Baum-Welch

First consider complete data

$$D = \{(x^1_{1:T}, z^1_{1:T+1}), \ldots, (x^N_{1:T}, z^N_{1:T+1})\}$$

Likelihood

$$\prod_{n=1}^{N} \prod_{t=1}^{T} p(x^n_t \mid z^n_t) \prod_{t=1}^{T+1} p(z^n_t \mid z^n_{t-1})$$

$$= \prod_n \prod_t \left[\prod_{s,h} B_{sh}^{\mathbb{I}(x^n_t = s,\, z^n_t = h)} \right] \prod_{h,\ell} A_{h\ell}^{\mathbb{I}(z^n_t = \ell,\, z^n_{t+1} = h)}$$

$$= \prod_{s,h} B_{sh}^{M_{sh}} \prod_{h,\ell} A_{h\ell}^{N_{h\ell}}$$

where $M_{sh} = \#(n,t) : x^n_t = s,\, z^n_t = h$ $\qquad N_h = \#(n,t) : z^n_t = h$

$N_{h\ell} = \#(n,t) : z^n_t = \ell,\, z^n_{t+1} = h$

So parameter MLE is

$B_{sh} = M_{sh} / \sum_s M_{sh} = M_{sh} / N_h$ & $A_{h\ell} = N_{h\ell} / N_h$

$$\sum_n E_{p(Z_{1:T+1} | x^n_{1:T}, \theta)} \left[\ell(\theta'; Z_{1:T+1}, x^n_{1:T}) \right]$$

$$= \sum_n E_{p(Z_{1:T+1} | x^n_{1:T}, \theta)} \log \prod_t \left[\prod_{k,s} B'^{\, \mathbb{I}(x^n_t = s, Z_t = k)}_{sk} \right.$$

$$\left. \prod_{k,\ell} A'^{\, \mathbb{I}(Z_t = k, Z_{t+1} = \ell)}_{\ell k} \right]$$

($\prod_t$ over x^n)

$$= \sum_n \sum_t \sum_{k,s} p(X_t = s, Z_t = k \mid x^n, \theta) \log B'_{sk}$$

$$\sum_n \sum_t \sum_{k,\ell} p(Z_t = k, Z_{t+1} = \ell \mid x^n, \theta) \log A'_{\ell k}$$

$$= \sum_k \sum_s \left[\sum_{n,t} p(X_t = s, Z_t = k \mid x^n, \theta) \right] \log B'_{sk}$$

$\bar{M}_{k,s}$

same k comp. ℓ for prob. mass hit diff

$$\sum_k \sum_\ell \left[\sum_{n,t} p(Z_t = k, Z_{t+1} = \ell \mid x^n, \theta) \right] \log A'_{\ell k}$$

$\bar{N}_{k,\ell}$

So exp. comp. log-lik. is max by

$B'_{sk} = \bar{M}_{sk} / \sum \bar{M}_{\cdot k} = \bar{M}_{sk} / \bar{N}_{\cdot}$ & $A'_{\cdot k} = \bar{N}_{\cdot k} / \bar{N}_{\cdot}$

$\sum_s \bar{M}_{sk} \quad \ldots / \bar{N}_k \quad \text{or} \quad \bar{M}_{kl} \quad \ldots / \bar{N}_k$

where $\bar{N}_k = \sum_s \bar{M}_{sk} = \sum_l \bar{N}_{kl}$