RBM 2hidden

thou

August 2018

1 Introduction

consider a RBM with 2 hidden unit h_1, h_2, the input is σ,which the dimension is N, the inverse temperature is β,so the joint probability is

$$P(\sigma, h_1, h_2) = \frac{1}{Z} \exp \frac{\beta}{\sqrt{N}} (h_1 \sum_{i=1}^{N} \xi_i^1 \sigma_i + h_2 \sum_{i=1}^{N} \xi_i^2 \sigma_i) \tag{1}$$

sum over the hidden unit, we can get,

$$P(\sigma) = \frac{1}{Z} [2 \cosh \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} (\xi_i^1 + \xi_i^2) \sigma_i + 2 \cosh \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} (\xi_i^1 - \xi_i^2) \sigma_i] \tag{2}$$

so we can get the partition function as

$$Z = \sum_{\sigma_i} [2 \cosh \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} (\xi_i^1 + \xi_i^2) \sigma_i + 2 \cosh \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} (\xi_i^1 - \xi_i^2) \sigma_i] \tag{3}$$

the final results is

$$Z = 2 \prod_{i=1}^{N} [2 \cosh \frac{\beta}{\sqrt{N}} (\xi_i^1 + \xi_i^2)] + 2 \prod_{i=1}^{N} [2 \cosh \frac{\beta}{\sqrt{N}} (\xi_i^1 - \xi_i^2)] \tag{4}$$

by using the equation for small x

$$\cosh x = \exp(\frac{x^2}{2}) \tag{5}$$

we simplify the paration function as:

$$Z = 2^{N+1} e^{\beta^2} (2 \cosh \frac{\beta^2}{N} \sum_{i=1}^{N} \xi_i^1 \xi_i^2) \tag{6}$$

the message passing for this systerm is:

$$P_{i\to a}(\xi_i^1,\xi_i^2)\propto\prod_{b\in\partial i/a}\hat{P}_{b\to i}(\xi_i^1,\xi_i^2) \tag{7}$$

and

$$\hat{P}_{b\to i}(\xi_i^1,\xi_i^2)\propto Z^{-1}\cosh\beta X\cosh\beta Y\prod_{j\in\partial b/i}P_{j\to b}(\xi_j^1,\xi_j^2) \tag{8}$$

where

$$X=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_i^1\sigma_i^a \tag{9}$$

and

$$Y=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_i^2\sigma_i^a \tag{10}$$

by use

$$\omega_1=\sum_{j\in\partial b/i}\frac{1}{\sqrt{N}}\xi_j^1\sigma_j^b \tag{11}$$

$$\omega_2=\sum_{j\in\partial b/i}\frac{1}{\sqrt{N}}\xi_j^2\sigma_j^b \tag{12}$$

by relaxed BP,

$$m_{b\to i}^1=\sum_{j\in\partial b/i}\frac{1}{\sqrt{N}}\sigma_j^n m_{j\to b}^1 \tag{13}$$

$$m_{b\to i}^2=\sum_{j\in\partial b/i}\frac{1}{\sqrt{N}}\sigma_j^n m_{j\to b}^2 \tag{14}$$

and

$$\hat{\sigma}_{b\to i}^1=1-\frac{1}{N}\sum_{j\in\partial b/i}(m_{j\to b}^1)^2 \tag{15}$$

$$\hat{\sigma}_{b\to i}^2=1-\frac{1}{N}\sum_{j\in\partial b/i}(m_{j\to b}^2)^2 \tag{16}$$

and the correlation is :

$$\rho_{b\to i}=\frac{\frac{1}{N}\sum_{j\in\partial b/i}(Q_{j\to b}-m_{j\to b}^1 m_{j\to b}^2)}{\hat{\sigma}_{b\to i}^1\hat{\sigma}_{b\to i}^2} \tag{17}$$

where

$$Q_{j\to b}=\frac{\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{++}+\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{--}-\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{+-}-\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{-+}}{\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{++}+\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{--}+\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{+-}+\prod_{a\in\partial j/b}\hat{P}_{a\to j}^{-+}} \tag{18}$$

we can get

$$\hat{P}_{b\to i} = \cosh^{-1}(\frac{\beta^2}{N}\sum_j Qj)e^{\frac{1}{2}\beta^2(\hat{\sigma}^1_{b\to i}+\hat{\sigma}^2_{b\to i})}[e^{\beta^2\rho_{b\to i}\hat{\sigma}^1_{b\to i}\hat{\sigma}^2_{b\to i}}\cosh\beta X+e^{-\beta^2\rho_{b\to i}\hat{\sigma}^1_{b\to i}\hat{\sigma}^2_{b\to i}}\cosh\beta Y] \tag{19}$$

where

$$X = \frac{1}{\sqrt{N}}(\xi_i^1+\xi_i^2)\sigma_i^b + m^1_{b\to i} + m^2_{b\to i} \tag{20}$$

$$Y = \frac{1}{\sqrt{N}}(\xi_i^1-\xi_i^2)\sigma_i^b + m^1_{b\to i} - m^2_{b\to i} \tag{21}$$

and

$$Q_j = \frac{\prod_{a\in\partial j}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j}\hat{P}^{--}_{a\to j}-\prod_{a\in\partial j}\hat{P}^{+-}_{a\to j}-\prod_{a\in\partial j/b}\hat{P}^{-+}_{a\to j}}{\prod_{a\in\partial j}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j}\hat{P}^{--}_{a\to j}+\prod_{a\in\partial j}\hat{P}^{+-}_{a\to j}+\prod_{a\in\partial j}\hat{P}^{-+}_{a\to j}} \tag{22}$$

$$m^1_{j\to b} = \frac{\prod_{a\in\partial j/b}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{+-}_{a\to j}-\prod_{a\in\partial j/b}\hat{P}^{-+}_{a\to j}-\prod_{a\in\partial j/b}\hat{P}^{--}_{a\to j}}{\prod_{a\in\partial j/b}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{--}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{+-}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{-+}_{a\to j}} \tag{23}$$

$$m^2_{j\to b} = \frac{\prod_{a\in\partial j/b}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{-+}_{a\to j}-\prod_{a\in\partial j/b}\hat{P}^{--}_{a\to j}-\prod_{a\in\partial j/b}\hat{P}^{+-}_{a\to j}}{\prod_{a\in\partial j/b}\hat{P}^{++}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{--}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{+-}_{a\to j}+\prod_{a\in\partial j/b}\hat{P}^{-+}_{a\to j}} \tag{24}$$