Chapter 5 - Probability distributions and statistics of SDEs

CSML Reading group

Yannis Zachos

January 29, 2021

Outline

1. Important definitions
2. SDE solution formulations
3. Probability density of solution
4. Moments of solution

Important notions I

$(\Omega, \mathbb{F}, \mathcal{P})$ well-defined probability space.
$\boldsymbol{x}(t)$: stochastic process.

- An Itô process solves the following SDE starting at $\boldsymbol{x}(0)$:

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

- The *available information* at time t about process $\boldsymbol{x}(t)$ denoted by $\{\mathcal{F}_t\} \subseteq \mathbb{F}$ is called a *filtration*.
- $\boldsymbol{x}(t)$ is called a *martingale* iff it has bounded expectation and it holds that $\mathbb{E}[\boldsymbol{x}(t)|\mathcal{F}_s] = \boldsymbol{x}(s) \;\; \forall t \geq s$.
- $\boldsymbol{x}(t)$ is a *Markov process* iff it is true that $p(\boldsymbol{x}(t)|\mathcal{F}_s) = p(\boldsymbol{x}(t)|\boldsymbol{x}(s)) \; \forall t \geq s$.

Important notions II

- The *generator* of an Itô process is a differential operator

$$\mathcal{A}_t(\bullet) = \frac{\partial(\bullet)}{\partial t} + \sum_i \frac{\partial(\bullet)}{\partial x_i} f_i(\boldsymbol{x}, t) + \frac{1}{2} \sum_{i,j} \left(\frac{\partial^2(\bullet)}{\partial x_i \partial x_j} \right) [\boldsymbol{L}(\boldsymbol{x}, t) \boldsymbol{Q} \boldsymbol{L}^T(\boldsymbol{x}, t)]_{i,j}$$

- A SDE solution is *weak* iff we can construct $\hat{\beta}(t)$, $\hat{\boldsymbol{x}}(t)$ such that the pair is a solution to the SDE.
- A solution to the *martingale problem* (MP) for generator $\mathcal{A}$ is a Markov process $\boldsymbol{x}(t)$ for which

$$h(\boldsymbol{x}(t)) - \int_0^t \mathcal{A} h(\boldsymbol{x}(s)) ds$$

is a martingale.

Existence and uniqueness of SDE solution

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

- **Theorem**:

  $\boldsymbol{x}(t)$ weak solution iff $\boldsymbol{x}(t)$ solves the MP.

- **Corollary**:

  existence of weak solution $\equiv$ existence of *some* solution to MP
  uniqueness in law $\equiv$ existence of *at most one* solution to MP

- Equivalence between weak solution and MP formulations of SDE.
- Benefits of using martingale formulation: theory of weak convergence, regular conditional probabilities, localisation.

Kolmogorov's forward equation

- We know solution at time t_0 in the form of $p(\boldsymbol{x}(t_0))$.
- What about time $t \geq t_0$?
- **Forward Kolmogorov equation**: $p(\boldsymbol{x}(t))$ of the solution solves the IVP with initial condition $p(\boldsymbol{x}(t_0))$:

$$\frac{\partial p(\boldsymbol{x}(t))}{\partial t} = -\sum_i \frac{\partial}{\partial x_i} [f_i(\boldsymbol{x}, t) p(\boldsymbol{x}(t))] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} \{[\boldsymbol{L}(\boldsymbol{x}, t) \boldsymbol{Q} \boldsymbol{L}(\boldsymbol{x}, t)]_{ij} p(\boldsymbol{x}(t))\}$$

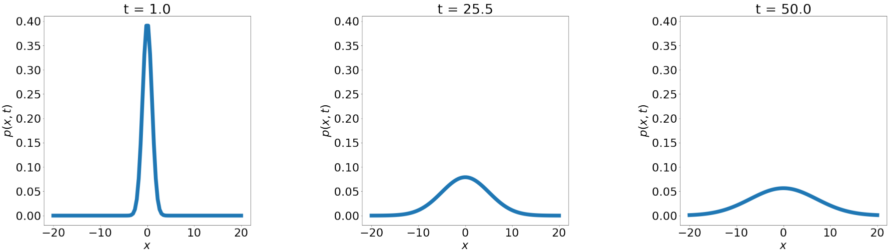
- This looks similar to applying the generator $\mathcal{A}$ to $p(\boldsymbol{x}(t))$.

$$\frac{\partial p}{\partial t} = \mathcal{A}^* p$$

with $\mathcal{A}^*$ adjoint operator of $\mathcal{A}$.

Applications of forward equation I

- **Example**: The SDE $dx = d\beta$ with constant diffusion $q = 2D$ reduces to the diffusion equation
$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}.$$
- whose solution given the initial condition $p(x(0)) = \delta(x)$ is
$$p(x(t)) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right)$$



Result for time-independent processes

- For time-independent processes like

$$d\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x})dt + \boldsymbol{L}(\boldsymbol{x})d\boldsymbol{\beta}$$

with diffusion matrix $\boldsymbol{Q} = q\boldsymbol{I}$ the forward equation satisfies

$$\frac{\partial p(\boldsymbol{x}(t))}{\partial t} = 0.$$

- If we can transform the SDE into

$$d\boldsymbol{x} = -\frac{1}{2}\nabla v(\boldsymbol{x})dt + \boldsymbol{L}(\boldsymbol{x})d\boldsymbol{\beta}$$

via $\boldsymbol{f}(\boldsymbol{x}) = -\nabla v(\boldsymbol{x})$ we can then use the following result:

- **Theorem**: The solution to the time-independent forward equation is

$$p(\boldsymbol{x}) = \frac{\exp(-v(\boldsymbol{x})/q)}{\int \exp(-v(\boldsymbol{x})/q)d\boldsymbol{x}},$$

which looks like the Boltzmann-Gibbs measure.

Applications of forward equation II

- **Example**: The Ornstein-Uhlenbeck process is

$$dx = -\lambda x dt + d\beta$$

  with $x(0) = x_0$.

- Let $v(x) = \lambda x^2$ since $-\frac{1}{2}\nabla v(x) = -\lambda x$. Therefore, the probability of the solution is

$$p(x) \propto \exp\left(-\frac{\lambda x^2}{q}\right),$$

  which resembles a Gaussian distribution with zero mean and $q/2\lambda$ variance.

- Another example of applying the forward equation is found in **"Stochastic modelling of urban structure"**.

Kolmogorov's backward equation

- We have seen how the solution probability propagates forward in time.
- How can we compute moments of the solution?
- **Theorem**: $\boldsymbol{u}(\boldsymbol{x}, t) = \mathbb{E}_x[h(\boldsymbol{x}(t))]$ solves the following initial value problem with initial condition $\boldsymbol{u}(\boldsymbol{x}, 0)$:

$$\frac{\partial \mathbb{E}_x[h(\boldsymbol{x}(t))]}{\partial t} = \mathcal{A}\, \mathbb{E}_x[h(\boldsymbol{x}(t))],$$

where

$$\mathcal{A}(\bullet) = \sum_i \frac{\partial(\bullet)}{\partial x_i} f_i(\boldsymbol{x}, t) + \frac{1}{2} \sum_{i,j} \left(\frac{\partial^2(\bullet)}{\partial x_i \partial x_j} \right) [\boldsymbol{L}(\boldsymbol{x}, t)\boldsymbol{Q}\boldsymbol{L}^T(\boldsymbol{x}, t)]_{i,j}$$

- Great! We can now use this to compute summary statistics of the Itô process.

Moments of Itô processes

- Backward Kolmogorov equation allows us to compute mean via $h(\boldsymbol{x}(t)) = x_u$ and covariance $h(\boldsymbol{x}, t) = x_u x_v - \mathbb{E}[x_u(t)]\mathbb{E}[x_v(t)]$, respectively.
- Mean $\boldsymbol{m}$ solves
$$\frac{d\boldsymbol{m}}{dt} = \mathbb{E}[\boldsymbol{f}(\boldsymbol{x}, t)]$$
while the covariance $\boldsymbol{P}$ solves
$$\frac{d\boldsymbol{P}}{dt} = \mathbb{E}[\boldsymbol{f}(\boldsymbol{x}, t)(\boldsymbol{x} - \boldsymbol{m})^T] + \mathbb{E}[(\boldsymbol{x} - \boldsymbol{m})\boldsymbol{f}^T(\boldsymbol{x}, t)] + \mathbb{E}[\boldsymbol{L}(\boldsymbol{x}, t)\boldsymbol{Q}\boldsymbol{L}^T(\boldsymbol{x}, t)]$$
- We need access to $p(\boldsymbol{x}(t))$ via the forward Kolmogorov equation which we cannot always solve.
- Free lunch only if solution to forward equation is Gaussian.

Examples

- **Example**: Ornstein-Uhlenbeck process: $dx = -\lambda x dt + d\beta$ with $x(0) = x_0$. We have

$$\frac{dm}{dt} = \mathbb{E}[-\lambda x] = -\lambda m$$

$$\frac{dP}{dt} = 2\mathbb{E}[-\lambda(x-m)^2] + \mathbb{E}[q] = -2\lambda P + q.$$

- **Example**: $dx = \sin(x)dt + d\beta$ has

$$\frac{dm}{dt} = \mathbb{E}[\sin(x)] \approx \mathbb{E}[x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots].$$

- Apart from $p(\boldsymbol{x}(t))$ we also need to compute higher order moments.
- The computation cost of the n-th moment for a d-dimensional state $\boldsymbol{x}$ is $\mathcal{O}(d^n)$ and may require us to compute expectation over infinite number of moment equations in the case of $h(x) = x^n$.

Itô processes as Markov processes

- Itô processes are Markovian and are characterised by their transition densities $p(\boldsymbol{x}(t)|\boldsymbol{x}(s))$.
- **Forward Kolmogorov equation**: $p(\boldsymbol{x}(t)|\boldsymbol{y}(s))$ solves the following PDE with $t \geq s$ and initial condition $\delta(\boldsymbol{x}(s) - \boldsymbol{y}(s))$:

$$\frac{\partial p(\boldsymbol{x}(t)|\boldsymbol{y}(s))}{\partial t} = \mathcal{A}^* \, p(\boldsymbol{x}(t)|\boldsymbol{y}(s))$$

- Same applies for the backward Kolmogorov equation.
- We can now factorise the joint distribution of the solution at arbitrary time points (where SDE can be discretised) as

$$p(\boldsymbol{x}_{t_0}, \ldots, \boldsymbol{x}_{t_T}) = p(\boldsymbol{x}_{t_0}) \prod_{k=1}^{T} p(\boldsymbol{x}(t_k)|\boldsymbol{x}(t_{k-1}))$$