Interest Rate Models

Ivan Matic

Department of Mathematics
Baruch College, CUNY
New York, NY 10010

Ivan.Matic@baruch.cuny.edu

Spring 2016

Outline

1. Simulating Brownian motion

2. Discretizing SDEs

3. Generating Monte Carlo paths for LMM

Monte Carlo methods for LMM

- There are two numerical schemes for generating Monte Carlo paths for LMM:
  (i) Euler's scheme and
  (ii) Milstein's scheme.
- Both of the methods rely on replacing the differentials with finite differences.
- The time is discretized and the distance between consecutive time points δt is taken to be small.
- Euler's scheme converges at a rate $(\delta t)^{\frac{1}{2}}$ while Milstein's scheme converges at the rate δt.

One factor Brownian motion

- The easiest way to simulate Brownian motion is to implement the random walk method.
- The outcome will be a Wiener process sampled at a finite set of event dates $t_0 < t_1 < \cdots < t_m$:

$$\begin{aligned} W(t_{-1}) &= 0, \\ W(t_n) &= W(t_{n-1}) + \sqrt{t_n - t_{n-1}} \cdot \xi_n, \quad n \in \{0, 1, \ldots, m\}. \end{aligned}$$

  t_{-1} is chosen to be 0.
- ξ_n are independent and $\xi_n \sim N(0,1)$.
- A good method of generating ξ_n is to generate a uniform normal u_n (e.g. by using Mersenne twister) and then set

$$\xi_n = N^{-1}(u_n).$$

- N is the cummulative distribution function of a normal (can be computed using Beasley-Springer-Moro algorithm).

One factor Brownian motion

- Another way for generating Brownian motion is the *spectral decomposition method*.
- Assuming that Z is a standard Brownian motion, let us denote by C the covariance matrix between the random variables $Z(t_1)$, $Z(t_2)$, $\ldots$, $Z(t_m)$.
- We have that $\mathrm{cov}\,(Z(t_i), Z(t_j)) = \min\{t_i, t_j\}$.
- The matrix C is positive semi-definite, hence it satisfies the following properties:
  (i) It has all non-negative eigenvalues $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_m$.
  (ii) It has m mutually orthogonal eigenvectors E_0, E_1, $\ldots$, E_m corresponding to the eigenvalues listed above.
  (iii) If $D = \mathrm{diag}\{\lambda_0, \lambda_1, \ldots, \lambda_m\}$ and Q the matrix whose columns are E_0, E_1, $\ldots$, E_m, then $Q^T Q = I$ and $D = Q^T C Q$.

One factor Brownian motion

- The equation $D = Q^T C Q$ is equivalent to $C = QDQ^T$, hence

$$C_{ij} = \sum_k \lambda_k E_{ki} E_{kj},$$

  where E_{ki} is the ith component of the vector E_k.
- Assume that $\xi_0, \ldots, \xi_m$ are iid standard normal and define

$$\hat{Z}_i = \sum_k \sqrt{\lambda_k} E_{ki} \xi_k, \tag{1}$$

  then the covariance matrix between $\hat{Z}_0, \ldots, \hat{Z}_m$ is precisely C.
- Thus we may simulate our Brownian motion as $Z(t_i) = \hat{Z}_i$. For computational efficiency we may truncate the summation in (1) to include only the terms $k \leq p$ for some p.

Multi-factor Brownian motion

- Our next task is to simulate d dimensional Brownian motion. We want to achieve $\mathbb{E}\left[dZ_a(t)dZ_b(t)\right] = \rho_{ab}\,dt$.
- In order to use the spectral method we need to ensure that ξ_k^a, ξ_k^b have covariance ρ_{ab} for every a and b.
- Let us denote by ρ the matrix with entries ρ_{ab}.
- Since ρ is symmetric non-negative definite, it has Cholesky decomposition $\rho = LL^T$ where L is a lower triangular matrix.
- For example, if $d = 2$, and if the covariance of two standard normal random variables is ρ_{12} we can write their covariance matrix as

$$\begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} = LL^T, \quad \text{where } L = \begin{bmatrix} 1 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} \end{bmatrix}.$$

Multi-factor Brownian motion

- If ξ is a d-dimensional vector of independent standard normal random variables, then the vector $L\xi$ represents exactly the d-dimensional vector of normal random variables with covariance ρ. Indeed the covariance matrix C satisfies

$$\begin{aligned} C &= \mathbb{E}\left[(L\xi)\cdot(L\xi)^T\right] = \mathbb{E}\left[L\xi\xi^T L^T\right] \\ &= L\mathbb{E}\left[\xi\xi^T\right] L^T = L \cdot I_{d\times d} \cdot L^T = \rho. \end{aligned}$$

Single equation

- The idea for numerical approach in solving stochastic differential equations is to simulate the paths of the state variables.
- The continuous time system is approximated by a discrete time stochastic sequence.
- We will first consider the one factor SDE

$$\begin{aligned} dX(t) &= A(t, X(t))\, dt + B(t, X(t))\, dW(t), \\ X(0) &= X_0. \end{aligned}$$

- The above equation can be written in an integral form as

$$X(t+\delta) = X(t) + \int_t^{t+\delta} A(s, X(s))\, ds + \int_t^{t+\delta} B(s, X(s))\, dW(s).$$

Single equation

- Applying Ito's formula to a twice continuously differentiable function f we obtain

$$\begin{aligned} df(t,X(t)) &= \mathcal{L}^0 f(t,X(t))\,dt + \mathcal{L}^1 f(t,X(t))\,dW(t), \quad \text{where} \\ \mathcal{L}^0 &= \frac{\partial}{\partial t} + A\frac{\partial}{\partial x} + \frac{1}{2}B^2\frac{\partial^2}{\partial x^2} \quad \text{and} \\ \mathcal{L}^1 &= B\frac{\partial}{\partial x}. \end{aligned}$$

- Applying the Ito formula to $A(t,X(t))$ gives us

$$A(s,X(s)) = A(t,X(t)) + \int_t^s \mathcal{L}^0 A(u,X(u))\,du + \int_t^s \mathcal{L}^1 A(u,X(u))\,dW(u).$$

- Assuming that $s-t$ is small we approximate

$$A(s,X(s)) \approx A(t,X(t)) + \mathcal{L}^0 A(t,X(t)) \int_t^s du + \mathcal{L}^1 A(t,X(t)) \int_t^s dW(u).$$

Single equation

- Therefore

$$\int_t^{t+\delta} A(u, X(u))\,du \approx \delta A(t, X(t)) + \mathcal{L}^0 A(t, X(t)) \int_t^{t+\delta} \int_t^s duds$$
$$+\mathcal{L}^1 A(t, X(t)) \int_t^{t+\delta} \int_t^s dW(u)\,ds.$$

- If we introduce $I_{0,0} = \int_t^{t+\delta} \int_t^s duds$ and $I_{1,0} = \int_t^{t+\delta} \int_t^s dW(u)\,ds$ we can re-write the previous equation as

$$\int_t^{t+\delta} A(u, X(u))\,du \approx \delta A(t, X(t)) + \mathcal{L}^0 A(t, X(t)) I_{0,0} + \mathcal{L}^1 A(t, X(t)) I_{1,0}.$$

Single equation

- In a similar way we obtain

$$B(s,X(s)) = B(t,X(t)) + \int_t^s \mathcal{L}^0 B(u,X(u))\,du + \int_t^s \mathcal{L}^1 B(u,X(u))\,dW(u).$$

- Therefore, if $s-t$ is small we have

$$B(s,X(s)) \approx B(t,X(t)) + \mathcal{L}^0 B(t,X(t)) \int_t^s du + \mathcal{L}^1 B(t,X(t)) \int_t^s dW(u).$$

- If we introduce $I_{0,1} = \int_t^{t+\delta} \int_t^s du dW(s)$ and $I_{1,1} = \int_t^{t+\delta} \int_t^s dW(u) dW(s)$ we obtain

$$\begin{aligned} \int_t^{t+\delta} B(u,X(u))\,dW(u) &\approx B(t,X(t))\Delta W + \mathcal{L}^0 B(t,X(t)) I_{1,0} \\ &+ \mathcal{L}^1 B(t,X(t)) I_{1,1}. \end{aligned}$$

Single equation

- We finally obtain

$$
\begin{aligned}
X(t+\delta) \quad &\approx \quad X(t) + \delta A(t, X(t)) + \mathcal{L}^0 A(t, X(t)) I_{0,0} + \mathcal{L}^1 A(t, X(t)) I_{1,0} \\
&\quad + B(t, X(t)) \Delta W + \mathcal{L}^0 B(t, X(t)) I_{0,1} + \mathcal{L}^1 B(t, X(t)) I_{1,1}.
\end{aligned}
$$

- The integrals $I_{0,0}$, $I_{0,1}$, $I_{1,0}$, and $I_{1,1}$ can be explicitly evaluated.

$$
\begin{aligned}
I_{0,0} &= \int_t^{t+\delta} \int_t^s du\, ds = \frac{\delta^2}{2}, \\
I_{1,0} &= \int_t^{t+\delta} \int_t^s dW(u)\, ds = \int_t^{t+\delta} \left(W(s) - W(t)\right)\, ds, \\
I_{0,1} &= \int_t^{t+\delta} \int_t^s du\, dW(s) = \int_t^{t+\delta} \int_u^{t+\delta} dW(s)\, du \\
&= \int_t^{t+\delta} \left(W(t+\delta) - W(u)\right)\, du = \delta \Delta W - I_{1,0}.
\end{aligned}
$$

Single equation

- We now calculate $I_{1,1} = \int_t^{t+\delta} \int_t^s \, dW(u) dW(s)$ applying Ito's formula to $\frac{1}{2}W^2$. We obtain

$$
\begin{aligned}
\frac{1}{2}W^2(t+\delta) - \frac{1}{2}W^2(t) &= \int_t^{t+\delta} W(s)\, dW(s) + \frac{1}{2}\int_t^{t+\delta} ds \\
&= \int_t^{t+\delta} W(s)\, dW(s) + \frac{\delta}{2}.
\end{aligned}
$$

- Therefore

$$
\begin{aligned}
I_{1,1} &= \int_t^{t+\delta} \left(W(s) - W(t)\right)\, dW(s) = \int_t^{t+\delta} W(s)\, dW(s) - W(t)\Delta W \\
&= \frac{1}{2}(\Delta W)^2 - \frac{\delta}{2}.
\end{aligned}
$$

Single equation

- Summary:

$$
\begin{aligned}
X(t+\delta) \quad \approx \quad & X(t) + A(t, X(t))\delta + B(t, X(t))\Delta W && (2)\\
& + \mathcal{L}^1 B(t, X(t)) I_{1,1} && (3)\\
& + \mathcal{L}^0 A(t, X(t)) I_{0,0} + \mathcal{L}^1 A(t, X(t)) I_{1,0} + \mathcal{L}^0 B(t, X(t)) I_{0,1}.
\end{aligned}
$$

$$
\mathcal{L}^0 = \frac{\partial}{\partial t} + A\frac{\partial}{\partial x} + \frac{1}{2}B^2\frac{\partial^2}{\partial x^2} \qquad \text{and} \qquad \mathcal{L}^1 = B\frac{\partial}{\partial x}.
$$

$$
\begin{aligned}
I_{0,0} &= \frac{\delta^2}{2}, & I_{1,0} &= \int_t^{t+\delta} (W(s) - W(t))\, ds\\
I_{0,1} &= \delta\Delta W - I_{1,0}, & I_{1,1} &= \frac{1}{2}(\Delta W)^2 - \frac{\delta}{2}.
\end{aligned}
$$

- Euler's scheme: Use the first line (2).
- Milstein's scheme: Use the first two lines (2) and (3).

Discretization

- Euler's approximation can be explicitly written as

$$X_{n+1} = X_n + A(t_n, X_n)\delta_n + B(t_n, X_n)\Delta W_n.$$

- Milstein's approximation is

$$\begin{aligned} X_{n+1} &= X_n + A(t_n, X_n)\delta_n + B(t_n, X_n)\Delta W_n \\ &\quad + \frac{1}{2} B(t_n, X_n) B_x(t_n, X_n)\left((\Delta W_n)^2 - \delta_n\right). \end{aligned}$$

- $\delta_n = t_{n+1} - t_n$ and ΔW_n are iid normal random variables with mean 0 and standard deviation δ_n.
- Notice that even if the terms $I_{0,0}$, $I_{0,1}$, and $I_{1,0}$ were included, they would not improve the accuracy by much. Earlier we made an approximation $A(u, X(u)) \approx A(t, X(t))$. This introduced an error of a bigger order.

Systems of SDEs

- Now the state variable is n-dimensional.

$$dX_i(t) = A_i(t, X(t))\, dt + \sum_{k=1}^{d} B_{ik}(t, X(t))\, dW_k(t).$$

  We assume that the components W_k are independent.
- In Integral form the SDE can be written as

$$X_i(t+\delta) = X_i(t) + \int_t^{t+\delta} A_i(s, X(s))\, ds + \sum_{k=1}^{d} \int_t^{t+\delta} B_{ik}(s, X(s))\, dW_k(s).$$

- Recall the multi-dimensional Ito's formula:

$$f(s, X(s)) = f(t, X(t)) + \int_t^s \mathcal{L}^0 f(u, X(u))\, du + \sum_{k=1}^{d} \int_t^s \mathcal{L}^k f(u, X(u))\, dW_k(u).$$

$$\mathcal{L}^0 = \frac{\partial}{\partial t} + \sum_{i=1}^{n} A_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \sum_{k=1}^{d} B_{ik} B_{jk} \frac{\partial^2}{\partial x_i \partial x_j}, \quad \mathcal{L}^k = \sum_{i=1}^{n} B_{ik} \frac{\partial}{\partial x_i}.$$

Systems of SDEs

- We apply Ito's formula to A_i:

$$\begin{aligned} A_i(s, X(s)) &= A_i(t, X(t)) + \int_t^s \mathcal{L}^0 A_i(u, X(u))\, du + \sum_{k=1}^{d} \int_t^s \mathcal{L}^k A_i(u, X(u))\, dW_k(u) \\ &\approx A_i(t, X(t)) + \mathcal{L}^0 A_i(t, X(t)) \int_t^s du + \sum_{k=1}^{d} \mathcal{L}^k A_i(t, X(t)) \int_t^s dW_k(u). \end{aligned}$$

- Let us introduce the notation

$$I_{0,0} = \int_t^{t+\delta} \int_t^s du ds, \qquad I_{k,0} = \int_t^{t+\delta} \int_t^s dW_k(u) ds$$

$$I_{0,k} = \int_t^{t+\delta} \int_t^s du dW_k(s), \qquad \begin{aligned} I_{k,\ell} &= \int_t^{t+\delta} \int_t^s dW_k(u) dW_\ell(s) \\ &= \int_t^{t+\delta} (W_k(s) - W_k(t))\, dW_\ell(s) \\ &= \int_t^{t+\delta} W_k\, dW_\ell - W_k(t) \Delta W_\ell. \end{aligned}$$

Systems of SDEs

- Then we obtain

$$\int_t^{t+\delta} A_i(s, X(s))\, ds \quad \approx \quad A_i(t, X(t))\delta + \mathcal{L}^0 A_i(t, X(t)) I_{0,0} + \sum_{k=1}^{d} \mathcal{L}^k A_i(t, X(t)) I_{k,0}.$$

- In a similar way we get

$$\begin{aligned} \int_t^{t+\delta} B_{ik}(s, X(s))\, dW_k(s) \quad &\approx \quad B_{ik}(t, X(t))\Delta W_k + \mathcal{L}^0 B_{ik}(t, X(t)) I_{0,k} \\ &\quad + \sum_{\ell=1}^{d} \mathcal{L}^\ell B_{ik}(t, X(t)) I_{\ell,k}. \end{aligned}$$

- The terms $I_{0,0}$, $I_{k,0}$ and $I_{0,k}$ are of order δ^2 and $\delta^{3/2}$ and that is smaller than the error we already made. On the other hand using Ito's product rule and the independence of W_k and W_ℓ we get

$$\begin{aligned} \Delta\left(W_k W_\ell\right) &= \int_t^{t+\delta} W_k\, dW_l + \int_t^{t+\delta} W_l\, dW_k + \int_t^{t+\delta} [W_k, W_\ell]\, dt \\ &= I_{k,\ell} + I_{\ell,k} + W_k(t)\Delta W_\ell + W_\ell(t)\Delta W_k + \delta \cdot 1_{k=\ell}. \end{aligned}$$

Integrability condition

- Therefore

$$I_{k,\ell} + I_{\ell,k} = \Delta W_k \cdot \Delta W_\ell - \delta \cdot 1_{k=\ell}.$$

- We now impose the following integrability condition

$$\mathcal{L}^\ell B_{ik} = \mathcal{L}^k B_{i\ell} \quad \text{for all} \quad k, \ell \in \{1, 2, \ldots, d\}.$$

- Therefore we obtain

$$\begin{aligned} X_i(t+\delta) \quad &\approx \quad X_i(t) + A_i(t, X(t))\delta + \sum_{k=1}^{d} B_{ik}(t, X(t))\Delta W_k \\ &\quad + \frac{1}{2} \sum_{k=1}^{d} \sum_{\ell=1}^{d} \mathcal{L}^\ell B_{ik}(t, X(t)) \left(\Delta W_k \cdot \Delta W_\ell - \delta \cdot 1_{k=\ell}\right). \end{aligned}$$

Integrability condition

- Euler's scheme is the following approximation:

$$X_{i,n+1} = X_{i,n} + A_{i,n}\delta_n + \sum_{k=1}^{d} B_{ik,n}\Delta W_{k,n}.$$

- Milstein's scheme is:

$$\begin{aligned} X_{i,n+1} &= X_{i,n} + \left(A_{i,n} - \frac{1}{2}\mathcal{L}^k B_{ik,n}\right)\delta_n + \sum_{k=1}^{d} B_{ik,n}\Delta W_{k,n} \\ &\quad + \frac{1}{2}\sum_{k=1}^{d}\sum_{\ell=1}^{d} \mathcal{L}^\ell B_{ik,n}\Delta W_{k,n}\cdot\Delta W_{\ell,n}. \end{aligned}$$

Discretizing LMM: Euler's scheme

- Choose a sequence of dates t_0, t_1, $\dots$, t_m and let us denote $\delta_n = t_{n+1} - t_n$.
- We will write L_{jn} as an approximation of $L_j(t_n)$.
- Denote $\Delta_{j,n} = \Delta_j(t_n, L_{jn})$ and $B_{jk,n} = B_{jk}(t_n, L_{jn})$.
- The Euler's scheme approximation results in the sequences L_{jn} that satisfy the following recursions

$$L_{i,n+1} \quad = \quad L_{i,n} + \Delta_{i,n}\delta_n + \sum_{k=1}^{d} B_{ik,n}\Delta W_{k,n}.$$

As before, $\Delta W_{k,n}$ is the discretized Brownian motion.

Discretizing LMM: Milstein's scheme

- LMM satisfies the integrability condition and the Milstein's scheme can be used.

$$
\begin{aligned}
L_{i,n+1} \quad &= \quad L_{i,n} + \left(\Delta_{i,n} - \frac{1}{2}\mathcal{L}^k B_{ik,n}\right)\delta_n + \sum_{k=1}^{d} B_{ik,n}\Delta W_{k,n} \\
&\quad + \frac{1}{2}\sum_{k=1}^{d}\sum_{\ell=1}^{d}\mathcal{L}^{\ell} B_{ik,n}\Delta W_{k,n}\cdot\Delta W_{\ell,n}.
\end{aligned}
$$

- Calculating the drift terms takes approximately 50% of total computation time.
- One idea to speed up the calculation is to freeze the values $F_j(t)$ at $F_{j,0}$ in drifts, pre-compute all of $\Delta_{j,0} = \Delta(t, F_{j,0})$, and use them throughout the calculations.
- This results in noticeable errors in pricing the longer dated options.

Efficient drift calculation

- Taking one more step in the low noise expansion beyond the frozen curve approximation produces sufficiently accurate results.
- We approximate $\Delta_j(t, F_j(t))$ with Ito's expansion in which we ignore the terms of order higher than $\frac{1}{2}$. More precisely,

$$\begin{aligned}
\Delta_j(t, F(t)) &\approx \Delta_{j,0} + \int_0^t \sum_{k=1}^{d} \mathcal{L}^k \Delta_j(s, F(s))\, dW_k(s) \\
&\approx \Delta_{j,0} + \sum_{k=1}^{d} \mathcal{L}^k \Delta_j(0, F(0))\, W_k(t)
\end{aligned}$$

- We replace the drifts with

$$\Delta_{j,\frac{1}{2}}(t) \equiv \Delta_{j,0} + \sum_{k=1}^{d} \mathcal{L}^k \Delta_j(0, F(0))\, W_k(t).$$

Efficient drift calculation

- Recall that under $\mathbb{Q}_{T_{k+1}}$-forward measure for $t < \min\{T_k, T_j\}$ we have

$$\Delta_j \quad = \quad C_j(t) \cdot \begin{cases} 0, & \text{if } j = k \\ -\sum_{i=j+1}^{k} \frac{\rho_{ij}\delta_i C_i(t)}{1+\delta_i F_i(t)}, & \text{if } j < k \\ \sum_{i=k+1}^{j} \frac{\rho_{ij}\delta_i C_i(t)}{1+\delta_i F_i(t)}, & \text{if } j > k. \end{cases}$$

- The coefficients $\mathcal{L}^a \Delta_j$ under the forward measure $\mathbb{Q}_k$ are:

$$\mathcal{L}^a \Delta_j = C_j \cdot \begin{cases} 0, & \text{if } j = k \\ -\sum_{i=j+1}^{k} \frac{\rho_{ij}\delta_i C_i(t)}{1+\delta_i F_i(t)} \left(U_{ja} \frac{\partial C_j}{\partial F_j} + U_{ia} \left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i} \right) \right), & \text{if } j < k \\ \sum_{i=k+1}^{j} \frac{\rho_{ij}\delta_i C_i(t)}{1+\delta_i F_i(t)} \left(U_{ja} \frac{\partial C_j}{\partial F_j} + U_{ia} \left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i} \right) \right), & \text{if } j > k. \end{cases}$$

- Under the spot measure:

$$dL_j(t) = C_j(t) \sum_{i=\gamma(t)}^{j} \frac{\rho_{ij}\delta_i C_i(t)}{1+\delta_i F_i(t)} \left(U_{ja} \frac{\partial C_j}{\partial F_j} + U_{ia} \left(\frac{\partial C_i}{\partial F_i} - \frac{\delta_i C_i}{1+\delta_i F_i} \right) \right).$$

Efficient drift calculation

- The order $\frac{3}{4}$ approximation uses the next order term in low noise expansion:

$$\Delta_{j,\frac{3}{4}} = \Delta_{j,0}(t, L_0) + \sum_{k=1}^{d} \Gamma_{jk} W_k(t) + \Omega_j t.$$

- The approximation of order $\frac{3}{4}$ offer excellent accuracy. It is possible to build better approximations by taking more terms in the Taylor expansion, however the benefit does not offset the computational costs.
- It turns out that $\frac{1}{2}$ approximation is the best balance between the accuracy and speed.

References

- Glasserman, P.: *Monte Carlo methods in financial engineering*, Springer (2004).
- Kloeden, P.E.; Platen, E: *Numerical solutions to stochastic differential equation*, Springer (1992).
- Lesniewski, A.: http://lesniewski.us/presentations.html