

Langevin MC

B. Han, T.M. Hodgson, M. Holden & M. Puza

February 14, 2019

Aim? Sample efficiently from high dimensional distributions of the form $\pi(x) = Z^{-1}e^{-U(x)}$ where Z is a normalising constant.

How? Use Langevin Dynamics

Wait, what LD is an SDE? Yup, use a variety of methods to simulate the dynamics of this SDE: ULA, MALA, LM, HOLA. Both tamed and coordinate-wise tamed!

Why does this work? The invariant measure of the Langevin SDE is exactly what we need!

What goes wrong with ULA? Good question, we can find that out to motivate introducing other methods!

1 Why Bother?

2 Langevin Equation

The Langevin Equation is a stochastic differential equation (SDE) originally developed to model random particle motion. Indeed, molecular dynamics continues to use and propose novel methods based on the equation. It is given by

$$dX_t = -\nabla U'(X) dt + \sqrt{2\gamma} dW_t,$$

where W_t is a d -dimensional Wiener process (Brownian motion). It can be shown that this equation has an invariant measure given by $\pi(x) = Z^{-1}e^{-U(x)}$ with Z a normalising constant – not sure if this is worth verifying ($\mathcal{L}^*\pi = 0$). What requirements on U are necessary for the SDE to be geometrically ergodic?

2.1 Discretisation: Euler-Maruyama/LM/HOLA

We can discretise the Langevin SDE so that paths can be simulated numerically. The classic method is to use the Euler-Maruyama (EM) method. This gives the following iterative scheme:

$$x_{n+1} = x_n - \gamma \nabla U(x) + \sqrt{2\gamma} Z_{n+1}$$

where $Z \sim N(0, 1)$. This is equivalent to $x_{n+1} = \xi$ where $\xi \sim N(x_n - \gamma \nabla U(x), 2\gamma I_n)$

When discretising an SDE, it is important to check that the invariant measure of the SDE is preserved.

In the Euler method, only a first order approximation is used. This can be extended by taking in to consideration more terms in the expansion. This leads to an iterative scheme defined as

2.2 Algorithms

2.2.1 ULA

+++ Error/Convergence bounds, non-asymptotic and asymptotic? +++

2.2.2 Leimkuhler-Matthews

Leimkuhler and Matthews propose a non-Markovian scheme to sample from the target distribution +++ Error/Convergence bounds +++

2.2.3 Higher Order Langevin Approximation

3 MALA

One way to reduce the issues with the ULA algorithm, Roberts + Tweedie proposed the Metropolis adjusted Langevin algorithm (MALA). Adding a Metropolis rejection step preserves the invariant measure and increases the class of potentials for which the algorithm converges. +++ Convergence/Error bounds — Non-asymptotic +++

4 Taming: Coordinate and otherwise

To deal with superlinear potentials was proposed by Roberts + Tweedie 1996 (MALTA), and also Brosse, Durmus, Moulines, Sabanis (201x) (we call tULA). These involve *taming* the potential by scaling it as follows.

$$G_\gamma = \frac{\nabla U(x)}{1 + \gamma \|\nabla U(x)\|}$$

5 Applications to Large Datasets

5.1 Stochastic Gradient Langevin Dynamics