FYS5429, MARCH 29, 2023

Probabilistic models:

Directed graph

$$t_0 \longrightarrow t_1 \longrightarrow t_2$$

$p(t_0)$ $\quad$ $p(t_1|t_0)$ $\quad$ $p(t_2|t_1)$

$$p(t_0, t_1, t_2) = p(t_0)\, p(t_1|t_0)\, p(t_2|t_1)$$

Bayesian statistics

undirected graph

$$t_0 \,—\, t_1 \,—\, t_2$$

$p(t_0)$ $\quad$ $p(t_1)$ $\quad$ $p(t_2)$

Markov chain + Monte Carlo sampling $= (MC)^2$

- $(MC)^2$
- Bayesian statistics
- probability model

$$p(x) = \frac{\bar{p}(x)}{Z}$$

normalization factor, partition statistical

$$Z = \int_{x \in D} \bar{p}(x)\,dx$$

$Z \leq |M| < \infty$

use Energy-based model (Boltzmann)

$$\bar{p}(x) = \exp(-E(x))$$

$E(x)$ known as energy function

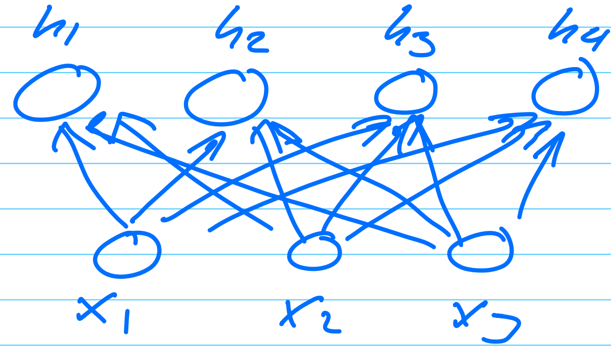
$\exp(x) > 0 \Rightarrow$ a probability which is always larger than zero.

Discrete $p(x)$

$$Z = \sum_{j \in D} \exp(-E(x_j))$$

$$p(x_i) = \frac{\exp(-E(x_i))}{Z}$$

$$E(x_i) = -b^T x_i - c^T h_i - x_i^T W h_i$$



hidden layer

visible/input layer

x could be a discrete variable (binary) or continuous, same with h.

Binary - Binary model
Gaussian - Binary — " —
⋮

Restricted Boltzmann machine

$$p(x) = \frac{\exp(-E(x))}{Z}$$

- RBM
- Variational AEs
- GANs = Generalized adversarial networks

Markov Chain Monte Carlo is the way we sample different variables x, h; sampling rules:

- Gibbs sampling
- Metropolis-Hasting sampling

For the modeling and optimization we will need to optimize b, c, W.

Define the Free energy

$$F(x) = -\log\left[\sum_h \exp(-E(x,h))\right]$$