

$$Q2 a) \quad P(\theta) \propto \theta^{\alpha_1-1} \dots \theta^{\alpha_K-1} \\ = \prod_{k=1}^K \theta_k^{\alpha_k-1}$$

$$P(x_i = j | \theta) = \theta_j$$

$$P(D|\theta) = \prod_{i=1}^n P(x_i = x_i | \theta) = \prod_{i=1}^n \theta_{x_i} = \prod_{j=1}^m \theta_j^{\sum I(x_i=j)}$$

$$\text{let } N = \sum I(x_i=j)$$

$$= \prod_{j=1}^m \theta_j^{\sum I(x_i=j)}$$

$$P(\theta|D) \propto P(\theta) P(D|\theta)$$

$$= \prod_{k=1}^K \theta_k^{\alpha_k-1} \prod_{j=1}^m \theta_j^N$$

$$= \prod_{j=1}^m \theta_j^N$$

$$P(D'|D) = \theta_{\text{pred}} = \int P(\theta|D) P(D'|D) d\theta$$

$$= \int \text{Dirichlet}(\alpha_1 + N_1, \dots, \alpha_K + N_K) \theta_K d\theta$$

$$= E(\theta_K | D)$$

$$= \frac{N_K + \alpha_K}{N + \sum \alpha_k}$$

$$\begin{aligned}
 b) \hat{\theta}_{MAP} &= \arg \max P(\theta | D) \\
 &= \arg \max P(D | \theta) P(\theta). \\
 &= \arg \max \log P(\theta) + \log P(D | \theta)
 \end{aligned}$$

$$\begin{aligned}
 &\log P(\theta | D) \\
 &= \log \prod_{k=1}^K \theta_k^{\alpha_k + N_k - 1}
 \end{aligned}$$

$$= \sum_{k=1}^K (\alpha_k + N_k - 1) \log \theta_k$$

$$\frac{\partial f}{\partial \theta_i} = \frac{\alpha_i + N_i - 1}{\theta_i}$$

$$\frac{\partial g}{\partial \theta_i} = 1$$

$$= \lambda \frac{\partial g}{\partial \theta_i}$$

$$\frac{\alpha_i + N_i - 1}{\theta_i} = \lambda$$

$$\sum (\alpha_i + N_i - 1) = \lambda \sum \theta_i$$

$$\text{MAP} = \frac{\alpha + N + 1}{N - K + \sum_{j=1}^K \alpha_j}$$