We consider an RBM with Gaussian input units.

The energy function suggested in [Salakhutdinov and Hinton, 2008] is

$$\text{Energy}(x, h) = \sum_j \frac{(x_j - b_j)^2}{2\sigma_j^2} - c^T h - \sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j}$$

where $W_{\cdot j}$ is the j-th column of W and c^T the transpose of c.

We will derive and justify all the equations needed for the RBM update algorithm, that is:

I) $\mathbb{P}(h|x)$

II) $\mathbb{P}(x|h)$ (or way to sample x from $\mathbb{P}(x|h)$)

III) $\dfrac{\partial \text{Energy}(x, h)}{\partial \theta}$

IV) $\mathbb{E}_{\mathbb{P}(h|x)}\left[\dfrac{\partial \text{Energy}(x,h)}{\partial \theta}\right]$

As with RBM, ~~with~~ we have:

$$\mathbb{P}(x, h) = \frac{e^{-\text{Energy}(x,h)}}{Z}$$

where $\cancel{Z = \sum_{x,h} e^{-\text{Energy}(x,h)}}$ is the "partition function".

$$Z = \iint e^{-\text{Energy}(x,h)} \, dx \, dh$$

iii) We have:

- $$\frac{\partial \, Energy(x,h)}{\partial b_k} = \frac{\partial}{\partial b_k}\left(\sum_j \frac{(x_j - b_j)^2}{2\sigma_j^2} - c^T h - \sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j}\right)$$

$$= \frac{\partial}{\partial b_k}\left(\sum_j \frac{(x_j - b_j)^2}{2\sigma_j^2}\right) - \underbrace{\cancel{\frac{\partial}{\partial b_k}(c^T h)}}_{=0} - \underbrace{\cancel{\frac{\partial}{\partial b_k}\left(\sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j}\right)}}_{=0}$$

$$= \sum_j \frac{\partial}{\partial b_k}\left(\frac{(x_j - b_j)^2}{2\sigma_j^2}\right) \underset{\substack{=0 \text{ if} \\ j \neq k}}{=} \frac{\partial}{\partial b_k}\left(\frac{(x_k - b_k)^2}{2\sigma_k^2}\right)$$

$$= \frac{-2(x_k - b_k)}{2\sigma_k^2} = \frac{-(x_k - b_k)}{\sigma_k^2}$$

$$\boxed{\frac{\partial \, Energy(x,h)}{\partial b_k} = \frac{b_k - x_k}{\sigma_k^2}}$$

- $$\frac{\partial \text{Energy}(x,h)}{\partial c_k} = \frac{\partial}{\partial c_k}\left(\underbrace{\cancel{\sum_j \frac{(x_j-b_j)^2}{2\sigma_j^2}}}_{=0} - c^T h - \underbrace{\cancel{\sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}_{=0} \right)$$

$$= \frac{\partial}{\partial c_k}\left(-\sum_j c_j h_j \right) = -\frac{\partial}{\partial c_k}\left(c_k h_k \right)$$

$$\boxed{\frac{\partial \text{Energy}(x,h)}{\partial c_k} = -h_k}$$

- $$\frac{\partial \text{Energy}(x,h)}{\partial W_{kl}} = \frac{\partial}{\partial W_{kl}}\left(\underbrace{\cancel{\sum_j \frac{(x_j-b_j)^2}{2\sigma_j^2}}}_{=0} - \underbrace{\cancel{c^T h}}_{=0} - \sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j} \right)$$

$$= -\frac{\partial}{\partial W_{kl}}\left(\sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j} \right)$$

$$= -\frac{\partial}{\partial W_{kl}}\left(\sum_j \sum_i h_i W_{ij} \frac{x_j}{\sigma_j} \right)$$

$$\overset{=0 \text{ if } (i,j)\neq(k,l)}{=} -\frac{\partial}{\partial W_{kl}}\left(h_k W_{kl} \frac{x_l}{\sigma_l} \right)$$

$$\boxed{\frac{\partial \text{Energy}(x,h)}{\partial W_{kl}} = h_k \frac{x_l}{\sigma_l}}$$

$$\frac{\partial \, Energy(x,h)}{\partial U_{kl}} = \frac{\partial \, Energy(x,h)}{\partial V_{kl}} = 0$$

since $U = V = 0$

- $$\mathbb{P}(h|x) = \frac{\mathbb{P}(h,x)}{\mathbb{P}(x)}$$

$$= \frac{e^{-Energy(x,h)} / \cancel{Z}}{\left[\sum_{\hat{h}} e^{-Energy(x,\hat{h})}\right] / \cancel{Z}}$$

$$= \frac{e^{-\cancel{\sum_j \frac{(x_j - b_j)^2}{2\sigma_j^2}} + c^T h + \sum_j h^T W x_\sigma}}{\sum_{\hat{h}} e^{-\cancel{\sum_j \frac{(x_j - b_j)^2}{2\sigma_j^2}} + c^T \hat{h} + \hat{h}^T W x_\sigma}}$$

où $x_\sigma = \begin{pmatrix} \frac{x_1}{\sigma_1} \\ \vdots \\ \frac{x_k}{\sigma_k} \end{pmatrix}$

$$= \frac{e^{c^T h + h^T W x_\sigma}}{\sum_{\hat{h}} e^{+c^T \hat{h} + \hat{h}^T W x_\sigma}}$$

$$= \frac{e^{\sum_j c_j h_j + h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}{\sum_{\hat{h}} e^{\sum_j c_j \hat{h}_j + \hat{h}^T W_{\cdot j} \frac{x_j}{\sigma_j}}}$$

Scanned by CamScanner

Moreover,

$$e^{\Sigma \cdots} = \prod e^{\cdots}$$

$$= \frac{\prod_j e^{c_j h_j + h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}{\sum_{\tilde{h}} \prod_j e^{c_j \tilde{h}_j + \tilde{h}^T W_{\cdot j} \frac{x_j}{\sigma_j}}}$$

$$= \prod_j \underbrace{\frac{e^{c_j h_j + h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}{\sum_{\tilde{h}} e^{c_j \tilde{h}_j + \tilde{h}^T W_{\cdot j} \frac{x_j}{\sigma_j}}}}_{= \mathbb{P}(h_j \mid x)}$$

$$\boxed{\mathbb{P}(h \mid x) = \prod_j \mathbb{P}(h_j \mid x)} \quad (1)$$

Moreover, it follows that

$$\mathbb{P}(h_j = 1 \mid x) = \frac{e^{h_j c_j + h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}{\sum_{\tilde{h}_j \in \{0,1\}} e^{c_j \tilde{h}_j + \tilde{h}_j^T W_{\cdot j} \frac{x_j}{\sigma_j}}}.$$

$$\boxed{\mathbb{P}(h_j=1|x)=\frac{1}{1+e^{-(c_j+W_{\cdot j}\frac{x_j}{\sigma_j})}}}$$

$$\mathbb{P}(h_j=1|x)=\operatorname{lgst}\left(c_j+W_j\frac{x}{\sigma}\right)$$

- $\mathbb{E}_{\mathbb{P}(h|x)}\left[\frac{\partial\, Energy(x,h)}{\partial\theta}\right]$

$$=\sum_h\frac{\partial\, Energy(x,h)}{\partial\theta}\,\mathbb{P}(h|x)$$

$$\boxed{\mathbb{E}_{\mathbb{P}(h|x)}\left[\frac{\partial\, Energy(x,h)}{\partial\theta}\right]\overset{1}{=}\sum_h\frac{\partial\, Energy(x,h)}{\partial\theta}\prod_i\mathbb{P}(h_i|x)}$$

- $\mathbb{P}(x|h)=\dfrac{\mathbb{P}(x,h)}{\int\mathbb{P}(x,h)\,dx}$

$$=\frac{e^{-Energy(x,h)}}{\int e^{-Energy(x,h)}\,dx}$$

$$= \frac{e^{-\sum_j \frac{(x_j-b_j)^2}{2\sigma_j^2} + c^T h + \sum_j h^T W_{\cdot j} \frac{x_j}{\sigma_j}}}{\int e^{-\sum_j \frac{(\tilde{x}_j-b_j)^2}{2\sigma_j^2} + c^T h + \sum_j h^T W_{\cdot j} \frac{\tilde{x}_j}{\sigma_j}} d\tilde{x}}$$

With the same reasoning as before:

$$= \frac{\cancel{e^{c^T h}} \prod_j e^{h^T W_{\cdot j} \frac{x_j}{\sigma_j} - \frac{(x_j-b_j)^2}{2\sigma_j^2}}}{\int \cancel{e^{c^T h}} \prod_j e^{h^T W_{\cdot j} \frac{\tilde{x}_j}{\sigma_j} - \frac{(\tilde{x}_j-b_j)^2}{2\sigma_j^2}} d\tilde{x}}$$

$$= \prod_j \frac{e^{h^T W_{\cdot j} \frac{x_j}{\sigma_j} - \frac{(x_j-b_j)^2}{2\sigma_j^2}}}{\int e^{h^T W_{\cdot j} \frac{\tilde{x}_j}{\sigma_j} - \frac{(\tilde{x}_j-b_j)^2}{2\sigma_j^2}} d\tilde{x}_j}$$

$$= \prod_j \left(\frac{e^{-\frac{(x_j-b_j-h^T W_{\cdot j}\sigma_j)^2}{2\sigma_j^2}}}{\int e^{-\frac{(\tilde{x}_j-b_j-h^T W_{\cdot j}\sigma_j)^2}{2\sigma_j^2}} d\tilde{x}_j} \right)$$

$$= \prod_j \mathcal{N}(b_j + h^T W_{\cdot j}\sigma_j, \sigma_j^2)$$

So, $\boxed{P(x_j \mid h) \sim \mathcal{N}(b_j + h^T W_{\cdot j}\sigma_j, \sigma_j)}$