Mean field limits for interacting particle systems, their inference, and applications

Part 1: Mean field limits and inference.

Susana Gomes (University of Warwick)

Woudschoten Conference 2023

Outline

Motivation

Systems of interacting particles

Long time behaviour

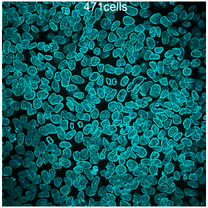
Inference for SDEs

Motivation

Motivation

Interacting particle systems are ubiquitous in the real-world, appearing in several application areas:

- Biology/Life sciences (flocks of birds, schools of fish, herds of sheep,...)
- Social sciences (crowd dynamics, opinion dynamics, ...)
- Cell dynamics
- Engineering (drones, robots, ...)
- Physics (molecular dynamics, movement of galaxies, ...)



Modelling approaches

There are several ways of modelling these types of systems. Today I will focus on (stochastic) **interacting particle systems**:

- Simple models for each particle (usually based on Newton's Laws).
- In common applications, we would have a very large number of particles.
- Analytically and computationally hard to tackle.
- To tackle this, it is common to consider macroscopic limits: model the density of agents as the number of particles $N \to \infty$ using a **mean-field approach**.

There are alternative models, e.g. **deterministic models**, **rational agents** (common in social sciences), **lattice based models** (common in biology, e.g. total exclusion or contact processes).

(Stochastic) Interacting Particle Systems

and their mean field limit

Basic Model

I will consider a class of first-order **weakly** interacting particle systems in one dimension:[1]

$$dX_t^i = V'\left(X_t^i\right)\,\mathrm{d}t + \frac{1}{N}\sum_{j\neq i} K\left(X_t^i - X_t^j\right)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i, \quad X_0^i = x_0^i, \quad i,j = 1,\ldots,N,$$

where

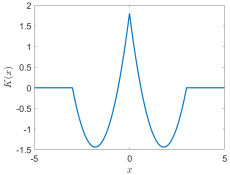
- X_t^i denotes the position of particle i at time t
- $V(\cdot)$ is a **confining potential**
- $K(\cdot)$ is an **interaction potential**, such that $K(0) = 0$ and $K'(0) = 0$.
- W_t^i are independent **Brownian motions** and σ is the strength of noise (sometimes I'll write β^{-1}, which is more common in physics contexts)
- x_0^i are initial positions which can be deterministic or stochastic (independently distributed with some chosen law)
- The scaling $\frac{1}{N}$ is the **mean-field scaling** and is critical for us, as it keeps the strength of interactions of order 1.

[1] This is for simplicity – similar results can be obtained in higher dimensions, and for second-order systems of the type $dX_t^i = V_t^i, \quad dV_t^i = K(X_t^i - X_t^j)\,\mathrm{d}t + \sqrt{2\sigma}dW_t^i$. Alternatively, one can also solve this SDE on a torus, and exclude the potential V, see Carrillo, Gvalani, Pavliotis and Schlichting, ARMA 2018.

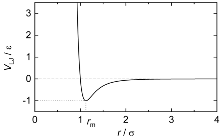
Some examples

There are several examples of potentials, which depend on the application.

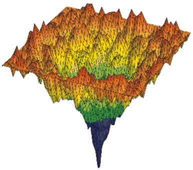
- aggregation potentials (attraction/repulsion) for interactions (common for cells, animals)



- Lennard-Jones interaction potentials (common in chemistry for molecular interactions)



- Protein folding examples (confining potentials)



In *rational agents*, the potentials can be, e.g., utility functions.

Empirical measure and N-particle distribution

To pass to the mean-field limit, it is important to define two measures:

- The **empirical measure**

$$\mu_N(t,x) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - X_t^i).$$

  - ⋆ contains all the information about the solution $(X_t^1, \dots, X_t^N)$.
  - ⋆ is a random probability measure[2]
  - ⋆ The stochastic behaviour only vanishes as $N \to \infty \Rightarrow$ important to quantify fluctuations if N remains finite[3] (we will discuss this later)

- The $N-$**particle or joint distribution**

$$F^N(t, x_1, \dots, x_n) = \mathrm{Law}(X_t^1, \dots, X_t^N)$$

  - ⋆ not experimentaly measurable, but
  - ⋆ its marginals contain statistical information on the process

$$F_k^N(t, x_1, \dots, x_k) = \int_{\mathbb{R}^{N-k}} F^N(t, x_1, \dots, x_N)\mathrm{d}x_{k+1}\cdots\mathrm{d}x^N$$

[2] In the deterministic case (no noise, $\sigma = 0$), this is a deterministic probability measure

[3] See [J. Worsfold, T. Rogers, P. Milewski, SIAM J. Appl. Math (2023)]

N Particle dynamics

Using Itô's formula, one can write a PDE for the evolution of F^N:

$$\partial_t F^N = -\sum_{i=1}^{N} \partial_{x_i} \left(V'(x_i) F^N + \sum_{i=1}^{N} \sum_{j \neq i} K(x_i - x_j) F^N \right) + \sigma \sum_{i=1}^{N} \Delta_{x_i} F^N.$$

Recall Itô's formula (for our case):

Let $(X_t : t \geq 0)$ solve

$$dX_t = a(X_t, t)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t.$$

Then, for a smooth function f, we have

$$df(X_t) = a(X_t, t) f'(X_t) dt + \sigma f''(X_t) dt + \sqrt{2\sigma} f'(X_t) dW_t.$$

To obtain the above PDE, we apply Itô's formula to a general function f and then compute expectations with respect to the law of the process, F^N.

The last term vanishes because it is an Itô integral of a deterministic function.

the relevant points.

The mean-field limit

To formally[4] pass to the limit, we use the **mean field ansatz**, i.e., that

$$F^N(t, x_1, \dots, x_N) = \prod_{i=1}^{N} \rho(t, x_i), \quad \text{and} \quad F^N(0, x_1, \dots, x_N) = \prod_{i=1}^{N} \rho_0(x_i).$$

Using this ansatz in the PDE for the evolution of the N-particle distribution, we can then integrate out $N-1$ variables $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N$ and obtain a PDE for the evolution of x_i:

$$\partial_t \rho = -\frac{\partial}{\partial x_i}\left(V'(x_i)\rho + \frac{N-1}{N}\rho \int_{\mathbb{R}} K(x_i - y)\, dy\right) + \sigma \partial_{x_i}^2 \rho.$$

Sending N to infinity, we obtain the **Fokker-Planck equation**

$$\partial_t \rho = -\frac{\partial}{\partial x_i}\left(V'(x_i)\rho + (K * \rho)\rho\right) + \sigma \partial_{x_i}^2 \rho,$$

where $*$ denotes convolution.

[4] Passing rigorously to this limit can be done using martingale techniques or other classical stochastic analysis results, see [P.-E. Jabin and Z. Wang, Mean Field Limits for Stochastic Particle Systems, Active Particles Volume 1, (2017)] and references therein. This formal derivation follows Urbain Vaes' PhD Thesis, 2019

Some relevant results

An alternative method consists of considering initial data $X_0^i = x_0^i$ i.i.d. with $\mathrm{Law}(x_0^i) = f_0$, and constructing a particle system coupled to the original SDE:

$$\mathrm{d}\bar{X}_t^i = V'(\bar{X}_t^i)\,\mathrm{d}t + (K * f_t)(\bar{X}_t^i)\,\mathrm{d}t + \sqrt{2\sigma}dW_t^i,, \quad X_0^i = x_0^i, \quad i,j = 1,\dots,N,$$

where the W_t^i are the **same Brownian motions** as before, and f_t is the law of $\bar{X}_t$.
This is known as the **McKean-Vlasov equation** (and is no longer an SDE because it depends on the law of the process).

One can check that f_t solves the Fokker–Planck equation on the previous page, and show that the empirical measure μ_N converges in law to $\fallingdotseq$ solving the Fokker–Planck equation.

Under appropriate conditions on K and V, it can be shown[5] that solutions to the McKean-Vlasov equation are **close to the solutions** of the original SDE, and use this to obtain bounds on the difference $|X_t^i - \bar{X}_t^i|^2$, as well as quantify large deviations.

[5] See [P.-E. Jabin and Z. Wang (2017)]

Long time behaviour

Example of a multi-well interacting potential

SNG, G.A. Pavliotis, J. Nonlinear Sci 28, 905-941, 2018
SNG, S. Kalliadasis, G.A. Pavliotis, P. Yatsyshin, Phys Rev E 99, 032109, 2019
(not discussed - 2^{nd} order problem) SNG, G.A. Pavliotis, U. Vaes, Multiscale Modelling and Simulation, 2020

A system of interacting particles

We consider a particular case of N weakly interacting particles given by:

$$dX_t^i = \left(-V'(X_t^i) - \theta \left(X_t^i - \frac{1}{N} \sum_{j=1}^{N} X_t^j \right) \right) dt + \sqrt{2\sigma}\, dW_t^i.$$

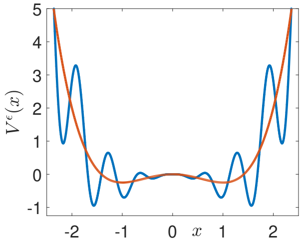
Here, the particles interact via their mean, with strength θ, i.e., in a quadratic Currie-Weiss potential $K(x) = \frac{x^2}{2}$.

We also consider multi-well confining potentials. For example:

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2}.$$

$$\begin{aligned} V_8(x) &= h\left(x^8 - 14x^6 + 49x^4 - 36x^2\right) \\ &= hx^2(x^2-1)(x^2-4)(x^2-9), \end{aligned}$$

$$V^\epsilon(x) = V_0(x) + \delta \frac{x^2}{2} \cos\left(\frac{x}{\epsilon}\right).$$



$N \to \infty$ and the McKean-Vlasov equation

Using the previous arguments and using the Law of Large Numbers, we can formally study the mean field limit::

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} X_t^j = \mathbb{E} X_t,$$

where $\mathbb{E}$ is taking with respect to the one-particle distribution. We pass to the limit $N \to \infty$ and obtain the McKean-Vlasov SDE for X_t

$$dX_t = -V'(X_t)\,dt - \theta(X_t - \mathbb{E}X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t.$$

This SDE has a corresponding nonlinear Fokker-Planck equation:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(V'(x)p + \theta\left(x - \int_{\mathbb{R}} xp(x,t)\,dx\right)p + \beta^{-1}\frac{\partial p}{\partial x}\right).$$

Its steady states allow us to investigate the long-time behaviour of this system.

Multiple invariant measures

Invariant measures of the McKean-Vlasov SDE are steady states of the Fokker-Planck equation:

$$\frac{\partial}{\partial x}\left(V'(x)p_\infty + \theta\left(x - \int_{\mathbb{R}} x p_\infty(x)\,dx\right)p_\infty + \beta^{-1}\frac{\partial p_\infty}{\partial x}\right) = 0.$$

This admits a one-parameter family of solutions:

$$p_\infty(x;\theta,\beta,m) = \frac{e^{-\beta\left(V(x)+\theta\left(\frac{1}{2}x^2 - xm\right)\right)}}{Z(\theta,\beta;m)}, \quad Z(\theta,\beta;m) = \int_{\mathbb{R}} e^{-\beta\left(V(x)+\theta\left(\frac{1}{2}x^2 - xm\right)\right)}\,dx,$$

subject to the constraint that they provide us with the correct formula for the first moment:

The selfconsistency equation

$$m = \int_{\mathbb{R}} x p_\infty(x;\theta,\beta,m)\,dx =: R(m;\theta,\beta).$$

Critical temperature

To find invariant measure(s) of the McKean-Vlasov dynamics we need to solve the following:

The selfconsistency equation

$$m = \int_{\mathbb{R}} x p_\infty(x; \theta, \beta, m)\, dx =: R(m; \theta, \beta).$$

For sufficiently small β, $m = 0$ is the only solution of the selfconsistency equation. However, for **nonconvex** confining potentials, there exists a **critical temperature**, β_C, at which this is no longer true.[6]

To find β_C, one can differentiate the selfconsistency equation at $m = 0$, and conclude that β_C is the solution of

$$\mathrm{Var}_{p_\infty}(\theta, \beta; m = 0) := \int_{\mathbb{R}} x^2 p_\infty(x; \theta, \beta, m = 0)\, dx = \frac{1}{\beta\theta}.$$

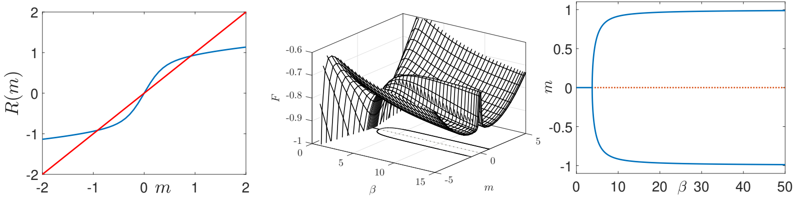
[6] [Dawson, J. Stat.Phys 1983, Tamura,J. Fac.Sci. Univ. Tokyo 1984, Shiino, Phys. Rev. A 1987, Tugaut, Stochastics 2013]

Numerical results: Bistable potential[7]

The simplest example we can consider is the bistable potential,

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2}.$$

For sufficiently large β, the selfconsistency equation has two solutions

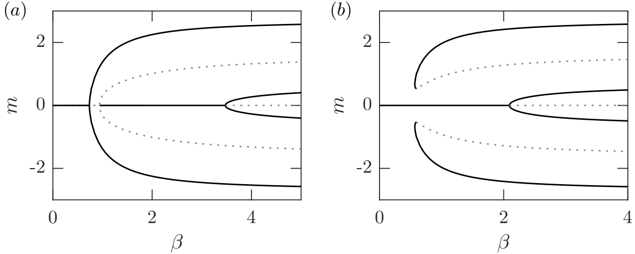


$R(m; 0.5, 10)$ against $y = x$ (left), bifurcation diagram of m as a function of β for $\theta = 0.5$ (middle), and free energy surface as a function of β and m (right).

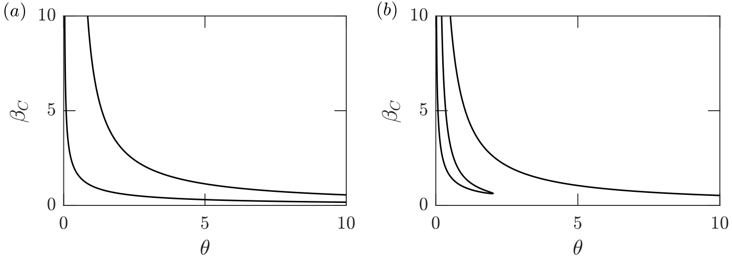
7 See, e.g., [Dawson, J. Stat. Phys 1983] for a detailed study of this case.

Numerical results: Multi-well potentials

Phase diagrams for the potential $V_8(x)$ for $h = 0.001$, and (left) $\theta = 1.5$, (right) $\theta = 2.5$.



Critical temperature β_C as a function of θ for (left) $V_6(x)$ and (right) $V_8(x)$.



Numerical results: Multiscale potentials

In this case, we need to distinguish between ϵ small but finite, and $\epsilon \to 0$. It also matters when we pass to the mean-field limit $N \to \infty$.

One can use homogenisation techniques to obtain an homogenised SDE (first $\epsilon \to 0$ then $N \to \infty$), or first pass to the mean field limit and then send $\epsilon \to 0$.

We can show[8] that if the oscillations are additive, then the two limits commute. Otherwise, we obtain different long-time behaviour.

This can be seen from the self-consistency equation

$\epsilon \to 0$ first

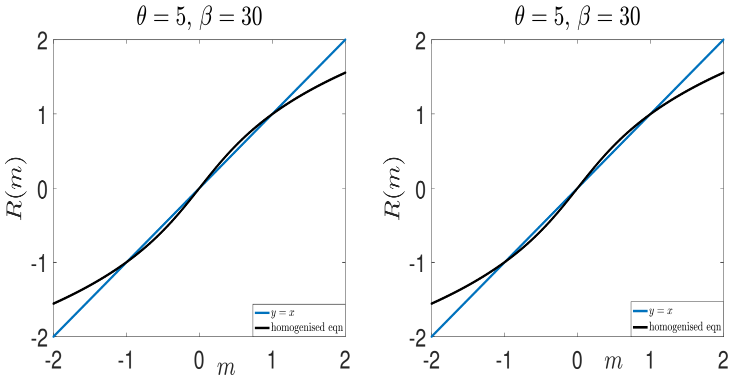
$$m = \int_{\mathbb{R}} \frac{x e^{-\beta(V_{\mathit{eff}}(x)+\psi(x))}\,dx}{Z}$$

$N \to \infty$ first

$$m = \int_0^L \int_{\mathbb{R}} \frac{x\ e^{-\beta(V_{\mathit{eff}}(x)+V_1(x,y))}\,dx\,dy}{\bar{Z}}$$

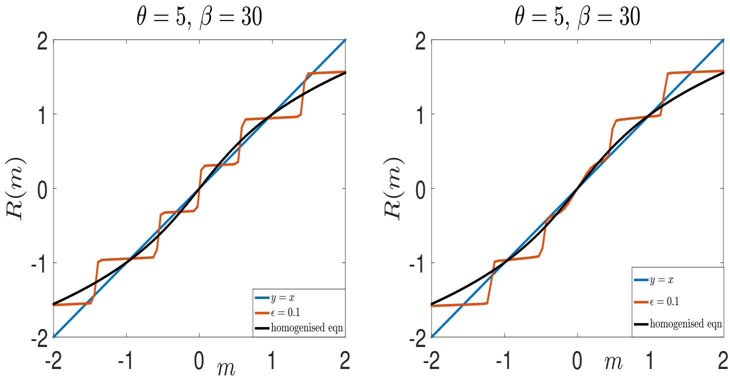
[8] see [SNG, G.A. Pavliotis, J. Nonlinear Sci 2018]

Numerical illustration



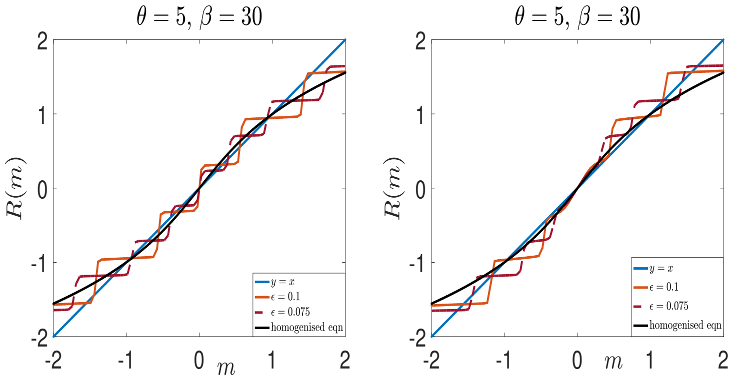
Plot of $R(m; \theta, \beta) = m$ and $R(m^\epsilon; \theta, \beta)$ for $\theta = 5,\ \beta = 30, \delta = 1$ and various values of ϵ for separable fluctuations (left) and multiplicative fluctiations (right).

Numerical illustration



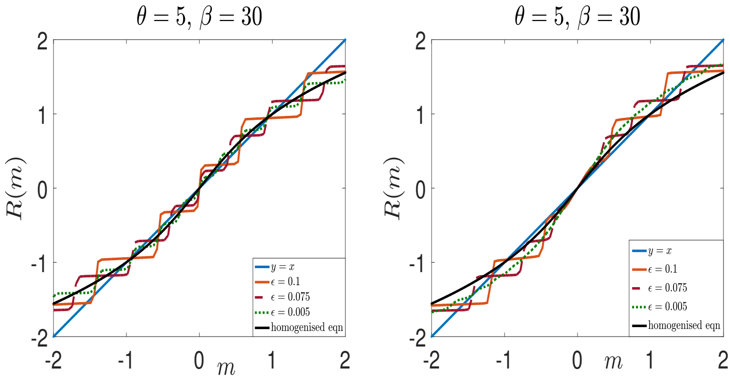
Plot of $R(m; \theta, \beta) = m$ and $R(m^{\epsilon}; \theta, \beta)$ for $\theta = 5,\ \beta = 30, \delta = 1$ and various values of ϵ for separable fluctuations (left) and multiplicative fluctiations (right).

Numerical illustration



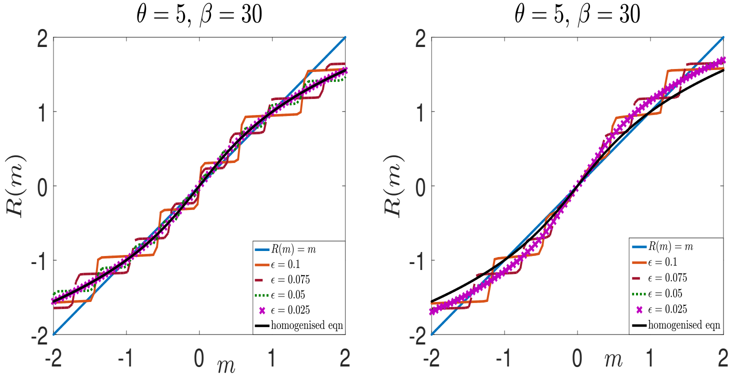
Plot of $R(m;\theta,\beta) = m$ and $R(m^\epsilon;\theta,\beta)$ for $\theta = 5,\ \beta = 30, \delta = 1$ and various values of ϵ for separable fluctuations (left) and multiplicative fluctiations (right).

Numerical illustration



Plot of $R(m; \theta, \beta) = m$ and $R(m^\epsilon; \theta, \beta)$ for $\theta = 5,\ \beta = 30, \delta = 1$ and various values of ϵ for separable fluctuations (left) and multiplicative fluctiations (right).

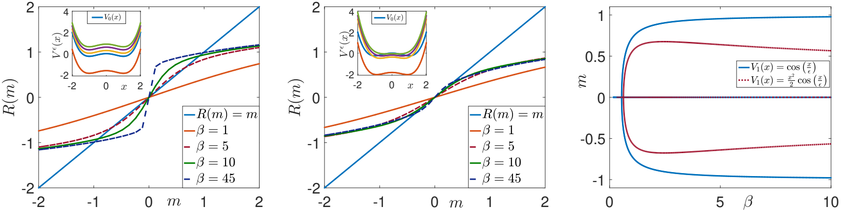
Numerical illustration



Plot of $R(m;\theta,\beta) = m$ and $R(m^{\epsilon};\theta,\beta)$ for $\theta = 5,\ \beta = 30, \delta = 1$ and various values of ϵ for separable fluctuations (left) and multiplicative fluctiations (right).

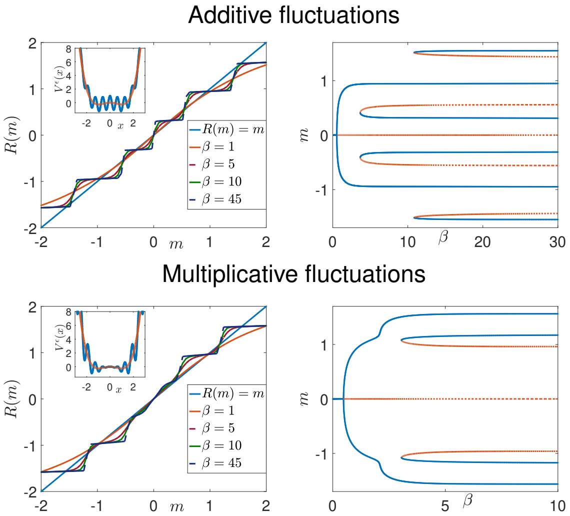
Homogenised bistable potential

The bistable potential maintains its two extra solutions... But now the homogenised potential depends on β.



$R(m; \theta, \beta)$ compared to $y = x$ for $\theta = 0.5, \delta = 1$ and various values of β for the homogenized bistable potentials with additive (left) and multiplicative (middle) fluctuations, and bifurcation diagram of m as a function of β for the additive (full line) and multiplicative (dashed line) fluctuations (right).

Finite ϵ: bistable potential, $\theta = 5$, $\delta = 1$, $\epsilon = 0.1$



$R(m^\epsilon; \theta, \beta)$ for various values of β (left), and bifurcation diagram of m as a function of β (right). Full lines are stable solutions, while dashed lines represent unstable ones.

Inference for SDEs

Parameter estimation - usual inverse problem setting

In several problems, one wants to estimate parameters present in our models (SDEs).
Consider an SDE that depends on a parameter, θ

$$dX_t = b(X_t;\theta)\,\mathrm{d}t + \mathrm{d}W_t,$$

where we assume we know the diffusion coefficient and $\sigma = 1$.

Intuitively, one would want to find the best value of θ given an observation of a trajectory $\mathbf{X}_t$. This would correspond to minimising the function

$$\Phi(\theta; X_t) = \int_0^T |\dot{X}_t - b(X_t;\theta)|^2.$$

However, $\mathbf{X}_t$ solves an SDE, so computing $\Phi(\theta;\mathbf{X}_t)$ is equivalent to *integrating the square of the derivative of a Brownian motion*!

Recall that...

The Brownian motion has unbounded variation - this means that it is not differentiable anywhere. In particular,

$$\mathbb{P}\left(\forall t > 0 : \limsup_{\Delta t \to 0} \left| \frac{W_{t+\Delta t} - W_t}{\Delta t} \right| = \infty \right) = 1.$$

For this reason, $\Phi(\theta;\cdot)$ is almost surely infinite, and one can't solve this inverse problem in the usual way.

Maximum likelihood inference

If the problem we are modelling involves noise, we need to do something better. We can fix this by defining the **maximum likelihood estimator**[9].

Assume we have a *random variable* X with probability distribution function $f(x;\theta)$, known up to *parameters* θ *that we want to estimate* from observations.

Example: $X \sim \mathcal{N}(\mu, \sigma^2)$

The parameters are the mean and variance, $\theta = (\mu, \sigma)$. f is

$$f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

[9] Casella and Berger, Statistical inference, 2002.

Maximum likelihood inference

If the problem we are modelling involves noise, we need to do something better. We can fix this by defining the **maximum likelihood estimator**[9].

Assume we have a *random variable X* with probability distribution function $f(x;\theta)$, known up to *parameters θ that we want to estimate* from observations.

Suppose that we have *J independent observations* of X. We define the likelihood function

$$L(\{x_j\}_{j=1}^{J};\theta) = \prod_{j=1}^{J} f(x_j;\theta).$$

Example: $X \sim \mathcal{N}(\mu, \sigma^2)$

The parameters are the mean and variance, $\theta = (\mu, \sigma)$. f is

$$f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

The likelihood function is

$$L(\{x_j\}_{j=1}^{J};\theta) = \frac{e^{-\sum_{j=1}^{J}\frac{(x_j-\mu)^2}{2\sigma^2}}}{(2\pi\sigma)^{J/2}}.$$

[9] Casella and Berger, Statistical inference, 2002.

Maximum likelihood inference

If the problem we are modelling involves noise, we need to do something better. We can fix this by defining the **maximum likelihood estimator**[9].

Assume we have a *random variable X* with probability distribution function $f(x;\theta)$, known up to *parameters* θ *that we want to estimate* from observations.

Suppose that we have *J independent observations* of X. We define the likelihood function

$$L(\{x_j\}_{j=1}^J;\theta) = \prod_{j=1}^J f(x_j;\theta).$$

The *maximum likelihood estimator* (MLE) is

$$\hat{\theta} = \arg\max L(\{x_j\}_{j=1}^J;\theta)$$

Example: $X \sim \mathcal{N}(\mu, \sigma^2)$

The parameters are the mean and variance, $\theta = (\mu, \sigma)$. f is

$$f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

The likelihood function is

$$L(\{x_j\}_{j=1}^J;\theta) = \frac{e^{-\sum_{j=1}^J \frac{(x_j-\mu)^2}{2\sigma^2}}}{(2\pi\sigma)^{J/2}}.$$

We maximise w.r.t. μ and σ^2

$$\hat{\mu} = \frac{1}{J}\sum_{j=1}^J x_j,\ \hat{\sigma^2} = \frac{1}{J}\sum_{j=1}^J (x_j - \hat{\mu})^2$$

[9] Casella and Berger, Statistical inference, 2002.

Maximum likelihood inference for SDEs

In our case, the observations are a *series of discrete observations of a stochastic process:* $\{X_{i\Delta t}\}_{i=1}^{N}$ which solves an SDE.

$$dX_t = b(X_t;\theta)\,dt + dW_t, \qquad \longleftrightarrow \qquad X_{(i+1)\Delta t} = X_{i\Delta t} + b(X_{i\Delta t};\theta)\Delta t + \Delta W_{i\Delta t}.$$

Using the fact that $\Delta W_{i\Delta t} \sim \mathcal{N}(0, \Delta t)$, we can see that

$$\mathbb{P}(X_{(i+1)\Delta t} | X_{i\Delta t}) \sim \mathcal{N}(X_{i\Delta t} + b(X_{i\Delta t};\theta)\Delta t, \Delta t),$$

and therefore, writing $f_i = f(X_{i\Delta t};\theta)$ we can write the *law of the process X_t*

$$p_X^N = \frac{1}{(\sqrt{2\pi\Delta t})^N} \exp\left(-\sum_{i=0}^{N-1}\left(\frac{1}{2\Delta t}(\Delta X_i)^2 + \frac{1}{2}(b_i)^2\Delta t - b_i\Delta X_i\right)\right).$$

Similarly, the distribution function for Brownian motion is

$$p_W^N = \prod_{i=0}^{N-1}\frac{1}{\sqrt{2\pi\Delta t}}\exp\left(-\frac{1}{2\Delta t}(\Delta W_i)^2\right) = \frac{1}{(\sqrt{2\pi\Delta t})^N}\exp\left(-\frac{1}{2\Delta t}\sum_{i=0}^{N-1}(\Delta W_i)^2\right).$$

MLE for SDEs (continued)

Now we can calculate the ratio of the laws of the two processes, evaluated at the path $\{X_n\}_{n=0}^{N-1}$:

$$\frac{p_X^N}{p_W^N} = \exp\left(-\frac{1}{2}\sum_{i=0}^{N-1} b_i^2 \Delta t + \sum_{i=0}^{N-1} b_i \Delta X_i\right).$$

Taking the limit as $N \to \infty$, we get the likelihood:

$$L\left(\{X_t\}_{t\in[0,T]};\theta, T\right) := \exp\left(\int_0^T b(X_s;\theta)\, dX_s - \frac{1}{2}\int_0^T b(X_s;\theta)^2\, ds\right).$$

Rigorously, this can be done using *Girsanov's theorem*: one can show that $\mathbb{P}^X$ is absolutely continuous with respect to the law of the Brownian motion $\mathbb{P}^W$,[10] and therefore the *likelihood function* is defined by the Radon-Nikodym derivative of $\mathbb{P}^X$ w.r.t. $\mathbb{P}^W$, i.e. $\frac{d\mathbb{P}^X}{d\mathbb{P}^W}$, which is given by the above expression.

The maximum likelihood estimator given the observed path $(X_t)_{t\in[0,T]}$ is given by

$$\hat{\theta} = \arg\max_{\theta\in\Theta} L(\{X_t\}_{t\in[0,T]};\theta).$$

[10] See Sørensen, International Statistical Review, 2004 or Pavliotis, Stochastic processes and applications, 2014 for a justification, or Liptser and Shiryaev, Statistics of random processes: I. General theory. Vol. 1., 2001 for a proof.

Example

Consider the stationary Ornstein-Uhlenbeck process

$$dX_t = -\alpha X_t\, dt + dW_t,$$

with $X_0 \sim \mathcal{N}\left(0, \frac{1}{2\alpha}\right)$. The log-likelihood function is

$$\log L = -\alpha \int_0^T X_t\, dX_t - \frac{\alpha^2}{2} \int_0^T X_t^2\, dt.$$

From this, the Maximum Likelihood estimator is

$$\hat{\alpha} = -\frac{\int_0^T X_t\, dX_t}{\int_0^T X_t^2\, dt}.$$

To evaluate this estimator, we use a trajectory: given a set of discrete equidistant observations $\{X_j\}_{j=0}^{J}$, we have, for $X_j = X_{j\Delta t}$ and $\Delta X_j = X_{j+1} - X_j$,

$$\hat{\alpha} = -\frac{\sum_{j=0}^{J-1} X_j \Delta X_j}{\sum_{j=0}^{J-1} |X_j|^2 \Delta t}.$$

One can show that this Maximum Likelihood estimator becomes asymptotically unbiased in the *large sample limit* $J \to +\infty$, for Δt fixed.

Discussion, and what to expect tomorrow

- Systems of interacting particles are ubiquitous in applications such as physics, biology, chemistry, life and social sciences. Other (non-discussed) applications include **particle swarm optimisation** or more recently **consensus optimisation**
- Their behaviour can be characterised by McKean–Vlasov or Fokker–Planck equations.
- The latter can be used to, e.g. explore long time behaviour and phase transitions of solutions.
  - ⋆ we saw examples of multi-well and multiscale potentials exhibiting phase transitions
  - ⋆ depending on the parameters, we also observe topology changes in the bifurcation diagrams.
- We also explored inference (parameter estimation) for SDEs (not dependent on law of the process)

Tomorrow:

We will look at inference for McKean–Vlasov equations (and/or the corresponding Fokker–Planck equation) in connection to two different applications.

Thank you for your attention!

Susana N. Gomes

https://warwick.ac.uk/fac/sci/maths/people/staff/gomes

susana.gomes@warwick.ac.uk

WARWICK
THE UNIVERSITY OF WARWICK