

Observed data

$$\mathbf{x} = (x_1, x_2, \dots, x_N)$$

Hidden variables

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$$

Two hidden states

$$1, 2$$

Two emission probability

$$P(X_i = x | Z_i = 1) = f_{\varphi_1}(x)$$

$$P(X_i = x | Z_i = 2) = f_{\varphi_2}(x)$$

Transition probability

$$P(Z_i = j | Z_{i-1} = k) = a_{jk}, j, k \in \{1, 2\}$$

initial probability

$$P(Z_1 = i) = \pi_i, i \in \{1, 2\}$$

Variables to be inferenced

$$\theta = \{\varphi, A, \pi\}$$

E step :

$$Q(\theta, \theta_s) = E_{Z|X; \theta_s}[\log(P(Z, X; \theta))]$$

$$\begin{aligned} P(Z, X; \theta) &= P(Z_1)P(X_1|Z_1)P(X_2|Z_2)P(Z_2|Z_1) \dots P(X_N|Z_N)P(Z_N|Z_{N-1}) \\ &= P(Z_1)P(X_1|Z_1) \prod_{i=2}^N P(X_i|Z_i)P(Z_i|Z_{i-1}) \end{aligned}$$

$$P(Z_1) = \pi_1^{\mathbb{1}(Z_1=1)} \pi_2^{\mathbb{1}(Z_1=2)}$$

$$P(X_i|Z_i) = f_{\varphi_1}(X_i)^{\mathbb{1}(Z_i=1)} f_{\varphi_2}(X_i)^{\mathbb{1}(Z_i=2)}$$

$$P(Z_i|Z_{i-1}) = a_{11}^{\mathbb{1}(Z_i=1)\mathbb{1}(Z_{i-1}=1)} a_{12}^{\mathbb{1}(Z_i=2)\mathbb{1}(Z_{i-1}=1)} a_{21}^{\mathbb{1}(Z_i=1)\mathbb{1}(Z_{i-1}=2)} a_{22}^{\mathbb{1}(Z_i=2)\mathbb{1}(Z_{i-1}=2)}$$

$$\begin{aligned} P(Z, X; \theta) &= \pi_1^{\mathbb{1}(Z_1=1)} \pi_2^{\mathbb{1}(Z_1=2)} \cdot \\ &\quad \prod_{i=1}^N f_{\varphi_1}(X_i)^{\mathbb{1}(Z_i=1)} f_{\varphi_2}(X_i)^{\mathbb{1}(Z_i=2)} \cdot \\ &\quad \prod_{i=2}^N a_{11}^{\mathbb{1}(Z_i=1)\mathbb{1}(Z_{i-1}=1)} a_{12}^{\mathbb{1}(Z_i=2)\mathbb{1}(Z_{i-1}=1)} a_{21}^{\mathbb{1}(Z_i=1)\mathbb{1}(Z_{i-1}=2)} a_{22}^{\mathbb{1}(Z_i=2)\mathbb{1}(Z_{i-1}=2)} \\ &= \prod_{i=1}^2 \pi_1^{\mathbb{1}(Z_1=i)} \prod_{i=1}^N \prod_{j=1}^2 f_{\varphi_j}(X_i)^{\mathbb{1}(Z_i=j)} \prod_{i=2}^N \prod_{j=1}^2 \prod_{k=1}^2 a_{jk}^{\mathbb{1}(Z_i=k)\mathbb{1}(Z_{i-1}=j)} \end{aligned}$$

$$\begin{aligned} \log(P(Z, X; \theta)) &= \sum_{i=1}^2 \mathbb{1}(Z_1 = i) \log(\pi_i) + \sum_{i=1}^N \sum_{j=1}^2 \mathbb{1}(Z_i = j) \log(f_{\varphi_j}(X_i)) + \\ &\quad \sum_{i=2}^N \sum_{j=1}^2 \sum_{k=1}^2 \mathbb{1}(Z_i = k) \mathbb{1}(Z_{i-1} = j) \log(a_{jk}) \end{aligned}$$

$$E_{Z|X; \theta_s} \left[\sum_{i=1}^2 \mathbb{1}(Z_1 = i) \log(\pi_i) \right] = \sum_{i=1}^2 E_{Z|X; \theta_s} [\mathbb{1}(Z_1 = i) \log(\pi_i)] = \sum_{i=1}^2 P(Z_1 = i | X; \theta_s) \log(\pi_i)$$

$$\begin{aligned} E_{Z|X; \theta_s} \left[\sum_{i=1}^N \sum_{j=1}^2 \mathbb{1}(Z_i = j) \log(f_{\varphi_j}(X_i)) \right] &= \sum_{i=1}^N \sum_{j=1}^2 E_{Z|X; \theta_s} [\mathbb{1}(Z_i = j) \log(f_{\varphi_j}(X_i))] \\ &= \sum_{i=1}^N \sum_{j=1}^2 P(Z_i = j | X; \theta_s) \log(f_{\varphi_j}(X_i)) \end{aligned}$$

$$\begin{aligned} E_{Z|X; \theta_s} \left[\sum_{i=1}^N \sum_{j=1}^2 \sum_{k=1}^2 \mathbb{1}(Z_i = k) \mathbb{1}(Z_{i-1} = j) \log(a_{jk}) \right] &= \sum_{i=1}^N \sum_{j=1}^2 \sum_{k=1}^2 E_{Z|X; \theta_s} [\mathbb{1}(Z_i = k) \mathbb{1}(Z_{i-1} = j) \log(a_{jk})] \\ &= \sum_{i=2}^N \sum_{j=1}^2 \sum_{k=1}^2 P(Z_i = k, Z_{i-1} = j | X; \theta_s) \log(a_{jk}) \end{aligned}$$

$$\text{Let } \gamma_{ij} = P(Z_i = j | X; \theta_s); \beta_{ijk} = P(Z_i = k, Z_{i-1} = j | X; \theta_s)$$

$$Q(\theta, \theta_s) = \sum_{i=1}^2 \gamma_{1i} \log(\pi_i) + \sum_{i=1}^N \sum_{j=1}^2 \gamma_{ij} \log(f_{\varphi_j}(X_i)) + \sum_{i=2}^N \sum_{j=1}^2 \sum_{k=1}^2 \beta_{ijk} \log(a_{jk})$$

M step :

For initial probability π

$$\frac{dQ(\theta, \theta_s)}{d\pi_1} = \frac{\gamma_{11}}{\pi_1} - \frac{\gamma_{12}}{1 - \pi_1} = 0$$

$$\gamma_{11} (1 - \pi_1) = \gamma_{12} \cdot \pi_1$$

$$\pi_1 = \frac{\gamma_{11}}{\gamma_{11} + \gamma_{12}}, \quad \pi_1 = \frac{\gamma_{12}}{\gamma_{11} + \gamma_{12}}$$

For emission probability

$$\frac{dQ(\theta, \theta_s)}{d\varphi_j} = \sum_{i=1}^N \gamma_{ij} \frac{d\log(f_{\varphi_j}(X_i))}{d\varphi_j} = 0$$

For transition probability

$$\frac{dQ(\theta, \theta_s)}{d\varphi_{i11}} = \sum_{i=2}^N \left(\frac{\beta_{i11}}{a_{11}} - \frac{\beta_{i12}}{1 - a_{11}} \right) = 0$$

$$\sum_{i=2}^N \beta_{i12} \cdot a_{11} = (1 - a_{11}) \sum_{i=2}^N \beta_{i11}$$

$$a_{11} = \frac{\sum_{i=2}^N \beta_{i11}}{\sum_{i=2}^N (\beta_{i11} + \beta_{i12})} = \frac{\sum_{i=2}^N \beta_{i11}}{\sum_{i=2}^N (\beta_{i11} + \beta_{i12})} = \frac{\sum_{i=2}^N \beta_{i11}}{\sum_{i=1}^{N-1} \gamma_{i1}}$$

$$a_{12} = \frac{\sum_{i=2}^N \beta_{i12}}{\sum_{i=1}^{N-1} \gamma_{i1}}$$