

2.1

$$P(z_i = k | x_i; \beta, \pi) = \frac{P(x_i | z_i = k; \beta, \pi) P(z_i = k; \pi)}{P(x_i; \beta, \pi)}$$

$$= \frac{P(x_i | z_i = k; \beta, \pi) P(z_i = k; \pi)}{\sum_{j=1}^K P(x_i | z_i = j; \beta, \pi) P(z_i = j; \pi)}$$

$$= \frac{\text{Multinomial}(x_i; \beta_k) \cdot \pi_k}{\sum_{j=1}^K \text{Multinomial}(x_i; \beta_j) \cdot \pi_j}$$

2.2

$$\beta, \alpha = \underset{\beta, \alpha}{\operatorname{argmax}} \sum_{i=1}^N \sum_{k=1}^K p(z_i=k | x_i; \beta, \alpha) \log P(x_i, z_i=k; \beta, \alpha)$$
$$\mathcal{L}(\beta, \alpha) = \sum_{i=1}^N \sum_{k=1}^K p(z_i=k | x_i; \beta, \alpha) \log P(x_i | z_i=k; \beta, \alpha) P(z_i=k; \alpha)$$
$$- \lambda_h \left(\sum_{n=1}^V \beta_{kn} - 1 \right) - \lambda \left(\sum_{j=1}^K \alpha_j - 1 \right)$$

$$\frac{\partial \mathcal{L}}{\partial \beta_{h,w}} = \frac{\partial}{\partial \beta_{h,w}} \left[\sum_{i=1}^N p(z_i=k | x_i; \beta, \alpha) \log P(x_i | z_i=k; \beta, \alpha) \right] - \lambda_h$$
$$\propto \frac{\partial}{\partial \beta_{h,w}} \left[\sum_{i=1}^N p(z_i=k | x_i) \sum_{n=1}^V \log \beta_{hn}^{x_{i,n}} \right] - \lambda_h$$
$$= \frac{\partial}{\partial \beta_{h,w}} \left[\sum_{i=1}^N p(z_i=k | x_i) \log \beta_{h,w}^{x_{i,w}} \right] - \lambda_h$$
$$= \frac{\sum_{i=1}^N p(z_i=k | x_i) \cdot x_{i,w}}{\beta_{h,w}} - \lambda_h = 0$$

$$\beta_{h,w} = \frac{\sum_{i=1}^N p(z_i=k | x_i) \cdot x_{i,w}}{\sum_{w=1}^V \sum_{i=1}^N p(z_i=k | x_i) \cdot x_{i,w}}$$

$$\frac{\partial \mathcal{L}}{\partial \pi_k} = \frac{\partial}{\partial \pi_k} \left[\sum_{i=1}^n P(z_i=k | x_i) \log P(z_i=k) \right] - \lambda$$

$$= \frac{\sum_{i=1}^n P(z_i=k | x_i)}{\pi_k} - \lambda = 0$$

$$\pi_k = \frac{\sum_{i=1}^N P(z_i=k | x_i)}{N}$$