Long time approximation and modified equations for SDEs

Tony Shardlow
shardlow@maths.man.ac.uk

The University of Manchester

August 2011

Outline

Outline

MANCHESTER 1824

The University of Manchester

What is a modified equation?

Consider an IVP for the ODE

$$\frac{dX}{dt} = f(X), \qquad X(0) = X_0$$

and its approximation by a sequence $X_n \approx X(t_n)$ for $t_n = n\Delta t$ and time step Δt. For example, we may consider X_n to be the explicit Euler approximation given by

$$X_{n+1} = X_n + \Delta t f(X_n).$$

Traditional forward analysis seeks to understand the error by studying

$$e_n := X_n - X(t_n)$$

and we know $\sup_{0 \leq t_n \leq T} |e_n| = \mathcal{O}(\Delta t)$ for explicit Euler.

Backward error analysis

An alternative is backward error analysis, where we ask if there is a model that the approximation X_n actually solves or better approximates. For example, can we find an IVP

$$\frac{dY}{dt} = \tilde{f}(Y), \qquad Y(0) = X_0,$$

for a modified drift $\tilde{f}$, such that

$$Y(t_n) = X_n?$$

Then the backward error is how close $\tilde{f}$ is to f.
The ODE for Y is known as the *modified differential equation*.

Simple example

Consider

$$\frac{dX}{dt} = \lambda X, \qquad X(0) = X_0$$

The explicit Euler method is

$$X_{n+1} = X_n(1 + \lambda \Delta t)$$

and

$$X_n = (1 + \lambda \Delta t)^n X_0$$

Now, assuming $\lambda \Delta t > -1$,

$$\log \frac{X_n}{X_0} = \frac{t_n}{\Delta t} \log(1 + \lambda \Delta t).$$

Then $X_n = Y(t_n)$, where

$$Y(t) := X_0 e^{t\, a(\Delta t)}, \qquad a(\Delta t) := \frac{1}{\Delta t} \log(1 + \lambda \Delta t).$$

Continued.

Then $Y(t_n) = X_n$ and

$$\frac{dY}{dt} = a(\Delta t)Y, \qquad Y(0) = X_0.$$

This is the modified equation for the explicit Euler method and is an ODE that describes the method's behaviour.
Usually, only approximate modified equations are available.
Expand $\log(1+x) = x - x^2/2 + ..$,

$$\begin{aligned}\log X_n/X_0 =& (t_n/\Delta t)(\lambda\Delta t - (\lambda\Delta t)^2/2 + (\lambda\Delta t)^3/3 + \dots) \\ =& \lambda t_n(1 - \lambda\Delta t/2 + \lambda\Delta t^2/3 -).\end{aligned}$$

Then,

$$X_n = e^{t_n(\lambda - \lambda^2\Delta t/2)}X_0 + \mathcal{O}(\Delta t^2)\,.$$

Continued.

Hence, explicit Euler approximation

$$X_n = e^{\lambda t(1-\lambda \Delta t/2)} X_0 + \mathcal{O}(\Delta t^2)$$

The first term is the solution to the IVP

$$\frac{dY}{dt} = \left(\lambda - \frac{\lambda^2 \Delta t}{2}\right) Y$$

and we have $\sup_{0 \le t_n \le T} |Y(t_n) - X_n| = \mathcal{O}(\Delta t^2)$.
This modified better describes what the method is doing than the original equation, but not exactly.
Approximate modified equations are more useful, as they generalise.

MANCHESTER 1824

The University of Manchester

Hamiltonian systems

The big success for backward error analysis is in understanding the approximation of Hamiltonian systems. For Hamiltonian $H(q, p)$, consider

$$\frac{dq}{dt} = H_p, \qquad \frac{dp}{dt} = -H_q$$

Often consider separable Hamiltonian $H(q, p) = \frac{1}{2}p^2 + V(q)$, for potential $V(q) = \frac{1}{2}q^2$. Then,

$$\frac{dq}{dt} = p, \quad \frac{dp}{dt} = -V'(q) = -q$$

Arise in statistical mechanics and often interested in long term calculations. Hamiltonian flow maps are symplectic and preserve areas in phase space (q, p).
Symplectic integrators do the same.

Symplectic methods

The symplectic Euler method is

$$
\begin{aligned}
q_{n+1} =& q_n + p_n \Delta t \\
p_{n+1} =& p_n - V'(q_{n+1}) \Delta t
\end{aligned}
$$

For this type of problem, the modified ODE is also Hamiltonian. For example $V(q) = q^2/2$

$$
q_{n+1} = q_n + p_n \Delta t; \quad p_{n+1} = p_n - q_{n+1} \Delta t
$$

Modified equation to first order is the Hamiltonian system with $H(q, p) - \Delta t pq/2$.
Gives an approximate statistical mechanics.
This construction can be done to arbitrarily high order.

The University of Manchester

MANCHESTER 1824

Outline

Basic problem and notations

Consider an Ito SDE with drift f and diffusion σ

$$dX = f(X)\,dt + \sigma(X)\,dW(t), \qquad X(0) = X_0,$$

where $W(t)$ is a standard Brownian motion.
Error estimates for finite time approximation in weak or strong sense usually diverge as time interval becomes large.

$$\text{error at step } n \leq K e^{K t_n} \Delta t^p$$

for a constant K and rate p.
What can be said as $t_n \to \infty$?

MANCHESTER 1824

The University of Manchester

Example: Euler/OU

Consider the SDE (OU process)

$$dX = -X\,dt + \sqrt{2}dW, \qquad X(0) = X_0.$$

Its invariant measure is $N(0,1)$.
Euler-Maruyama method for $\Delta W_n = W(t_{n+1}) - W(t_n)$,

$$\begin{aligned} X_{n+1} =& X_n(1-\Delta t) + \sqrt{2}\Delta W_n \\ X_n =& (1-\Delta t)^n X_0 + \sum_{i=0}^{n-1}(1-\Delta t)^{n-1-i}\sqrt{2}\Delta W_i. \end{aligned}$$

We know $(1-\Delta t)^n = e^{a(\Delta t)t_n}$ (case $\lambda = -1$) and

$$\sum_{i=0}^{n-1}(1-\Delta t)^{n-1-i}\sqrt{2}\Delta W_i = \sum_{i=0}^{n-1} e^{a(\Delta t)t_{n-1-i}}\sqrt{2}\Delta W_i$$

Mean and variance

Mean zero and variance

$$\sum_{i=0}^{n-1} e^{2a(\Delta t)t_{n-1-i}} 2\Delta t = 2\Delta t \frac{1 - e^{2a(\Delta t)t_n}}{1 - e^{2a(\Delta t)\Delta t}}.$$

then

$$\mathrm{Var}(X_n) \to \frac{2\Delta t}{1 - e^{2a(\Delta t)\Delta t}}, \quad \text{as } n \to \infty.$$

Thus, Euler method has invariant measure
$N(0, 2\Delta t/(1 - e^{2a(\Delta t)\Delta t}))$.
Note $a(\Delta t) = -1 + \mathcal{O}(\Delta t)$, so that

$$\frac{2\Delta t}{1 - e^{2a(\Delta t)\Delta t}} = \frac{2\Delta t}{2\Delta t + \mathcal{O}(\Delta t^2)} = 1 + \mathcal{O}(\Delta t).$$

Invariant measure of method accurate to $\mathcal{O}(\Delta t)$.

Approximation of invariant measure

Talay made a general result about the last observation in a series of papers.
Consider Milstein method in $d = 1$

$$X_{n+1} = X_n + f(X_n)\Delta t + \sigma(X_n)\Delta W_n + \frac{1}{2}\sigma'(X_n)\sigma(X_n)(\Delta W_n^2 - \Delta t).$$

Order one in weak and strong sense.
Let SDE be ergodic with invariant measure π. Talay gives general conditions that imply

$$\mathbf{E}\Big[\phi(X_n)\Big] = \int_{\mathbb{R}} \phi(x)\, d\,\pi(x) + \mathcal{O}(\Delta t)\,.$$

for continuous functions ϕ of polynomial growth.
Also shows true for class of second order methods.
Later, extended his results to non-smooth situation.

Talay: Hamiltonians

Consider

$$
\begin{aligned}
dq =& H_p \, dt \\
dp =& \Big[- H_q - F(q,p) H_p \Big] \, dt + \sigma dW(t)
\end{aligned}
$$

subject to smoothness assumptions and, in particular, need that F be strictly positive.

This gives standard Langevin equation for $H(q,p) = \frac{1}{2}p^2 + V(q)$ with $F(q,p) = 1$.

Talay shows geometric convergence to invariance measure and approximation properties of the implicit Euler method.

Other results

Shardlow-Stuart: geometric ergodicity + finite time approximation gives approximation.

- For process $X(t)$ with $X(0) = X_0$, suppose **Geometric ergodicity**

$$|\mathbf{E}\phi(X(t)) - \pi(\phi)| \leq K_1 \mathcal{V}(X_0) e^{-k_1 t},$$

  for test functions ϕ, invariant measure π, constants $K_1, k_1 > 0$, Lyapunov fn. $\mathcal{V}$.
  Weak convergence for constants $K_2, k_2, p > 0$

$$|\mathbf{E}\phi(X(t_n)) - \mathbf{E}\phi(X_n)| \leq K_2 \mathcal{V}(X_0)\, \Delta t^p\, e^{k_2 t_n},$$

- If this holds for enough ϕ, numerical averages are close to exact averages.
- Applies in many situations (including SPDE), but rate is not optimal.

Outline

Dissipative Particle Dynamics

Mesoscopic model of fluids comprising particles with position momentum $(\mathbf{q}_i, \mathbf{p}_i)$. Forces on Particle i,

- **Pair potential** V soft, short range:

$$-a_{ij} V'(q_{ij})\hat{\mathbf{q}}_{ij}.$$

- **Dissipation** compactly supported $w^D(q)$, parameter λ,

$$-\lambda w^D(q_{ij})(\hat{\mathbf{q}}_{ij} \cdot \mathbf{p}_{ij})\hat{\mathbf{q}}_{ij}.$$

- **Noise** $w^R(q)^2 = w^D(q)$, parameter σ:

$$\sigma w^R(q_{ij})\hat{\mathbf{q}}_{ij} \frac{d\beta_{ij}(t)}{dt},$$

with β_{ij} for $i < j$ independent BMs and $\beta_{ij} = \beta_{ji}$.

Full DPD equations

$\hat{\mathbf{q}}$ unit vector in direction $\mathbf{q}$; q length of $\mathbf{q}$; $\mathbf{q}_{ij} = \mathbf{q}_i - \mathbf{q}_j$.
Repulsion $a_{ij} \geq 0$, Dissipation λ. Noise σ.

$$V(r) = \begin{cases} \frac{1}{2}(1 - \frac{r}{r_c})^2, & r < r_c, \\ 0, & r \geq r_c, \end{cases} \quad w^D(r) = \begin{cases} (1 - \frac{r}{r_c})^2, & r < r_c, \\ 0, & r \geq r_c. \end{cases}$$

$$\begin{aligned} d\mathbf{q}_i =& \mathbf{p}_i \, dt \\ d\mathbf{p}_i =& -\sum_{j \neq i} a_{ij} V'(q_{ij}) \hat{\mathbf{q}}_{ij} \, dt - \lambda \sum_{j \neq i} w^D(q_{ij}) (\hat{\mathbf{q}}_{ij} \cdot \mathbf{p}_{ij}) \hat{\mathbf{q}}_{ij} \, dt \\ & + \sigma \sum_{j \neq i} w^R(q_{ij}) \hat{\mathbf{q}}_{ij} \, d\beta_{ij}(t), \end{aligned}$$

Periodic boundary conditions on $\mathbf{q}$. $\beta_{ij} = \beta_{ji}$.

Properties

1. Momentum/angular momentum is const.
2. Let

$$H(\mathbf{q}_1,\dots,\mathbf{q}_N,\mathbf{p}_1,\dots,\mathbf{p}_N) = \tfrac{1}{2}\sum_i \mathbf{p}_i^2 + \tfrac{1}{2}\sum_{i\neq j} V(q_{ij})$$

If $\sigma^2 = 2\lambda k_B T$, where k_B is Boltzmann's constant and T is temperature, and $w^D = (w^R)^2$, then

$$\rho = \frac{1}{Z}\exp\left[\frac{-H(\mathbf{q}_1,\dots,\mathbf{q}_N,\mathbf{p}_1,\dots,\mathbf{p}_N)}{k_B T}\right]$$

is invariant measure.

Existence and uniqueness of solutions

1. The DPD system is a system of DEs with phase space $\mathbb{R}^{2dN}$ with coefficients that are smooth for distinct particles ($\mathbf{q}_i \neq \mathbf{q}_j$).
   Thus, for initial conditions with distinct particles, the solution exists for a small time interval.
2. If $\mathbf{q}_{ij}(t) = 0$ and $\mathbf{p}_{ij}(t) \neq 0$, the particles immediately separate and solution is again well defined.
3. Difficulties if $\mathbf{q}_{ij}(t) = 0$ and $\mathbf{p}_{ij}(t) = 0$ for some $i \neq j$.

Simple case

Consider

$$d\mathbf{q} = \mathbf{p}\, dt \qquad d\mathbf{p} = \hat{\mathbf{q}}\, d\beta(t).$$

where $\hat{\mathbf{q}} = \mathbf{q}/q$ and $q = \|\mathbf{q}\|$.
Further simplify to the case $d = 1$,

$$dq = p\, dt, \qquad dp = \operatorname{sgn}(q) d\beta(t).$$

This contains many of the technical difficulties of the DPD system. For $q_0 \neq 0$,

$$q(t) = q_0 + \int_0^t p(s)\, ds, \qquad p(t) = \operatorname{sgn}(q_0)\beta(t).$$

is a solution until $q(t) = 0$. Difficulty when $q_0 = p_0 = 0$.

Density

$$dq = p\,dt \qquad dp = d\beta(t),$$

with initial data $q(0) = x$ and $p(0) = y$. (drop sgn(q)). Denote the probability of reaching $dp \times dq$ from (x, y) by $P_t(x, y; p, q)$. Then,

$$\frac{\partial P}{\partial t} = \frac{1}{2}\frac{\partial^2 P}{\partial y^2} + y\frac{\partial P}{\partial x}.$$

$$P_t(x, y; p, q)$$
$$= \frac{\sqrt{3}}{\pi t^2} \exp\left[\frac{-(q-x-yt)^2}{t^3/6} + \frac{(q-x-yt)(p-y)}{t^2/6} - \frac{(p-y)^2}{t/2}\right]$$

$(0,0)$ unreachable

Let

$$G(x,y) = \int_0^\infty P_t(x,y;0,0)\,dt$$

and note $G(x,y) < \infty$ for $(x,y) \neq 0$.
Let $M_t = G(q(t),p(t))$; this is a super martingale as

$$\begin{aligned}\mathbf{E} M_t =& \mathbf{E}\int_0^\infty P_s(q(t),p(t);0,0)\,ds\\ =& \int_t^\infty P_s(x,y;0,0)\,ds\\ \leq& \int_0^\infty P_s(x,y;0,0)\,ds = M_0.\end{aligned}$$

So M_t is finite almost surely (if $(x,y) \neq 0$).
Thus, $(q(t),p(t))$ cannot reach $(q,p) = (0,0)$.

Uniqueness for DPD

- For $d = 1$ DPD, can show $q_{ij} = p_{ij} = 0$ is unreachable a.s. Using change of variables and McKean's argument.
- For 2d, must show that the solution of

$$d\mathbf{q} = \mathbf{p}\,dt, \qquad d\mathbf{p} = \hat{\mathbf{q}}\,d\beta(t)$$

  for non zero initial data does not hit the origin a.s. ?
- For numerical method given later any dimension, bad states are unreachable.

The University of Manchester

MANCHESTER 1824

Proving geometric ergodicity

Consider Markov process $x(t)$ with $x(0) = y$.
Minorization condition For a set C, $\exists T > 0$, measure ν with $\nu(C) > 0$ s.t.

$$P_T(y, A) \geq \nu(A), \qquad y \in C, \quad \text{Borel sets } A,$$

($P_T(y, A)$=prob of reaching set A from y in time T).
Drift condition $\exists$ Lyapunov function $\mathcal{V}$ and $T > 0$ s.t.

$$\mathbf{E}\mathcal{V}(x(T)) - \mathcal{V}(y) \leq -\alpha\mathcal{V}(y) + \beta\mathbf{1}_C(y),$$

where $\mathbf{1}_C$ is indicator function on C, $\alpha, \beta > 0$.
If two conditions hold for $C = \{\mathcal{V} \leq K\}$, then $x(t)$ converges geometrically convergence to a unique invariant measure.

Proving geometric ergodicity

Consider Markov process $x(t)$ with $x(0) = y$.

Minorization condition For a set C, $\exists T > 0$, measure ν with $\nu(C) > 0$ s.t.

$$P_T(y, A) \geq \nu(A), \qquad y \in C, \quad \text{Borel sets } A,$$

($P_T(y, A)$=prob of reaching set A from y in time T).

Drift condition $\exists$ Lyapunov function $\mathcal{V}$ and $T > 0$ s.t.

$$\mathbf{E}\mathcal{V}(x(T)) - \mathcal{V}(y) \leq -\alpha\mathcal{V}(y) + \beta\mathbf{1}_C(y),$$

where $\mathbf{1}_C$ is indicator function on C, $\alpha, \beta > 0$.

If two conditions hold for $C = \{\mathcal{V} \leq K\}$, then $x(t)$ converges geometrically convergence to a unique invariant measure.

Proving geometric ergodicity

Consider Markov process $x(t)$ with $x(0) = y$.

Minorization condition For a set C, $\exists T > 0$, measure ν with $\nu(C) > 0$ s.t.

$$P_T(y, A) \geq \nu(A), \qquad y \in C, \quad \text{Borel sets } A,$$

($P_T(y, A)$=prob of reaching set A from y in time T).

Drift condition $\exists$ Lyapunov function $\mathcal{V}$ and $T > 0$ s.t.

$$\mathbf{E}\mathcal{V}(x(T)) - \mathcal{V}(y) \leq -\alpha\mathcal{V}(y) + \beta\mathbf{1}_C(y),$$

where $\mathbf{1}_C$ is indicator function on C, $\alpha, \beta > 0$.

If two conditions hold for $C = \{\mathcal{V} \leq K\}$, then $x(t)$ converges geometrically convergence to a unique invariant measure.

Hypoelliptic case

Consider

$$dq = p\,dt, \quad dp = (-\lambda p - V'(q))\,dt + \sigma\,d\beta(t).$$

Minorization condition

- Existence of a continuous density. Hormander Theorem.
- Show that $P_T(y, A) > 0$. Control Theory. Noise non-degenerate in p equation.

Drift condition

- Ito's formula on $\mathcal{V}(q,p) = \frac{1}{2}p^2 + V(q)$. Use uniform dissipation $-\lambda p$. Infact, in this case prove,

$$\frac{d}{dt}\mathbf{E}\mathcal{V}(x(t)) \le -\alpha\mathcal{V}(x(t)) + \beta\mathbf{1}_C(x(t)).$$

The University of Manchester

MANCHESTER 1824

Hypoelliptic case

Consider

$$dq = p\,dt, \quad dp = (-\lambda p - V'(q))\,dt + \sigma\, d\beta(t).$$

Minorization condition

- Existence of a continuous density. Hormander Theorem.
- Show that $P_T(y, A) > 0$. Control Theory. Noise non-degenerate in p equation.

Drift condition

- Ito's formula on $\mathcal{V}(q,p) = \frac{1}{2}p^2 + V(q)$. Use uniform dissipation $-\lambda p$. Infact, in this case prove,

$$\frac{d}{dt}\mathbf{E}\mathcal{V}(x(t)) \leq -\alpha\mathcal{V}(x(t)) + \beta\mathbf{1}_C(x(t)).$$

MANCHESTER 1824

The University of Manchester

Minorization condition

- Hormander Thm requires C^∞ smooth coefficients.
- Not hypoelliptic for all y as noise switches off if $q_{ij} > r_c$. This makes control methods hard.

Drift condition

- Dissipation switches off if $q_{ij} > r_c$. In fact, can show

$$\frac{d}{dt}\mathbf{E}\mathcal{V}(x(t)) = \tfrac{1}{2}\sum_{i\neq j} w^D(q_{ij})(\sigma^2 - \lambda \mathbf{p}_{ij}\cdot\mathbf{p}_{ij})$$

where $\mathcal{V}(x) = 1 + \frac{1}{2}\sum_i \mathbf{p}_i\cdot\mathbf{p}_i + \sum_{i\neq j} a_{ij}V(q_{ij})$.

Difficulties for DPD

Minorization condition

- Hormander Thm requires C^∞ smooth coefficients.
- Not hypoelliptic for all y as noise switches off if $q_{ij} > r_c$. This makes control methods hard.

Drift condition

- Dissipation switches off if $q_{ij} > r_c$. In fact, can show

$$\frac{d}{dt}\mathbf{E}\mathcal{V}(x(t)) = \tfrac{1}{2}\sum_{i\neq j} w^D(q_{ij})(\sigma^2 - \lambda \mathbf{p}_{ij}\cdot\mathbf{p}_{ij})$$

where $\mathcal{V}(x) = 1 + \frac{1}{2}\sum_i \mathbf{p}_i\cdot\mathbf{p}_i + \sum_{i\neq j} a_{ij} V(q_{ij})$.

Plots of energy

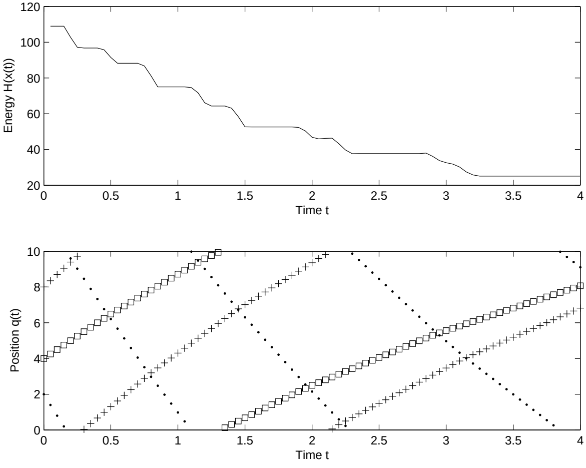

Figure: Plots of $\mathcal{V}(x(t))$ and $q_i(t)$ for initial data $(2, 4, 8, -12, 5, 7)$ on spatial domain $[0, 10]$ with $\lambda = \sigma = 1$, $a_{ij} = 0$, and $r_c = 1$.

Working with DPD

Drift condition

- Prove that collisions of two particles gives rise to geometric decay on a time interval length dt.
- Prove that collisions of particles happen sufficiently often.

Minorization condition

- Coefficients are smooth except at bad points $\mathbf{q}_i = \mathbf{q}_j$. Apply Hormander on a non-characteristic subdomain with absorbing BCs (Cattiaux, 1991), to get lower bound on transition density.
- Assume domain length $L < Nr_c$, then at least one pair will interact and can start a control argument.

Working with DPD

Drift condition

- Prove that collisions of two particles gives rise to geometric decay on a time interval length dt.
- Prove that collisions of particles happen sufficiently often.

Minorization condition

- Coefficients are smooth except at bad points $\mathbf{q}_i = \mathbf{q}_j$. Apply Hormander on a non-characteristic subdomain with absorbing BCs (Cattiaux, 1991), to get lower bound on transition density.
- Assume domain length $L < Nr_c$, then at least one pair will interact and can start a control argument.

MANCHESTER 1824

The University of Manchester

Theorem

For $d = 1$, suppose that $L < Nr_c$ and $\sigma, \lambda > 0$. There exists a probability measure π on a state space

$$S = \Big\{ (q_1, \dots, q_N, p_1, \dots, p_N) \in \mathcal{T}^N \times R^N \colon \frac{1}{N}\sum_{i=1}^{N} q_i = \bar{q},$$

$$\sum_{i=1}^{N} p_i = 0, \quad |p_{ij}| + |q_{ij}| > 0 \text{ if } i \neq j \Big\}$$

such that

$$|\mathbf{E}\phi(x(t)) - \pi(\phi)| \leq K(1 + \sum p_i^2)e^{-kt}$$

for all measurable $\phi \colon S \to R$ with $|\phi| \leq \mathcal{V}$.

In two/three dimensions?

Many difficulties:

- Pathwise uniqueness.
- Geometry harder, as particles may move parallel to one another and not collide.
- A single pair interaction gives noise in one dimension. So need at least two pair interactions for noise to span the two dimensions of $\mathbf{p}$.

Outline

Example: OU

For λ, σ constants, consider

$$dX = \lambda X\, dt + \sigma\, dW, \qquad X(0) = X_0.$$

Explicit Euler method:

$$\begin{aligned} X_{n+1} =& (1+\lambda\Delta t)X_n + \sigma \Delta W_n \\ X_n =& (1+\lambda\Delta t)^n X_0 + \sigma \sum_{i=0}^{n-1} (1+\lambda\Delta t)^{n-1-i} \Delta W_i \\ =& e^{a(\Delta t)t_n} X_0 + \sigma \sum_{i=0}^{n-1} e^{a(\Delta t)t_{n-1-i}} \Delta W_i. \end{aligned}$$

Is this the solution of some SDE?

$$dY = f_{\Delta t}(Y)\, dt + \sigma_{\Delta t}(Y) dW(t), \qquad Y(0) = X_0.$$

Modified SDE for OU

Guess the correct form is

$$dY = a(\Delta t) Y \, dt + b(\Delta t) \, dW(t), \qquad Y(0) = X_0.$$

where $a(\Delta t)$ is already derived for deterministic case. Then,

$$\begin{aligned} Y(t_n) =& e^{a(\Delta t) t_n} X_0 + b(\Delta t) \int_0^{t_n} e^{a(\Delta t)(t_n - s)} \, dW(s) \\ =& e^{a(\Delta t) t_n} X_0 + b(\Delta t) \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} e^{a(\Delta t)(t_n - s)} \, dW(s). \end{aligned}$$

As Y, X_n are Gaussian, to get distribution of $Y(t_n)$ correct, need to get mean and variance correct.
Mean is correct, as $a(\Delta t)$ is correct and Ito integral mean zero.

Correct variance

$$\mathrm{Var}\left(b(\Delta t)\int_{t_i}^{t_{i+1}} e^{a(\Delta t)(t_n-s)}\,dW(s)\right)$$
$$=b(\Delta t)^2\int_{t_i}^{t_{i+1}} e^{2a(\Delta t)(t_n-s)}\,ds$$

equal to

$$\mathrm{Var}(\sigma e^{a(\Delta t)t_{n-1-i}}\Delta W_i)=\sigma^2 e^{2a(\Delta t)t_{n-1-i}}\Delta t.$$

So $b(\Delta t)$ given by solution of

$$b(\Delta t)^2\int_{t_i}^{t_{i+1}} e^{2a(\Delta t)(t_n-s)}\,ds=\sigma^2 e^{2a(\Delta t)t_{n-1-i}}\Delta t.$$

$$b(\Delta t)^2\int_{0}^{\Delta t} e^{2a(\Delta t)(\Delta t-s)}\,ds=\sigma^2\Delta t.$$

Can find modified equation in weak sense for OU/Explicit Euler.

Modified equations for SDEs

Consider an Ito SDE with drift f and diffusion σ

$$dX = f(X)\,dt + \sigma(X)\,dW(t), \qquad X(0) = X_0,$$

and look for a **modified SDE** that best fits the numerics

$$\begin{aligned} dY =& \left[f(Y) + \Delta t\,\tilde{f}(Y)\right] dt + \left[\sigma(Y) + \Delta t\,\tilde{\sigma}(Y)\right] dW(t), \\ Y(0) =& X_0. \end{aligned}$$

To determine the modified term, we ask for an increase in the order of the error in the weak sense. That is, find $\tilde{f}(Y)$ and $\tilde{\sigma}(Y)$ such that

$$\mathbf{E}\phi(X_n) - \mathbf{E}\phi(Y(t_n)) = \mathcal{O}(\Delta t^p)\,,$$

where p improves on the weak order of the method X_n (e.g., $p = 2$ for Explicit Euler.).

MANCHESTER 1824

The University of Manchester

Weak Convergence

- For explicit Euler in one dimension,

$$X_{n+1} = X_n + f(X_n)\Delta t + \sigma(X_n)\Delta W_n,$$

  where ΔW_n are independent $N(0, \Delta t)$ (Gaussian mean 0, variance Δt).
- Basic idea: show consistency conditions on moments.
- To gain convergence of order p, require approximation to $\mathcal{O}(\Delta t^{p+1})$ of $\mathbf{E}\phi(X(\Delta t))$, for polynomials ϕ upto degree $2p+1$.

Weak consistency conditions

- To gain a modified equation of second order, must satisfy **five conditions** but only have **two free parameters** $\tilde{f}$ and $\tilde{\sigma}$.
- Solution not guaranteed for any method!
- Long calculation: compute moments $\mathbf{E} X_n^p$ and $\mathbf{E} X(t_n)^p$ for $p = 1, \ldots, 5$ and see if there is a solution.
- If the noise is multiplicative, the consistency equations have no solution. There is no modified equation.

MANCHESTER 1824

The University of Manchester

Modified equation for explicit Euler

If we look at the additive case, all five conditions hold for

$$dY = \left[f(Y) - \Delta t\left(\frac{1}{2}f'(Y)f(Y) + \frac{1}{4}f''(Y)\sigma^2\right)\right]\Delta t + \sigma\left(1 - \Delta t f'(Y)/2\right)dW(t),$$

For example, if $f(Y) = \lambda Y$ and $\sigma(Y) = \sigma$,

$$dY = \left(\lambda - \Delta t\frac{1}{2}\lambda^2\right)Y\,dt + \sigma\left(1 - \Delta t\frac{\lambda}{2}\right)dW(t).$$

is weak second order close to the explicit Euler method.
In general, cannot go to even higher order as the $f'(Y)$ term causes the modified equation to have multiplicative noise.

Extensions: explicit Euler

- In $\mathbf{R}^2$, there are now **twenty** moments conditions. It **is possible** to find a second order modified SDE for explicit Euler in this case.
- ∞-modified equation for Gaussian cases (like OU and explicit Euler). Already derived.
- Zygalakis, developed ∞ modified equation and introduces a technique based on the generator, which simplifies calculations.

MANCHESTER 1824

The University of Manchester

Milstein's method

Zygalakis has shown there is a second order modified equation for Milstein's method. It is

$$\begin{aligned} dY = &\left[f(Y) - \frac{\Delta t}{2}\left(f(Y)f'(Y) + \frac{1}{2}\sigma(Y)^2 f''(Y)\right)\right] dt \\ &+ \left[\sigma(Y) - \frac{\Delta t}{2}\left(\sigma(Y)f'(Y)\right.\right. \\ &+ f(Y)\sigma'(Y) + \frac{\sigma(Y)}{2}\sigma'(Y)^2 + \frac{\sigma(Y)^2\sigma''(Y)}{2} \\ &\left.\left.- \frac{1}{2}\sigma'(Y)\right)\right] dW(t). \end{aligned}$$

There also a version in multiple dimensions. For GBM,

$$\begin{aligned} dY = &\left[r - \frac{r^2\Delta t}{2}\right] Y\,dt + \left[\sigma - \Delta t\left(r\sigma + \frac{\sigma^3}{4}\right)\right] Y\,dW(t) \\ &+ \frac{\Delta t}{4}\sigma\,dW(t). \end{aligned}$$

Notice the additive term.

Kurtz asymptotic formula for Euler

Consider

$$\begin{aligned} dZ =& f'(X)Z\,dt + \sigma'(X)ZdW(t) \\ &+ \frac{1}{\sqrt{2}}(\sigma'(X)\sigma(X) + f'(X)\sigma(X))d\beta(t), \end{aligned}$$

where $\beta(t), W(t)$ are independent Brownian motions and $Z(0) = 0$.
Then $X_n = X(t_n) + \Delta t^{1/2} Z(t_n) + o(\Delta t^{1/2})$. For OU, the equation for $Z(t)$ is

$$dZ = \lambda Z\,dt + \frac{1}{\sqrt{2}}\sigma\lambda\,d\beta(t).$$

GBM

Geometric Brownian Motion SDE

$$dX = rX\,dt + \sigma X\,dW(t), \qquad X(0) = X_0,$$

Then if X_n is the Euler approximation,
$X_n = X(t_n) + \Delta t^{1/2} Z(t_n) + o(\Delta t^{1/2})$, where

$$dZ = rZ\,dt + \sigma Z\,dW(t) + \frac{1}{\sqrt{2}}(\sigma^2 + \sigma r)X\,d\beta(t).$$

Is this really backward error analysis?

The behaviour of Euler is described in terms of two SDEs

It is not obvious from the form of $Z(t)$ why weak error is order one.

Prefer an SDE of the same type so can compare drift and diffusion.

MANCHESTER 1824

The University of Manchester

Outline

Back to long time approximation

Can we use the modified equation to understand the long time approximation of SDEs?
For Ornstein-Uhlenbeck, the modified SDE

$$dY = a(\Delta t) Y \, dt + b(\Delta t) Y \, dW(t)$$

are good for all time; i.e.,

- correct distribution at time t_n and
- correct invariant measure.

For multiplicative noise problems, we cannot even find a modified equation.

MANCHESTER 1824

The University of Manchester

Milstein method

- Milstein method has order two modified equation
- Talay's work gives error analysis of invariant measure for SDE approximation by Milstein's method.

Expect that invariant measure of Milstein's method is equal invariant measure of modified equation to second order.
if Talay's argument allows a dependence on Δt in drift and diffusion, could be made in to a general theorem.

Langevin equation

- Consider now the following SDE for position q, momentum p, dissipation λ, diffusion σ, and potential V:

$$dq = p\,dt, \quad dp = (-\lambda p - V'(q))\,dt + \sigma dW(t)$$

and the following generalisation of **symplectic Euler**

$$\begin{aligned} q_{n+1} =& q_n + p_{n+1}\Delta t, \\ p_{n+1} =& p_n - (\lambda p_n + V'(q_n))\Delta t + \sigma \Delta W_n. \end{aligned}$$

Modified SDE for Langevin?

Modified equation

$$
\begin{aligned}
dq &= \left(\tilde{H}_p(q,p) - \frac{\lambda \Delta t p}{2}\right) dt + \frac{\sigma \Delta t}{2} dW(t) \\
dp &= -\tilde{H}_q(q,p)\, dt - \lambda\left(1 + \frac{1}{2}\Delta t \lambda + \frac{1}{2}\Delta t V'(q)\right) p\, dt \\
&\quad + \sigma\left(1 + \frac{\lambda \Delta t}{2}\right) dW(t),
\end{aligned}
$$

where the modified Hamiltonian

$$
\tilde{H}(q,p) = \frac{1}{2}p^2 + V(q) - \frac{1}{2}\Delta t V'(q) p.
$$

Structure of modified equation

The invariant measure of

$$\begin{aligned} dp =& H_p\, dt \\ dq =& -(H_q + \lambda H_p)\, dt + \sigma dW(t) \end{aligned}$$

is $e^{-2\lambda H(q,p)/\sigma^2}$. For $H = \frac{1}{2}p^2 + V(q)$, this system is

$$\begin{aligned} dp =& p\, dt \\ dq =& -(V'(q) + \lambda p)\, dt + \sigma dW(t) \end{aligned}$$

the Langevin equation. We'd like the modified SDE to have this structure, just as the symplectic method had a modified equation with a Hamiltonian structure.
It does not.
Zygalakis has developed a method, whose 1st order modified equation is.

MANCHESTER 1824

The University of Manchester

Outline

Conclusions

1. Long time error analysis for SDEs well developed, especially Talay and coworkers.
2. Euler additive SDE, can write down modified equation to high order.
3. Milstein has second order modified equation.
4. For Langevin equation, symplectic Euler has modified equation but not appropriate structure. Zygalakis gives a method where Langevin structure is found in modified equation.
5. Is there a pathwise modified equation? Certainly need to step outside SDEs of the same type to do this.
6. Perhaps, use rough path space analysis to consider SDEs forced by some appropriate rough path?
7. Relate modifed equation to asymptotic analysis.