Yucheng Yang 2333448896 HW4

problem 1:

1.1

let

$$\frac{a_k}{q_k^*} + \lambda + \lambda_k = 0$$

$$\Downarrow$$

$$q_k^* = -\frac{a_k}{\lambda + \lambda_k}$$

And because of the complementary slackness condition.

$\lambda_k q_k^* = 0$, and $\lambda_k = 0$

$$\therefore \sum_{k=1}^{k} q_k^* = -\sum_{k=1}^{k} \frac{a_k}{\lambda} = 1$$

$$\therefore \quad q_k^* = \frac{a_k}{\sum_{k'} a_{k'}}$$

1.2 Lagrangian:

$$L(q, \lambda, \lambda_1, \cdots, \lambda_k) = \sum_{k=1}^{k} (q_k b_k - q_k \ln q_k) + \lambda\left(\sum_{k=1}^{k} q_k - 1\right) + \sum_{k=1}^{k} \lambda_k q_k$$

and $\lambda \neq 0$ $\lambda_k \geq 0$

let

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

$$q_k = \exp(b_k - 1 + \lambda + \lambda_k) \propto e^{b_k + \lambda_k = 0}.$$

$$\therefore \quad q_k \propto e^{b_k}$$

1.3

$$q^*_{nk} = \frac{P(x_n, z_n = k; \theta^{(t)})}{\sum_{k=1}^{K} P(x_n, z_n = k \mid \theta^{(t)})} = \frac{P(x_n, z_n = k; \theta^{(t)})}{P(x_n \mid \theta^{(t)})}$$

$$= P(z_n = k \mid x_n, \theta^{(t)})$$

Problem 2:

To solve this w_k, we only need to solve $\sum_n \sum_k \gamma_{nk} \ln w_k$.

Like in 1.1 with $a_k = \sum_n \gamma_{nk}$

we got $w_k = \frac{\sum_n \gamma_{nk}}{\sum_k \sum_n \gamma_{nk}} = \frac{\sum_n \gamma_{nk}}{N}$

To find μ_k and σ_k

$$\arg\max_{\mu_k, \sigma_k} \sum_n \gamma_{nk} \ln N(x_n \mid \mu_k, \sigma_k) = \arg\max_{\mu_k, \sigma_k} \sum_n \gamma_{nk} \left(\ln \frac{1}{\sigma_k} - \frac{(x_n - \mu_k)^2}{2\sigma_k^2}\right)$$

$$\rightarrow \sum_n \gamma_{nk}(x_n - \mu_k) = 0$$

$$\mu_k = \frac{\sum_n \gamma_{nk} x_n}{\sum_n \gamma_{nk}}.$$

$$\sum_n \gamma_{nk}\left(-\frac{1}{\sigma_k} + \frac{(x_n - \mu_k)^2}{\sigma_k^3}\right) = 0$$

$$\therefore \sigma_k^2 = \frac{\sum_n \gamma_{nk}(x_n - \mu_k)^2}{\sum_n \gamma_{nk}}$$

problem 3.

3.1

$$\gamma_0' = P(z_n=0 \mid x_n>0;\pi,\lambda) = \frac{P(x_n>0 \mid z_n=0;\pi,\lambda)P(z_n=0;\pi,\lambda)}{P(x_n>0;\pi,\lambda)} = 0$$

$$\gamma_1' = P(z_n=1 \mid x_n>0;\pi,\lambda) = \frac{P(x_n>0 \mid z_n=1;\pi,\lambda)P(z_n=1;\pi,\lambda)}{P(x_n>0;\pi,\lambda)} = 1$$

$$\gamma_0 = P(z_n=0 \mid x_n=0;\pi,\lambda) = \frac{P(x_n=0 \mid z_n=0;\pi,\lambda)P(z_n=0;\pi,\lambda)}{P(x_n=0;\pi,\lambda)} = \frac{1-\pi}{1+\pi e^{-\lambda}-\pi}$$

$$\gamma_1 = P(z_n=1 \mid x_n=0;\pi,\lambda) = \frac{P(x_n=0 \mid z_n=1;\pi,\lambda)P(z_n=1;\pi,\lambda)}{P(x_n=0;\pi,\lambda)} = \frac{\pi e^{-\lambda}}{1+\pi e^{-\lambda}-\pi}$$

3.2

$$\begin{aligned} Q(\pi,\lambda) &= E_{z|x;\pi,\lambda}(\ln p(X,Z;\pi,\lambda)) \\ &= \sum_{n=1}^{N}\sum_{z_n\in\{0,1\}} P(z_n \mid x_n;\pi^t,\lambda^t)\ln p(x_n,z_n;\pi,\lambda) \\ &= \sum_{n:x_n>0}\left(0\ln 0 + \ln\frac{\pi\lambda^{x_n}e^{-\lambda}}{x_n!}\right) + \sum_{n:x_n=0}\left(\gamma_0\ln(1-\pi)+\gamma_1\ln(\pi e^{-\lambda})\right) \end{aligned}$$

3.3

$$\frac{\partial Q(\pi,\lambda)}{\partial\pi} = \sum_{n:x_n>0}\frac{1}{\pi} + N_0\left(-\frac{\gamma_0}{1-\pi}+\frac{\gamma_1}{\pi}\right)$$

$$\frac{\partial Q(\pi,\lambda)}{\partial\lambda} = \sum_{n:x_n>0}\left(\frac{x_n}{\lambda}-1\right) + \sum_{n:x_n=0}(-\gamma_1)$$

3.4 In order to find π^{new} and λ^{new}

we just find π^* and λ^* in 3.3 and

plugging it back to 3.1, and we can find the solution.