Haniyeh Ehsani Oskouie/ 306300374

Problem1:

1. E-step according to the slides: $p(z|w,d) = \dfrac{\beta_{zw}\,\theta_{dz}}{\sum_{z'}\beta_{z'w}\,\theta_{dz'}}$

$\beta_{zw}: p(w|z)$ $\qquad$ $\theta_{dz}: p(z|d)$

Given $\theta_{11}^{(0)} = 0.3$, $\theta_{21}^{(0)} = 0.4$, $\beta_1^{(0)} = (1,0,0,0)$, $\beta_2^{(0)} = (0, 0.4, 0.3, 0.3)$,
we have: $\theta_{12}^{(0)} = 1-0.3 = 0.7$ $\qquad$ $\theta_{22}^{(0)} = 1-0.4 = 0.6$

$$P(z=1|w=A,d_1) = \frac{\beta_{11}^{(0)}\theta_{11}^{(0)}}{\beta_{11}^{(0)}\theta_{11}^{(0)} + \beta_{21}^{(0)}\theta_{12}^{(0)}} = \frac{(1)(0.3)}{(1)(0.3)+(0)(0.7)} = 1$$

$$P(z=1|w=B,d_1) = \frac{\beta_{12}^{(0)}\theta_{11}^{(0)}}{\beta_{12}^{(0)}\theta_{11}^{(0)} + \beta_{22}^{(0)}\theta_{12}^{(0)}} = \frac{(0)(0.3)}{(0)(0.3)+(0.4)(0.7)} = 0$$

$p(z=1|w=C,d_1) = 0$ because $\beta_1^{(0)}$ is 0 everywhere except for $\beta_{11}^{(0)}$

$p(z=1|w=D,d_1) = 0$ similarly.

2. M-step according to the slides:

$$\beta_{zw} = \frac{\sum_d p(z|w,d)\,c(w,d)}{\sum_{w',d} p(z|w',d)\,c(w',d)} \qquad \theta_{dz} = \frac{\sum_w P(z|w,d)\,c(w,d)}{N_d}$$

$$\beta_{11} = \frac{1(4)+1(2)}{1(4)+1(2)+0} = 1 \qquad \beta_{12} = \frac{0(3)+0(2)}{1(4)+1(2)+0} = 0$$

$$\theta_{11} = \frac{1(4)+0(3)+0(2)+0(1)}{4+3+2+1} = 0.4$$

$$\theta_{12} = \frac{0(4)+1(3)+1(2)+1(1)}{4+3+2+1} = 0.6$$

Problem 2:

1.

$$p(z_i = k \mid n_i; \beta, \pi) = \frac{p(n_i \mid z_i = k; \beta, \pi)\, p(z_i = k; \beta, \pi)}{\sum_z p(n_i \mid z_i = z; \beta, \pi)\, p(z_i = z; \beta, \pi)}$$

$$\Rightarrow p(z_i = k \mid n_i; \beta, \pi) = \frac{\frac{(\sum_n n_{in})!}{\prod_n n_{in}!} \prod_n \beta_{kn}^{n_{in}} \pi_k}{\frac{(\sum_n n_{in})!}{\prod_n n_{in}!} \sum_z \prod_n \beta_{zn}^{n_{in}} \pi_z}$$

$$= \frac{\prod_n \beta_{kn}^{n_{in}} \pi_k}{\sum_z \prod_n \beta_{zn}^{n_{in}} \pi_z} \Rightarrow p(z_i = k \mid n_i; \beta, \pi) \propto \prod_n \beta_{kn}^{n_{in}} \pi_k$$

2. we should find β and π such that they maximize the following

$$\text{log likelihood} = \log p(n_{1:n}, z_{1:n}; \pi, \beta_{1:k})$$

$$= \sum_i \log p(n_i, z_i; \pi, \beta_{1:k}) = \sum_i \log\left(p(z_i = k)\, p(n_i \mid z_i = k)\right)$$

$$= \sum_i \log p(z_i = k) + \log p(n_i \mid z_i = k)$$

$$= \sum_i \log \pi_{z_i} + \log \prod_n \beta_{z_i n}^{n_{in}} = \sum_i \log \pi_{z_i} + \sum_n n_{in} \log \beta_{z_i n}$$

$$J = \sum_{i,n} n_{in} \log \beta_{z_i n} + \sum_i \lambda\left(\sum_n \beta_{z_i n} - 1\right) \rightarrow \text{Lagrangian}$$

$$J = \sum_{in} \sum_k \mathbb{1}(z_i = k)\, n_{in} \log \beta_{kn} + \sum_k \lambda\left(\sum_n \beta_{kn} - 1\right)$$

$$\frac{dJ}{d\beta_{kn}} = \sum_i \frac{x_{in}\,\mathbb{1}(z_i=k)}{\beta_{kn}} + \lambda = 0$$

$$\Rightarrow \sum_i x_{in}\,\mathbb{1}(z_i=k) + \lambda\,\beta_{kn} = 0$$

Sum on $n \rightarrow \sum_{n,i} x_{in}\,\mathbb{1}(z_i=k) + \lambda \underbrace{\left(\sum_n \beta_{kn}\right)}_{1} = 0$

$$\Rightarrow \lambda = -\sum_{n,i} x_{in}\,\mathbb{1}(z_i=k)$$

$$\Rightarrow \beta_{kn} = \frac{\sum_i x_{in}\,\mathbb{1}(z_i=k)}{\sum_{n,i} x_{in}\,\mathbb{1}(z_i=k)}$$

$J = \sum_i \sum_k \log \pi_k\,\mathbb{1}(z_i=k) + \lambda\left(\sum_k \pi_k - 1\right) \rightarrow$ lagrangian

$$\frac{dJ}{\pi_j} = \frac{\mathbb{1}(z_i=j)}{\pi_j} + \lambda \cdot 1 = 0 \Rightarrow \mathbb{1}(z_i=j) + \lambda\,\pi_j = 0$$

summation on π_js $\rightarrow \sum_j \mathbb{1}(z_i=j) + \lambda \underbrace{\left(\sum_j \pi_j\right)}_{1} = 0$

$$\Rightarrow \lambda = -\sum_j \mathbb{1}(z_i=j)$$

$$\Rightarrow \pi_j = \frac{\mathbb{1}(z_i=j)}{\sum_j \mathbb{1}(z_i=j)}$$

instead of $\mathbb{1}(z_i=j)$ use $p(z_i=k \mid x_i; \beta, \pi)$
as estimited in previous part