comp540 Homework6

Xiaoye Steven Sun (xs6), Wanyi Ye (wy13)

April 12th 2018

1 Hidden Markov Models

1.1

Observables $O = \{$ Low, Medium, High $\}$
Hidden States $S = \{$ Healthy, Unhealthy $\}$
Parameters $\lambda = [\pi, a, b]$:
Initial State Distribution:

$$\pi = [0.5, 0.5]$$

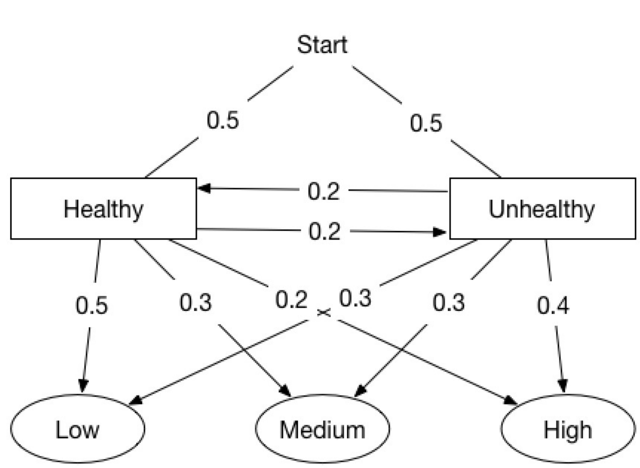
Transition Matrix:

$$a = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

Emission Matrix:

$$b = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

Diagram see below



1.2

1.2 Let $P(t_2 = \text{Healthy} \mid \text{low}, \text{low})$ = the probability that the patient is healthy at $t=2$ given that test readings $t_1 = \text{low}$, $t_2 = \text{low}$

Define $\alpha_t(i) = P(e_1 \cdots e_t, X_t = S_i)$

$t=0$ · $\alpha_0(H) = 0.5$ $\alpha_0(U) = 0.5$

$t=1$ $\alpha_1(H) = b_j e_1 \cdot \sum_{i=1}^{2} \alpha_0(S_i) a_{ij} = 0.5 \cdot (0.5 \cdot 0.8 + 0.5 \cdot 0.2)$

$= 0.25$

$\alpha_1(U) = b_j e_1 \cdot \sum_{i=1}^{2} \alpha_0(S_i) a_{ij} = 0.3 \cdot (0.5 \cdot 0.2 + 0.5 \cdot 0.8)$

$= 0.15$

$t=2$ $\alpha_2(H) = \cancel{0.5(\alpha_1(H) \cdot a_{ij} + \alpha_1(U) \cdot a_{i})}$

$0.5 \times (0.25 \cdot 0.8 + 0.15 \times 0.2) = 0.115$

$\alpha_2(U) = 0.3(0.25 \cdot 0.2 + 0.15 \cdot 0.8) = 0.051$

Hence $P(t_2 = \text{Healthy} \mid \text{low}, \text{low}) = \dfrac{\alpha_2(H)}{\alpha_2(H) + \alpha_2(U)} = 0.69$

$$\delta_2(U) = \max \begin{cases} \delta_1(H) \cdot a_{H \to U} \cdot b_{U \to low} = 0.2 \cdot 0.2 \cdot 0.3 = 0.012 \\ \delta_1(U) \cdot a_{U \to U} \cdot b_{U \to low} = 0.12 \cdot 0.8 \cdot 0.3 = 0.0288 \end{cases}$$

$= 0.0288$

We can see from above that the sequence {Healthy, Healthy, Healthy} have the max probability when $t=2$.

1.3

1.3 Define $\delta_t(i) = \max P(x_1 \cdots x_{t-1}, x_t = S_i, e_1 \cdots e_t)$

$\delta_0(H) = \pi(H) = 0.5 = \pi(U) = \delta_0(U)$

$$\delta_1(H) = \max \begin{cases} \delta_0(H)\cdot a_{H\to H}\, b_{H\to low} = 0.5\cdot 0.8\cdot 0.3 = \overset{0.2}{\cancel{0.12}} \\ \delta_0(U)\cdot a_{U\to H}\, \cancel{b}\, b_{U\to low} = 0.5\cdot 0.2\cdot 0.3 = 0.03 \end{cases}$$

$= \cancel{0.12}\ 0.2$

$$\delta_1(U) = \max \begin{cases} \delta_0(H)\cdot a_{H\to U}\cdot b_{U\to low} = 0.5\cdot 0.2\cdot 0.3 = 0.12 \\ \delta_0(U)\cdot a_{U\to U}\cdot b_{U\to low} = 0.5\cdot 0.8\cdot 0.3 = 0.03 \end{cases}$$

$= 0.12$

$$\delta_2(H) = \max \begin{cases} \delta_1(H)\cdot a_{H\to H}\cdot b_{H\to low} = 0.2\cdot 0.8\cdot 0.5 = 0.08 \\ \delta_1(U)\cdot a_{U\to H}\, b_{H\to low} = \overset{0.12}{\cancel{0.03}}\cdot 0.2\cdot 0.5 = \overset{0.012}{\cancel{0.003}} \end{cases}$$

$= 0.08$

2 EM for mixtures of Bernoullis

2.1

Proof

The likelihood of mixture of Bernoullis is ~~G~~ given by:

$$p(x|\mu_k) = \prod_{j=1}^{m} \mu_{kj}^{x_j}(1-\mu_{kj})^{1-x_j}$$

Introduce Z, a latent variable, note that

$r_k = p(Z_k|X, \mu, \pi)$

Then $p(X,Z|\mu) = \sum_{k=1}^{n} p(x|\mu_k)\cdot p(z_k)$

$$= \sum_{k=1}^{n} r_k \prod_{j=1}^{m} \mu_{kj}^{x_j}\cdot(1-\mu_{kj})^{1-x_j}$$

Set the derivative of μ_{kj} to zero, we will have:

$$\frac{\partial}{\partial \mu_{kj}} E_z[\log p(X,Z|\mu)] = \sum_{i=1}^{m} r_k^{(i)} \left(\frac{x_j^{(i)}}{\mu_{kj}} - \frac{1-x_j^{(i)}}{1-\mu_{kj}}\right)$$

$$= \frac{\sum_{i=1}^{m} r_k^{(i)} x_j^{(i)} - r_k^{(i)}\mu_{kj}}{\mu_{kj}(1-\mu_{kj})} = 0$$

Hence $\mu_{kj} = \dfrac{\sum_{i=1}^{m} r_k^{(i)} x_j^{(i)}}{\sum_{i=1}^{m} r_k^{(i)}}$

2.2

Proof the mixture of Bernoullis has a Beta (α, β) prior

$$\frac{\partial}{\partial \mu_{kj}} E_z[\log P(x,z|\mu)] = \frac{\left(\sum_{i=1}^{m} r_k^{(i)} x_j^{(i)}\right) + \alpha - 1}{\mu_{kj}} - \frac{\sum_{i=1}^{m} r_k^{(i)} (1 - x_j^{(i)}) + \beta - 1}{1 - \mu_{kj}} = 0$$

Then

$$\mu_{kj} = \frac{\sum_{i=1}^{m} r_k^{(i)} x_j^{(i)} + \alpha - 1}{\sum_{i=1}^{m} r_k^{(i)} + \alpha + \beta - 2}$$

3 Principal Components Analysis

Proof

u is the unit vector that gives the projection direction

Let x' denote the projection of x

$$x' = d\, u \quad \text{where } d = |x| \cos\theta = |x| \frac{u \cdot x}{|u||x|}$$

$\because$ $u \cdot x = u^T x$ and $|u| = 1$

$\therefore$ $x' = u^T x u = f_u(x)$

$$\arg\min_{u: uu^T=1} \sum_{i=1}^{m} \| x^{(i)} - f_u(x^{(i)}) \|^2$$

$$= \arg\min_{u: uu^T=1} \sum_{i=1}^{m} \| x^{(i)} - u^T x^{(i)} u \|^2$$

$$= \arg\min_{u: uu^T=1} \sum_{i=1}^{m} (x^{(i)} - u^T x^{(i)} u)^T (x^{(i)} - u^T x^{(i)} u)$$

$$= \arg\min_{u: uu^T=1} \sum_{i=1}^{m} (x^{(i)T} - u^T x^{(i)T} u)(x^{(i)} - u^T x^{(i)} u)$$

$$= \arg\min_{u: uu^T=1} \sum_{i=1}^{m} x^{(i)T} x^{(i)} - 2(u^T x^{(i)})^2 + (u^T x^{(i)})^2$$

$$= \arg\min_{u: uu^T=1} \sum_{i=1}^{m} - (u^T x^{(i)})^2$$

$$= \arg\min_{u: uu^T=1} u^T \left(\sum_{i=1}^{m} x^{(i)} x^{(i)T} \right) u.$$