STAT 403 5/23/25

Lecture 22

MH in practice

HW6 : now due 5/28

Lecture Notes VIII – Markov Chain Monte Carlo

Marina Meila
mmp@stat.washington.edu

Department of Statistics
University of Washington

April 2025

Notation ✓

Gibbs sampling ✓

The detailed balance ✓

Metropolis-Hastings sampling ←

Metropolis-Hastings algorithm

$S(x, x') = Pr(x' | x)$

How to choose S?

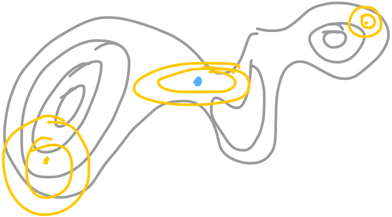
In Proposal distribution $S(x, x') \propto$ transition probability $x \to x'$
no need to be normalized, no need to be symmetric

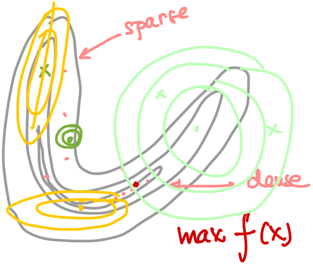
1. Initialize x^0 with some arbitrary values
2. For $t = 1, 2, \ldots$ we will sample sequentially $x^t | x^{t-1}$ as follows
   - 2.1 Sample $x' \sim S(x^{t-1}, x')$
   - 2.2 Compute acceptance probability

$$a(x^{t-1}, x') \;=\; \min\left(1, \frac{P(x')S(x', x^{t-1})}{P(x^{t-1})S(x^{t-1}, x')}\right) = \begin{cases} 1 & \text{if } \quad > \\ & \text{otherwise} \end{cases} \tag{10}$$

   - 2.3 $x^t = \begin{cases} x' & \text{w.p. } a \\ x^{t-1} & \text{w.p. } 1 - a \end{cases}$
   - 2.4 Every T steps (where T is a LARGE number), output x^t

↓ for n samples need nT MH steps

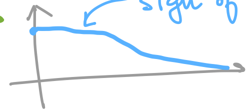




Hamiltonian MC

$S(x^t, x')$ — centered around x^t

- large variance "spread" ⇒ Prob[accept] ≪ 1 → optimum by trial and err.
- small variance ⇒ mixing slow



How to choose "mixing time" T ?

Auto-covariance function

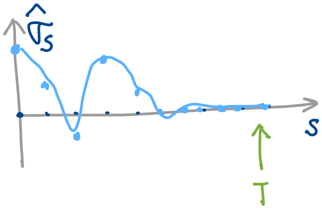
$x^1, x^2, \dots x^t, \dots \sim p$ desired $x \in \mathbb{R}$

$\hat\gamma_s = \mathrm{cov}[x^t, x^{t+s}]$ $\hat\sigma^2 = \hat\gamma_0$

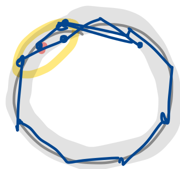
$\hat\sigma^2 = \mathrm{Var}\, X$ from $x^{1:N}$

$\hat\mu = E[x^t] = 0$ w.l.o.g.

$$\hat\gamma_s = \frac{1}{N-s}\sum_{t=1}^{N-s} x^t x^{t+s}$$



$\hat\gamma_s \to 0$ for large s



Does it satisfy the detailed balance?

- If x′ rejected ✓ $\Rightarrow x^t = x^{t-1}$
- If x′ accepted $\Rightarrow x^t = x'$

$$
\begin{aligned}
\mathrm{P}(x', x) &= S(x', x)a(x', x) && (11)\\
P(x)\mathrm{P}(x, x') &= P(x)\mathrm{S}(x, x')\min\left(1, \frac{P(x')\mathrm{S}(x', x)}{P(x)\mathrm{S}(x, x')}\right) && (12)\\
&= \min\left(P(x')\mathrm{S}(x', x), \P(x)\mathrm{S}(x, x')\right) && (13)\\
&= P(x')\mathrm{P}(x', x) \qquad \text{by symmetry} && (14)
\end{aligned}
$$

Recap: What we need to be able to do MH sampling

- To calculate $\mathrm{P}(x)/\mathrm{P}(x')$ but not P itself (okay not to have Z)
- To calculate $\mathrm{S}(x, x')/\mathrm{S}(x', x)$
- To sample from $\mathrm{S}(x, x')$