CS 8803 MCM - Markov Chain Monte Carlo Methods
Homework 3
"Due" Wednesday, April 21

You do not have to hand this homework in, but can if you would like to. We are going to prove all the necessary lemma to show that simulated tempering can give an exponential speedup. Please prove all of the following lemmas, and try to fill in bounds on the mixing rate, where appropriate.

Let $G = K_n$ be the complete graph with n vertices, and let $\beta > 0$ be the inverse temperature. The state space $\Omega = \{\pm 1^n\}$ assigns spins to each vertex, and the stationary distribution of the Ising model is

$$\pi(\sigma) = e^{\beta\left(\sum_{i \neq j} \sigma_i \sigma_j\right)}/Z,$$

where Z is the normalizing constant.

We are considering the Markov chain M that changes one spin at a time and converges to π.

Lemma 1 *When β is sufficiently large, M requires exponential time to converge to the stationary distribution.*

This motivates the need for a new algorithm. We start by showing that M is fast if we restrict to the positive or negative half of the state space, regardless of β.

Lemma 2 *Let $\Omega^+ \subseteq \Omega$ be the set of configurations σ such that the number of vertices assigned + is at least the number assigned -. Show that M restricted to Ω^+ is rapidly mixing for any β.*

To prove this we will use the decomposition theorem. We will decompose $\Omega^+ = \cup \Omega_i$ where the union is over $n/2 \leq i \leq n$ and Ω_i is the set of configurations with exactly i +'s.

Lemma 3 *The 2-step Markov chain M^2 restricted to Ω_i is rapidly mixing.* (Hint: Use coupling.)

Lemma 4 *The projection has $n/2$ vertices, where*

$$\pi_p(i) = \binom{n}{i} e^{\beta\left(\binom{i}{2} + \binom{n-i}{2}\right)}/Z_p.$$

The projection chain that uses Metropolis probabilities to move from i to $i-1$ or $i+1$ is rapidly mixing.

Now use Lemmas 3 and 4 and the Decomposition Theorem to prove Lemma 2. Clearly the same proof shows that M restricted to $\Omega^- = \Omega \setminus \Omega^+$ is also rapidly mixing.

We now introduce the tempering chain $\widehat{M}$. The state space is

$$\widehat{\Omega} = \Omega \times [m],$$

where $[m] = \{0, ..., m\}$. The stationary distribution of the tempering chain is $\widehat{\pi}(\sigma, i) = \frac{1}{m+1}\pi_i(\sigma)$, where

$$\pi_i(\sigma) = exp(\beta_i(\sum_{i \neq j} \sigma_i \sigma_j)/Z$$

and $\beta_i = \beta * i/m$. The Markov chain $\widehat{M}$ either performs level moves (pick i and use M with distribution π_i) or temperature moves (pick $(i, d) \in [m] \times \{+1, -1\}$ and move from (σ, i) to $(\sigma, i + d)$ with the appropriate Metropolis probabilities if $i + d \in [m]$.

We will prove the following main theorem:

Theorem 5 *The Markov chain $\widehat{M}$ is rapidly mixing on $\widehat{\Omega}$ for any β.*

Again we will use decomposition. We will start by partitioning $\widehat{\Omega}$ into $2m + 1$ pieces by dividing each Ω_i (except the bottom one when i=0) into Ω_i^+ and Ω_i^-.

Notice that the restriction of $\widehat{M}$ to the bottom distribution is rapidly mixing (just use coupling), and the restriction to every other set in the partition is rapidly mixing by Lemma 2.

To analyze the projection chain arising from this decomposition of $\widehat{\Omega}$, we consider first the Metropolis chain on the $2m + 1$ points in the projection. The center point has probability $1/(m + 1)$ and all other point have probability $1/2(m + 1)$. It is easy to show that the Metropolis chain (that assigns probability 1 or 1/2 to each transition) is rapidly mixing.

Show that we can now use the Decomposition Theorem for a second time to prove Theorem 5.

If you are feeling really ambitious, consider the generalized distributions that are not necessarily symmetric (such as the exponential distribution on state space $\{-m_1, ..., m_2\}$ with

$$\pi(i) = e^{|i|}/Z$$

where

$$Z = \sum_{i=-m_1}^{m_2} e^{|i|}.$$

The easiest way to do this is to allow moves from the left-hand point (representing Ω_i^+ and Ω_i^-) and then use the comparison theorem.