JIN CHAO QIN

4195629854

Problem 1 1.1 Lagrangian is

$$L(\{a_k\},\{q_k\},\{\lambda_k\},\eta) = \sum_{k=1}^{K} a_k \ln q_k - \lambda_k q_k + \eta(1-\sum_{k=1}^{K} q_k)$$

$$\frac{\partial L}{\partial q_k} = \frac{a_k}{q_k} - \lambda_k - \eta = 0$$

$$a_k = (\lambda_k+\eta) q_k$$

$$\sum_{k=1}^{K} a_k = (\lambda_k+\eta)\sum_{k=1}^{K} q_k$$

$$\text{so, } q_k^* = \frac{a_k}{\lambda_k+\eta} = \frac{a_k}{\sum_{k=1}^{K} a_k} \quad \forall k$$

1.2 $L(\{q_k\},\{b_k\},\{\lambda_k\},\eta) =$

$$\sum_{k=1}^{K}(q_k b_k - q_k \ln q_k) - \lambda_k q_k + \eta(1-\sum_{k=1}^{K} q_k)$$

$$\frac{\partial L}{\partial q_k} = b_k - (\ln q_k + 1) - \lambda_k - \eta = 0$$

$$\sum_{k=1}^{K} e^{b_k-\lambda_k-\eta-1} = \sum_{k=1}^{K} q_k = 1$$

$$\cancel{b_k - \lambda_k - \eta - 1 = 0}$$

$$\lambda_k q_k = 0 \Rightarrow \lambda_k = 0$$

$$\sum_{k=1}^{K} e^{b_k} = e^{\eta+1}$$

$$\eta = \ln \sum_{k=1}^{K} e^{b_k} - 1$$

$$q_k^* = e^{b_k - \ln\sum_{k=1}^{K} e^{b_k}}$$

$$= \frac{e^{b_k}}{\sum_{k=1}^{K} e^{b_k}}$$

Problem 2

2.1 from 1.1 let $a_k = \sum_n r_{nk}$, $q_k = \omega_k$

$$\omega_k^* = \frac{\sum_n r_{nk}}{\sum_{k=1}^{K}\sum_n r_{nk}} = \frac{\sum_n r_{nk}}{N}$$

For $\mu_k \Sigma_k$, $L = \arg\max \sum_n \sum_k r_{nk} \left(-\cancel{\ln}\frac{D}{2}\ln 2\pi - D\ln\sigma_k - \frac{1}{2\sigma_k^2}(x_n - \mu_k)^2\right)$

$$\frac{\partial L}{\partial \mu_k} = \sum_n r_{nk} \frac{1}{\sigma_k^2}(x_n - \mu_k) = 0$$

$$\mu_k = \frac{\sum_n r_{nk} x_n}{\sum_n r_{nk}}$$

$$\frac{\partial L}{\partial \sigma_k} = -\sum_n \frac{D \cdot r_{nk}}{\sigma_k} + \sum_n \frac{r_{nk}}{\sigma_k^3}(x_n - \mu_k)^2 = 0$$

$$\sigma_k^2 = \frac{\sum_n r_{nk} \|x_n - \mu_k\|_2^2}{D \sum_n r_{nk}}$$

2.2 $$\arg\max_{q_n \in \Delta} E_{z_n \sim q_n}[\ln p(x_n, z_n; \theta^t)] - E_{z_n \sim q_n}[\ln q_n]$$

$$= \arg\max_{q_n \in \Delta} q_n \ln p(x_n, z_n; \theta^t) - q_n \ln q_n$$

from 1.2

$$q_n^{(t)}(k) = \frac{p(x_n, z_n; \theta^t)}{\sum_{k=1}^{K} p(x_n, z_n; \theta^t)} = \frac{p(x_n, z_n = k; \theta^t)}{p(x_n; \theta^t)}$$

$$= p(z_n = k \mid x_n; \theta^t)$$

2.3 $p(z_n=k|x_n)=\dfrac{w_k N(x_n|\mu_k,\Sigma_k)}{\sum w_k N(x_n|\mu_k,\Sigma_k)}$

let $w=\frac{1}{K}$,

$$p(z_n=k|x_n)=\frac{e^{-\frac{1}{2\sigma^2}\|x_n-\mu_k\|_2^2}}{\sum_K e^{-\frac{1}{2\sigma^2}\|x_n-\mu_k\|_2^2}}$$

let $\sigma \to 0$

$\lim_{\sigma\to 0} p(z_n=k|x_n)=1$ if $\|x_n-\mu_k\|_2^2=\min\|x_n-\mu_j\|_2^2$ $j\in[1,K]$

$\lim_{\sigma\to 0} p(z_n=k|x_n)=0$ if $\|x_n-\mu_k\|_2^2 \mathrel{!=} \min\|x_n-\mu_j\|_2^2$ $j\in[1,K]$

so, $p(z_n=k|x_n)=\mathbb{I}[k==\operatorname{argmin}\|x_n-\mu_j\|_2^2]$ $j\in[1,K]$

Problem 3 3.1 $P(X_{T+1}=s \mid O_{1:T}=o_{1:T})$

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

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

$= P(X_{T+1}=s \mid X_T=s')\, \alpha_{s'}(T)$

$= a_{s,s'}\, \alpha_{s'}(T)$

3.2

(a) $\cancel{\alpha_s(2) = b_{s,o}}\ b_{s,o_t} \sum_{s'} \alpha_{s's}\alpha_{s'}(t-1)$

$\cancel{= P(X_2=B \mid X_1=A)}$

$\cancel{\alpha_s(2) =}\ \alpha_A(1) = \pi_A b_{A,0} = 0.7 \times 0.4 = 0.28$

$\alpha_B(1) = \pi_B b_{B,0} = 0.3 \times 0.7 = 0.21$

$\alpha_A(2) = P(X_{t+1}=A \mid X_t=A)\, \alpha_A(1)\, P(O_t=on \mid X_t=A)$

$+ P(O_t=on \mid X_t=A)\, P(X_{t+1}=A \mid X_t=B)\, \alpha_B(1)$

$= 0.2 \times 0.28 \times 0.6 + 0.6 \times 0.21 \times 0.7$

$= 0.1218$

$\alpha_B(2) = P(O_t=on \mid X_t=B)\, P(X_{t+1}=B \mid X_t=A)\, \alpha_A(1)$

$+ P(O_t=on \mid X_t=B)\, P(X_{t+1}=B \mid X_t=B)\, \alpha_B(1)$

$= 0.3 \times (0.7 \times 0.28 + 0.2 \times 0.21)$

$= 0.0714$

$\cancel{\delta_A(2) = P(X_t=A) = P(X}\ P(X_1=A)P(X_2=A \mid X_1=A)P(X_3=A \mid X_2=A)$

$\cancel{+P(}$

$P(X_t=A) = \sum_{s_1,s_2} P(X_1=s_1)P(X_2=s_2 \mid X_1=s_1)P(X_3=A \mid X_2=s_2)$

$= 0.7\times0.2\times0.2 + 0.7\times0.7\times0.7 + 0.3\times0.2\times0.7 + 0.3\times0.7\times0.2$

$= 0.455$

$P(X_t=B) = 0.7\times0.2\times0.7 + 0.7\times0.7\times0.2 + 0.3\times0.7\times0.7 + 0.3\times0.2\times0.2$

$= 0.355$

$$P(X_3=A \mid O_{1:2}) = \frac{\alpha_A(2)\, P(X_3=A \mid X_2=A) + \alpha_B(2)\, P(X_3=A \mid X_2=B)}{\alpha_A(2)+\alpha_B(2)}$$

$$= 0.3848$$

$$P(X_3=B \mid O_{1:2}) = \frac{\alpha_A(2)\, P(X_3=B \mid X_2=A) + \alpha_B(2)\, P(X_3=B \mid X_2=B)}{\alpha_A(2)+\alpha_B(2)}$$

$$= 0.5152$$

$P(X_3=\text{End} \mid O_{1:2}) = 1 - 0.3848 - 0.5152 = 0.1$

most likely state is B

~~most likely state at t=3 is A~~

(b) $\delta_A(1) = 0.28$

$\delta_B(1) = 0.21$

$\delta_A(2) = P(O_2 = on | X_2 = A) \max\{P(X_2=A|X_1=A)\delta_A(1), P(X_2=A|X_1=B)\delta_B(1)\}$

$= 0.6 \max\{0.2 \times 0.28, 0.7 \times 0.21\}$

$= 0.0882$

$\Delta_A(2) = B$

$\delta_B(2) = P(O_2 = on | X_2 = B) \max\{P(X_2=B|X_1=A)\delta_A(1), P(X_2=B|X_1=B)\delta_B(1)\}$

$= 0.3 \max\{0.7 \times 0.28, 0.2 \times 0.21\}$

$= 0.0588$

$\Delta_B(2) = A$

$\delta_A(3) = P(O_3 = on | X_3 = A) \max\{P(X_3=A|X_2=A)\delta_A(2), P(X_3=A|X_2=B)\delta_B(2)\}$

$= 0.6 \max\{0.2 \times 0.0882, 0.7 \times 0.0588\}$

$= 0.0247$

$\Delta_A(3) = B$

$\delta_B(3) = P(O_3 = on | X_3 = B) \max\{P(X_3=B|X_2=A)\delta_A(2), P(X_3=B|X_2=B)\delta_B(2)\}$

$= 0.3 \max\{0.7 \times 0.0882, 0.2 \times 0.0588\}$

$= 0.0185 < 0.0247$

$\Delta_B(3) = A$

so, most likely sequence is ABA

3.3 let $P(X_1 = s) = w_s$

$P(X_{t+1} = s' | X_t = s) = w_{s'}$

$P(O_t | X_t = s) = \cancel{N(\mu_k, \Sigma_k)}\ N(\mu_t, \Sigma_t)$

$P(X_2 = s' | O_1, O_2) = P(X_2 = s' | O_2) = \frac{P(X_2 = s, O_2)}{P(O = O_2)} = \frac{P(X_2 = s)P(O_2 | X_2 = s)}{P(O = O_2)}$

$= \frac{w_s \times N(\mu_t, \Sigma_t)}{\sum_k w_k \times N(\mu_t, \Sigma_t)} = P(z = s | O = O_2)$