Stats 318: Lecture # 11

Agenda: MCMC

- Cluster algorithms
- The Swendsen-Wang algorithm
- The Wolff's algorithm
- Application to the Ising model

Gibbs sampling

- Wish to sample from $\pi(x_1, \ldots, x_n)$
- Partition coordinates into subgroups $g_1, \ldots, g_k$: $\pi(x_{g_1}, \ldots, x_{g_k})$
- Pick a subset and resample group of coordinates X_g given other coordinates

e.g. 2 blocks

Repeat

- Sample from $\pi(x_{g_1} \mid x_{g_2})$
- _____ $\pi(x_{g_2} \mid x_{g_1})$

Until convergence

Phase transition in Ising model

$$\pi(x) \;\propto\; \exp\left(\beta \sum_{v \sim v'} x_v x_{v'}\right) \qquad\qquad \beta = \frac{1}{T}$$

- When temperature is high, all spins behave nearly independently (no long-range correlation)
- When temperature is below a critical temp T_c, all the spins tend to be aligned (cooperative performance)
- Single site update Gibbs sampler slows down rapidly once T is approaching T_c or is below T_c : e.g.

$$\mathbb{P}(X_v = -1 \mid \text{neighbors} = +1) = (1 + e^{2\beta \#\text{neighbors}})^{-1} \ll 1$$

- Swendsen & Wang algorithm almost completely eliminates critical slow down

Data augmentation

$$\pi(x) \;\propto\; e^{\beta \sum_{v \sim v'} x_v x_{v'}} = \prod_{v \sim v'} e^{\beta x_v x_{v'}}$$

Augment state space by introducing bond variables $\{u_e\}$ taking values in $\left[0, e^{\beta}\right]$ and consider

$$\pi(x,u) \;\alpha \prod_{e \,=\, (v,v') \in E} 1\left\{0 \le u_e \le e^{\beta x_v x_{v'}}\right\}$$

- Marginal distribution is $\pi(x)$
- Conditional distributions
  - $\pi(u|x)$ is a product of uniform dist. (with range depending on two neighboring spins)
  - $\pi(x|u)$

$$\text{: if } \begin{cases} u_e > e^{-\beta} & \Longrightarrow x_v = x_{v'} \\ u_e \le e^{-\beta} & \Longrightarrow \text{no constraints} \end{cases} \qquad \therefore \quad \pi(x \mid u) \;\propto \prod_{e : u_e > e^{-\beta}} \mathbb{1}(x_v = x_{v'})$$

Clustering

- Based on u : cluster sites according to whether they have a mutual bond : i.e. $u_e > e^{-\beta}$
- All sites in a cluster should take identical values
- Conditional on clusters, all configurations that do not violate the cluster homogeneity are equally likely

Simpler augmented model

- Only use auxiliary bonding variable $b_e = 1\{u_e > e^{-\beta}\}$

$$\pi(x, b) \propto \prod_{e: x_v \neq x_{v'}} 1(b_e = 0) \prod_{e: x_v = x_{v'}} \Big[1(b_e = 0) + (e^{2\beta} - 1)1(b_e = 1)\Big]$$

- Clustering done by connecting all neighboring sites with bond value 1
- Conditional on b, spin value of one cluster independent of the other clusters

Swendsen-Wang algorithm [1987]

Iterate between sampling from $\pi(b \mid x)$ & $\pi(x \mid b)$

1. For given spin configuration x, form bond variables

$$b_e = \begin{cases} 1 \text{ wp} & 1 - e^{-2\beta} \quad \text{if } x_v = x_{v'} \\ 0 & e^{-2\beta} \end{cases}$$

$$b_e = 0 \qquad \text{if } x_v \neq x_{v'}$$

2. Conditional on b, update spins by sampling uniformly on all compatible configurations; i.e. clusters take on iid $\mathrm{Ber}(1/2)$ values (every cluster flipped wp $1/2$)
   Clusters connect neighboring sites with bond value 1

Numerical example

Swendsen-Wang algorithm (click link)

Wolff's algorithm

Wolff (1989) introduced a modification which significantly enhances the SW algorithm

- For a given configuration x, randomly pick a site v & grow recursively a bonded set from it
  - Check all the unchecked neighboring sites of a current set C^{old}; add a bond between a neighboring site and C^{old} in the same way as in the SW algorithm
  - Add newly bonded sites to C^{old} & form C^{new}
  - Stop recursion when there is no unchecked site to add. Final cluster is C
- Flip all the signs in C (no random sampling)

Difference with SW

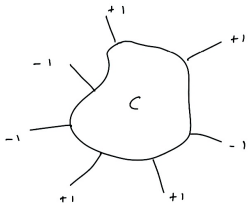
In each iteration

- Only one cluster is constructed
- All spins are changed

Interpretation as metropolis algorithm

Suppose C has $n+m$ neighboring sites

n links to -1 $\quad$ m links to $+1$

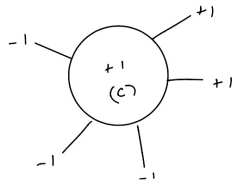


If C is all $1 \Longrightarrow$ flipping to -1

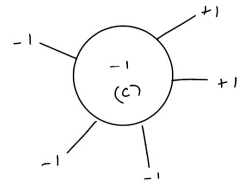
Probability ratio

$$\frac{\mathbb{P}(\text{new})}{\mathbb{P}(\text{old})} = \frac{e^{-\beta \sum_v x_v}}{e^{+\beta \sum_v x_v}} = \frac{e^{-\beta(m-n)}}{e^{\beta(m-n)}} = e^{-2\beta(m-n)}$$

Consider process of building C (proposal)



old



new

$$\frac{\mathbb{P}(\text{old} \longrightarrow \text{new})}{\mathbb{P}(\text{new} \longrightarrow \text{old})} = \frac{\mathbb{P}(\text{breaking } m \text{ bonds})}{\mathbb{P}(\text{breaking } n \text{ bonds})} = \frac{e^{-2\beta m}}{e^{-2\beta n}} = e^{-2\beta(m-n)}$$

Conclusion: $$\frac{\mathbb{P}(\text{new})\mathbb{P}(\text{new} \longrightarrow \text{old})}{\mathbb{P}(\text{old})\mathbb{P}(\text{old} \longrightarrow \text{new})} = e^{-2\beta(m-n)} e^{2\beta(m-n)} = 1$$

Metropolis always accepts the move!