Exercise Sheet 10

1 Mixture Density Networks

a) $$\frac{\partial E^q}{\partial \alpha_i} = \frac{\partial}{\partial \alpha_i} - \log\left\{\sum_{i=1}^{m} \alpha_i(x^q)\phi_i(t^q|x^q)\right\}$$

$$= -\frac{\phi_i}{\sum_i \alpha_i \phi_i} = -\frac{\alpha_i}{\alpha_i}\frac{\phi_i}{\sum_i \alpha_i \phi_i} = -\frac{\pi_i}{\alpha_i}$$

$$\frac{\partial E^q}{\partial \mu_{ik}} = \frac{\partial E^q}{\partial \phi_i}\left(\frac{\partial \phi_i}{\partial \mu_{ik}}\right) = \frac{\alpha_i}{-\sum_i \alpha_i \phi_i} \cdot \phi_i\left(-\frac{\mu_{ik}(x) - t_k}{\sigma_i(x^2)}\right)$$

$$= \pi_i\left(\frac{\mu_i(x) - t_k}{\sigma_i(x^2)}\right)$$

b) $$\frac{\partial E^q}{\partial z_i^\alpha} = \frac{\partial E^q}{\partial \alpha_j} \cdot \frac{\partial \alpha_j}{\partial z_i^\alpha} = \sum_j \frac{\partial E^q}{\partial \alpha_j} \cdot \frac{\partial \alpha_j}{\partial z^\alpha}$$

$u = e^{z_i\alpha}$

$v = \sum_t e^{z_i\alpha}$

$f' = \left(\frac{u}{v}\right)' = \frac{uv' - vu'}{v^2}$

$$f' = \frac{\sum_j e^{z_k^\alpha} \mathbb{I}(i=k) e^{z_i\alpha} - e^{z_i\alpha} e^{z_j\alpha}}{\left(\sum_j e^{z_j^\alpha}\right)^2} = \delta(i=k)\alpha_i - \alpha_i\alpha_j$$

$$\frac{\partial E^q}{\partial z_i^\alpha} = \sum_j -\frac{\pi_j}{\alpha_j}\mathbb{I}(i=j)\alpha_i - \alpha_i\alpha_j) = -\pi_i + \sum_j \pi_j \alpha_i$$

$$= -\pi_i + \alpha_i = \alpha_i - \pi_i$$

2 Conditional RBM

a) ~~p~~ i) $p(h_k = 1 \mid x, y) = p(h_k = 1, x, y) / p(x, y)$

$$= \frac{\sum_{h_{-k}} p(h_k = 1, h_{-k}, x, y)}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} p(h_k = q, h_{-k}, x, y)}$$

$$= \frac{\sum_{h_{-k}} \frac{1}{Z} e^{x^T W_{:,k} h_k + y^T U_{:,k} h_k - E(x, y, h_{-k})}}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} \frac{1}{Z} e^{x^T W_{:,k} q + y^T U_{:,k} q - E(x, y, h_{-k})}}$$

$$= \frac{e^{x^T W_{:,k} + y^T U_{:,k}}}{1 + e^{x^T W_{:,k} + y^T U_{:,k}}} = \mathrm{sigm}(x^T W_{:,k} + y^T U_{:,k})$$

ii) $p(y_j = 1 \mid h, x) = \mathrm{sigm}(U_{j,:}^T h) = \frac{p(y_j, h, x)}{p(h, x)}$

$$= \frac{\sum_{y_{-j}} p(y_j = 1, y_{-j}, h, x)}{\sum_{q \in \{0,1\}} \sum_{y_{-j}} p(y_j = q, y_{-j}, h, x)} = \frac{e^{U_{j,:}^T h}}{1 + e^{U_{j,:}^T h}} = \mathrm{sigm}(U_j^T$$

b) $p(x, y) = \sum_h p(x, y, h) = \sum_h \frac{1}{Z} e^{-E(x, y, h)}$

$$= \sum_h \frac{1}{Z} \prod_k e^{-E(x, y, h_k)} = \frac{1}{Z} \prod_k (1 + e^{-E(x, y, h_k = 1)})$$

$$= \frac{1}{Z} e^{\sum_k \log(1 + e^{-E(x, y, h_k = 1)})} = \frac{1}{Z} e^{-F(x, y)}$$