Numerical integration of SDE's

1 Introduction

A self-contained presentation of numerical methods for stochastic differential equations including MATLAB codes can be found in the review paper [4]. More details can be found in [1]. A good starting reference for *ordinary differential equations* is [3].

2 Integration by Taylor series of ODE's

Problems involving ordinary differential equations (ODE's) can always be reduced to the study of sets of first-order differential equations. For example, the second-order equation

$$\frac{d^2\xi_t}{dt^2} = f(t)\frac{d\xi_t}{dt} + g(t) \tag{2.1}$$

can be turned into the first order system

$$\begin{aligned} \frac{d\xi_t}{dt} &= \zeta_t \\ \frac{d\zeta_t}{dt} &= f(t)\,\zeta_t + g(t) \end{aligned} \tag{2.2}$$

Thus it is sufficient to discuss numerical schemes for the first order system of equations

$$\frac{d\boldsymbol{\xi}_t}{dt} = \boldsymbol{f}\left(\boldsymbol{\xi}_t, t\right) \tag{2.3}$$

in a time interval for $t > t_o$ and with initial condition

$$\boldsymbol{\xi}_{t_o} = \boldsymbol{x}_o \tag{2.4}$$

We will suppose in what follows the vector field $\boldsymbol{f}(\boldsymbol{x}, t)$ smooth (i.e. analytic) in $\mathbb{R}^d \times \mathbb{R}$. The equation (2.3) can be couched into the integral form

$$\boldsymbol{\xi}_t = \boldsymbol{x}_o + \int_{t_o}^{t} ds\, \boldsymbol{f}\left(\boldsymbol{\xi}_s, s\right) \tag{2.5}$$

which we can integrate by iteration using the relation

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{\xi}_t, t) &= \boldsymbol{g}(\boldsymbol{x}_o, t_o) + \int_{t_o}^{t} d\boldsymbol{g}\left(\boldsymbol{\xi}_s, s\right) \\ &= \boldsymbol{g}(\boldsymbol{x}_o, t_o) + \int_{t_o}^{t} ds\left[\boldsymbol{g}\left(\boldsymbol{\xi}_s, s\right) \cdot \left(\partial_{\boldsymbol{\xi}_s}\boldsymbol{g}\right)\left(\boldsymbol{\xi}_s, s\right) + \left(\partial_s \boldsymbol{g}\right)\left(\boldsymbol{\xi}_s, s\right)\right] \end{aligned} \tag{2.6}$$

satisfied by any smooth vector field $\boldsymbol{g}$. For example, iterating twice we get into

$$\begin{aligned}\boldsymbol{\xi}_t &= \boldsymbol{x}_o + (t-t_o)\, f(\boldsymbol{x}_o, t_o) + \int_{t_o}^{t} df(\boldsymbol{\xi}_u, u) \\ &= \boldsymbol{x}_o + (t-t_o)\, f(\boldsymbol{x}_o, t_o) + \frac{(t-t_o)^2}{2}\frac{df}{dt}(\boldsymbol{x}_o, t_o) + \int_{t_o}^{t} du_1 \int_{t_o}^{u_1} du_2 \int_{t_o}^{u_2} du_3 \frac{d^2 f}{du_3^3}\left(\boldsymbol{\xi}_{u_3}, u_3\right) \end{aligned} \tag{2.7}$$

In other words, smoothness guarantees that integrating (2.3) is equivalent to generate the coefficients of the Taylor expansion of $\boldsymbol{f}$ around the point $(\boldsymbol{x}_o\,, t_o)$. This result can be used to construct numerical integration schemes of ODE's.

3 Euler scheme

The simplest integration scheme is the *Euler method*. First we partition the *finite* time interval $T := t - t_o$ into n sub-interval of equal size

$$\delta t := \frac{t-t_o}{n} \tag{3.1}$$

so that

$$t_k = t_o + k\,\delta\, t \quad \& \quad t = t_n = t_o + n\,\delta\, t \tag{3.2}$$

The quantity δt is often referred to as the *mesh* size of the discretization. If δt is "sufficiently" small i.e. n is large enough we can approximate

$$\boldsymbol{x}_{t_{k+1}} \simeq \boldsymbol{x}_{t_k} + \boldsymbol{f}\left(\boldsymbol{x}_{t_k}\,, t_k\right)\, \delta\, t \tag{3.3}$$

The symbol $\simeq$ here means that the left hand side equals the right hand side if we neglect terms of order $O\left(\delta t^2\right)$. In such a case, we can estimate the flow generated by the ordinary differential equation with the one of the discrete map

$$\boldsymbol{y}_{k+1} = \boldsymbol{y}_k + \boldsymbol{f}(\boldsymbol{y}_k\,, t_k)\,\delta t \tag{3.4}$$

In one dimension or for each vector component of $\boldsymbol{x} \in \mathbb{R}^d$, the Euler scheme is the following recursion algorithm:

Algorithm 1 Euler

```
xvar = x_o
tvar = t_o
tfin =t
npartinions = n
mesh = (tfin-tinit)/n
for k = 1, n do
   xvar = xvar + f(xvar,tvar) * mesh
   tvar= tvar+mesh
   print  tvar xvar
end for
```

The statement **print** means that the outcome of the calculation is sent to some output (e.g. stored into a data file). The *local discretization error*

$$l_k = x_{t_k} - y_k \tag{3.5}$$

and the *global discretization error*

$$e_n = x_{t_n} - y_n \tag{3.6}$$

are standard measures of the accuracy of the approximation. A reliable numerical integration should provide an estimate for this errors. This is usually done by checking the convergence of the results to a given value versus the mesh size. In order to examine the dependence of the *global discretization error* upon the mesh it is convenient to consider exponential changes of the mesh size

$$\delta t_i = \frac{t - t_o}{n^i} \qquad i = 1, 2, \ldots \tag{3.7}$$

and define

$$\boldsymbol{y}_{k+1}^{(i)} = \boldsymbol{y}_k^{(i)} + \boldsymbol{f}\left(\boldsymbol{y}_k^{(i)}, t_k\right) \tag{3.8}$$

Then one can study

$$\Delta_i = \ln ||\boldsymbol{y}_{n^{i+1}}^{(i+1)} - \boldsymbol{y}_{n^i}^{(i)}|| \tag{3.9}$$

versus $\ln \delta t_i$. The reason for introducing logarithms is that variations of order of magnitude in the Δ_i are reflected in change of slope in logarithmic scale. An alternative way to proceed to estimate errors, is to compare at fixed mesh the results of the Euler scheme and of an higher order scheme. The *order of a scheme* is defined as follows. A method is said to converge with *order* $\gamma \in \mathbb{N}$ if there exists a constant $K < \infty$ such that the global discretization error satisfies the bound

$$||e_{n+1}|| < K\,(\delta t)^\gamma \qquad \forall\, \delta t \in [0, \delta_\star t] \tag{3.10}$$

The Euler scheme can be proved to have order 1. Intuitively the order of the scheme can be thought as specified by the highest order term of the Taylor series matching the increment of the discrete map defining the approximation scheme in the limit of vanishing mesh size.

3.1 Limitations of the Euler scheme: stiffness

There are several reasons that Eulers method is not recommended for practical use, among them,

1. the method is not very accurate when compared to other at equivalent mesh size.
2. the method is not very stable.

The second pathology arises in the treatment of *"stiff"* systems of differential equations. Dictionary definitions of the word "stiff" refer to concept like "being not easily bent", "rigid" and "stubborn". In the context of ODE's a problem is said to be stiff if [3]

> *A problem is stiff if the solution being sought varies slowly, but there are nearby solutions that vary rapidly, so the numerical method must take small steps to obtain satisfactory results.*

[2] provides the following example.

$$\begin{aligned} \frac{dx_1}{dt} &= a\,x_1 + b\,x_2 \\ \frac{dx_2}{dt} &= -(a+c)\,x_1 - (b+c)\,x_2 \end{aligned} \tag{3.11}$$

with

$$c = O(1) \;>\; 0 \qquad \& \qquad b - a = O(10^3) \;>\; 0 \tag{3.12}$$

Independently of the value of the parameters the system is explicitly integrable. The orthonormal change of variables

$$X := \frac{x_1 + x_2}{\sqrt{2}} \quad \& \quad x = \frac{x_1 - x_2}{\sqrt{2}} \tag{3.13}$$

partially diagonalizes the system

$$\begin{array}{l} \dfrac{d}{dt}(x_1 + x_2) = -c\,(x_1 + x_2) \\ \dfrac{d}{dt}(x_1 - x_2) = (2\,a + c)\,x_1 + (2\,b + c)\,x_2 \end{array} \qquad \Rightarrow \qquad \begin{array}{l} \dfrac{d}{dt}X = -c\,X \\ \dfrac{d}{dt}x = (a + b + c)\,X + (a - b)\,x \end{array} \tag{3.14}$$

Remark: a systematic theory for the analytic integration of linear ODE's with constant coefficients

$$\frac{d\boldsymbol{x}}{dt} = \mathsf{A}\boldsymbol{x} \tag{3.15}$$

proceeds from similarity transformations

$$\boldsymbol{x} = \mathsf{O}\,\boldsymbol{y} \qquad \det \mathsf{O} \neq 0 \tag{3.16}$$

such that

$$\frac{d\boldsymbol{y}}{dt} = \mathsf{O}^{-1}\,\mathsf{A}\,\mathsf{O}\boldsymbol{y} \quad | \quad \mathsf{O}^{-1}\,\mathsf{A}\,\mathsf{O} = \operatorname{diag}\mathsf{A} \tag{3.17}$$

where diag means the full diagonalization of A or at least its reduction to Jordan form. The variables (3.13) reduce the equivalent of the matrix A for the system (3.11) to triangular form so that the system can be readily integrated

$$\begin{aligned} X_t &= X_o\,e^{-c\,t} \\ x_t &= x_o\,e^{(a-b)\,t} + (a + b + c)\,X_o \int_0^t ds\,e^{(a-b)\,(t-s)} e^{-c\,s} \end{aligned} \tag{3.18}$$

Performing the integral gives

$$x_t = x_o\,e^{(a-b)\,t} + \frac{a + b + c}{b - a - c} X_o \left\{ e^{-c\,t} - e^{(a-b)\,t} \right\} \tag{3.19}$$

Going back to the original variables the solution of (3.11) versus initial conditions $(x_{1;o}\,, x_{2;o})$ reads

$$\begin{aligned} x_1 &= \frac{X_t + x_t}{\sqrt{2}} = \frac{b(x_{1;o} + x_{2;o})\,e^{-c\,t} - [(a + c)\,x_{1;o} + b\,x_{2;o}]e^{(a-b)\,t}}{b - a - c} \\ x_2 &= \frac{X_t + x_t}{\sqrt{2}} = \frac{-\,(a + c)\,(x_{1;o} + x_{2;o})\,e^{-c\,t} + [(a + c)\,x_{1;o} + b\,x_{2;o}]e^{(a-b)\,t}}{b - a - c} \end{aligned} \tag{3.20}$$

Using now the hypotheses (3.12) with

$$c = \frac{1}{\tau} \qquad \& \qquad b - a = \frac{1000}{\tau} \tag{3.21}$$

we see that in order to observe the decay of the exponential $e^{(a-b)\,t}$ we need a mesh size

$$\delta t \;\ll\; \frac{\tau}{1000} \tag{3.22}$$

Failing to satisfy (3.22) may compromise the stability of the numerical integration scheme. The reason is that for any matrix A the components of the map

$$\boldsymbol{y}_{k+1} = \mathsf{A}\,\boldsymbol{y}_k \tag{3.23}$$

with solution

$$\boldsymbol{y}_n = \mathsf{A}^n\,\boldsymbol{y}_o \tag{3.24}$$

tends to zero as n tends to infinity if the *largest eigenvalue* of A has magnitude less than unity. The Euler scheme for

$$\frac{d\boldsymbol{x}}{dt} = -\mathsf{C}\boldsymbol{x} \tag{3.25}$$

with C a positive definite matrix corresponds to the map

$$\boldsymbol{y}_{k+1} = (1 - \mathsf{C}\,\delta t)\,\boldsymbol{y}_k \tag{3.26}$$

Denoting by $c_\star$ the largest eigenvalue of C

$$c_\star := \max \mathrm{sp}\mathsf{C} \tag{3.27}$$

the condition for $||\boldsymbol{y}_k||$ to be bounded is therefore that

$$\max \mathrm{sp}\left\{1 - \mathsf{C}\,\delta t\right\} \,<\, 1 \qquad \Rightarrow \qquad \delta t \,<\, \frac{2}{c_\star} \tag{3.28}$$

The example shows the source of the instability of the Euler scheme: the mesh size must be carefully chosen in order to achieve convergence. Other integration schemes maybe, however, less sensitive to the mesh size even in the presence of "stiff" problems. A nice article on stiff systems can be found at
`http://www.scholarpedia.org/article/Stiff_systems`

4 Euler-Maruyama and Milstein scheme

We can adapt the above considerations to stochastic differential equations. Consider the Ito model

$$\begin{aligned} d\boldsymbol{\xi}_t &= \boldsymbol{b}\left(\boldsymbol{\xi}_t, t\right) dt + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_t, t\right)[d\boldsymbol{w}_t] \\ \boldsymbol{\xi}_{t_o} &= \boldsymbol{x}_o \end{aligned} \tag{4.1}$$

with as usual

$$\sigma^i\left(\boldsymbol{\xi}_t, t\right)[d\boldsymbol{w}_t] := \sigma^{ij}\left(\boldsymbol{\xi}_t, t\right) dw_t^j \tag{4.2}$$

4.1 Euler-Maruyama scheme

The simplest integration scheme is the *Euler-Maruyama*

$$\boldsymbol{\xi}_{t_{n+1}} = \boldsymbol{\xi}_{t_n} + \boldsymbol{b}\left(\boldsymbol{\xi}_{t_n}, t_n\right)\,\delta t + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_{t_n}, t_n\right)[\boldsymbol{\eta}_{t_n}]\sqrt{\delta t} \tag{4.3}$$

As in the ODE case implies a uniform partition of the interval $[t_o, t]$ of mesh

$$\delta t = \frac{t - t_n}{N} \tag{4.4}$$

At the n-th time step t_n, the future state of the system $\boldsymbol{\xi}_{t_{n+1}}$ is computed using the present state $\boldsymbol{\xi}_{t_n}$ and by sampling the value of the random variable $\boldsymbol{\eta}_{t_n}$ belonging to a sequence $\left\{\boldsymbol{\eta}_{t_i}\right\}_{i=1}^N$ of *independent identically distributed Gaussian* random variables with zero mean and unit variance.

4.2 Milstein scheme

At variance with the ODE case, the Euler scheme *does not* correspond to a truncation of the Taylor series of the solution of the stochastic differential equation at order dt. Namely, if we iterate (4.1) with the help of Ito's lemma, we get into

$$\begin{aligned}\boldsymbol{\xi}_t = \boldsymbol{x}_o + \boldsymbol{b}\left(\boldsymbol{\xi}_{t_o}, t_o\right)\,(t-t_o) + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_{t_o}, t_o\right)[\boldsymbol{w}_t - \boldsymbol{w}_{t_o}] \\ + \int_{t_o}^{t} ds_1 \int_{t_o}^{s_1} d\boldsymbol{b}\left(\boldsymbol{\xi}_{s_1}, s_2\right) + \int_{t_o}^{t}\int_{t_o}^{s_1} (d\boldsymbol{\sigma})\left(\boldsymbol{\xi}_{s_2}, s_2\right)[dw_{s_1}]\end{aligned} \tag{4.5}$$

where now

$$(d\boldsymbol{\sigma})\left(\boldsymbol{\xi}_s, s\right) = ds\left[\partial_s + \boldsymbol{b}\cdot\partial_{\boldsymbol{\xi}_s} + \frac{\sigma^{ik}\left(\boldsymbol{\xi}_s, s\right)\sigma^{jk}\left(\boldsymbol{\xi}_s, s\right)}{2}\partial_{\xi_s^i}\partial_{\xi_s^i}\right]\boldsymbol{\sigma}\left(\boldsymbol{\xi}_s, s\right) + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_s, s\right)[dw_s]\cdot\partial_{\boldsymbol{\xi}_s}\boldsymbol{\sigma}\left(\boldsymbol{\xi}_s, s\right) \tag{4.6}$$

The stochastic integral on the right hand side of (4.5) brings therefore about a term of the order $O(dt)$. In particular we can single out the $O(dt)$ term

$$\int_{t_o}^{t}\int_{t_o}^{s_1} \boldsymbol{\sigma}\left(\boldsymbol{\xi}_{s_2}, s_2\right)[dw_{s_2}]\cdot\partial_{\boldsymbol{\xi}_{s_2}}\boldsymbol{\sigma}\left(\boldsymbol{\xi}_{s_2}, s_2\right)[dw_{s_1}] \equiv \int_{t_o}^{t}\int_{t_o}^{s_1} \sigma^{jk}\left(\boldsymbol{\xi}_{s_2}, s_2\right) dw_{s_2}^k \partial_{\xi_{s_2}^j}\sigma^{il}\left(\boldsymbol{\xi}_{s_2}, s_2\right) dw_{s_1}^l \tag{4.7}$$

and evaluate it by iteration

$$\begin{aligned}&\int_{t_o}^{t}\int_{t_o}^{s_1} \sigma^{jk}\left(\boldsymbol{\xi}_{s_2}, s_2\right) dw_{s_2}^k \partial_{\xi_{s_2}^j}\sigma^{il}\left(\boldsymbol{\xi}_{s_2}, s_2\right) dw_{s_1}^l = \\ &\qquad \int_{t_o}^{t} dw_{s_1}^l \int_{t_o}^{s_1} dw_{s_2}^k \left\{\sigma^{jk}\left(\boldsymbol{\xi}_{t_o}, t_o\right)\partial_{\xi_{t_o}^j}\sigma^{il}\left(\boldsymbol{\xi}_{t_o}, t_o\right) + \int_{t_o}^{s_2} d[\sigma^{jk}\left(\boldsymbol{\xi}_{s_3}, s_3\right)\partial_{\xi_{s_3}^j}\sigma^{il}\left(\boldsymbol{\xi}_{s_3}, s_3\right)]\right\}\end{aligned} \tag{4.8}$$

Of the two terms on the right hand side, the second is readily order $O(dt^{3/2})$ and it can be neglected for the purpose of deriving the Taylor expansion up to order $O(dt)$. In order to analyze the first term, we can decompose it into index symmetric and antisymmetric parts

$$\int_{t_o}^{t} dw_{s_1}^l \int_{t_o}^{s_1} dw_{s_2}^k = \int_{t_o}^{t}\int_{t_o}^{s_1} \frac{dw_{s_1}^l dw_{s_2}^k + dw_{s_1}^k dw_{s_2}^l}{2} + \int_{t_o}^{t}\int_{t_o}^{s_1} \frac{dw_{s_1}^l dw_{s_2}^k - dw_{s_1}^k dw_{s_2}^l}{2} \tag{4.9}$$

The decomposition is useful because the index symmetric part is integrable

$$\int_{t_o}^{t}\int_{t_o}^{s_1} \frac{dw_{s_1}^l dw_{s_2}^k + dw_{s_1}^k dw_{s_2}^l}{2} = \frac{w_t^l w_t^k - w_{t_o}^l w_{t_o}^k}{2} - \frac{(t-t_o)\delta^{lk}}{2} \tag{4.10}$$

The antisymmetric part is instead vanishing in mean and can be therefore argued to pertain (in mean sense) to order higher than $O(dt)$ in the Taylor expansion of the solution. Gathering the above results, we conclude

$$\begin{aligned}\boldsymbol{\xi}_t = \boldsymbol{x}_o + \boldsymbol{b}\left(\boldsymbol{\xi}_{t_o}, t_o\right)\,(t-t_o) + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_{t_o}, t_o\right)[\boldsymbol{w}_t - \boldsymbol{w}_{t_o}] \\ + \left[\frac{(w_t^l - w_{t_o}^l)(w_t^k - w_{t_o}^k)}{2} - \frac{(t-t_o)\delta^{lk}}{2}\right]\sigma^{jk}\left(\boldsymbol{\xi}_{t_o}, t_o\right)\partial_{\xi_{t_o}^j}\sigma^{il}\left(\boldsymbol{\xi}_{t_o}, t_o\right) + \ldots\end{aligned} \tag{4.11}$$

where $\ldots$ stand for terms that we can neglect in the derivation of the $O(dt)$ integration scheme. After introducing, as above, a partition of the time interval mesh δt we may approximate paths of (4.1) by the map

$$\boldsymbol{\xi}_{t_{n+1}} = \boldsymbol{\xi}_{t_n} + \boldsymbol{b}\left(\boldsymbol{\xi}_{t_n}, t_n\right)\,\delta t + \boldsymbol{\sigma}\left(\boldsymbol{\xi}_{t_n}, t_n\right)[\boldsymbol{\eta}_{t_n}]\sqrt{\delta t} + \frac{\delta t}{2}\left(\eta_{t_n}^l \eta_{t_n}^k - \delta^{lk}\right)\sigma^{jk}\left(\boldsymbol{\xi}_{t_n}, t_n\right)\partial_{\xi_{t_n}^j}\sigma^{il}\left(\boldsymbol{\xi}_{t_n}, t_n\right) \tag{4.12}$$

where again $\boldsymbol{\eta}_{t_n}$ is the n-th element of the sequence $\left\{\boldsymbol{\eta}_{t_i}\right\}_{i=1}^{N}$ of *independent identically distributed Gaussian* random variables with zero mean and unit variance. Such an approximation is usually referred to as Milstein scheme.

References

[1] P.E. Kloeden, E. Platen, H. Schurz,
Numerical solution of SDE through computer experiments
Springer, (1994). 1

[2] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery,
Numerical recipes in C: the art of scientific computing
Cambridge University Press (1992) and `www.nr.com`. 3

[3] C. Moler,
Numerical Computing with MATLAB MathWorks web-book,
`http://www.mathworks.com/moler/index_ncm.html`. 1, 3

[4] D.J. Higham,
An algorithmic introduction to numerical simulation of stochastic differential equations,
SIAM Review, Education Section, 43, 2001, 525-546 and
`http://personal.strath.ac.uk/d.j.higham/papers/algsde.pdf` 1