CSCI 567 lec 18

Oct 26th, 2017

$$\sum_{m=1}^{M} \log \pi_{x_1^m} = \sum_{i=1}^{N} k_i \log \pi_i = \sum_{i=1}^{N-1} k_i \log \pi_i + k_N \log\left(1-\sum_{i=1}^{N-1}\pi_i\right)$$

$x_1^m \in \{1, \ldots, N\}$

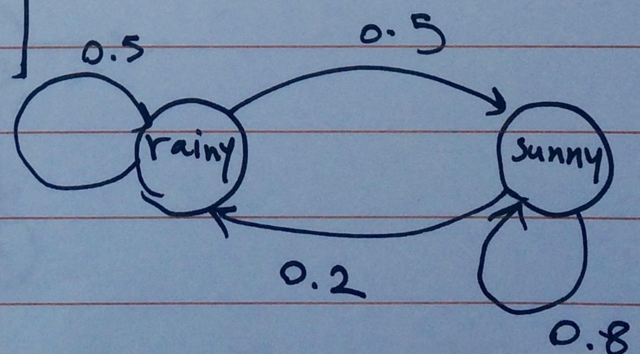
k_i: # ~~data~~ sequences with initial state i

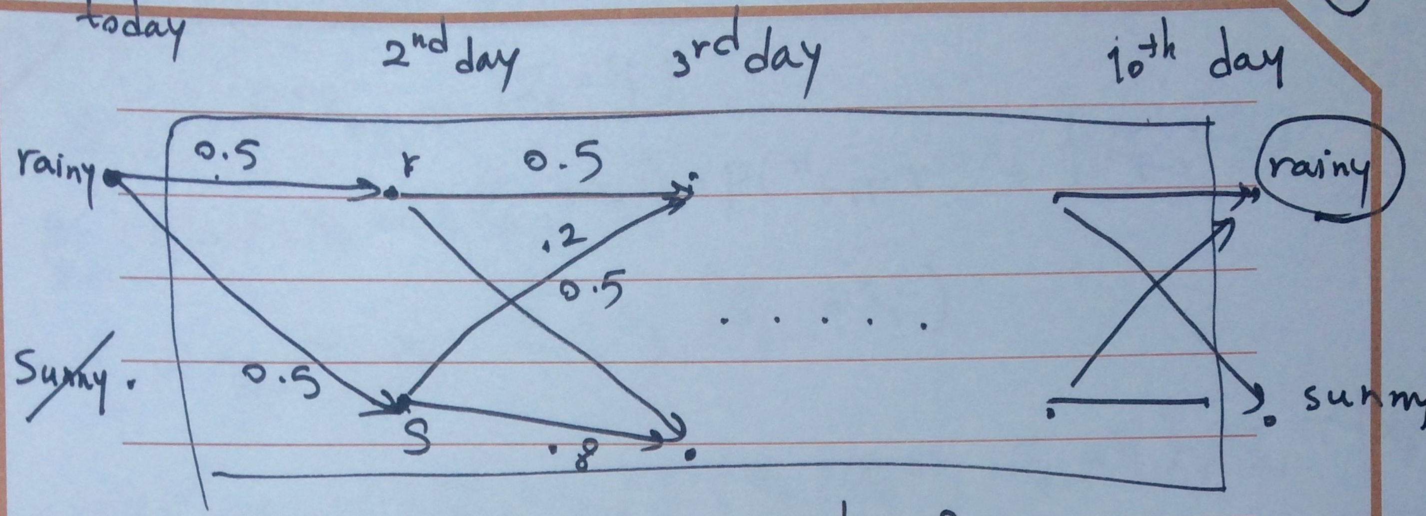
$$\pi_1 + \pi_2 + \cdots + \pi_N = 1$$

$$\pi_1, \ldots, \pi_{N-1} \rightarrow \pi_N = 1 - \sum_{i=1}^{N-1} \pi_i$$

$$\pi_i^* = \frac{k_i}{M}$$

$$A = \begin{array}{c|cc} & \text{rainy} & \text{sunny} \\ \hline \text{rainy} & 0.5 & 0.5 \\ \text{sunny} & 0.2 & .8 \end{array}$$



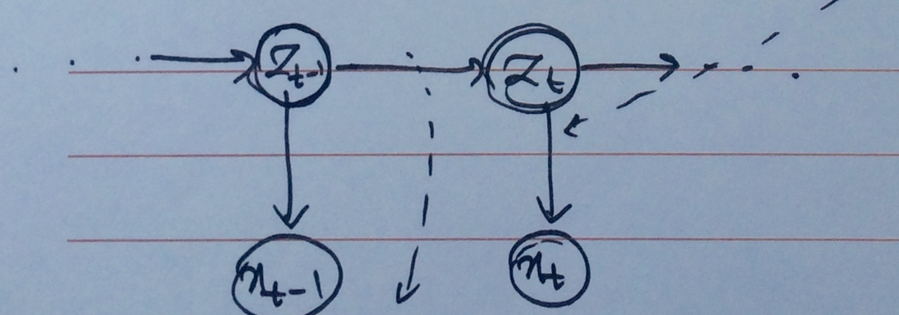


$$[\pi_0^r \ \ \pi_0^r]\ \underbrace{AA\cdots A}_{10} = [\overset{1}{\pi_0^r} \ \ \overset{0}{\pi_0^s}]\ A^{10}$$

$$= [\pi_{10}^r \ \ \pi_{10}^s]$$

$$P(A/B) = \frac{P(A,B)}{B}$$

$$P(x_t \mid Z_t = s_j,\ x_{1:t-1}) = P(x_t \mid Z_t = s_j)$$



$$P(Z_t = s_j \mid Z_{t-1} = s_i,\ x_{1:t-1})$$

$$= P(Z_t = s_j \mid Z_{t-1} = s_i)$$

Derivation on page 37:

$$\beta_t(j) = P(x_{t+1:T} \mid Z_t = s_j)$$

$$\beta_{t-1}(i) = P(x_{t:T} \mid Z_{t-1} = s_i) = P(x_{t+1:T}, x_t \mid Z_{t-1} = s_i)$$

$$= \sum_j P(x_{t+1:T}, x_t, Z_t = s_j \mid Z_{t-1} = s_i)$$

$$= \sum_j \underbrace{P(Z_t = s_j \mid Z_{t-1} = s_i)}_{= a_{ij}} \underbrace{P(x_t \mid Z_t = s_j, Z_{t-1} = s_i)}_{\substack{\downarrow \text{ independance} \\ = P(x_t \mid Z_t = s_j)}} \cdot P(x_{t+1:T} \mid x_t, Z_t = s_j, Z_{t-1} = s_i)$$

$$= \sum_j a_{ij}\, P(x_t \mid Z_t = s_j) \cdot \underbrace{P(x_{t+1:T} \mid x_t, Z_t = s_j, Z_{t-1} = s_i)}_{\downarrow \text{ independence}}$$

$$= \sum_j a_{ij}\, P(x_t \mid Z_t = s_j) \cdot P(x_{t+1:T} \mid Z_t = s_j)$$

$$= \sum_j a_{ij}\, P(x_t \mid Z_t = s_j)\, \beta_t(j)$$

✓

Derivation on page 39:

$$\delta_t(j) = \max_{z_1,\dots,z_{t-1}} p(Z_1=z_1,\dots,Z_{t-1}=z_{t-1}, Z_t=s_j, x_{1:t})$$

$$= \max_i \max_{z_1,\dots,z_{t-2}} p(Z_1=z_1,\dots,Z_{t-2}=z_{t-2}, Z_{t-1}=s_i, Z_t=s_j, x_{1:t-1}, x_t)$$

$$= \max_i \underbrace{\max_{z_1,\dots,z_{t-2}} p(Z_1=z_1,\dots,Z_{t-2}=z_{t-2}, Z_{t-1}=s_i, x_{1:t-1})}_{\triangleq\, \delta_{t-1}(i)} \times p(Z_t=s_j, x_t \mid Z_1=z_1,\dots,Z_{t-2}=z_{t-2}, Z_{t-1}=s_i, x_{1:t-1})$$

$$= \max_i \delta_{t-1}(i)\, p(Z_t=s_j \mid Z_1=z_1,\dots,Z_{t-2}=z_{t-2}, Z_{t-1}=s_i, x_{1:t-1}) \times p(x_t \mid Z_t=s_j, Z_1=z_1,\dots Z_{t-1}=s_i, x_{1:t-1})$$

$$= \max_i \delta_{t-1}(i)\, p(Z_t=s_j \mid Z_{t-1}=s_i) \cdot p(x_t \mid Z_t=s_j)$$

$$= \max_i \delta_{t-1}(i)\, a_{ij}\, p(x_t \mid Z_t=s_j)$$

pag 16

Define K_i = # of sequences starting with i

M = # of sequences

$$\sum_m \log \pi_{x_1^m} = \sum_{i=1}^{N} k_i \log \pi_i$$

$$\sum_{i=1}^{N} \pi_i = 1 \quad , \quad \sum_{i=1}^{N} k_i = M$$

$$\rightarrow \sum_m \log \pi_{x_1^m} = \sum_{i=1}^{N} k_i \log \pi_i = \sum_{i=1}^{N-1} k_i \log \pi_i + \underbrace{(M - \sum_{i=1}^{N-1} k_i)}_{= k_N} \log (1 - \sum_{i=1}^{N-1} \pi_i)$$

$$= \sum_{i=1}^{N-1} k_i \log \pi_i + k_N \log (1 - \sum_{i=1}^{N-1} \pi_i)$$

$$\frac{\partial}{\partial \pi_i} \sum_m \log \pi_{x_1^m} = \frac{k_i}{\pi_i} + k_N \frac{-1}{1 - \sum_{i=1}^{N-1} \pi_i} = 0$$

$$\rightarrow \frac{k_i}{\pi_i} - \frac{k_N}{\pi_N} = 0 \rightarrow \pi_i = k_i \frac{\pi_N}{k_N} \quad \forall i = 1, \dots, N-1$$

$$\sum_{i=1}^{N} \pi_i = 1 \rightsquigarrow \frac{\pi_N}{k_N} \sum_{i=1}^{N-1} k_i + \pi_N = 1 \longrightarrow \pi_N = \frac{k_N}{M}$$

$$\Rightarrow \pi_i = k_i \cdot \frac{k_N / M}{k_N} = \frac{k_i}{M} \qquad \forall i = 1, \dots, N$$

$$= \frac{\text{\# of seq. starting with } i}{\text{\# of sequences.}}$$

and similarly, you can derive a_{ij} with knowing that

$$\sum_j a_{ij} = 1 \rightarrow a_{iN} = 1 - \sum_{j=1}^{N-1} a_{ij} \quad , \dots .$$