Final Solution
SENG 474/CSC 578D

Question 1

(a)

$$M^TM = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \tag{1}$$

(b) Compute the eigenvector and eigenvalue of M^TM. We have:

$$\det(M^TM - \lambda I) = 0 \Rightarrow \lambda^2 - 7\lambda + 6 = 0 \Rightarrow \lambda_1 = 6, \lambda_2 = 1 \tag{2}$$

The two corresponding eigenvectors are $\mathbf{v}_1 = [\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}]^T$ and $\mathbf{v}_2 = [\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}]^T$.

Question 2

(a)

$$\mathbf{v}_1 = \frac{M\mathbf{v}_0}{||M\mathbf{v}_0||} = [1/3, 2/3, 2/3]^T \tag{3}$$

and

$$\mathbf{v}_2 = \frac{M\mathbf{v}_1}{||M\mathbf{v}_1||} = [1/3, 2/3, 2/3]^T \tag{4}$$

(b) Power Iteration algorithm of part (a) stop at $\mathbf{v}_1$, so it is the eigenvector and the largest eigenvalue is:

$$\mathbf{v}_1^T M \mathbf{v}_1 = 9 \tag{5}$$

Question 3

(a) The probability of observing the network is $f(p =)p^7(1-p)^3$. Thus, we have:

$$\frac{\partial f(p)}{\partial p} = 7p^6(1-p)^3 - 3p^7(1-p)^2 = p^6(1-p)^3(7-10p) = 0 \Leftrightarrow p = 7/10 = 0.7 \tag{6}$$

(b) The probability of observing the network is:

$$f(p_X, p_Y) = p_X^4(1-p_X)p_Y^2(p_X + p_Y - p_Xp_Y)(1-\epsilon)^2 \tag{7}$$

$f(p_X, p_Y)$ is an increasing function of p_Y, thus, it is maximum when $p_Y = 1$. So we should chose p_X that maximize $p_X^4(1-p_X)$, which solves to $p_X = 4/5 = 0.8$.

Question 4

(a) The Laplaian matrix:

$$L = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \tag{8}$$

(b) The eigenvector $\mathbf{v}_2$ would satisfy:

$$L\mathbf{v}_2 = \mathbf{v}_2 \tag{9}$$

that solves to $[1, 0, -1]^T$

Question 5

(a) We have:

$$\begin{aligned} Betweenness(DC) &= 3 \\ Betweenness(CA) &= 1 \\ Betweenness(CB) &= 1 \\ Betweenness(AB) &= 0 \end{aligned} \tag{10}$$

(b) Final betweeness:

$$\begin{aligned} Betweenness(AB) &= 1 \\ Betweenness(AC) &= 2 \\ Betweenness(BC) &= 2 \\ Betweenness(CD) &= 3 \end{aligned} \tag{11}$$

Question 6

(a) We have:

$$U_0V_0 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 6 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix} \tag{12}$$

Thus,

$$RMSE = \sqrt{\frac{(1-2)^2 + (6-4)^2 + (5-1)^2 + (6-3)^2 + (1-2)^2 + (3-4)^2}{6}} = \sqrt{\frac{16}{3}} \tag{13}$$

(b)

$$\begin{bmatrix} 1 & x_1 \\ 2 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2+x_1 & 2+2x_1 \\ 2 & 4+x_2 & 4+2x_2 \\ 1 & 2+x_3 & 2+2x_3 \end{bmatrix} \tag{14}$$

Thus,

$$RMSE = \sqrt{\frac{(1-2)^2 + (2x_1-2)^2 + (x_2-3)^2 + (2x_2+1)^2 + (1-2)^2 + (x_3-2)^2}{6}} \tag{15}$$

which is minimum when $x_1 = 1, x_2 = -1, x_3 = 2$.

Question 7

(a) The balanced algorithm in the worst case sells $xxyyzz$ to $BCBC$, thus has revenue 4, while the optimal algorithm has revenue 6, so the competitive ratio is $4/6 = 2/3$.

(b) Suppose that A has budget 1001, B has budget 1000, A bids 1\$ for x while B bids 1000. The query sequence is just x. The balance algorithm will sell x to A, making 1\$ of revenue, while the optimal algorithm would sell x to B. Thus the competitive ratio is $1/1000 = 10^{-3}$.

Question 8

(a) The centroids are $(3, 9), (6, 5)$ and $(7, 9)$.

(b) Three clusters are:

$$\begin{aligned} C_1 &= \{(3,9),(4,10),(4,8)\} \\ C_2 &= \{(7,9),(6,10),(6,8)\} \\ C_3 &= \{(4,6),(4,5),(6,6),(6,5)\} \end{aligned} \tag{16}$$

Question 9

(a) The transition matrix:

$$M = \begin{bmatrix} 0 & 0 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 1/2 \end{bmatrix} \tag{17}$$

(b)

$$\mathbf{v}_1 = M\mathbf{v}_0 = \begin{bmatrix} 0 & 0 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 1/2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/6 \\ 1/3 \\ 1/2 \end{bmatrix} \tag{18}$$

$$\mathbf{v}_2 = M\mathbf{v}_1 = \begin{bmatrix} 0 & 0 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 1/2 \end{bmatrix} \begin{bmatrix} 1/6 \\ 1/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/6 \\ 7/12 \end{bmatrix} \tag{19}$$

Question 10

(a) The page rank score of t_1

$$y = \frac{x}{1-\beta^2} + \frac{\beta}{1+\beta}\frac{m}{M} = 1.5837 \tag{20}$$

assume that $(1-\beta)/N$ contribution to y is ignored. (If you count $(1-\beta)/N$ contribution of y, then the number just change very little, something like 1.5838)

(b) Let y be the page rank score of t_1 (and t_2). Then each supporting page gets:

$$s = \beta \frac{2y}{2m} + \frac{(1-\beta)}{N} = \frac{\beta y}{m} + \frac{(1-\beta)}{N} \tag{21}$$

Each supporting page contribute $s/2$ to each page t_1, t_2. Thus, t_1 gets:

$$\frac{\beta}{2}(\frac{\beta y}{m} + \frac{(1-\beta)}{N}) \tag{22}$$

from each supporting page. Since there are $2m$ supporting pages link to t_1, we have:

$$y = x + 2m\frac{\beta}{2}(\frac{\beta y}{m} + \frac{(1-\beta)}{N}) \tag{23}$$

which solves to:

$$y = \frac{x}{1-\beta^2} + \frac{\beta}{\beta+1}\frac{m}{N} \tag{24}$$

Conclusion: the linking does not change the page ranks of t_1, t_2.