Lecture 4: Numerical Solution of SDEs, Itô–Taylor Series, Gaussian Approximations

Simo Särkkä

Aalto University, Finland

November 18, 2014

Contents

Overview of Numerical Methods

- Gaussian approximations:
  - Approximations of mean and covariance equations.
  - Gaussian assumed density approximations.
  - Statistical linearization.
- Numerical simulation of SDEs:
  - Itô–Taylor series.
  - Euler–Maruyama method and Milstein's method.
  - Stochastic Runge–Kutta (next week).
- Other methods (not covered in lectures):
  - Approximations of higher order moments.
  - Approximations of Fokker–Planck–Kolmogorov PDE.

Theoretical mean and covariance equations

- Consider the stochastic differential equation (SDE)

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\, \mathrm{d}t + \mathbf{L}(\mathbf{x}, t)\, \mathrm{d}\beta.$$

- The mean and covariance differential equations are

$$\begin{aligned} \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} &= \mathrm{E}\left[\mathbf{f}(\mathbf{x}, t)\right] \\ \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \mathrm{E}\left[\mathbf{f}(\mathbf{x}, t)\,(\mathbf{x} - \mathbf{m})^{\mathsf{T}}\right] + \mathrm{E}\left[(\mathbf{x} - \mathbf{m})\,\mathbf{f}^{\mathsf{T}}(\mathbf{x}, t)\right] \\ &\quad + \mathrm{E}\left[\mathbf{L}(\mathbf{x}, t)\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}(\mathbf{x}, t)\right] \end{aligned}$$

- Note that the expectations are w.r.t. $p(\mathbf{x}, t)$!

Gaussian approximations [1/5]

- The mean and covariance equations explicitly:

$$\frac{d\mathbf{m}}{dt} = \int \mathbf{f}(\mathbf{x}, t)\, p(\mathbf{x}, t)\; d\mathbf{x}$$

$$\frac{d\mathbf{P}}{dt} = \int \mathbf{f}(\mathbf{x}, t)\,(\mathbf{x} - \mathbf{m})^\mathsf{T}\, p(\mathbf{x}, t)\; d\mathbf{x} + \int (\mathbf{x} - \mathbf{m})\, \mathbf{f}^\mathsf{T}(\mathbf{x}, t)\, p(\mathbf{x}, t)\; d\mathbf{x}$$
$$+ \int \mathbf{L}(\mathbf{x}, t)\, \mathbf{Q}\, \mathbf{L}^\mathsf{T}(\mathbf{x}, t)\, p(\mathbf{x}, t)\; d\mathbf{x}.$$

- In Gaussian assumed density approximation we assume

$$p(\mathbf{x}, t) \approx \mathrm{N}(\mathbf{x} \mid \mathbf{m}(t), \mathbf{P}(t)).$$

Gaussian approximations [2/5]

Gaussian approximation I

Gaussian approximation to SDE can be obtained by integrating the following differential equations from the initial conditions $\mathbf{m}(0) = \mathrm{E}[\mathbf{x}(0)]$ and $\mathbf{P}(0) = \mathrm{Cov}[\mathbf{x}(0)]$ to the target time t:

$$
\begin{aligned}
\frac{d\mathbf{m}}{dt} &= \int \mathbf{f}(\mathbf{x}, t) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, d\mathbf{x} \\
\frac{d\mathbf{P}}{dt} &= \int \mathbf{f}(\mathbf{x}, t) \, (\mathbf{x} - \mathbf{m})^\mathsf{T} \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, d\mathbf{x} \\
&\quad + \int (\mathbf{x} - \mathbf{m}) \, \mathbf{f}^\mathsf{T}(\mathbf{x}, t) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, d\mathbf{x} \\
&\quad + \int \mathbf{L}(\mathbf{x}, t) \, \mathbf{Q} \, \mathbf{L}^\mathsf{T}(\mathbf{x}, t) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, d\mathbf{x}.
\end{aligned}
$$

Gaussian approximations [3/5]

Gaussian approximation I (cont.)

If we denote the Gaussian expectation as

$$\mathrm{E}_{\mathrm{N}}[\mathbf{g}(\mathbf{x})] = \int \mathbf{g}(\mathbf{x}) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x}$$

the mean and covariance equations can be written as

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} &= \mathrm{E}_{\mathrm{N}}[\mathbf{f}(\mathbf{x}, t)] \\
\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \mathrm{E}_{\mathrm{N}}[(\mathbf{x} - \mathbf{m}) \, \mathbf{f}^{\mathsf{T}}(\mathbf{x}, t)] + \mathrm{E}_{\mathrm{N}}[\mathbf{f}(\mathbf{x}, t) \, (\mathbf{x} - \mathbf{m})^{\mathsf{T}}] \\
&\quad + \mathrm{E}_{\mathrm{N}}[\mathbf{L}(\mathbf{x}, t) \, \mathbf{Q} \, \mathbf{L}^{\mathsf{T}}(\mathbf{x}, t)].
\end{aligned}$$

Gaussian approximations [4/5]

Theorem

Let $\mathbf{f}(\mathbf{x}, t)$ *be differentiable with respect to* $\mathbf{x}$ *and let* $\mathbf{x} \sim \mathrm{N}(\mathbf{m}, \mathbf{P})$. *Then the following identity holds:*

$$\int \mathbf{f}(\mathbf{x}, t)\,(\mathbf{x} - \mathbf{m})^{\mathsf{T}}\,\mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P})\,\mathrm{d}\mathbf{x} = \left[\int \mathbf{F}_x(\mathbf{x}, t)\,\mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P})\,\mathrm{d}\mathbf{x}\right] \mathbf{P},$$

where $\mathbf{F}_x(\mathbf{x}, t)$ *is the Jacobian matrix of* $\mathbf{f}(\mathbf{x}, t)$ *with respect to* $\mathbf{x}$.

Gaussian approximation II

Gaussian approximation to SDE can be obtained by integrating the following differential equations from the initial conditions $\mathbf{m}(0) = \mathrm{E}[\mathbf{x}(0)]$ and $\mathbf{P}(0) = \mathrm{Cov}[\mathbf{x}(0)]$ to the target time t:

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} &= \mathrm{E}_\mathrm{N}[\mathbf{f}(\mathbf{x}, t)] \\
\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \mathbf{P}\,\mathrm{E}_\mathrm{N}[\mathbf{F}_x(\mathbf{x}, t)]^\mathsf{T} + \mathrm{E}_\mathrm{N}[\mathbf{F}_x(\mathbf{x}, t)]\,\mathbf{P} + \mathrm{E}_\mathrm{N}[\mathbf{L}(\mathbf{x}, t)\,\mathbf{Q}\,\mathbf{L}^\mathsf{T}(\mathbf{x}, t)],
\end{aligned}$$

where $\mathrm{E}_\mathrm{N}[\cdot]$ denotes the expectation with respect to $\mathbf{x} \sim \mathrm{N}(\mathbf{m}, \mathbf{P})$.

Classical Linearization [1/2]

- We need to compute following kind of Gaussian integrals:

$$\mathrm{E}_{\mathrm{N}}[\mathbf{g}(\mathbf{x}, t)] = \int \mathbf{g}(\mathbf{x}, t) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x}$$

- We can borrow methods from filtering theory.
- Linearize the drift $\mathbf{f}(\mathbf{x}, t)$ around the mean $\mathbf{m}$ as follows:

$$\mathbf{f}(\mathbf{x}, t) \approx \mathbf{f}(\mathbf{m}, t) + \mathbf{F}_x(\mathbf{m}, t) \, (\mathbf{x} - \mathbf{m}),$$

- Approximate the expectation of the diffusion part as

$$\mathbf{L}(\mathbf{x}, t) \approx \mathbf{L}(\mathbf{m}, t).$$

Classical Linearization [2/2]

Linearization approximation of SDE

Linearization based approximation to SDE can be obtained by integrating the following differential equations from the initial conditions $\mathbf{m}(0) = \mathrm{E}[\mathbf{x}(0)]$ and $\mathbf{P}(0) = \mathrm{Cov}[\mathbf{x}(0)]$ to the target time t:

$$\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} = \mathbf{f}(\mathbf{m}, t)$$

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{P}\,\mathbf{F}_x^{\mathsf{T}}(\mathbf{m}, t) + \mathbf{F}_x(\mathbf{m}, t)\,\mathbf{P} + \mathbf{L}(\mathbf{m}, t)\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}(\mathbf{m}, t).$$

- Used in extended Kalman filter (EKF).

Cubature integration [1/3]

- Gauss–Hermite cubatures:

$$\int \mathbf{f}(\mathbf{x}, t) \, \mathrm{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x} \approx \sum_i W^{(i)} \, \mathbf{f}(\mathbf{x}^{(i)}, t).$$

- The sigma points (abscissas) $\mathbf{x}^{(i)}$ and weights $W^{(i)}$ are determined by the integration rule.
- In multidimensional Gauss-Hermite integration, unscented transform, and cubature integration we select:

$$\mathbf{x}^{(i)} = \mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i.$$

- The matrix square root is defined by $\mathbf{P} = \sqrt{\mathbf{P}} \, \sqrt{\mathbf{P}}^\mathsf{T}$ (typically Cholesky factorization).
- The vectors ξ_i are determined by the integration rule.

Cubature integration [2/3]

- In Gauss–Hermite integration the vectors and weights are selected as cartesian products of 1d Gauss–Hermite integration.
- Unscented transform uses:

$$\xi_0 = 0$$
$$\xi_i = \begin{cases} \sqrt{\lambda + n}\, e_i & , \quad i = 1, \ldots, n \\ -\sqrt{\lambda + n}\, e_{i-n} & , \quad i = n+1, \ldots, 2n, \end{cases}$$

and $W^{(0)} = \lambda/(n + \kappa)$, and $W^{(i)} = 1/[2(n + \kappa)]$ for $i = 1, \ldots, 2n$.

- Cubature method (spherical 3rd degree):

$$\xi_i = \begin{cases} \sqrt{n}\, e_i & , \quad i = 1, \ldots, n \\ -\sqrt{n}\, e_{i-n} & , \quad i = n+1, \ldots, 2n, \end{cases}$$

and $W^{(i)} = 1/(2n)$ for $i = 1, \ldots, 2n$.

Cubature integration [3/3]

Sigma-point approximation of SDE

Sigma-point based approximation to SDE:

$$
\begin{aligned}
\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} &= \sum_i W^{(i)} \, \mathbf{f}(\mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i, t) \\
\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \sum_i W^{(i)} \, \mathbf{f}(\mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i, t) \, \xi_i^\mathsf{T} \sqrt{\mathbf{P}}^\mathsf{T} \\
&\quad + \sum_i W^{(i)} \sqrt{\mathbf{P}} \, \xi_i \, \mathbf{f}^\mathsf{T}(\mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i, t) \\
&\quad + \sum_i W^{(i)} \, \mathbf{L}(\mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i, t) \, \mathbf{Q} \, \mathbf{L}^\mathsf{T}(\mathbf{m} + \sqrt{\mathbf{P}} \, \xi_i, t).
\end{aligned}
$$

- Use in (continuous-time) unscented Kalman filter (UKF) and (continuous-time) cubature-based Kalman filters (GHKF, CKF, etc.).

Taylor series of ODEs vs. Itô-Taylor series of SDEs

- Taylor series expansions (in time direction) are classical methods for approximating solutions of deterministic ordinary differential equations (ODEs).
- Largely superseded by Runge–Kutta type of derivative free methods (whose theory is based on Taylor series).
- Itô-Taylor series can be used for approximating solutions of SDEs—direct generalization of Taylor series for ODEs.
- Stochastic Runge–Kutta methods are not as easy to use as their deterministic counterparts
- It is easier to understand Itô-Taylor series by understanding Taylor series (for ODEs) first.

Taylor series of ODEs [1/5]

- Consider the following ordinary differential equation (ODE):

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}(t), t), \quad \mathbf{x}(t_0) = \text{given},$$

- Integrating both sides gives

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau.$$

- If the function **f** is differentiable, we can also write $t \mapsto \mathbf{f}(\mathbf{x}(t), t)$ as the solution to the differential equation

$$\frac{\mathrm{d}\mathbf{f}(\mathbf{x}(t), t)}{\mathrm{d}t} = \frac{\partial \mathbf{f}}{\partial t}(\mathbf{x}(t), t) + \sum_i f_i(\mathbf{x}(t), t) \, \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}(t), t).$$

Taylor series of ODEs [2/5]

- The integral form of this is

$$\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\mathbf{x}(t_0), t_0) + \int_{t_0}^{t} \left[\frac{\partial \mathbf{f}}{\partial t}(\mathbf{x}(\tau), \tau) + \sum_i f_i(\mathbf{x}(\tau), \tau) \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}(\tau), \tau) \right]$$

- Let's define the linear operator

$$\mathcal{L}\mathbf{g} = \frac{\partial \mathbf{g}}{\partial t} + \sum_i f_i \frac{\partial \mathbf{g}}{\partial x_i}$$

- We can now rewrite the integral equation as

$$\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\mathbf{x}(t_0), t_0) + \int_{t_0}^{t} \mathcal{L}\, \mathbf{f}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\tau.$$

Taylor series of ODEs [3/5]

- By substituting this into the original integrated ODE gives

$$
\begin{aligned}
\mathbf{x}(t) &= \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau \\
&= \mathbf{x}(t_0) + \int_{t_0}^{t} [\mathbf{f}(\mathbf{x}(t_0), t_0) + \int_{t_0}^{\tau} \mathcal{L}\, \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau] \, \mathrm{d}\tau \\
&= \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0) + \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}\, \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau \, \mathrm{d}\tau.
\end{aligned}
$$

- The term $\mathcal{L}\, \mathbf{f}(\mathbf{x}(t), t)$ solves the differential equation

$$
\begin{aligned}
\frac{\mathrm{d}[\mathcal{L}\, \mathbf{f}(\mathbf{x}(t), t)]}{\mathrm{d}t} &= \frac{\partial[\mathcal{L}\, \mathbf{f}(\mathbf{x}(t), t)]}{\partial t} + \sum_i f_i(\mathbf{x}(t), t) \, \frac{\partial[\mathcal{L}\, \mathbf{f}(\mathbf{x}(t), t)]}{\partial x_i} \\
&= \mathcal{L}^2\, \mathbf{f}(\mathbf{x}(t), t).
\end{aligned}
$$

Taylor series of ODEs [4/5]

- In integral form this is

$$\mathcal{L}\,\mathbf{f}(\mathbf{x}(t),t) = \mathcal{L}\,\mathbf{f}(\mathbf{x}(t_0),t_0) + \int_{t_0}^{t} \mathcal{L}^2\,\mathbf{f}(\mathbf{x}(\tau),\tau)\,\mathrm{d}\tau.$$

- Substituting into the equation of $\mathbf{x}(t)$ then gives

$$\begin{aligned}
\mathbf{x}(t) &= \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0),t)\,(t-t_0) \\
&\quad + \int_{t_0}^{t}\int_{t_0}^{\tau} [\mathcal{L}\,\mathbf{f}(\mathbf{x}(t_0),t_0) + \int_{t_0}^{\tau} \mathcal{L}^2\,\mathbf{f}(\mathbf{x}(\tau),\tau)\,\mathrm{d}\tau]\,\mathrm{d}\tau\,\mathrm{d}\tau \\
&= \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0),t_0)\,(t-t_0) + \frac{1}{2}\,\mathcal{L}\,\mathbf{f}(\mathbf{x}(t_0),t_0)\,(t-t_0)^2 \\
&\quad + \int_{t_0}^{t}\int_{t_0}^{\tau}\int_{t_0}^{\tau} \mathcal{L}^2\,\mathbf{f}(\mathbf{x}(\tau),\tau)\,\mathrm{d}\tau\,\mathrm{d}\tau\,\mathrm{d}\tau
\end{aligned}$$

Taylor series of ODEs [5/5]

- If we continue this procedure ad infinitum, we obtain the following Taylor series expansion for the solution of the ODE:

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0) + \frac{1}{2!}\,\mathcal{L}\,\mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0)^2 + \frac{1}{3!}\,\mathcal{L}^2\,\mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0)^3 + \ldots$$

  where

$$\mathcal{L} = \frac{\partial}{\partial t} + \sum_i f_i\,\frac{\partial}{\partial x_i}$$

- The Taylor series for a given function $\mathbf{x}(t)$ can be obtained by setting $\mathbf{f}(t) = \mathrm{d}\mathbf{x}(t)/\mathrm{d}t$.

Itô-Taylor series of SDEs [1/5]

- Consider the following SDE

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}(t), t)\ \mathrm{d}t + \mathbf{L}(\mathbf{x}(t), t)\ \mathrm{d}\boldsymbol{\beta}.$$

- In integral form this is

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau)\ \mathrm{d}\tau + \int_{t_0}^{t} \mathbf{L}(\mathbf{x}(\tau), \tau)\ \mathrm{d}\boldsymbol{\beta}(\tau).$$

- Applying Itô formula to $\mathbf{f}(\mathbf{x}(t), t)$ gives

$$\begin{aligned}
\mathrm{d}\mathbf{f}(\mathbf{x}(t), t) &= \frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial t}\ \mathrm{d}t + \sum_u \frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial x_u}\ f_u(\mathbf{x}(t), t)\ \mathrm{d}t \\
&+ \sum_u \frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial x_u}\ [\mathbf{L}(\mathbf{x}(t), t)\ \mathrm{d}\boldsymbol{\beta}(\tau)]_u \\
&+ \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{f}(\mathbf{x}(t), t)}{\partial x_u\, \partial x_v}\ [\mathbf{L}(\mathbf{x}(t), t)\, \mathbf{Q}\, \mathbf{L}^{\mathsf{T}}(\mathbf{x}(t), t)]_{uv}\, \mathrm{d}t
\end{aligned}$$

Itô-Taylor series of SDEs [2/5]

- Similarly for $\mathbf{L}(\mathbf{x}(t), t)$ we get via Itô formula:

$$
\begin{aligned}
\mathrm{d}\mathbf{L}(\mathbf{x}(t), t) &= \frac{\partial \mathbf{L}(\mathbf{x}(t), t)}{\partial t}\, \mathrm{d}t + \sum_u \frac{\partial \mathbf{L}(\mathbf{x}(t), t)}{\partial x_u}\, f_u(\mathbf{x}(t), t)\, \mathrm{d}t \\
&+ \sum_u \frac{\partial \mathbf{L}(\mathbf{x}(t), t)}{\partial x_u}\, [\mathbf{L}(\mathbf{x}(t), t)\, \mathrm{d}\boldsymbol{\beta}(\tau)]_u \\
&+ \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{L}(\mathbf{x}(t), t)}{\partial x_u\, \partial x_v}\, [\mathbf{L}(\mathbf{x}(t), t)\, \mathbf{Q}\, \mathbf{L}^{\mathsf{T}}(\mathbf{x}(t), t)]_{uv}\, \mathrm{d}t
\end{aligned}
$$

Itô-Taylor series of SDEs [3/5]

- In integral form these can be written as

$$
\begin{aligned}
\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\mathbf{x}(t_0), t_0) &+ \int_{t_0}^{t} \frac{\partial \mathbf{f}(\mathbf{x}(\tau), \tau)}{\partial t} \, \mathrm{d}\tau + \int_{t_0}^{t} \sum_u \frac{\partial \mathbf{f}(\mathbf{x}(\tau), \tau)}{\partial x_u} f_u(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau \\
&+ \int_{t_0}^{t} \sum_u \frac{\partial \mathbf{f}(\mathbf{x}(\tau), \tau)}{\partial x_u} [\mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\boldsymbol{\beta}(\tau)]_u \\
&+ \int_{t_0}^{t} \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{f}(\mathbf{x}(\tau), \tau)}{\partial x_u \, \partial x_v} [\mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathbf{Q} \, \mathbf{L}^\mathsf{T}(\mathbf{x}(\tau), \tau)]_{uv} \, \mathrm{d}\tau \\
\mathbf{L}(\mathbf{x}(t), t) = \mathbf{L}(\mathbf{x}(t_0), t_0) &+ \int_{t_0}^{t} \frac{\partial \mathbf{L}(\mathbf{x}(\tau), \tau)}{\partial t} \, \mathrm{d}\tau + \int_{t_0}^{t} \sum_u \frac{\partial \mathbf{L}(\mathbf{x}(\tau), \tau)}{\partial x_u} f_u(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau \\
&+ \int_{t_0}^{t} \sum_u \frac{\partial \mathbf{L}(\mathbf{x}(\tau), \tau)}{\partial x_u} [\mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\boldsymbol{\beta}(\tau)]_u \\
&+ \int_{t_0}^{t} \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{L}(\mathbf{x}(\tau), \tau)}{\partial x_u \, \partial x_v} [\mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathbf{Q} \, \mathbf{L}^\mathsf{T}(\mathbf{x}(\tau), \tau)]_{uv} \, \mathrm{d}\tau
\end{aligned}
$$

Itô-Taylor series of SDEs [4/5]

- Let's define operators

$$\mathcal{L}_t \, \mathbf{g} = \frac{\partial \mathbf{g}}{\partial t} + \sum_u \frac{\partial \mathbf{g}}{\partial x_u} \, f_u + \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{g}}{\partial x_u \, \partial x_v} \, [\mathbf{L} \, \mathbf{Q} \, \mathbf{L}^\mathsf{T}]_{uv}$$

$$\mathcal{L}_{\beta,v} \, \mathbf{g} = \sum_u \frac{\partial \mathbf{g}}{\partial x_u} \, \mathbf{L}_{uv}, \qquad v = 1, \ldots, n.$$

- Then we can conveniently write

$$\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\mathbf{x}(t_0), t_0) + \int_{t_0}^{t} \mathcal{L}_t \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau + \sum_v \int_{t_0}^{t} \mathcal{L}_{\beta,v} \mathbf{f}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\beta_v(\tau)$$

$$\mathbf{L}(\mathbf{x}(t), t) = \mathbf{L}(\mathbf{x}(t_0), t_0) + \int_{t_0}^{t} \mathcal{L}_t \mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\tau + \sum_v \int_{t_0}^{t} \mathcal{L}_{\beta,v} \mathbf{L}(\mathbf{x}(\tau), \tau) \, \mathrm{d}\beta_v(\tau)$$

Itô-Taylor series of SDEs [5/5]

- If we now substitute these into equation of $\mathbf{x}(t)$, we get

$$\begin{aligned}
\mathbf{x}(t) &= \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0) + \mathbf{L}(\mathbf{x}(t_0), t_0)\,(\boldsymbol{\beta}(t) - \boldsymbol{\beta}(t_0)) \\
&+ \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}_t \mathbf{f}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\tau\, \mathrm{d}\tau + \sum_v \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}_{\beta,v} \mathbf{f}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\beta_v(\tau)\, \mathrm{d}\tau \\
&+ \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}_t \mathbf{L}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\tau\, \mathrm{d}\boldsymbol{\beta}(\tau) + \sum_v \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}_{\beta,v} \mathbf{L}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\beta_v(\tau)\, \mathrm{d}\boldsymbol{\beta}(\tau).
\end{aligned}$$

- This can be seen to have the form

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0) + \mathbf{L}(\mathbf{x}(t_0), t_0)\,(\boldsymbol{\beta}(t) - \boldsymbol{\beta}(t_0)) + \mathbf{r}(t)$$

- $\mathbf{r}(t)$ is a remainder term.
- By neglecting the remainder we get the Euler–Maruyma method.

Euler–Maruyama method

Euler-Maruyama method

Draw $\hat{\mathbf{x}}_0 \sim p(\mathbf{x}_0)$ and divide time $[0, t]$ interval into K steps of length Δt. At each step k do the following:

1. Draw random variable $\Delta\boldsymbol{\beta}_k$ from the distribution (where $t_k = k\,\Delta t$)

$$\Delta\boldsymbol{\beta}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}\,\Delta t).$$

2. Compute

$$\hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \mathbf{f}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta t + \mathbf{L}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta\boldsymbol{\beta}_k.$$

Order of convergence

- Strong order of convergence γ:

$$\mathrm{E}\left[|\mathbf{x}(t_n) - \hat{\mathbf{x}}(t_n)|\right] \le K\,\Delta t^{\gamma}$$

- Weak order of convergence α:

$$|\,\mathrm{E}\left[g(\mathbf{x}(t_n))\right] - \mathrm{E}\left[g(\hat{\mathbf{x}}(t_n))\right]| \le K\,\Delta t^{\alpha},$$

  for any function g.

- Euler–Maruyama method has strong order $\gamma = 1/2$ and weak order $\alpha = 1$.
- The reason for $\gamma = 1/2$ is the following term in the remainder:

$$\sum_{v} \int_{t_0}^{t} \int_{t_0}^{\tau} \mathcal{L}_{\beta,v} \mathbf{L}(\mathbf{x}(\tau), \tau)\, \mathrm{d}\beta_v(\tau)\, \mathrm{d}\boldsymbol{\beta}(\tau).$$

Milstein's method [1/4]

- If we now expand the problematic term using Itô formula, we get

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \mathbf{f}(\mathbf{x}(t_0), t_0)\,(t - t_0) + \mathbf{L}(\mathbf{x}(t_0), t_0)\,(\boldsymbol{\beta}(t) - \boldsymbol{\beta}(t_0)) + \sum_v \mathcal{L}_{\beta,v} \mathbf{L}(\mathbf{x}(t_0), t_0) \int_{t_0}^{t} \int_{t_0}^{\tau} \mathrm{d}\beta_v(\tau)\, \mathrm{d}\boldsymbol{\beta}(\tau) + \text{remainder}.$$

- Notice the iterated Itô integral appearing in the equation:

$$\int_{t_0}^{t} \int_{t_0}^{\tau} \mathrm{d}\beta_v(\tau)\, \mathrm{d}\boldsymbol{\beta}(\tau).$$

- Computation of general iterated Itô integrals is non-trivial.
- We usually also need to approximate the iterated Itô integrals – different ways for strong and weak approximations.

Milstein's method [2/4]

Milstein's method

Draw $\hat{\mathbf{x}}_0 \sim p(\mathbf{x}_0)$, and at each step k do the following:

1. Jointly draw the following:

$$\Delta\boldsymbol{\beta}_k = \boldsymbol{\beta}(t_{k+1}) - \boldsymbol{\beta}(t_k)$$
$$\Delta\chi_{v,k} = \int_{t_k}^{t_{k+1}} \int_{t_k}^{\tau} \mathrm{d}\beta_v(\tau)\, \mathrm{d}\boldsymbol{\beta}(\tau).$$

2. Compute

$$\hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \mathbf{f}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta t + \mathbf{L}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta\boldsymbol{\beta}_k + \sum_v \left[\sum_u \frac{\partial \mathbf{L}}{\partial x_u}(\hat{\mathbf{x}}(t_k), t_k)\, \mathbf{L}_{uv}(\hat{\mathbf{x}}(t_k), t_k)\right] \Delta\chi_{v,k}.$$

Milstein's method [3/4]

- The strong and weak orders of the above method are both 1.
- The difficulty is in drawing the iterated stochastic integral jointly with the Brownian motion.
- If the noise is additive, that is, $\mathbf{L}(\mathbf{x}, t) = \mathbf{L}(t)$ then Milstein's algorithm reduces to Euler–Maruyama.
- Thus in additive noise case, the strong order of Euler–Maruyama is 1 as well.
- In scalar case we can compute the iterated stochastic integral:

$$\int_{t_0}^{t} \int_{t_0}^{\tau} \mathrm{d}\beta(\tau) \, \mathrm{d}\beta(\tau) = \frac{1}{2} \left[(\beta(t) - \beta(t_0))^2 - q \, (t - t_0) \right]$$

Milstein's method [4/4]

Scalar Milstein's method

Draw $\hat{x}_0 \sim p(x_0)$, and at each step k do the following:

1. Draw random variable $\Delta\beta_k$ from the distribution (where $t_k = k\,\Delta t$)

$$\Delta\beta_k \sim \mathrm{N}(0, q\,\Delta t).$$

2. Compute

$$\hat{x}(t_{k+1}) = \hat{x}(t_k) + f(\hat{x}(t_k), t_k)\,\Delta t + L(x(t_k), t_k)\,\Delta\beta_k + \frac{1}{2}\frac{\partial L}{\partial x}(\hat{x}(t_k), t_k)\,L(\hat{x}(t_k), t_k)\,(\Delta\beta_k^2 - q\,\Delta t).$$

Higher Order Methods

- By taking more terms into the expansion, can form methods of arbitrary order.
- The high order iterated Itô integrals will be increasingly hard to simulate.
- However, if **L** does not depend on the state, we can get up to strong order 1.5 without any iterated integrals.
- For that purpose we need to expand the following terms using the Itô formula (see the lecture notes):

$$\mathcal{L}_t \mathbf{f}(\mathbf{x}(t), t)$$

$$\mathcal{L}_{\beta,v} \mathbf{f}(\mathbf{x}(t), t).$$

Strong Order 1.5 Itô–Taylor Method

Strong Order 1.5 Itô–Taylor Method

When $\mathbf{L}$ and $\mathbf{Q}$ are constant, we get the following algorithm. Draw $\hat{\mathbf{x}}_0 \sim p(\mathbf{x}_0)$, and at each step k do the following:

1. Draw random variables $\Delta\zeta_k$ and $\Delta\beta_k$ from the joint distribution

$$\begin{pmatrix} \Delta\zeta_k \\ \Delta\beta_k \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{Q}\,\Delta t^3/3 & \mathbf{Q}\,\Delta t^2/2 \\ \mathbf{Q}\,\Delta t^2/2 & \mathbf{Q}\,\Delta t \end{pmatrix} \right).$$

2. Compute

$$\hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \mathbf{f}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta t + \mathbf{L}\,\Delta\beta_k + \mathbf{a}_k\,\frac{(t - t_0)^2}{2} + \sum_v \mathbf{b}_{v,k}\,\Delta\zeta_k$$

$$\mathbf{a}_k = \frac{\partial \mathbf{f}}{\partial t} + \sum_u \frac{\partial \mathbf{f}}{\partial x_u} f_u + \frac{1}{2} \sum_{uv} \frac{\partial^2 \mathbf{f}}{\partial x_u\,\partial x_v} [\mathbf{L}\,\mathbf{Q}\,\mathbf{L}^\mathsf{T}]_{uv}$$

$$\mathbf{b}_{v,k} = \sum_u \frac{\partial \mathbf{f}}{\partial x_u} \mathbf{L}_{uv}.$$

Weak Approximations

- If we are only interested in the statistics of SDE solutions, weak approximations are enough.
- In weak approximations iterated Itô integrals can be replaced with simpler approximations with right statistics.
- These approximations are typically non-Gaussian – e.g., a simple weak Euler–Maruyama scheme is

$$\hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \mathbf{f}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta t + \mathbf{L}(\hat{\mathbf{x}}(t_k), t_k)\,\Delta\hat{\boldsymbol{\beta}}_k,$$

where

$$\mathrm{P}(\Delta\hat{\beta}_k^j = \pm\sqrt{\Delta t}) = \frac{1}{2}.$$

- For details, see Kloeden and Platen (1999).

Summary

- Gaussian approximations of SDEs can be formed by assuming Gaussianity in the mean and covariance equations.
- The resulting equations can be numerically solved using linearization or cubature integration (sigma-point methods).
- Itô–Taylor series is a stochastic counterpart of Taylor series for ODEs.
- With first order truncation of Itô–Taylor series we get Euler–Maruyama method.
- Including additional stochastic term leads to Milstein's method.
- Computation/approximation of iterated Itô integrals can be hard and needed for implementing the methods.
- In additive noise case we get a simple 1.5 strong order method.
- Weak approximations are simpler and enough for approximating the statistics of SDEs.