Jeff Gould
Math 611
HW 10

1a) suppose we know $Z(t)$ for $t=0,1,\dots,T$

MLE: $\max_{\theta}$

$$\max_{\theta} \log\left(\pi_{Z(0)} M_{Z(0)Z(1)} M_{Z(1)Z(2)} \cdots M_{Z(T-1)Z(T)}\, g_{Z(0)}(Y(0)) \cdots g_{Z(T)}(Y(T))\right)$$

$$= \max_{\theta} \underbrace{\log(\pi_{Z(0)}) + \sum_{t=0}^{T-1} \log M_{Z(t)Z(t+1)} + \sum_{t=0}^{T} \log g_{Z(t)}(Y(t))}_{L(\theta)}$$

$\nabla L(\theta) = 0$, solve for θ

$$\frac{\partial}{\partial M_{ij}} L(\theta) = \sum_{\substack{t=0 \\ Z(t)=i \\ Z(t+1)=j}}^{T-1} \frac{1}{M_{ij}} = \frac{1}{M_{ij}} \sum_{\substack{t=0 \\ Z(t)=i, \\ Z(t+1)=j}}^{T-1} 1 = \frac{N_{ij}}{M_{ij}}$$

, note $\sum_{j=1}^{n} M_{ij} = 1$
$n=2$ in this example

$$G(M_{i1}, M_{i2}) = \sum_{j=1}^{2} M_{ij} = 1$$

$$\begin{pmatrix} \frac{\partial L}{\partial M_{i1}}(\theta) \\ \frac{\partial L}{\partial M_{i2}}(\theta) \end{pmatrix} = \lambda \nabla G = \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\frac{N_{ij}}{M_{ij}} = \lambda$, for $j=1,2$

$$M_{ij} = \frac{N_{ij}}{\lambda} \rightarrow 1 = \sum_{j=1}^{2} M_{ij} = \frac{\sum_{j=1}^{2} N_{ij}}{\lambda} = \frac{N_{i\bullet}}{\lambda}, \Rightarrow \lambda = N_{i\bullet}$$

$$\Rightarrow M_{ij} = \frac{N_{ij}}{N_{i\bullet}} \blacksquare$$

---

g: constraint - if $g_j(i) = P(Y(t)=i \mid Z(t)=j)$, then $\sum_{i=1}^{6} g_j(i) = 1$

$$\frac{\partial L(\theta)}{\partial g_j(i)} = \sum_{\substack{t=0 \\ Z(t)=j \\ Y(t)=i}}^{T} \frac{1}{g_{Z(t)}(Y(t))} = \frac{1}{g_j(i)} \sum_{t=0}^{T} I(Z(t)=j \text{ AND } Y(t)=i)$$

$$\begin{pmatrix} \frac{\partial L(\theta)}{\partial g_j(1)} \\ \vdots \\ \frac{\partial L(\theta)}{\partial g_j(6)} \end{pmatrix} = \lambda \nabla G = \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\frac{\sum_{t=0}^{T} I(Z(t)=j \text{ AND } Y(t)=i)}{g_j(i)} = \lambda$$

$$1 = \sum_{i=1}^{6} g_j(i) = \frac{\sum_{t=1}^{6} I(Z(t)=j)}{\lambda}$$

$$\Rightarrow g_j(i) = \frac{\sum_{t=0}^{T} I(Z(t)=j \text{ AND } Y(t)=i)}{\sum_{t=0}^{T} I(Z(t)=j)} \blacksquare$$

$$\frac{\partial L(\theta)}{\partial \pi} = \frac{1}{\pi_{Z(0)}}$$

$\pi = (\pi_1, \pi_2)$

$\max \log \pi_{Z(0)} \rightarrow$ Clearly set $\pi_i = \begin{cases} 1 & \text{if } Z(0)=i \\ 0 & \text{else} \end{cases}$

1a) continued

$P(Z(t)=i \mid Y(0), \ldots Y(\text{zero}))$ • Let $\alpha_t(i) = P(Y(0), \ldots, Y(t), Z(t)=i)$

$\cancel{\beta_t}\ \beta_t(i) = P(Y(t+1), Y(t+2), \ldots, Y(T) \mid Z(t)=i)$

$$\hookrightarrow = \frac{P(Z(t)=i, Y(0), \ldots, Y(T))}{P(Y(0), \ldots, Y(T))} \leftarrow D$$

$$= \frac{P(Y(0), Y(1), \ldots, Y(t), Z(t)=i) \cdot P(Y(t+1), \ldots, Y(T) \mid Z(t)=i)}{D}$$

$$= \frac{\alpha_t(i)\beta_t(i)}{D} \Rightarrow \text{Forward-Backwards Algorithm}$$

1b)

Let θ' be the current parameters

$$Q(\theta', \theta) = E_{\theta'}[\log P(\underbrace{Z(0), \ldots,}_{\theta'} Z(T), Y(0), \ldots, Y(T) \mid \theta)]$$

$$\max_{\theta} Q(\theta', \theta) = \max_{\theta} E[\log P(\quad)]$$

$$= \sum_{Z(0)=1}^{2} \sum_{Z(1)=1}^{2} \sum_{Z(2)=1}^{2} \cdots \sum_{Z(T)=1}^{2} [\log(\pi_{Z(0)}) + \sum_{t=0}^{T-1} \log M_{Z(t)Z(t+1)} + \sum_{t=0}^{T} \log g_{Z(t)}(Y(t))] \cdot P(Z(0), Z(1), \ldots, Z(T) \mid Y(0), Y(1), \ldots, Y(T), \theta')$$

set $\nabla_\theta Q(\theta', \theta) = 0$

$$\frac{\partial}{\partial M_{ij}} Q(\theta', \theta) = \sum_{Z(0)=1}^{2} \sum_{Z(1)=1}^{2} \cdots \sum_{Z(T)=1}^{2} \sum_{t=0}^{T-1} [\frac{1}{M_{ij}} I(Z(t)=i, Z(t+1)=j) \cdot P(Z(0), Z(1), \ldots Z(T) \mid Y(0), \ldots, Y(T), \theta')]$$

$$= \sum_{t=0}^{T-1} \sum_{Z(t)=1}^{2} \sum_{Z(t+1)=1}^{2} \Sigma \cdots \Sigma [\frac{1}{M_{ij}} I(Z(t)=i, Z(t+1)=j) \cdot P(\text{---})]$$

$$= \sum_{t=0}^{T-1} \sum_{Z(t)=1}^{2} \sum_{Z(t+1)=1}^{2} \frac{1}{M_{ij}} I(Z(t)=i, Z(t+1)=j) P(Z(t)=i, Z(t+1)=j \mid Y(0), \ldots, Y(T), \theta')$$

$\cdot\, p(Z(t)=i, Z(t+1)=j \mid Y(0),\dots,Y(T),\theta')$

$$= \underbrace{p(Y(0),\dots,Y(t),Z(t)=i)}_{\alpha_t(i)}\ \underbrace{p(Y(t+2),\dots Y(T)\mid Z(t+1)=j)}_{\beta_{t+1}(j)}\ \overbrace{p(Y(t+1)\mid Z(t+1)=j)}^{g_j(Y(t+1))} \cdot \underbrace{p(Z(t+1)=j \mid Z(t)=i)}_{M'_{ij}}$$

$$= \alpha_t(i)\, g_j(Y(t+1))\,\beta_{t+1}(j)\, M'_{ij} = \psi_t(i,j)$$

$$\Rightarrow \frac{\partial Q(\theta',\theta)}{\partial M_{ij}} = \sum_{t=0}^{T-1} \frac{1}{M_{ij}}\, I(Z(t)=i, Z(t+1)=j)\ \psi_t(i,j)$$

following the same steps from 1a)

$$\Rightarrow M_{ij} = \frac{\sum_{t=0}^{T-1} \psi_t(i,j)}{\sum_{t=0}^{T-1}\sum_{k=1}^{2} \psi_t(i,k)} \quad \blacksquare$$

---

define $\gamma_t(j) = p(Z(t)=j \mid Y(0),\dots Y(T),\theta')$

$$\frac{\partial}{\partial g_j(i)} Q(\theta',\theta) = \sum_{\substack{t=0\\ Y(t)=i}}^{T} \left(\frac{1}{g_{Z(t)}(Y(t))}\right) \cdot p(Z(t)=j \mid Y(0),\dots,Y(T)\,\theta')$$

$$= \sum I(Y(t)=i)\ \gamma_t(j)$$

$\rightarrow$ constraint $\sum_{j=1}^{6} \gamma_t(j) = 1$

- Lagrange as before

$$\rightarrow \frac{\sum_{t=0}^{T} I(Y(t)=i)\,\gamma_t(j)}{\sum_{t=0}^{T} \gamma_t(j)} \quad \blacksquare$$

$$\frac{\partial}{\partial \pi_i} Q(\theta',\theta) = 0$$

$$\left.\begin{array}{l} \pi_i = p(Z(0)=i \mid Y(0),\dots,Y(T),\theta') \\ \sum_{j=1}^{6} p(Z(0)=j \mid Y(0),\dots,Y(T),\theta') = 1 \end{array}\right\} = \underline{p(Z(0)=i \mid Y(0),\dots,Y(T),\theta')}$$