Sheet 08

(1) a) (i) $\frac{\partial E^q}{\partial \alpha_i} = -\frac{\pi_i}{\alpha_i}$

$$\frac{\partial E^q}{\partial \alpha_i} = -\frac{1}{\sum_j \alpha_j \phi_j}\phi_i = -\frac{1}{\alpha_i}\frac{\alpha_i \phi_i}{\sum_j \alpha_j \phi_j} = -\frac{\pi_i}{\alpha_i}$$

(ii) $\frac{\partial E^q}{\partial \mu_{ik}} = \frac{\partial E^q}{\partial \phi_i} \cdot \frac{\partial \phi_i}{\partial \mu_{ik}}$

$$\frac{\partial E^q}{\partial \phi_i} = -\frac{1}{\sum_j \alpha_j \phi_j}\alpha_i \qquad \frac{\partial \phi_i}{\partial \mu_{ik}} = \phi_i \cdot \left(\frac{t_k - \mu_{ik}}{\sigma_i^2}\right)$$

$$\Rightarrow \frac{\partial E^q}{\partial \mu_{ik}} = \frac{\alpha_i \phi_i}{\sum_j \alpha_j \phi_j}\left(\frac{\mu_{ik} - t_k}{\sigma_i^2}\right) = \pi_i \left(\frac{\mu_{ik} - t_k}{\sigma_i^2}\right)$$

b) $$\frac{\partial \alpha_j}{\partial z_i^\alpha} = \frac{\delta_{ij}\exp(z_i^\alpha)\cdot\left(\sum_k \exp(z_k^\alpha)\right) - \exp(z_j)\exp(z_i)}{\left(\sum_k \exp(z_k^\alpha)\right)^2} = \delta_{ij}\alpha_i - \alpha_i \alpha_j$$

$$\frac{\partial E_q}{\partial z_i^\alpha} = \sum_j \frac{\partial E^q}{\partial \alpha_j}\frac{\partial \alpha_j}{\partial z_i^\alpha} = \sum_j \left(-\frac{\pi_j}{\alpha_j}\delta_{ij}\alpha_i + \frac{\pi_j}{\alpha_j}\alpha_i \alpha_j\right)$$

$$= -\pi_i + \sum_j \pi_j \alpha_i = \underline{\underline{\alpha_i - \pi_i}}$$

② a) (i) $p(h_k = 1 \mid x, y) = \dfrac{p(h_k = 1, x, y)}{p(x, y)}$

with $\sum_{h_{-k}}$ meaning: sum over all h but h_k

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

$$= \frac{\sum_{h_{-k}} \exp[x^T W_{:,k} 1 + y^T U_{:,k} 1 - E(x, y, h_{-k})]/Z}{\sum_q \sum_{h_{-k}} \exp[x^T W_{:,k} q + y^T U_{:,k} q - E(x, y, h_{-k})]/Z}$$

$$= \frac{\exp[x^T W_{:,k} 1 + y^T U_{:,k} 1] \cdot \underbrace{\cancel{\sum_{h_{-k}} \exp(-E(x, y, h_{-k}))/Z}}_{A}}{\sum_{q\in\{0,1\}} \exp(x^T W_{:,k} q + y^T U_{:,k} q) \cdot \cancel{A}}$$

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

Same procedure for (ii)

b) $p(x, y) = \sum_h p(x, y, h) = \sum_h \exp(-E(x, y, h)) = \sum_h \exp(-\sum_k E(x, y, h_k))$

$= \sum_h \prod_k \exp(-E(x, y, h_k))$

$= \prod_k \sum_{q\in\{0,1\}} \exp(-E(x, y, h_k = q))$

$= \prod_k (1 + \exp(-E(x, y, h_k = 1))) = \exp(\log(\prod_k (1 + \exp(-E(x, y, h_k = 1)))))$

$$= \exp\Big(\underbrace{\sum_k \log(1 + \exp(-E(x, y, h_k = 1)))}_{F(x,y)}\Big)$$

□