Stochastic Numerical Analysis

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Multi-dimensional SDEs

So far we have considered scalar SDEs with a single driving Wiener path.

Now we consider the generalisation to a vector SDE

$$\mathrm{d}S_t = a(S_t, t)\,\mathrm{d}t + b(S_t, t)\,\mathrm{d}W_t$$

where $a(S,t)$ is a vector, $b(S,t)$ is a matrix and W_t is a vector of Wiener paths.

In some cases, the different components of W_t are independent, so over a time interval of length h we have

$$\mathbb{E}[\Delta W \Delta W^T] = h\ I$$

where I is the identity matrix.

Multi-dimensional SDEs

In other cases we may have a correlation between the components so that

$$\Delta W \sim N(0, h\,\Sigma)$$

meaning that the components of ΔW have a joint Normal distribution with covariance

$$\mathbb{E}[\Delta W \, \Delta W^T] = h\,\Sigma$$

ΔW can be simulated by letting $\Delta W = \sqrt{h}\,L\,Z$ where Z is a vector of independent $N(0,1)$ random variables and $L\,L^T = \Sigma$, since

$$\mathbb{E}[L\,Z\,Z^T L^T] \;=\; L\,\mathbb{E}[Z\,Z^T]\,L^T \;=\; L\,L^T \;=\; \Sigma$$

Itô isometry

The generalisation of the Itô isometry is

$$
\begin{aligned}
\mathbb{E}\left[\left(\int_0^T\left(\sum_j f_j\,\mathrm{d}W_j\right)\right)^2\right] &= \mathbb{E}\left[\int_0^T \sum_{j,k} f_j\, f_k \Sigma_{jk}\ \mathrm{d}t\right] \\
&\equiv \mathbb{E}\left[\int_0^T f^T \Sigma\, f\ \mathrm{d}t\right]
\end{aligned}
$$

since

$$
\mathbb{E}[\mathrm{d}W_j\ \mathrm{d}W_k] = \Sigma_{jk}\mathrm{d}t.
$$

Itô lemma

The generalisation of the Itô lemma for $f(S,t)$ is

$$\begin{aligned} \mathrm{d}f &= \frac{\partial f}{\partial t}\,\mathrm{d}t + \sum_j \frac{\partial f}{\partial S_j}\,\mathrm{d}S_j \\ &\quad + \tfrac{1}{2}\sum_{j,k} \frac{\partial^2 f}{\partial S_j \partial S_k} \left(\sum_{l,m} b_{jl} b_{km} \Sigma_{lm} \right) \mathrm{d}t \end{aligned}$$

which comes from the fact that

$$\begin{aligned} \mathbb{E}[\,\mathrm{d}S_j\ \mathrm{d}S_k\,] &= \sum_{l,m} b_{jl} b_{km}\ \mathbb{E}[\mathrm{d}W_l\ \mathrm{d}W_m] \\ &= \sum_{l,m} b_{jl} b_{km} \Sigma_{lm}\ \mathrm{d}t \end{aligned}$$

Euler-Maruyama method

For the vector SDE

$$\mathrm{d}S = a(S,t)\ \mathrm{d}t + b(S,t)\ \mathrm{d}W$$

the Euler-Maruyama approximation is again

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\, h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

Very simple – this is one of the great attractions of the Euler-Maruyama method.

The strong error remains $O(h^{1/2})$, and the weak error is still $O(h)$ in most applications.

Milstein Method

In the vector case, the SDE

$$\mathrm{d}S_i = a_i(S,t)\,\mathrm{d}t + \sum_j b_{ij}(S,t)\,\mathrm{d}W_j$$

corresponds to the integral equation:

$$S_i(t) = S_i(0) + \int_0^t a_i(S(s),s)\,\mathrm{d}s + \sum_j \int_0^t b_{ij}(S(s),s)\,\mathrm{d}W_j(s)$$

Milstein Method

An asymptotic expansion gives

$$S_i(t) - S_i(0) \approx \sum_j b_{ij}(S(0),0)\, W_j(t)$$

and hence

$$\begin{aligned} b_{ij}(S(t),t) \quad &\approx \quad b_{ij} + \sum_l \frac{\partial b_{ij}}{\partial S_l}\, (S_l(t) - S_l(0)) \\ &\approx \quad b_{ij} + \sum_{k,l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk}\, W_k(t) \end{aligned}$$

with b and its derivatives evaluated at $(S(0),0)$.

Milstein Method

This then leads to

$$
\begin{aligned}
S_i(h) \;\approx\;& S_i(0) + a_i\, h + \sum_j b_{ij}\, W_j(h) \\
&+ \sum_{j,k,l} \frac{\partial b_{ij}}{\partial S_l}\; b_{lk} \int_0^h W_k(t)\; \mathrm{d}W_j(t)
\end{aligned}
$$

where a, b and its derivatives all evaluated at $(S(0), 0)$.

The problem now is to evaluate the iterated Itô integral

$$
I_{jk} \equiv \int_0^h W_k(t)\; \mathrm{d}W_j(t)
$$

Lévy areas

Itô calculus gives us

$$\mathrm{d}(W_j\, W_k) = W_j\; \mathrm{d}W_k(t) + W_k\; \mathrm{d}W_j(t) + \Sigma_{jk}\; \mathrm{d}t$$

where Σ_{jk} is the correlation between $\mathrm{d}W_j$ and $\mathrm{d}W_k$.

Hence,

$$W_j(h)\, W_k(h) - \Sigma_{jk}\, h = I_{kj} + I_{jk}$$

If we define the Lévy area to be

$$A_{jk} \;=\; I_{kj} - I_{jk} \;=\; \int_0^h W_j(t)\; \mathrm{d}W_k(t) - W_k(t)\; \mathrm{d}W_j(t)$$

then

$$I_{jk} = \tfrac{1}{2}\left(W_j(h)\, W_k(h) - \Sigma_{jk}\, h - A_{jk}\right)$$

Lévy areas

The problem is that it is hard to simulate the Lévy areas:

- conditional distribution depends on $W_j(h)$ and $W_k(h)$ so can't simply invert a cumulative distribution function
- Lyons & Gaines have an efficient technique in 2-dimensions but for higher dimensions, need to simulate Brownian motion within each timestep to approximate the Lévy area

Two notes for later use:

- $\mathbb{E}[A_{jk} \mid \Delta W] = 0$
- $\mathbb{E}[A_{jk}^2] = h^2$ if $\mathbb{E}[\,\mathrm{d}W_j \, \mathrm{d}W_k\,] = 0$.

Milstein method

The Milstein method is

$$
\begin{aligned}
\widehat{S}_{i,n+1} &= \widehat{S}_{i,n} + a_{i,n}\,h + \sum_j b_{ij,n}\,\Delta W_{j,n} \\
&\quad + \tfrac{1}{2} \sum_{j,k,l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk,n} \left(\Delta W_{j,n}\,\Delta W_{k,n} - \Sigma_{jk}\,h - A_{jk,n} \right)
\end{aligned}
$$

with

$$
A_{jk,n} = \int_{t_n}^{t_{n+1}} (W_j(t) - W_j(t_n))\,\mathrm{d}W_k(t) - (W_k(t) - W_k(t_n))\,\mathrm{d}W_j(t)
$$

The strong error is $O(h)$, but the problem is the Lévy areas.

Milstein method

However, using $A_{jk} = -A_{kj}$,

$$\begin{aligned}
\sum_{j,k,l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk}\, A_{jk,n} &= -\sum_{j,k,l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk}\, A_{kj,n} \\
&= -\sum_{j,k,l} \frac{\partial b_{ik}}{\partial S_l}\, b_{lj}\, A_{jk,n} \\
&= \tfrac{1}{2} \sum_{j,k,l} \left(\frac{\partial b_{ij}}{\partial S_l}\, b_{lk} - \frac{\partial b_{ik}}{\partial S_l}\, b_{lj} \right) A_{jk,n}
\end{aligned}$$

and so the Lévy areas are not needed <u>if</u>, for all i, j, k,

$$\sum_{l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk} - \frac{\partial b_{ik}}{\partial S_l}\, b_{lj} = 0.$$

Milstein method

If b is a non-singular diagonal matrix, so each component of $S(t)$ is driven by a separate component of $W(t)$, the commutativity condition reduces to

$$\frac{\partial b_{ij}}{\partial S_k}\, b_{kk} - \frac{\partial b_{ik}}{\partial S_j}\, b_{jj} = 0$$

- if either $i\!=\!j\!=\!k$, or $i\!\neq\!j$ and $i\!\neq\!k$, this is satisfied
- if $i\!=\!j$ and $i\!\neq\!k$, it requires $\dfrac{\partial b_{ii}}{\partial S_k} = 0$
- if $i\!=\!k$ and $i\!\neq\!j$, it requires $\dfrac{\partial b_{ii}}{\partial S_j} = 0$
- hence, OK provided b_{ii} depends only on S_i

Clark & Cameron paper

We now come to an important result in the paper:

"The maximum rate of convergence of discrete approximations for stochastic differential equations", J.M.C. Clark & R.J. Cameron, pp 162-171 in Lecture Notes in Control and Information Sciences, Volume 25, 1980.

Their analysis considers a very simple 2-dimensional SDE, and proves that **no** numerical approximation is capable of better than $O(h^{1/2})$ strong convergence if it is based solely on the discrete Brownian increments ΔW_n.

Implication: in general you can't do better than Euler-Maruyama method.

Clark & Cameron paper

Their model SDE is

$$\begin{aligned} \mathrm{d}X_1 &= \mathrm{d}W_1 \\ \mathrm{d}X_2 &= X_1\, \mathrm{d}W_2 \end{aligned}$$

with $X_1(0) = X_2(0) = 0$ and independent $\mathrm{d}W_1, \mathrm{d}W_2$

– could hardly be simpler!

They don't consider a numerical approximation, but instead look at how much we know about the analytic solution given knowledge only of the Brownian increments.

Clark & Cameron paper

Over a time interval $[0,h]$, can integrate to obtain

$$\begin{aligned} X_1(h) &= W_1(h) \\ X_2 &= \int_0^h X_1 \, \mathrm{d}W_2 \end{aligned}$$

Now, Itô's lemma gives $\mathrm{d}(W_1 W_2) = W_1\, \mathrm{d}W_2 + W_2\, \mathrm{d}W_1$

$$\Longrightarrow \quad W_1(h)\; W_2(h) = \int_0^h (W_1\, \mathrm{d}W_2 + W_2\, \mathrm{d}W_1)$$

$$\Longrightarrow \quad \int_0^h W_1\, \mathrm{d}W_2 = \tfrac{1}{2} W_1(h)\; W_2(h) + \tfrac{1}{2} A$$

where A is Lévy area for time interval $[0,h]$.

Clark & Cameron paper

Repeating this for N "timesteps" of size $h = T/N$ we obtain

$$
\begin{aligned}
X_1(T) &= \sum_n \Delta W_{1,n} \\
X_2(T) &= \sum_n \Big(X_1(t_n)\Delta W_{2,n} + \tfrac{1}{2}\Delta W_{1,n}\Delta W_{2,n} + \tfrac{1}{2} A_n \Big)
\end{aligned}
$$

Now recall that $\mathbb{E}[A_n \mid \Delta W_n] = 0$, and $\mathbb{E}[A_n^2] = h^2$, and also that

$$\mathbb{V}[Y] \;=\; \mathbb{E}[Y^2] - \left(\mathbb{E}[Y]\right)^2$$

for any random variable Y.

Hence, for **any** numerical approximation $\widehat{X}$ we have ...

Clark & Cameron paper

$$
\begin{aligned}
\mathbb{E}\left[(X_2(T) - \widehat{X}_2(T))^2\right] &= \mathbb{E}\left[\,\mathbb{E}[(X_2(T) - \widehat{X}_2(T))^2 \mid \Delta W]\,\right] \\
&\geq \mathbb{E}\left[\,\mathbb{V}[X_2(T) \mid \Delta W]\,\right] \\
&= \tfrac{1}{4}\,\mathbb{E}\left[\sum_{n=0}^{N-1} \mathbb{V}[A_n \mid \Delta W]\right] \\
&= \tfrac{1}{4}\,\mathbb{E}\left[\sum_{n=0}^{N-1} \mathbb{E}[A_n^2 \mid \Delta W]\right] \\
&= \tfrac{1}{4}\left[\sum_{n=0}^{N-1} \mathbb{E}[A_n^2]\right] \\
&= \tfrac{1}{4}\,T\,h.
\end{aligned}
$$

Clark & Cameron paper

Hence, the minimum RMS error is $O(h^{1/2})$, and the best numerical approximation is the one which gives

$$\widehat{X}_2(T) = \mathbb{E}[X_2(T) \mid \Delta W]$$

which corresponds to neglecting the Lévy areas A_n.

As well as proving that in general you can't achieve a better order of strong convergence than the Euler-Maruyama method (when using only Brownian increments) it also shows the Milstein approximation without the Lévy areas gives the best asymptotic accuracy.