Recitation: Homework 4
Markov Chain Monte Carlo

10-418/10-618: ML for Structured Data

October 24, 2022

1. Recap of Gibbs Sampling and MH algorithm

   MH algorithm summary

   - Draws a sample x' from $Q(x'|x)$, where x is the previous sample.
   - The new sample x' is accepted or rejected with some probability $A(x'|x) = \min(1, \frac{P(x')Q(x|x')}{P(x)Q(x'|x)})$
   - In case that Q is symmetric ,i.e. $Q(x|x') = Q(x'|x)$ (Gaussian, etc.), the acceptance probability simplifies to $\min(1, \frac{P(x')}{P(x)})$

   pseudo-code for M-H algorithm

   1. Initialize starting state $x^{(0)}$, set t =0
   2. Burn-in: while samples have "not converged":
      - $x = x^{(t)}$
      - t = t+1
      - sample $x* \sim Q(x*|x)$ (draw proposal)
      - sample $u \sim Uniform(0, 1)$ (draw acceptance threshold)
      - if $u < A(x*|x)$: $x^{(t)} = x*$ (accept, make state transition)
      - else: $x^{(t)} = x$ (reject, stay in current state)
   3. Takes samples from $P(x)$: after observing convergence, do the same as 2 to sample from the distribution.

   Gibbs sampling

   - Let $\mathbf{x}^{(1)}$ be the initial assignment to variables.
   - Set t = 1
   - while true:
     - for i = 1...J:
       - sample $\mathbf{x}_i^{(t+1)} \sim p(\mathbf{x}_i|\{\mathbf{x}_j^{(t)}(j \neq i))$
       - set $\mathbf{x}_i^{(t+1)}$ to $\mathbf{x}_i^{(t)}$
       - t = t+1

2. Consider $X_1, ..., X_n$ being i.i.d. Poisson(λ). Show that a Gamma(α, β) prior on λ is a conjugate prior, and find the posterior distribution.

   Likelihood:
   $$L(\lambda) = \prod_{i=1}^{n} \frac{\exp(-\lambda)\lambda^{x_i}}{x_i!} = \frac{\exp(-n\lambda)\lambda^{\sum_i x_i}}{\prod_i x_i!}$$
   Prior:
   $$p(\lambda) \sim Gamma(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\exp(-\beta\lambda)$$
   Posteior:
   $$p(\lambda) \propto L(\lambda)p(\lambda) \propto \lambda \exp(-(\beta+n)\lambda)\lambda^{\sum_i x_i+\alpha-1}$$
   So $p(\lambda)$ is Gamma($\sum_i x_i + \alpha$,$n + \beta$)

3. Gibbs sampling can proceed either rotationally (sweeping through indices i) or randomly (by sampling i). For the purposes of this problem consider the version where i is sampled randomly with probability π_i. **Show that Gibbs sampling satisfies detailed balance.**

   Detailed balance means that for each pair of states x and x', (1) arriving at x then x' and (2) arriving at x' then x are equiprobable. That is,
   $$S(x' \leftarrow x)p(x) = S(x \leftarrow x')p(x').$$

   First, let's consider the transition probability $S(x' \leftarrow x)$. Since Gibbs sampling samples from the full conditionals, this probability is given by:
   $$S(x' \leftarrow x) = \pi_i p(x'_i|x_{\setminus i})$$

   Next, let's compute the left hand side and right hand sides of the detailed balance equation separately.

   LHS:
   $$\begin{aligned} S(x' \leftarrow x)p(x) &= \pi_i p(x'_i|x_{\setminus i})p(x) \\ &= \pi_i p(x'_i|x_{\setminus i})p(x_i|x_{\setminus i})p(x_{\setminus i}) \end{aligned}$$

   RHS:
   $$\begin{aligned} S(x \leftarrow x')p(x') &= \pi_i p(x_i|x'_{\setminus i})p(x') \\ &= \pi_i p(x_i|x'_{\setminus i})p(x'_i|x'_{\setminus i})p(x'_{\setminus i}) \\ &= \pi_i p(x_i|x_{\setminus i})p(x'_i|x_{\setminus i})p(x_{\setminus i}) \\ &= S(x' \leftarrow x)p(x) \end{aligned}$$

where the second to last step follows from the observation that $x'_{\setminus i} = x_{\setminus i}$ because Gibbs sampling holds the other variables constant when updating the ith variable. Thus, detailed balance holds.

Note: to prove detailed balance for the version of Gibbs sampling where we sweep through indices i, we would consider the update after a full sweep.