1.1) primal constrained

$$\max_{\vec{q}} L(\vec{q}) = \sum_k a_k \ln q_k \quad \text{st} \quad q_k \geq 0, \forall k; \ \sum_k q_k \leq 1; \ \sum_k q_k \geq 1$$

primal unconstrained

$$\max_{\vec{q}} \min_{\vec{\lambda},\alpha,\mu} L(\vec{q},\vec{\lambda},\alpha,\mu) = \max \min \left[\sum_k a_k \ln q_k + \sum_k \lambda_k q_k + \alpha\left(\sum_k q_k - 1\right) + \mu\left(1 - \sum_k q_k\right)\right]$$

dual unconstrained

$$\min \max \left[\sum_k a_k \ln q_k + \sum_k \lambda_k q_k + \alpha\left(\sum_k q_k - 1\right) + \mu\left(1 - \sum_k q_k\right)\right]$$

$$\nabla_{\vec{q}} L = \frac{d}{d\vec{q}}\left[\vec{a}^T \ln \vec{q} + \vec{\lambda}^T \vec{q} + \alpha(\vec{q}^T \vec{1} - 1) + \mu(1 - \vec{q}^T \vec{1})\right]$$

$$= \vec{a} \oslash \vec{q} + \vec{\lambda} + \alpha - \mu = 0$$

$$\vec{a} \oslash \vec{q} = \mu - \alpha - \vec{\lambda}$$

$$\vec{q} = \vec{a} \oslash (\mu - \alpha - \vec{\lambda}) \Rightarrow q_k^* = \frac{a_k}{\mu - \alpha - \lambda_k}$$

complementary slackness

$$\sum_k \lambda_k^* q_k^* = 0 \Rightarrow \sum_k \frac{\lambda_k^* a_k}{\mu - \alpha - \lambda_k} = 0 \quad \text{since } a_k \in \mathbb{R}^+, \ \lambda_k^* = 0, \forall k$$

$$q_k^* = \frac{a_k}{\mu - \alpha} \Rightarrow \frac{q_k^*}{\sum_k q_k^*} = \frac{a_k/(\mu-\alpha)}{\sum a_k/(\mu-\alpha)} = \frac{a_k/\cancel{(\mu-\alpha)}}{\cancel{(\mu-\alpha)}\sum a_k} = \frac{a_k}{\sum_k a_k}$$

1-2) primal constrained

$$\max \sum_k (q_k b_k - q_k \ln q_k) \quad \text{st} \quad q_k \geq 0, \forall k; \ \sum_k q_k = 1$$

primal unconstrained

$$\max \min \left[\sum_k (q_k b_k - q_k \ln q_k) + \sum_k \lambda_k q_k + \alpha\left(\sum_k q_k - 1\right) + \mu\left(1 - \sum_k q_k\right)\right]$$

dual unconstrained

$$\nabla_{\vec{q}} L = \nabla_{\vec{q}}\left[\vec{q}^T \vec{b} - \vec{q}^T \ln \vec{q} + \vec{\lambda}^T \vec{q} + \alpha \vec{q}^T \vec{1} - \mu \vec{q}^T \vec{1}\right]$$

$$= \vec{b} - (\ln \vec{q} + 1) + \vec{\lambda} + \alpha - \mu = 0$$
$$\vec{b} + \vec{\lambda} + \alpha - \mu - 1 = \ln \vec{q}$$
$$\vec{q} = e^{\vec{b}+\vec{\lambda}+\alpha-\mu-1} \Rightarrow q_k^* = e^{b_k + \lambda_k + \alpha - \mu - 1}$$

complementary slackness

$$\sum_k \lambda_k^* e^{b_k + \lambda_k + \alpha - \mu - 1} = 0 \Rightarrow \lambda_k^* = 0$$

$$q_k^* = e^{b_k + \alpha - \mu - 1} \Rightarrow \frac{q_k^*}{\sum_k q_k^*} = \frac{e^{b_k+\alpha-\mu-1}}{\sum_k e^{b_k+\alpha-\mu-1}} = \frac{\cancel{e^{\alpha-\mu-1}} e^{b_k}}{\cancel{e^{\alpha-\mu-1}} \sum_k e^{b_k}} = \frac{e^{b_k}}{\sum_k e^{b_k}}$$

2.1) $\max\min \sum_n \sum_k \gamma_{nk} \ln w_k + \sum_n \sum_k \gamma_{nk} \ln(N(\vec{x}_n | \vec{\mu}_k, \Sigma_k)) + \sum_k \lambda_k w_k + \alpha(\sum_k w_k - 1) + \mu(1 - \sum_k w_k)$

$$a_k = \sum_n \gamma_{nk} \Rightarrow w_k^* = \frac{a_k}{\sum_k a_k} = \frac{\sum_n \gamma_{nk}}{\sum_n \sum_k \gamma_{nk}}$$

$$N(\vec{x}_n | \vec{\mu}_k, \Sigma_k) = \sum_k \frac{1}{(\sqrt{2\pi})^D |\Sigma_k|^{\frac{1}{2}}} \exp(-\tfrac{1}{2}(\vec{x}_n - \mu_k)^T \Sigma_k^{-1} (\vec{x}_n - \mu_k))$$

$$\nabla_{\mu_k} \Rightarrow \boxed{\sum_n \gamma_{nk} \frac{1}{N} \frac{dN}{d\mu_k}} \Rightarrow \sum_k \frac{A}{e^B} \cdot \frac{1}{A} \frac{d(e^B)}{d\mu} = \sum_k \frac{1}{e^B} \cdot e^B \frac{dB}{d\mu} = \sum_k \frac{dB}{d\mu}$$

$$A = (\sqrt{2\pi})^D |\Sigma_k|^{\frac{1}{2}}$$
$$B = -\tfrac{1}{2}(\vec{x}_n - \vec{\mu}_k)^T \Sigma_k^{-1} (\vec{x}_n - \vec{\mu}_k)$$

$$\sum_k \frac{dB}{d\mu} = \Sigma_k^{-1}(\vec{x}_n - \mu_k) \Rightarrow \sum_n \gamma_{nk} \cancel{\Sigma_k^{-1}} (x_n - \mu_k) = 0$$

$$\sum_n \gamma_{nk} x_n - \sum_n \gamma_{nk} \mu_k = 0$$
$$\mu_k^* = \frac{\sum_n \gamma_{nk} \vec{x}_n}{\sum_n \gamma_{nk}}$$

$$\Sigma_k = \sigma^2 I$$

2.2) $b_k = \ln(p(\vec{x}_n, z_n; \theta^{(t)}))$ $\qquad q_k^* = \frac{e^{b_k}}{\sum_k e^{b_k}} = \frac{p(\vec{x}_n, z_n; \theta^{(t)})}{\sum_n p(\vec{x}_n, z_n; \theta^{(t)})}$

2.3) $\sigma^2 \to 0$, $w_k = \frac{\sum_n \gamma_{nk}}{N}$ s.t. $\gamma_{nk} \in \{0,1\}$

3.1) $P(X_{T+1} = s \mid O_{1:T} = o_{1:T}) = \sum_{s'} a_{s's} P(X_T = s' \mid O_{1:T} = o_{1:T})$

$$= \frac{\sum_{s'} a_{s's} P(X_T = s', O_{1:T} = o_{1:T})}{P(O_{1:T} = o_{1:T})} = \frac{\sum_{s'} a_{s's} \alpha_{s'}(T)}{P(O_{1:T} = o_{1:T})}$$