1

$$\mathbb{E}_q\left[\log p(\pi \mid \alpha)\right] = \mathbb{E}_q\left[\log \frac{\Gamma(\alpha K)}{\prod_{i=1}^{K}\Gamma(\alpha)} \prod_{i=1}^{K} (\pi_i)^{(\alpha-1)}\right]$$

$$= \log \Gamma(\alpha K) - \sum_{i=1}^{K} \log \Gamma(\alpha) + \sum_{i=1}^{K} (\alpha-1)\, \mathbb{E}_q\left[\log \pi_i\right]$$

$$\cancel{\mathbb{E}_q\left[\log p(\mu_K \mid 0, \tau^2)\right] = -\frac{1}{2}\log(2\pi\tau^2) - \frac{\mathbb{E}_q[\mu_K^2]}{2\tau^2}}$$

$$\mathbb{E}_q\left[\log p(\mu_K \mid 0, \tau^2)\right] = \mathbb{E}_q\left[\log\left(\frac{1}{\sqrt{2\pi\tau^2}}\, e^{-\frac{\mu_K^2}{2\tau^2}}\right)\right]$$

$$= -\frac{1}{2}\log(2\pi\tau^2) - \frac{\mathbb{E}_q[\mu_K^2]}{2\tau^2}$$

$$\mathbb{E}_q\left[\log p(z_m \mid \pi)\right] = \mathbb{E}_q\left[\log \prod_{i=1}^{K} (\pi_i)^{(z_{m,i})}\right] = \sum_{i=1}^{K} \phi_{m,i}\, \mathbb{E}_q\left[\log \pi_i\right]$$

$$\mathbb{E}_q\left[\log p(x_m \mid z_m, \mu, \sigma^2)\right] = \mathbb{E}_q\left[\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x_m-\mu_{z_m})^2}{2\sigma^2}}\right)\right]$$

$$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{x_m^2}{2\sigma^2} + \frac{\mathbb{E}_q[\mu_{z_m}]\, x_m}{\sigma^2} - \frac{\mathbb{E}_q[\mu_{z_m}^2]}{2\sigma^2}$$

$$\mathbb{E}_q\left[\log \pi_i\right] = \Psi(\gamma_i) - \Psi\left(\sum_{j=1}^{K} \gamma_j\right)$$

$$\mathbb{E}_q\left[\mu_K^2\right] = \tilde{\sigma}_K^2 + \tilde{\mu}_K^2 \quad , \quad \mathbb{E}_q\left[\mu_K\right] = \tilde{\mu}_K$$

$$\mathbb{E}_q\left[\mu_{z_m}\right] = \sum_{i=1}^{K} \phi_{m,i}\, \tilde{\mu}_i$$

2

$$\mathbb{E}_q\left[\log q(\pi \mid \gamma)\right] = \cancel{\log\Gamma(\gamma K) - \sum_{i=1}^{K}\log\Gamma(\gamma) + \sum_{i=1}^{K}(\gamma-1)\mathbb{E}_q[\log \pi_i]}$$

$$= \mathbb{E}_q\left[\log \frac{\Gamma\left(\sum_{i=1}^{K}\gamma_i\right)}{\prod_{i=1}^{K}\Gamma(\gamma_i)} \prod_{i=1}^{K}(\pi_i)^{(\gamma_i-1)}\right]$$

$$= \log\Gamma\left(\sum_{i=1}^{K}\gamma_i\right) - \sum_{i=1}^{K}\log\Gamma(\gamma_i) + \sum_{i=1}^{K}(\gamma_i-1)\mathbb{E}_q[\log \pi_i]$$

$$\mathbb{E}_q\left[\log q(\mu_k \mid \tilde{\mu}_k, \tilde{\sigma}_k^2)\right] = \mathbb{E}_q\left[\log\left(\frac{1}{\sqrt{2\pi\tilde{\sigma}_k^2}} e^{-\frac{(\mu_k-\tilde{\mu}_k)^2}{2\tilde{\sigma}_k^2}}\right)\right]$$

$$= -\frac{1}{2}\log(2\pi\tilde{\sigma}_k^2) - \frac{1}{2}$$

$$\mathbb{E}_q\left[\log q(z_n \mid \phi_n)\right] = \mathbb{E}_q\left[\log \prod_{i=1}^{K}(\phi_{n,i})^{z_{n,i}}\right] = \sum_{i=1}^{K}\phi_{n,i}\log\phi_{n,i}$$

$$L_{\tilde{\mu}_k} = -\frac{\mathbb{E}_q[\mu_k^2]}{2\tau^2} + \sum_{n=1}^{N}\left(\frac{2\mathbb{E}_q[\mu_{z_n}]x_n}{2\sigma^2} - \frac{\mathbb{E}_q[\mu_{z_n}^2]}{2\sigma^2}\right)$$

$$= -\frac{\cancel{\tilde{\sigma}_k^2 +}\tilde{\mu}_k^2}{2\tau^2} + \sum_{n=1}^{N}\left(\frac{2\phi_{n,k}\tilde{\mu}_k x_n}{2\sigma^2} - \frac{\phi_{nk}\tilde{\mu}_k^2}{2\sigma^2}\right)$$

$$dL_{\tilde{\mu}_k} = -\frac{2\mu_k}{2\tau^2} + \sum_{n=1}^{N}\frac{2\phi_{nk}x_n}{2\sigma^2} - \mu_k\sum_{n=1}^{N}\frac{2\phi_{nk}}{2\sigma^2} = 0$$

$$\Leftrightarrow \sum_{n=1}^{N}\frac{\phi_{nk}x_n}{\sigma^2} = \mu_k\left(\frac{1}{\tau^2} + \sum_{n=1}^{N}\frac{\phi_{nk}}{\sigma^2}\right) \Leftrightarrow \mu_k = \frac{\frac{1}{\sigma^2}\sum_{n=1}^{N}\phi_{nk}x_n}{\frac{1}{\tau^2} + \frac{1}{\sigma^2}\sum_{n=1}^{N}\phi_{nk}}$$

③

$$L_{\tilde{\sigma}_k^2} = -\frac{\tilde{\sigma}_k^2}{2\Gamma^2} - \sum_{m=1}^{N} \frac{\phi_{m,k}\tilde{\sigma}_k^2}{2\sigma^2} + \frac{1}{2}\log \tilde{\sigma}_k^2$$

$$\frac{dL_{\tilde{\sigma}_k^2}}{\tilde{\sigma}_k^2} = -\frac{1}{\Gamma^2} - \sum_{m=1}^{N} \frac{\phi_{m,k}}{\sigma^2} + \frac{1}{\tilde{\sigma}_k^2} = 0$$

$$\Leftrightarrow \tilde{\sigma}_k^2 = \frac{1}{\frac{1}{\Gamma^2} + \frac{1}{\sigma^2}\sum_{m=1}^{N} \phi_{m,k}}$$

$$L_{\gamma_k} = \cancel{(\alpha-1)\left(\Psi(\gamma_k) - \Psi\left(\sum_{j=1}^{K}\gamma_j\right)\right) + \phi_{m,k}\left(\Psi\right.}$$

$$(\alpha-1)\,\mathbb{E}[\log \pi_i] + \sum_{m=1}^{N} \phi_{m,k}\,\mathbb{E}[\log \pi_i] - \log\Gamma\left(\sum_{j=1}^{K}\gamma_j\right)$$

$$+ \log\Gamma(\gamma_k) - (\gamma_k - 1)\,\mathbb{E}[\log \pi_i]$$

$$= \underbrace{\mathbb{E}[\log \pi_i]}\left(\alpha - \gamma_k + \sum_{m=1}^{N}\phi_{m,k}\right) + \log\Gamma(\gamma_k) - \log\Gamma\left(\sum_{j=1}^{K}\gamma_j\right)$$

$$= \Psi(\gamma_k) - \Psi\left(\sum_{j=1}^{K}\gamma_k\right) = \frac{\Gamma'(\gamma_k)}{\Gamma(\gamma_k)} - \frac{\Gamma'\left(\sum_{j=1}^{K}\gamma_j\right)}{\Gamma\left(\sum_{j=1}^{K}\gamma_j\right)}$$

$$\cancel{\frac{\Gamma''\Gamma - \Gamma'\Gamma'}{\Gamma\Gamma} = \frac{\Gamma''}{\Gamma} - \Psi\Psi}$$

(4)

$$L_{\phi_{m,k}} = \phi_{mk} \mathbb{E}_q[\log \pi_k] + \frac{\phi_{m,k} \tilde{\mu}_k x_m}{\sigma^2} - \frac{\phi_{mk}(\tilde{\sigma}_k^2 + \tilde{\mu}_k^2)}{2\sigma^2}$$

$$- \phi_{mk} \log \phi_{mk} + \lambda \left(\sum_{j=1}^{K} \phi_{mj} - 1 \right)$$

$$\frac{d L_{\phi_{mk}}}{d \phi_{mk}} = \mathbb{E}_q[\log \pi_k] + \frac{\tilde{\mu}_k x_m}{\sigma^2} - \frac{\tilde{\sigma}_k^2 + \tilde{\mu}_k^2}{2\sigma^2} - \log \phi_{mk} - 1 + \lambda \stackrel{!}{=} 0$$

$$\Leftrightarrow \phi_{m,k} = \exp\Big(\underbrace{\mathbb{E}_q[\log \pi_k] + \frac{\tilde{\mu}_k x_m}{\sigma^2} - \frac{\tilde{\sigma}_k^2 + \tilde{\mu}_k^2}{2\sigma^2}}_{A_k} - 1 \Big) \exp(\lambda)$$

$$\Leftrightarrow \phi_{m,k} = \exp(A - 1) \exp(\lambda)$$

$$\sum_{j=1}^{K} \exp(A_j - 1) \exp(\lambda) = 1 \Leftrightarrow \exp(\lambda) = \frac{1}{\sum_{j=1}^{K} \exp(A_j - 1)}$$

$$\phi_{m,k} = \frac{\exp(A_k - 1)}{\sum_{j=1}^{K} \exp(A_j - 1)} = \frac{\exp(A_k) \cancel{\exp(-1)}}{\cancel{\exp(-1)} \sum_{j=1}^{K} \exp(A_j)} = \frac{\exp(A_k)}{\sum_{j=1}^{K} \exp(A_j)}$$

$$\propto \exp(A_k) = \exp\left(\mathbb{E}_q[\log \pi_k] + \frac{\tilde{\mu}_k x_m}{\sigma^2} - \frac{\tilde{\sigma}_k^2 + \tilde{\mu}_k^2}{2\sigma^2} \right)$$

(1)

$$\cancel{p(w,\alpha) = \mathcal{N}(w|0, \alpha^{-1} I) \prod}$$

$$p(w,\alpha) = \mathcal{N}(w|0, \mathrm{diag}(\alpha)^{-1}) \prod_{j=1}^{D} \mathrm{Ga}(\alpha_j|a_0,b_0)$$

$$p(y_n|w) = \mathcal{N}(y_n| w^T x_n, \sigma^2)$$

$$q(w,\alpha) = q(w)\prod_{j=1}^{D} q(\alpha_j) = \mathcal{N}(w|m_N, S_N)\prod_{j=1}^{D} \mathcal{N}(\alpha_j|a_{Nj}, b_{Nj})$$

$$p(y,w,\alpha|X) = \mathcal{N}(w|0,\mathrm{diag}(\alpha)^{-1})\left(\prod_{j=1}^{D} \mathrm{Ga}(\alpha_j|a_0,b_0)\right) \prod_{n=1}^{N} \mathcal{N}(y_n|w^T x_n, \sigma^2)$$

$$L(q) = \mathbb{E}_q[\log p(y,w,\alpha|X)] - \mathbb{E}_q[\log q(w,\alpha)]$$

$$= \mathbb{E}_q[\log \mathcal{N}(w|0,\mathrm{diag}(\alpha)^{-1})] + \sum_{j=1}^{D} \mathbb{E}_q[\log \mathrm{Ga}(\alpha_j|a_0,b_0)]$$

$$+ \sum_{n=1}^{N} \mathbb{E}_q[\log \mathcal{N}(y_n|w^T x_n, \sigma^2)] - \mathbb{E}_q[\log \mathcal{N}(w|m_N, S_N)]$$

$$- \sum_{j=1}^{D} \mathbb{E}_q[\log \mathrm{Ga}(\alpha_j|a_{Nj}, b_{Nj})]$$

2

$$\mathbb{E}_q\left[\log \mathcal{N}(w \mid 0, \operatorname{diag}(\alpha)^{-1})\right] = \cancel{\mathbb{E}_q\left[\log\left(\det\left(2\pi \operatorname{diag}(\alpha)^{-1}\right)^{-\frac{1}{2}} e^{-\frac{1}{2}(w^T \operatorname{diag}(\alpha) w)}\right)\right]}$$

$$\cancel{= -\frac{1}{2}\mathbb{E}_q\left[\log\left(\det\left(2\pi \operatorname{diag}(\alpha)^{-1}\right)\right)\right] - \frac{1}{2}\mathbb{E}_q}\left[w^T \operatorname{diag}(\alpha) w\right]$$

$$= \mathbb{E}_q\left[\log\left(\frac{1}{(2\pi)^{D/2}} \frac{1}{|\operatorname{diag}(\alpha)^{-1}|^{1/2}} \exp\left(-\frac{1}{2} w^T \operatorname{diag}(\alpha) w\right)\right)\right]$$

$$= -\frac{D}{2}\log(2\pi) - \frac{1}{2}\mathbb{E}_q\left[\log|\operatorname{diag}(\alpha)^{-1}|\right] - \frac{1}{2}\mathbb{E}_q\left[w^T \operatorname{diag}(\alpha) w\right]$$

$$= -\frac{D}{2}\log(2\pi) - \frac{1}{2}\mathbb{E}_q\left[\log\frac{1}{|\operatorname{diag}(\alpha)|}\right] - \frac{1}{2}\mathbb{E}_q\left[w^T \operatorname{diag}(\alpha) w\right]$$

$$= -\frac{D}{2}\log(2\pi) + \frac{1}{2}\mathbb{E}_q\left[\log|\operatorname{diag}(\alpha)|\right] - \frac{1}{2}\mathbb{E}_q\left[w^T \operatorname{diag}(\alpha) w\right]$$

$$= -\frac{D}{2}\log(2\pi) + \frac{1}{2}\sum_{j=1}^{D}\mathbb{E}_q\left[\log \alpha_j\right] - \frac{1}{2}\mathbb{E}_q\left[\sum_{j=1}^{D} w_j^2 \alpha_j\right]$$

$$= -\frac{D}{2}\log(2\pi) + \frac{1}{2}\sum_{j=1}^{D}\left(\psi(a_{Nj}) - \log b_{Nj}\right) - \frac{1}{2}\sum_{j=1}^{D}\frac{a_{Nj}}{b_{Nj}}\left(m_{Nj}^2 + V_{Njj}\right)$$

3

$$\mathbb{E}_q\left[\log \mathcal{N}(y_n \mid w^T x_n, \sigma^2)\right]$$

$$= \mathbb{E}_q\left[\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(y_n - w^T x_n)^2}{2\sigma^2}}\right)\right]$$

$$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2} y_n^2 + \frac{\mathbb{E}_q[w^T] x_n \cdot y_n}{\sigma^2} - \frac{1}{2\sigma^2}\overbrace{\mathbb{E}_q[w^T w]\, x_n^T x_n}^{x_n^T \mathbb{E}_q[w^T w] x_n}$$

$$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{y_n^2}{2\sigma^2} + \frac{m_N^T x_n y_n}{\sigma^2} - \frac{\cancel{2}\, x_n^T (m_N^T m_N + V_N) x_n}{2\sigma^2}$$

$$L_{[a_{Nj}]} = \frac{1}{2}\psi(a_{Nj}) - \frac{1}{2}\frac{a_{Nj}}{b_{Nj}}(m_{Nj}^2 + V_{Njj}) + (a_0 - 1)\psi(a_{Nj})$$

$$- b_0 \frac{a_{Nj}}{b_{Nj}} - \cancel{\ln\Gamma(a_{Nj})} + \cancel{\ln\Gamma(a_{Nj})} - (a_{Nj} - 1)\psi(a_{Nj})$$

$$= \psi(a_{Nj})\left(\frac{1}{2} + a_0 \cancel{-1} - a_{Nj} \cancel{+1}\right) - a_{Nj}\left(\frac{m_{Nj}^2 + V_{Njj}}{2 b_{Nj}} + \frac{b_0}{b_{Nj}}\right)$$

$$a_{Nj} = a_0 + \frac{1}{2}$$

4

$$L[b_{NJ}] = -\frac{1}{2}\log b_{NJ} - \frac{1}{2}\frac{a_{NJ}}{b_{NJ}}\left(m_{NJ}^2 + V_{NJJ}\right) - (a_0 - 1)\log b_{NJ}$$

$$- b_0\frac{a_{NJ}}{b_{NJ}} - \log b_{NJ}$$

$$= \left(-\frac{1}{2} - a_0\right)\log b_{NJ} - \frac{1}{b_{NJ}}\left(b_0 a_{NJ} + \frac{1}{2}a_{NJ}\left(m_{NJ}^2 + V_{NJJ}\right)\right)$$

$$= \left(-\frac{1}{2} - a_0\right)\log b_{NJ} - b_{NJ}^{-1}\left(b_0 a_{NJ} + \frac{1}{2}a_{NJ}\left(m_{NJ}^2 + V_{NJJ}\right)\right)$$

$$\frac{dL_{[b_{NJ}]}}{db_{NJ}} = \frac{1}{b_{NJ}}\left(-\frac{1}{2} - a_0\right) + b_{NJ}^{-2}\left(b_0 a_{NJ} + \frac{1}{2}a_{NJ}\left(m_{NJ}^2 + V_{NJJ}\right)\right)$$

$$\Leftrightarrow \quad b_{NJ} = \frac{b_0 \cancel{a_{NJ}} + \frac{1}{2}\cancel{a_{NJ}}\left(m_{NJ}^2 + V_{NJJ}\right)}{\cancel{\underbrace{a_0 + \frac{1}{2}}_{= a_{NJ}}}}$$

$$\Leftrightarrow \quad b_{NJ} = b_0 + \frac{1}{2}\left(m_{NJ}^2 + V_{NJJ}\right)$$