$$\log p_\theta(v) = \log \frac{1}{Z} \sum_h e^{-E(v,h)}$$

$$P(v,h) = \frac{e^{-E(v,h)}}{Z}$$

$$P(v) = \sum_h \frac{e^{-E(v,h)}}{Z}$$

$$= \log\left(\sum_h e^{-E_\theta(v,h)}\right) - \log\left(\sum_v \sum_h e^{-E_\theta(v,h)}\right)$$

$$Z = \sum_h \sum_v e^{-E(v,h)}$$

$$\frac{\partial}{\partial\theta} \log p_\theta(v) = \frac{\partial}{\partial\theta} \log\left(\sum_h e^{-E_\theta(v,h)}\right) - \frac{\partial}{\partial\theta} \log\left(\sum_h \sum_v e^{-E_\theta(v,h)}\right)$$

$$\frac{\partial}{\partial\theta} \log\left(\sum_h e^{-E_\theta(v,h)}\right) = \frac{-\sum_h e^{-E_\theta(v,h)} \frac{\partial}{\partial\theta} E_\theta(v,h)}{\sum_h e^{-E_\theta(v,h)}}$$

$$= \frac{-\sum_h e^{-E_\theta(v,h)} \frac{\partial}{\partial\theta} E_\theta(v,h)}{Z\,p(v)}$$

$$P(v,h) = \frac{e^{-E(v,h)}}{Z}$$

$$P(v) = \frac{1}{Z} \sum_h e^{-E(v,h)}$$

$$\sum_h e^{-E(v,h)} = P(v)\,Z$$

$$= -\sum_h \frac{e^{-E_\theta(v,h)}}{Z\,P(v)} \frac{\partial}{\partial\theta} E_\theta(v,h)$$

$$= -\sum_h \frac{P(v,h)}{P(v)} \frac{\partial}{\partial\theta} E_\theta(v,h)$$

$$= -\sum_h P(h|v) \frac{\partial}{\partial\theta} E_\theta(v,h)$$

$$= + \mathbb{E}_{h \sim P(h|v)}\left[-\frac{\partial}{\partial\theta} E_\theta(v,h)\right]$$

$$\frac{\partial}{\partial\theta} \log\left(\sum_h \sum_v e^{-E_\theta(v,h)}\right) = \frac{-\sum_h \sum_v e^{-E_\theta(v,h)} \frac{\partial}{\partial\theta} E_\theta(v,h)}{\boxed{\sum_h \sum_v e^{-E_\theta(v,h)}}}$$

$$= -\sum_h \sum_v \frac{e^{-E_\theta(v,h)}}{Z} \frac{\partial}{\partial\theta} E_\theta(v,h)$$

$$= -\sum_h \sum_v P(v,h) \frac{\partial}{\partial\theta} E_\theta(v,h)$$

$$= \mathbb{E}_{(v,h) \sim P(v,h)}\left[-\frac{\partial}{\partial\theta} E_\theta(v,h)\right]$$

$$\frac{\partial}{\partial\theta} \log P(v) = \mathbb{E}_{h \sim P(h|v)}\left[-\frac{\partial}{\partial\theta} E_\theta(v,h)\right] - \mathbb{E}_{(v,h) \sim P(v,h)}\left[-\frac{\partial}{\partial\theta} E_\theta(v,h)\right]$$

$$\frac{\partial}{\partial w_{ij}} \log p(v) = \mathbb{E}_{h \sim p(h|v)}\left[-\frac{\partial}{\partial w_{ij}} E_\theta(v,h)\right] - \mathbb{E}_{(v,h) \sim p(v,h)}\left[-\frac{\partial}{\partial w_{ij}} E_\theta(v,h)\right]$$

$$-\frac{\partial}{\partial w_{ij}} E_\theta(v,h) = \frac{\partial}{\partial w_{ij}}\left[v^T W h + b^T v + c^T h\right]$$

$$= \frac{\partial}{\partial w_{ij}}\left[\sum_r \sum_s v_r h_s w_{rs}\right]$$

$$= v_i h_j$$

$$\frac{\partial}{\partial w_{12}}\left(a_1 \cancel{w_{11}} + a_2 w_{12} + \cancel{a_3 w_{21}} + \cancel{a_4 w_{22}}\right)$$

$a \in \mathbb{R}^{m \times 1}$, $B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{n \times 1}$

$$\cdot\ a^T B C = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i c_j B_{ij}$$

$$[a_1\ a_2 \ldots a_m]\begin{bmatrix} B_{11} & B_{1\cdot} \\ \cdot & \cdot \end{bmatrix}\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$\frac{\partial}{\partial w_{ij}} \log p(v) = v_i \mathbb{E}_{h \sim p(h|v)}\left[h_j\right] - \mathbb{E}_{(v,h) \sim p(v,h)}\left[v_i h_j\right]$$

$$\sum_{h \in \{0,1\}^m} p(h|v)\, v_i h_j = \sum_{h_0 \in \{0,1\}} \cdots \sum_{h_{m-1}} p(h_0|v) \cdots p(h_{m-1}|v)\, h_j$$

$$\sum_{h_j \in \{0,1\}} p(h_j|v) = 1$$

$$= \sum_{h_j \in \{0,1\}} h_j\, p(h_j|v)$$

$$= 0 \times p(h_j = 0|v) + 1 \times \underline{p(h_j = 1|v)}$$

$$= \sigma(v^T w_j + c_j)$$

$$\frac{\partial}{\partial w_{ij}} \log p(v) = v_i \sigma(v^T w_j + c_j) - \mathbb{E}_{(v,h) \sim p(v,h)}\left[v_i h_j\right]$$