MCAA lecture 11

Graph coloring: optimization algorithm

$G = (V, E)$ $|V| = N$

q^N possible colorings $x = (x_v, v \in V) \in \{1..q\}^V$

proper coloring: no two adjacent vertices share the same color

Aim today: to find a proper coloring

(no more $q > 3\Delta$)

Define $f(x) = \#$ edges with a conflict

(so $f(x) = 0$ iff x is a proper coloring)

$$\pi_\infty(x) = \begin{cases} 1/Z_\infty & \text{if } x = \text{proper coloring} \\ 0 & \text{otherwise} \end{cases}$$

where Z_∞ = number of proper colorings

$$\rightsquigarrow \pi_\beta(x) = \frac{1}{Z_\beta} \cdot \exp(-\beta f(x)) \qquad \beta > 0$$

<u>inverse temperature</u>

Metropolis algorithm

a) <u>base chain</u>: choose $v \in V$ and $c \in \{1..q\}$ unif. at random and recolor $x_v = c$

$$\varphi_{xy} = \begin{cases} \frac{1}{Nq} & \text{if } y \sim x \quad (d(x,y)=1) \\ \frac{1}{q} & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

(irred. aper. symm.) ✓

b) <u>acceptance probabilities</u> (next page)

$$\pi_\beta(x) = \frac{e^{-\beta f(x)}}{Z_\beta}$$

$$a_{xy} = \min\left(1, \frac{\pi_\beta(y)}{\pi_\beta(x)}\right)$$

$$= \begin{cases} 1 & \text{if } \pi_\beta(y) \geq \pi_\beta(x) \\ \pi_\beta(y)/\pi_\beta(x) & \text{if } \pi_\beta(y) < \pi_\beta(x) \end{cases}$$

$$= \begin{cases} 1 & \text{if } f(y) \leq f(x) \\ \exp(-\beta(f(y)-f(x))) & \text{if } f(y) > f(x) \end{cases}$$

c) Metropolis chain: $P_{xy} = \begin{cases} \varphi_{xy}\, a_{xy} & \text{if } y \sim x \\ \varphi_{xx} + \dots & \text{if } y = x \end{cases}$

How to choose β? A theoretical answer

Sampling from π_β does not always lead to global minimum of f. Set ε such that

$$1-\varepsilon = \sum_{\substack{x\in S \\ x = \text{global min}}} \pi_\beta(x)$$

Let $f_0 = \min\limits_{x\in S} f(x)$, $f_1 = \min\limits_{\substack{x\in S \\ f(x)\neq f_0}} f(x)$, $f_2 = \min\limits_{\substack{x\in S \\ f(x)\neq f_0, f_1}} f(x)$

N_k = the number x's s.t. $f(x) = f_k$

$$1-\varepsilon = \sum_{x=\text{global min}} \pi_\beta(x) = \frac{N_0 e^{-\beta f_0}}{Z_\beta}$$

$$Z_\beta = \sum_{x\in S} e^{-\beta f(x)} = \sum_{k\geq 0} N_k\, e^{-\beta f_k}$$

$$= N_0 e^{-\beta f_0} + N_1 e^{-\beta f_1} + N_2 e^{-\beta f_2} + \ldots$$

$$\simeq N_0 e^{-\beta f_0} + N_1 e^{-\beta f_1} \quad (\beta \text{ large})$$

So $$1-\varepsilon \simeq \frac{N_0 e^{-\beta f_0}}{N_0 e^{-\beta f_0} + N_1 e^{-\beta f_1}} = \frac{1}{1+\frac{N_1}{N_0} e^{-\beta(f_1-f_0)}}$$

$$\simeq 1 - \frac{N_1}{N_0} e^{-\beta(f_1-f_0)} \quad \leadsto \boxed{\beta \simeq \frac{1}{f_1-f_0} \log\left(\frac{N_1}{\varepsilon N_0}\right)}$$

In practice : simulated annealing

β = inverse temperature

β low: high temperature regime

$\pi_\beta \sim$ uniform : exploration of S

slow increase $\downarrow$

β high: low temperature regime

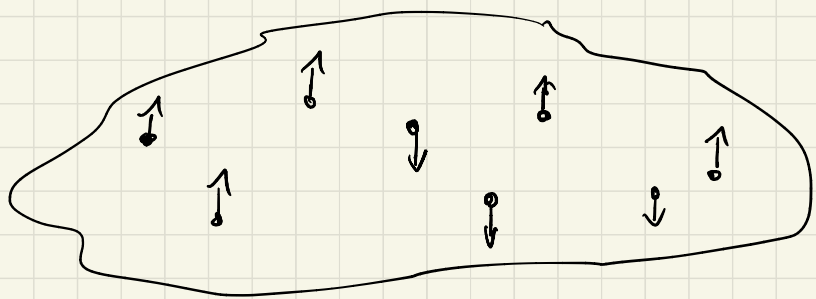
$\pi_\beta \sim \pi_\infty$

Ising model

$G = (V, E)$, $|V| = N$

spin configurations: $\sigma = (\sigma_v, v \in V)$

$S = \{+1, -1\}^V$ $\qquad$ $\sigma_v \in \{+1, -1\}$



$\uparrow$ h = external magnetic field

Hamiltonian (energy):

$$H(\sigma) = -\sum_{(v,w)\in E} \underbrace{J_{vw}}_{\text{interaction "strength"}} \sigma_v \sigma_w - \sum_{v \in V} h_v \sigma_v$$

external magnetic field

Physical principle:

Spins tend to minimize their energy

Gibbs distribution:

$$\mu_\beta(\sigma) = \frac{\exp(-\beta H(\sigma))}{Z_\beta}$$

where $Z_\beta = \sum_{\sigma \in S} \exp(-\beta H(\sigma))$

and $\beta = \frac{1}{T}$ real inverse temperature

(β low $\longleftrightarrow$ μ_β ~ uniform)

(β high $\Longleftrightarrow$ μ_β concentrated on states of low energy)

Interactions:

- $J_{vw} > 0 \quad \forall v,w$: σ_v & σ_w tend to align to each other
(ferromagnetic model)
- $J_{vw} < 0 \quad \forall v,w$: σ_v & σ_w tend to go in opposite directions
(antiferromagnetic model)

Magnetization:

- $m(\sigma) = \frac{1}{N} \sum_{v \in V} \sigma_v$

- $\langle m \rangle_\beta = \langle \frac{1}{N} \sum_{v \in V} \sigma_v \rangle_\beta$

$=$ average magnetization at a given inverse temperature $\beta > 0$.

$$= \sum_{\sigma \in S} m(\sigma)\, \mu_\beta(\sigma) = \sum_{\sigma \in S} m(\sigma) \cdot \frac{e^{-\beta H(\sigma)}}{Z_\beta}$$

The Curie-Weiss model

$G=(V,E)$ = complete graph [mean field model]

("The" Ising model: G = grid, nearest neighbour interactions)

$J_{vw} \equiv J/N > 0 \quad \forall (v,w) \in E$ (ferromagnetic)

$h_v = h \in \mathbb{R} \quad \forall v \in V$

In this case:

$$\mathcal{H}(\sigma) = -\sum_{(v,w)\in E} \frac{J}{N}\sigma_v\sigma_w - \sum_{v\in V} h\cdot\sigma_v$$

$$= -\frac{J}{2N}\Big(\underbrace{\sum_{v,w\in E}\sigma_v\sigma_w}_{N^2 m(\sigma)^2} - \underbrace{\sum_{v\in V}\sigma_v^2}_{N}\Big) - h\cdot\underbrace{\sum_{v\in V}\sigma_v}_{N\cdot m(\sigma)}$$

$$= -\frac{JN}{2}\left(m(\sigma)^2 - \frac{1}{N}\right) - hN\cdot m(\sigma)$$

$$= -N\left(+\frac{J}{2}\underline{\underline{m(\sigma)^2}} - \frac{J}{2N} + h\,\underline{\underline{m(\sigma)}}\right)$$

depends only on $m(\sigma)$

Fix now $m \in [-1, +1]$:

$$\mu_\beta\left(\left\{\sigma \in S : \frac{1}{N}\sum_{v \in V} \sigma_v = m\right\}\right)$$

$$= \sum_{\substack{\sigma \in S: \\ m(\sigma) = m}} \frac{\exp(-\beta H(\sigma))}{Z_\beta}$$

$$= \frac{1}{Z_\beta} \sum_{\substack{\sigma \in S: \\ m(\sigma) = m}} \exp\left(+\beta N\left(\frac{J m(\sigma)^2}{2} - \frac{J}{2N} + h m(\sigma)\right)\right)$$

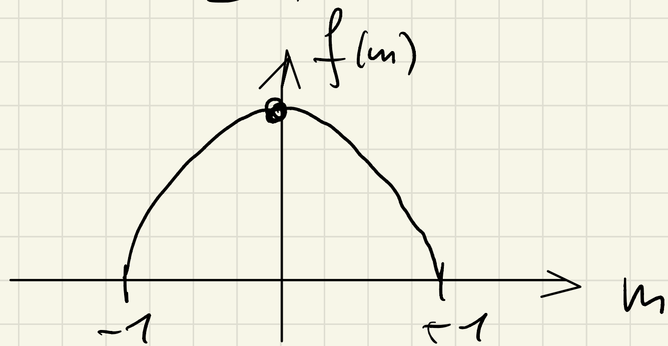
$$= \frac{1}{Z_\beta} \cdot \underbrace{\#\{\sigma : m(\sigma) = m\}}_{\simeq \exp\left(N\, h_0\left(\frac{1+m}{2}\right)\right)} \cdot \underbrace{\exp\left(\beta N\left(\frac{J m^2}{2} - \frac{J}{2N} + hm\right)\right)}$$

where $h_0(p) = -p\log p - (1-p)\log(1-p)$

So $\mu_\beta(\{\sigma : m(\sigma) = m\}) \doteq \exp\Big(N \underbrace{\Big(h_0\big(\tfrac{1+m}{2}\big) + \beta\big(\tfrac{J m^2}{2} + hm\big)\Big)}_{= f(m)}\Big)$

$$f(m) = h_0\left(\frac{1+m}{2}\right) + \beta\left(\frac{Jm^2}{2} + hm\right)$$

<u>$h=0$</u> :

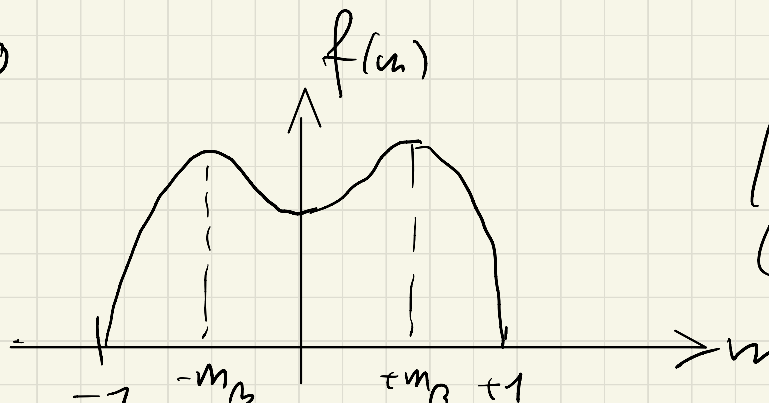


$\beta J < 1$

(high temp.)

maximum in $m=0$

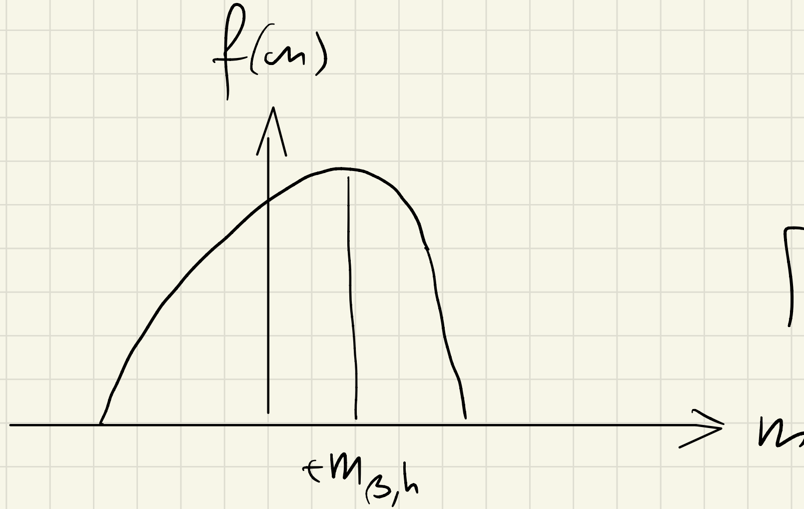
$\langle m \rangle_\beta = 0$



$\beta J > 1$

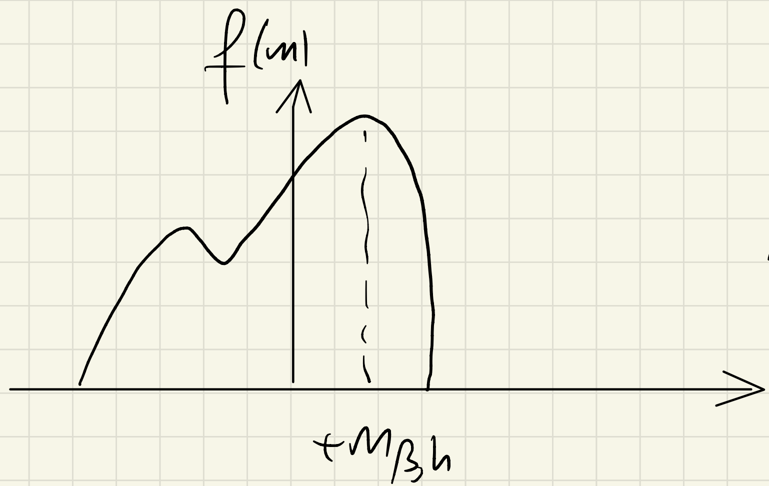
(low temp.)

$\underline{h > 0}$:



$\beta J < 1$

$\langle m \rangle_\beta > 0$



$\beta J > 1$