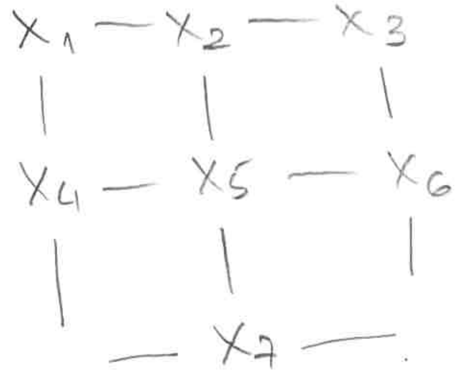
Gibbs

$\underline{X} = (X_1 \dots X_n) \in \{\pm 1\}^n = \Omega \qquad |\Omega| = 2^n$

$V = \{1, \dots n\}$ variables, $E = \{(i,j)\} \subseteq V \times V$ edges



$\pi(\underline{x}) = \frac{1}{Z} e^{-\beta \varphi(\underline{x})}$

$\beta > 0$

$$\varphi(\underline{x}) = \sum_{i=1}^{n} h_i x_i + \sum_{\substack{(i,j) \in E \\ i \sim j}} h_{ij} x_i x_j$$

$h_{ij} > 0$

for simplicity $\beta = 1$ $\qquad \varphi(\underline{x}) \nearrow \Rightarrow P(\underline{x}) \searrow$

ex Prior: $P_0(\underline{x}) = \frac{1}{Z_0} e^{-\sum h_{ij} x_i x_j} \qquad h_{ij} > 0$

Likelihood $P(\underline{y} | \underline{x}) = \frac{1}{Z_1} e^{-\sum_i \underbrace{y_i}_{h_i} x_i} \qquad y_i \in \{\pm 1\}$

$$P(\underline{y}, \underline{x}) = \frac{1}{Z_0 Z_1} e^{-\sum y_i x_i - \sum_{i \sim j} h_{ij} x_i x_j}$$

$$P(\underline{x} | \underline{y}) = \frac{P(\underline{y}, \underline{x})}{P(\underline{y})} = \frac{e^{- \dots}}{Z}$$

$$Z = \sum_{\underline{x}} e^{-\varphi(\underline{x})} \qquad \text{partition fction}$$

Ex: What if we wanted $P(\underline{y})$?

2

Tractable $\varphi(\underline{x})$ || $\pi(\underline{x})/\pi(\underline{x}')$ tractable

Intract: Z || $\pi(\underline{x})$ intract.

Init. $\underline{x}^0$

For $t = 1, 2, \ldots$

pick $i \in V$

$$p_i = \frac{\pi_i^+}{\pi_i^- + \pi_i^+}$$

$x_i^t \leftarrow 1$ w.p p_i

$\leftarrow 0$ w.p $1-p_i$

if "mixed": output x^t, set $x^0 \leftarrow x^t \ldots$

$\underline{x}_{-i} = \quad \in \{\pm 1\}^{n-1}$

$\underline{x}_{-i,+} = (x_i = +, \underline{x}_{-i} = \underline{x}_{-i}^t)$ $\Big\} \in \{\pm 1\}^n$

$\underline{x}_{-i,-}$

$$\frac{\pi_i^+}{\pi_-} = e^{-2h_i - 2\sum_{j \sim i} h_{ij} x_j^{t-1}}$$

$$p_i = \frac{\pi_+/\pi_-}{1 + \pi_+/\pi_-} = \frac{\pi(\underline{x}^t)}{\pi(\underline{x}_{-i}^t)} = \frac{\pi(\underline{x}^t)}{\pi(\underline{x}_{-i}^{t-1})}$$

High level | Example

Algo

if ergodic $\Rightarrow$

$\tilde{\pi} = \pi$

Mixing time

Mixing: 1) $t > 10^5 \ldots 10^6 \ldots$

2) Proof: construction of $T_{\underline{x}, \underline{x}'}$ (conductance)

3) auto corr

4) "perfect sampling"

Detailed balance $\Rightarrow$ π stationary MH

$$\pi(\underline{x})\, T_{\underline{x},\underline{x}'} = \pi(\underline{x}')\, T_{\underline{x}'\underline{x}} \qquad Pr(\underline{x}^{t+1}) = \sum_{\underline{x}^t} \pi(\underline{x}^t)\, T_{\underline{x}^t x^{t+1}} = \sum_{\underline{x}^t} \pi(x^{t+1})\, T_{x^{t+1}, x^t} = \pi(x^{t+1})$$

Gibbs stationary

$$Pr(\underline{x}^t) = \sum_{\underline{x}^{t-1}} \pi(\underline{x}^{t-1})\, T_{\underline{x}^{t-1}, x^t} =$$

$$= \sum_{\underline{x}} \pi(x_i \mid \underline{x}_{-i})\, \pi(\underline{x}_{-i}) \begin{cases} \dfrac{\pi(\underline{x}^{t})}{\pi(\underline{x}^t_{-i})} & \underline{x}^{t-1}_{-i} = \underline{x}^t_{-i} \\ 0 & \text{otherw} \end{cases}$$

$$= \underbrace{\sum_{x_i^{t-1}} \pi(x_i^{t-1} \mid \underline{x}_{-i})}_{1}\, \pi(\underline{x}_{-i})\, \frac{\pi(\underline{x}^t)}{\pi(\underline{x}^t_{-i})} = \pi(\underline{x}^t)$$

MH Algorithm

$$S(\underline{x}, \underline{x}') = S(\underline{x}', \underline{x})$$

Init $\underline{x}^0$

for $t = 1, 2 \ldots$

$\underline{x}' \sim S(\underline{x}', \underline{x}^{t-1})$ proposal distribution

$$a(\underline{x}, \underline{x}') = \min\left(1, \frac{\pi(\underline{x}')\, S(x', x)}{\pi(\underline{x}^t)\, S(x, x')}\right) \begin{cases} 1 & \pi(\underline{x}') \geq \pi(\underline{x}^t) \\ \dfrac{\pi(\underline{x}')\, S_{x x'}}{\pi(\underline{x}^t)\, S_{x' x}} & \text{otherwise} \end{cases}$$

$$\underline{x}^t = \begin{cases} \underline{x}' & \text{w.p. } a \\ \underline{x}^{t-1} & \text{w.p. } 1-a \end{cases}$$

$$\pi(\underline{x})\, T_{\underline{x},\underline{x}'} = \pi(\underline{x})\, \underbrace{S(\underline{x},\underline{x}')\, a(\underline{x},\underline{x}')}_{T_{\underline{x}\underline{x}'} \text{ when } \underline{x}' \neq \underline{x}} = S(\underline{x},\underline{x}') \min(\pi(\underline{x}), \pi(\underline{x}')) = \pi(\underline{x}')\, S(\underline{x}', \underline{x})$$

$\pi(\underline{x})\, T_{\underline{x}\underline{x}} = \pi(\underline{x})\, T_{\underline{x}\underline{x}}$ trivial when $\underline{x}' = \underline{x}$

<u>Generalization</u> $\quad S(\underline{x},\underline{x}') \neq S(\underline{x}',\underline{x})$

$$\bar{a}(\underline{x},\underline{x}') = \min\left(1, \frac{\pi(\underline{x}')\, S(\underline{x}',\underline{x})}{\pi(\underline{x})\, S(\underline{x},\underline{x}')}\right)$$

Rejectionless: $\quad W = \sum_{\underline{x}'} S(\underline{x},\underline{x}')\, a(\underline{x},\underline{x}')$ very small

$$P(n) = W(1-W)^{n-1}$$

n ← number rejection

if W known, sample $n \sim P(n)$ count state $\underline{x}$ n times