Metropolis-Hastings

Sometimes Gibbs samples are too correlated

Gibbs is doing local and small steps in the sample space.

Apply rejection sampling to Markov Chains

We are given a family of markov chains, instead of a predefined markov chain. The idea is to pick the best converging candidate.

^ idea of metropolis hastings.

Metropolis-Hastings

For k = 1, 2, …

- Sample x' from a wrong $Q(x^k \to x')$
- Accept proposal x' with probability $A(x^k \to x')$
- Otherwise stay at x^k

A is a 'critic' function. It can tell you whether to go or not. Critic makes sure that markov chain doesn't go to undesired locations.

^^ This means >> (x^{k+1} stays at x^k) $x^{k+1} = x^k$

This the transition probability. It depends on the Q and the critic function

$$T(x \to x') = Q(x \to x')A(x \to x')$$

for all $x \neq x'$

$$T(x \to x) = Q(x \to x)A(x \to x)$$

We could have proposed to move anywhere else, and have this rejected by A >>.

$$+ \sum_{x' \neq x} Q(x \to x')(1 - A(x \to x'))$$

^NOTICE here that the critic is given a 1 -.

How to choose A: $\pi(x') = \sum_x \pi(x)T(x \to x')$

The critic must be CHOSEN by the criteria that the distribution we want to sample from is a stationary function.

The sample distribution is ideal when it is stationary since we know it has already conveged.

Detailed Balance

Lets unravel the 'stationary' requirement here to make life easier and more specified..
We introduce the concept of detailed balance equation.

recall that A markov-chain distribution pi is stationary if the below holds:

$$\pi(x') = \sum_x \pi(x) T(x \to x')$$

Detailed Balance

If going from x to x' is like going reverse (This is the detailed balance condition)

If $$\pi(x)T(x \to x') = \pi(x')T(x' \to x)$$

Then $$\pi(x') = \sum_x \pi(x)T(x \to x')$$

Proof $$\sum_x \pi(x)T(x \to x') = \sum_x \pi(x')T(x' \to x)$$

$$= \pi(x') \sum_x T(x' \to x)$$

$$= \pi(x')$$

^^^^^^^^^^^^^^^^^^^^^^^^

This is 1!

In plain english:

If the detailed balance condtion holds for a distribution pi, then pi is definitely a stationary distribution.

Metropolis-Hastings

For k = 1, 2, …

- Sample x' from a wrong $Q(x^k \to x')$
- Accept proposal x' with probability $A(x^k \to x')$
- Otherwise stay at x^k

$$x^{k+1} = x^k$$

$$T(x \to x') = Q(x \to x')A(x \to x') \quad \text{for all } x \neq x'$$

$$T(x' \to x') = Q(x' \to x')$$

How to choose A:

$$\pi(x)T(x \to x') = \pi(x')T(x' \to x)$$

THis is a somewhat clearer definition of a critic.
Define a critic such that the markov chain that was created
by critic will indeed create this stationary distribution pi.