Computational tools III: Markov Chain Monte Carlo and Metropolis - Hastings method

Consider a Markov chain with transition probability $T(x|y)$ for going from y to x.

The update of the marginal distribution given by

$$p_{t+1}(x) = \int T(x|y)p_t(y)\,dy$$

This can be written as

$$p_{t+1}(x) - p_t(x) = \int T(x|y)p_t(y)\,dy - p_t(x)\int T(y|x)\,dy$$

Hence, the *stationary distribution* must fulfil

$$0 = \int T(x|y)p(y)\,dy - \int p(x)T(y|x)\,dy$$

If the transition probability $T(y|x)$ is constructed in such a way that we have

$$T(x|y)p(y) = p(x)T(y|x)$$

we say that the Markov chain fulfills **detailed balance**. The chain is also known as a reversible Markov chain.

The Metropolis - Hastings method

- Define a **proposal distribution** $q(x'|x)$.

- Given a state $x = x_t$ at *step* t generate a new state x' with probability distribution $q(x'|x)$.

- Define **acceptance ratio**

$$A(x'; x) = \min\left(1, \frac{p(x')q(x|x')}{p(x)q(x'|x)}\right)$$

- **Accept** new state $x_{t=1} = x'$ with probability $A(x'; x)$

  Reject new state, ie **keep old** state $x_{t+1} = x$ with probability $1 - A(x'; x)$

Analysis

- We see that this defines a Markov chain with transition probability (assume x' was accepted)

$$T(x'|x) = A(x';x)q(x'|x) \ .$$

- It fulfills detailed balance

$$p(x)T(x'|x) = p(x)q(x'|x)\,\text{min}\left(1, \frac{p(x')q(x|x')}{p(x)q(x'|x)}\right) =$$

$$\text{min}\Big(q(x'|x)p(x), q(x|x')p(x')\Big) = p(x')q(x|x')A(x;x') = p(x')T(x|x')$$

with the stationary distribution $p(x)$.

- Note, that only ratios of probabilities $\frac{p(x')}{p(x)}$ are required. Hence, normalization constants of probabilities are not needed.

- This general method depends on clever choices of proposals q.

Independence sampler

Let $q(x'|x) = q(x')$ independent of x in the Metropolis method.

Then the acceptance probability is

$$A(\mathbf{x}'; \mathbf{x}) = \min\left\{\frac{p(x')q(x)}{p(x)q(x')}, 1\right\}$$

This is similar to a rejection method, but samples are dependent. Again, q should be similar to p to achieve good acceptance rates.

Random walk sampler

This method can be easily applied to continuos states. As the proposal, one often chooses a move

$$\mathbf{x}' = \mathbf{x} + \sqrt{\rho}\mathbf{z}$$

where $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$. This is a **symmetric proposal** with $q(x'|x) = q(x|x')$

The acceptance probability is then

$$A(\mathbf{x}'; \mathbf{x}) = \min \left\{ \frac{p(x')}{p(x)}, 1 \right\}$$

With this form of A, one speaks of a **Metropolis sampler**.

The choice of ρ is important. For large ρ acceptance will be unlikely. Small ρ will lead to high acceptance rates but too a very slow **diffusion**.

Example: Two dimensional Gaussian with $\rho = 1$ and $\rho = 0.1$ (1000 samples).