{n+1} = X^n + \sigma X^n \Delta W + \frac{1}{2}\sigma^2 (\Delta W^2 X^n - h X^{n+1}).$$

This, however, is not very satisfatory in practice.

The idea of balanced method is that the geometric Brownian motion is always positive. Hence, when $\Delta W < 0$, it is better to use implicit scheme.

This motivates the following scheme:

$$X^{n+1} = X^n + \sigma X^n \Delta W + \sigma (X^n - X^{n+1}) |\Delta W|. \tag{1}$$

This scheme can guarantee the positivity of the solution can it behaves well for long time.

A natural question is: do we need the correction so that the mean is captured correctly? In fact, for this case, we do not need.

$$X^{n+1} = \frac{X^n (1 + \sigma \Delta W + \sigma |\Delta W|)}{1 + \sigma |\Delta W|}$$

Formal expansion shows that the key error term is $-\sigma^2 X^n |\Delta W| \Delta W$. Due to the symmetry, the mean of this term is zero. Hence, there is no need to do correction.

The general balanced method for $dX = b\,dt + \sigma dW$ is given by

$$X^{n+1} = X^n + b(X^n) h + \sigma(X^n) \Delta W + C_n (X^n - X^{n+1}),$$

where

$$C_n = c_0(X^n) h + \sum_{r=1}^{m} c_r(X^n) |\Delta W_r|.$$

Theorem 1. *If The matrices c_i are all positive definite, then the balanced method has strong order* $1/2$.

2.3 Fully implicit schemes

As we have seen the fully implicit scheme like

$$X^{n+1} = X^n + b(X^n)h + \sigma(X^n)\Delta W$$

is troublesome due to the unboundedness of ΔW.

Below, we focus on $d = 1$ to resolve this issue. The general cases can be found in section 1.3.6 in the book of Milstein. Let

$$\Delta W = \sqrt{h}\xi.$$

Now, we aim to truncate this variable which is still symmetric:

$$\zeta_h = \begin{cases} A_h & \xi > A_h, \\ \xi & |\xi| \le A_h, \\ -A_h, & \xi < -A_h. \end{cases}$$

We desire

$$\mathbb{E}|\zeta_h - \xi|^2 \le h^k,\ \ k \ge 1.$$

Clearly, taking

$$A_h = \sqrt{2m|\ln h|}$$

will suffice.

Let us check this.

Lemma 1. *With the truncation above, one has*

$$\mathbb{E}|\zeta_h - \xi|^2 < h^m,$$

and

$$|\mathbb{E}\xi^2 - \mathbb{E}\zeta_h^2| \le (1 + 2\sqrt{2m|\ln h|})h^m.$$

With this approximation, we now get the implicit scheme

$$X^{n+1} = X^n + b(X^{n+1})h - \sigma(X^n)\sigma'(X^n)h + \sigma(X^{n+1})\sqrt{h}\zeta_h.$$

Similarly,

$$X^{n+1} = X^n + b(X^{n+1})h - \beta\sigma(X^n)\sigma'(X^n)h + \sigma((1-\beta)X^n + \beta X^{n+1})\sqrt{h}\zeta_h.$$

Theorem 2. *These schemes have strong convergence order* $1/2$*.*

To verify this, we essentially check the conditions in the convergence theorem. However, the verification is kind of involved. Here, we take a simpler example to illustrate this.

$$b = 0,\ \sigma(x) = \sigma x.$$

The method becomes (taking $m = 1$ in ζ_h)

$$X^{n+1} = X^n - \sigma^2 X^n h + \sigma X^{n+1}\sqrt{h}\zeta_h.$$

Hence, we have

$$X^{n+1} = \frac{X^n - \sigma^2 X^n h}{1 - \sigma\sqrt{h}\zeta_h}$$

Hence, if

$$2h|\ln h| < 1/\sigma^2,$$

the denominator is guaranteed to be positive.

We now verify the two conditions in the convergence theorem. Compare this to E-M.

$$|\mathbb{E}\frac{x - \sigma^2 xh}{1 - \sigma\sqrt{h}\zeta_h} - \mathbb{E}(x + \sigma x\Delta W)| \le C|x|\sigma^2 h|\mathbb{E}\zeta_h^2 - 1| \le C|x|\sigma^2 h^2(1 + C\sqrt{|\ln h|})$$

Hence, $p_1 = 2 - \epsilon$ for any $\epsilon > 0$.

Moreover,

$$\frac{x - \sigma^2 xh}{1 - \sigma\sqrt{h}\zeta_h} - (x + \sigma x\Delta W) = \frac{(\zeta_h\xi - 1)\sigma^2 hx + x\sigma\sqrt{h}(\zeta_h - \xi)}{1 - \sigma\sqrt{h}\zeta_h}$$

Hence, the second moment is controlled by

$$C\mathbb{E}|(\zeta_h\xi - 1)\sigma^2 hx + x\sigma\sqrt{h}(\zeta_h - \xi)|^2 = C(\mathbb{E}|(\zeta_h\xi - 1)\sigma^2 hx|^2 + \mathbb{E}|x\sigma\sqrt{h}(\zeta_h - \xi)|^2)$$

The first term is like h^2. The second term is $|x|^2\sigma^2 h^2$ as well. Taking the square root, we find $p_2 = 1$.

Hence, the convergence order is $1/2$.

3 Stochastic stability and stiff systems

3.1 The notion of stochastic stability

The theory here is an analogy of the stability region for ODE schemes. The model problem for which we apply the scheme is the geometric Brownian motion

$$dX = \lambda X dt + \mu X dW.$$

(Similar to $dX = \lambda X dt$ for ODEs.)

There are several notions of stability. Here, we consider two of them.

Definition 1. *Given $\lambda \in \mathbb{C}$ and $\mu \in \mathbb{C}$, we say the GBM is mean-square stable if $\lim_{t\to\infty} \mathbb{E}|X_t|^2 = 0$. We say it is asymptotically stable if $\mathbb{P}(\lim_{t\to\infty} |X_t| = 0) = 1$.*

The first is satisfies if $Re(\lambda) + \frac{1}{2}|\mu| < 0$ while the second is satisfied if $Re(\lambda - \frac{1}{2}|\mu|^2) < 0$.

We now move to numerical methods. What is the stability condition for Euler-Maruyama scheme? We have the relation

$$X^{n+1} = (1 + \lambda h + \mu \Delta W_n) X^n \Rightarrow \mathbb{E}|X^{n+1}|^2 = (|1 + \lambda k|^2 + |\mu|^2 h)\mathbb{E}|X^n|^2.$$

Denote

$$z = \lambda h, \quad y = |\mu|^2 h.$$

We need

$$|1 + z|^2 + y < 1,$$

for mean-square stable.

For asymptotic stability, one needs

$$\mathbb{E} \log |1 + \lambda k + \mu \sqrt{k} N(0, 1)| < 0.$$

In fact, we need

$$\prod_{i=1}^{n} (1 + \lambda k + \mu \sqrt{k} z_i) \to 0, a.s.$$

We need

$$\limsup_{n\to\infty} Re \sum_{i=1}^{n} \log(1 + \lambda k + \mu \sqrt{k} z_i) = -\infty, a.s.$$

The claim follows by strong LLN.

Hence, for a general numerical method, one can introduce the analogue of stability regions, i.e. the domain for $(z, y) = (\lambda h, |\mu|^2 h)$ such that the numerical solutions are mean-square stable.

3.2 Stiff problems and implicit schemes

We first look at the determinstic ODEs. In the so-called stiff problems, we care about a slowly varying solution while solutions nearby are rapidly varying with much smaller time scales. In some typical physical applications, the transition to the equilibrium solution is fast but the equilibrium solution itself changes slowly. We care the equilibrium solution instead of the fast transition.

For stiff problems, designing numerical schemes is challenging since the fast transition corresponds to negative eigenvalues with large absolute value. For the method to be stable, we need $k\lambda$ to fall into the stability region. However, for explicit methods, the intersection of the stability region and negative real axis usually has a finite length. The explicit schemes requires that k to be very small for stiff problems.

The issue is that we don't care the fast transition, i.e. we only care the smaller eigenvalues but the eigenvalues for fast transition put restrictions. We hope $k \sim 1/|\lambda_{slow}|$ instead of $1/|\lambda_{fast}|$.

Example in the book...

A scheme is said to be **A-stable** if its stability region contains the whole left half plane. Clearly, if we use A-stable schemes, we won't face instability even if our k is large. For deterministic problems, schemes like backward Euler is A-stable.

What happens about the stochastic stability for implicit schemes? This is a possible course project. Figure out the current status in literature regarding the stability of the balanced scheme, the fully implicit schemes.