SDEs and filtering

Erik Lindström

Centre for Mathematical Sciences
Lund University

LU/LTH & DTU

LUND UNIVERSITY

Overview

General

Generalize classic ODE

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu(t, X_t)$$

to stochastic ODE

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu(t, X_t) + \sigma(t, X_t)v_t,$$

where v_t is a formal time derivative of a continuous time random walk.

Such object cannot be defined!!!

This is an applied course. We try to preserve the relevant properties.

Interpretation of μ and σ

- $\mu = \lim_{h\to 0} \frac{1}{h}\mathbf{E}[X_{t+h} - X_t | X_t]$.
- $\sigma\sigma^T = \lim_{h\to 0} \frac{1}{h} Var[X_{t+h} - X_t | X_t]$.
- Thus $\frac{\mathrm{d}X}{\mathrm{d}t} = \mu(t, X_t) + "noise"$,
- and $\mathbf{E}[noise] = 0$.

These choices conditions have farreaching consequences!

Problems (P) and solutions (S)

- P: The time derivative of a BM cannot be defined $\nexists v_t$
- S: Thus, we define stoch. integral equations instead
- $\int \mathrm{d}X_u = \int \mu(u, X_u)\mathrm{d}u + \int \sigma(u, X_u)\mathrm{d}W_u$

- P: The BM has unbounded variation - integral does not converge.
- S: But the BM has bounded quadratic variation.
- Thus, define the stoch integral as limit in $L^2(d\mathbb{P} \times dt)$.

Integrals

- $\int \mu(u, X_u)\mathrm{d}u$ is a standard **Riemann integral**.
- $\int \mu(u, X_u)\mathrm{d}u = \lim \sum \mu(\tau, X_\tau)(u_{n+1} - u_n)$, where $u_n \leq \tau \leq u_{n+1}$.
- $\int \sigma(u, X_u)\mathrm{d}W_u$ is a **Ito integral**.
- $\int \sigma(u, X_u)\mathrm{d}W_u = \lim \sum \sigma(u_n, X_{u_n})(W_{u_{n+1}} - W_{u_n})$.
- **To do:** Calculate $\int W_u \mathrm{d}W_u$.
- **To do:** Calculate $\mathbf{E}[\int_t^T W_u \mathrm{d}W_u | \mathcal{F}_t]$.

Chain rule for SDEs

- ODEs are solved by transforming, and thereafter integrating the system.
- The same holds for SDEs.
- Let $\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t$
- Introduce $Y_t = F(t, X_t)$. What about $\mathrm{d}Y_t$?
- $$\begin{aligned}\mathrm{d}Y_t &= \left(F_t + \mu F_x + \frac{1}{2}\sigma^2 F_{xx}\right)\mathrm{d}t + \sigma F_x \mathrm{d}W_t && (1)\\ &= \left(F_t + \mathcal{A}F\right)\mathrm{d}t + \sigma F_x \mathrm{d}W_t. && (2)\end{aligned}$$

Look at handouts for more theoretical background.

Monte Carlo

We can easily simulate from an SDE.

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t \tag{3}$$

Convergence is defined as

- Weak convergence of $\mathcal{O}(h^\beta)$ is defined for $h < h_0$

$$\left|\mathbf{E}[g(X_T)] - \mathbf{E}[g(X_T^h)]\right| \leq Dh^\beta \tag{4}$$

- Strong convergence of $\mathcal{O}(h^\alpha)$ is defined as

$$\mathbf{E}\left[\left|X_T - X_T^h\right|\right] \leq Ch^\alpha \tag{5}$$

Popular schemes

- The Euler-Maruyama scheme is given by

$$X_{t+h} = X_t + \mu(t, X_t)h + \sigma(t, X_t)\delta W_t^h \tag{6}$$

where $\delta W_t^h = W_{t+h} - W_t$.
- Weak convergence $\mathcal{O}(h)$, strong convergence $\mathcal{O}(\sqrt{h})$
- The Milstein scheme is given by

$$X_{t+h} = X_t + \mu(t, X_t)h + \sigma(t, X_t)\delta W_t^h + \frac{1}{2}\sigma(t, X_t)\sigma'(t, X_t)\left((\delta W_t^h)^2 - h\right) \tag{7}$$

- Weak convergence $\mathcal{O}(h)$, strong convergence $\mathcal{O}(h)$

Fokker-Planck

The transition density $p(x_t|x_s)$ is derived from the Fokker-Planck equation

$$\frac{\partial}{\partial t}p(x_t|x_s) = \mathcal{A}^* p(x_t|x_s) \tag{8}$$

where $\mathcal{A}^*$ is the adjoint operator.

Filtering

- Assume that we have a latent (hidden) process X, in discrete or continuous time.
- X is assumed to be Markov.
- Observations Y are generated from X
- and perturbated by noise.

$$\begin{aligned} x_{n+1} &\sim p(x_{n+1}|x_n) \\ y_{n+1} &= h(t_{n+1}, x_{n+1}, \eta_{n+1}) \end{aligned}$$

Filter equations

What do we know about x? Recursive formulas can be derived for the filter density

$$p(x_n|y_{1:n}) = \frac{p(y_n|x_n)p(x_n|y_{1:n-1})}{\int p(y_n|x_n)p(x_n|y_{1:n-1})\mathrm{d}x_n}, \tag{9}$$

where the predictive density $p\left(x_n|y_{1:n-1}\right)$ is given by

$$p\left(x_n|y_{1:n-1}\right) = \int p\left(x_n|x_{n-1}\right) p\left(x_{n-1}|y_{1:n-1}\right) \mathrm{d}x_{n-1}. \tag{10}$$

Filter algorithms

- Kalman filter
- HMM filter
- (Iterated) Extended Kalman filter
- Unscented Kalman filter
- Gaussian sum filters
- Ensemble Kalman filter
- Seq. Monte Carlo filters

Approx. Gaussian Filters

Treat all distributions as Gaussians, $p(x|\cdot) = \phi(x, \cdot, \cdot)$

- Approximate Kalman filters (think QML)
- Easy to implement
- Nearly optimal (projections in Hilbert space)
- Historical reasons

Non-Gaussian Filters

Treat all distributions as countable sums of Dirac measures
$p(x|\cdot) = \frac{\sum_j \omega_j \delta(x-x_j)}{\sum \omega_l} = \sum_j \frac{\omega_j}{\sum \omega_l} \delta(x - x_j)$

- Use when the drift is highly non-linear
- Non-Gaussian innovations
- E.g. Stochastic volatility
- Performance is critical, and computational resources are abundant.

Proofs in handouts.

Parameter estimation

Still an open question

- Direct maximization not optimal...
- EM
- Iterated filtering
- PMMH

Boat navigation

One possible project on this course was navigation in a lake, based on only a sonar and a chart. This film is by a former student on this course, Jonas Hallgren

https://www.youtube.com/watch?v=bNDcXDilSHA

Information: Red star = origin, Red triangle = true pos. of the boat, Blue star = particle, Green star-ring = current guess of location.

Feedback

Send feedback, questions and information about typos to `erikl@maths.lth.se`.