Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Question 1

Simulate a random walker on the graph represented by the adjacency matrix A.

```
# Random Walker
# Simulates a random walker on the graph
# represented by the adjacency matrix A

# X := graph object returned by ER
# t := maximum time

randomWalker <- function(X, t = 1e4){
  # A := adjacency matrix
  A <- X$graph
  # p := number of vertices in graph
  p <- X$p
  # randomly choose starting vertex
  v <- sample.int(p, 1)
  # pi := stationary probability vector
  pi <- rep(0, p)
  # sum pi_j = 1
  pi[v] <- 1
  for(time in 2:t){
    # randomly choose neighbor of v
    v <- sample(which(A[,v] == 1), 1)
    # update position of v
    pi[v] <- (1 + time*pi[v])/(time + 1)
    # update all other positions
    pi[-v] <- (time*pi[-v])/(time + 1)
    plot(pi, pch = 19, xlim = c(1,p),
         ylim = c(0, 1), xlab = "Vertex ID",
         ylab = "Stationary Probability",
         main = "Random Walker")
    Sys.sleep(0.05)
  }
  # d := degree sequence
  d <- colSums(A)
  # P := the stochastic matrix which represents
  # the transition probabilities for the random walker
  # compute this stochastic matrix by simply taking the
  # adjacency matrix A that represents the graph G and
  # diving each column by the vertex degree vector d.
  P <- sweep(A,2,d,"/")
  return(P)
}
```

Question 2

Why does P represent the correct probability transition matrix?

P is computed by simply taking the adjacency matrix A that represents the graph G and diving each column by the vertex degree vector d.

```
Notice that if
> A
     [,1] [,2] [,3] [,4] [,5]
[1,]    0    0    1    1    1
[2,]    0    0    0    1    1
[3,]    1    0    0    0    1
[4,]    1    1    0    0    1
[5,]    1    1    1    1    0
then
> P
          [,1] [,2] [,3]      [,4] [,5]
[1,] 0.0000000  0.0  0.5 0.3333333 0.25
[2,] 0.0000000  0.0  0.0 0.3333333 0.25
[3,] 0.3333333  0.0  0.0 0.0000000 0.25
[4,] 0.3333333  0.5  0.0 0.0000000 0.25
[5,] 0.3333333  0.5  0.5 0.3333333 0.00
```

P is basically a matrix of relative frequencies. If an arbitrary vector v has a degree of d then the probability that the random walker will move from v to any one of it's neighbors is $\frac{1}{d}$ since A represents a simple graph G.

Question 3

Show that the j^{th} element π_j of $\vec{\pi}$ is equal to

$$\pi_j = \frac{d_j}{\sum_{k=1}^{p} d_k}$$

where d_j is the degree of vertex v_j.

We can represent P as the following matrix where p is the number of vertices in the graph and d_j is the degree of vertex v_j. Additionally, $A[i,j]$ is the entry of the adjacency matrix in the $(i,j)^{th}$ position.

$$P = \begin{bmatrix} \frac{A[1,1]}{d_1} & \frac{A[1,2]}{d_2} & \frac{A[1,3]}{d_3} & \cdots & \frac{A[1,p]}{d_p} \\ \frac{A[2,1]}{d_1} & \frac{A[2,2]}{d_2} & \frac{A[2,3]}{d_3} & \cdots & \frac{A[2,p]}{d_p} \\ \frac{A[3,1]}{d_1} & \frac{A[3,2]}{d_2} & \frac{A[3,3]}{d_3} & \cdots & \frac{A[3,p]}{d_p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{A[p,1]}{d_1} & \frac{A[p,2]}{d_2} & \frac{A[p,3]}{d_3} & \cdots & \frac{A[p,p]}{d_p} \end{bmatrix}$$

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Note that $\sum_{k=1}^{p} d_k$ is just the total number of degrees. We are given that $P\pi = \pi$.
Let P be the matrix described above. Recall that we can write d_j as the sum of the row j of the adjacency matrix A. Then

$$P\pi = \begin{bmatrix} \frac{A[1,1]}{d_1} & \frac{A[1,2]}{d_2} & \frac{A[1,3]}{d_3} & \cdots & \frac{A[1,p]}{d_p} \\ \frac{A[2,1]}{d_1} & \frac{A[2,2]}{d_2} & \frac{A[2,3]}{d_3} & \cdots & \frac{A[2,p]}{d_p} \\ \frac{A[3,1]}{d_1} & \frac{A[3,2]}{d_2} & \frac{A[3,3]}{d_3} & \cdots & \frac{A[3,p]}{d_p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{A[p,1]}{d_1} & \frac{A[p,2]}{d_2} & \frac{A[p,3]}{d_3} & \cdots & \frac{A[p,p]}{d_p} \end{bmatrix} \begin{bmatrix} \frac{d_1}{\sum_{k=1}^{p} d_k} \\ \frac{d_2}{\sum_{k=1}^{p} d_k} \\ \frac{d_3}{\sum_{k=1}^{p} d_k} \\ \vdots \\ \frac{d_p}{\sum_{k=1}^{p} d_k} \end{bmatrix}$$

I will only demonstrate the first multiplication, it will be similar for the others.

$$\begin{bmatrix} \frac{A[1,1]}{d_1} & \frac{A[1,2]}{d_2} & \cdots & \frac{A[1,p]}{d_p} \end{bmatrix} \begin{bmatrix} \frac{d_1}{\sum_{k=1}^{p} d_k} \\ \frac{d_2}{\sum_{k=1}^{p} d_k} \\ \vdots \\ \frac{d_p}{\sum_{k=1}^{p} d_k} \end{bmatrix} = \frac{A[1,1]d_1}{d_1 \sum_{k=1}^{p} d_k} + \frac{A[1,2]d_2}{d_2 \sum_{k=1}^{p} d_k} + \cdots + \frac{A[1,p]d_p}{d_p \sum_{k=1}^{p} d_k}$$

$$= \frac{A[1,1]}{\sum_{k=1}^{p} d_k} + \frac{A[1,2]}{\sum_{k=1}^{p} d_k} + \cdots + \frac{A[1,p]}{\sum_{k=1}^{p} d_k} = \frac{A[1,1] + A[1,2] + \cdots + A[1,p]}{\sum_{k=1}^{p} d_k} = \frac{d_1}{\sum_{k=1}^{p} d_k}$$

For the other rows of π we would calculate them similarly.
I hope now it is obvious that the given vector $\pi = \begin{bmatrix} \frac{d_1}{\sum_{k=1}^{p} d_k} & \frac{d_2}{\sum_{k=1}^{p} d_k} & \cdots & \frac{d_p}{\sum_{k=1}^{p} d_k} \end{bmatrix}^T$ satisfies the equation $P\pi = \pi$.
Therefore $\pi_j = \frac{d_j}{\sum_{k=1}^{p} d_k}$ for $j \in \{1, 2, \cdots, p\}$.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Question 4

Show for $r = 1, 2, 3$ that the $(i, j)^{th}$ entry in A^r is the number of paths from vertex v_i to vertex v_j of length r.

Let $r = 1$. Then since we are not given any special characteristics of A,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1p} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2p} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \cdots & a_{pp} \end{bmatrix}$$

The number of paths from vertex v_i to vertex v_j of length 1 is trivially equal to a_{ij} since $a_{ij} = 1$ only if vertex v_i to vertex v_j are neighbors, that is, vertex v_i to vertex v_j are connected by an edge.
Let $r = 2$. Then

$$A^2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1p} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2p} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \cdots & a_{pp} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1p} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2p} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \cdots & a_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \cdots & \sum_{k=1}^{p} a_{ik}a_{kj} \\ \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \cdots & \sum_{k=1}^{p} a_{ik}a_{kj} \\ \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \cdots & \sum_{k=1}^{p} a_{ik}a_{kj} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \sum_{k=1}^{p} a_{ik}a_{kj} & \cdots & \sum_{k=1}^{p} a_{ik}a_{kj} \end{bmatrix}$$

where i is the row and j is the column respectively.

Hence the $(i, j)^{th}$ entry in A^2 is equal to $\sum_{k=1}^{p} a_{ik}a_{kj}$. Recall that a_{ij} is the number of edges between vertex v_i and vertex v_j. That is, a_{ij} is the number of walks of length $r = 1$ between vertex v_i and vertex v_j. Since a walk of length $r = 2$ must be at a vertex k after $r - 1$ steps, $a_{ik}a_{kj}$ is the number of walks of length $r = 2$ between vertex v_i and vertex v_j.

We could easily prove this by induction at this point. For $r = 3$, this will be very similar to the case above. The abstract matrix multiplication will get to be messy so allow the $(i, j)^{th}$ entry in A^2 to be equal to b_{ij}. Then the $(i, j)^{th}$ entry in A^3 is equal to $\sum_{k=1}^{p} b_{ik}a_{kj}$ where b_{ij} is equal to the number of walks of length $r = 2$ between vertex v_i and vertex v_j. Since a walk of length $r = 3$ must be at a vertex k after $r - 1$ steps, $b_{ik}a_{kj}$ is the number of walks of length $r = 3$ between vertex v_i and vertex v_j.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Question 5

Show that the only solution v^* to the following equation

$$v^* = \lim_{n\to\infty} (\frac{1}{\lambda} A)^n v$$

is when v^* is the dominant eigenvector.

If an $p \times p$ matrix A is diagonalizable, then we can write some vector v^* as a linear combination of the eigenvectors of A. Suppose $v_1, v_2, \cdots, v_p$ are p linearly independent eigenvectors of A and $\lambda_1, \lambda_2, \cdots, \lambda_p$ are the p corresponding eigenvalues with $\lambda_1 |>| \lambda_2 |\geq \cdots \geq| \lambda_p |$. Then we can write

$$v^* = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p$$

If we apply A to v^* then since $Av = \lambda v$,

$$Av^* = \alpha_1 Av_1 + \alpha_2 Av_2 + \cdots + \alpha_p Av_p$$

$$= \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_n \lambda_p v_p$$

$$= \lambda_1 (\alpha_1 v_1 + \alpha_2 \frac{\lambda_2}{\lambda_1} v_2 + \cdots + \alpha_p \frac{\lambda_p}{\lambda_1} v_p)$$

If we apply the process above indefinitely then

$$A^n v^* = \lambda_1^n (\alpha_1 v_1 + \alpha_2 \frac{\lambda_2}{\lambda_1^n} v_2 + \cdots + \alpha_p \frac{\lambda_p}{\lambda_1^n} v_p)$$

Then

$$\frac{A^n v^*}{\lambda_1^n} = \alpha_1 v_1 + \alpha_2 \frac{\lambda_2}{\lambda_1^n} v_2 + \cdots + \alpha_p \frac{\lambda_p}{\lambda_1^n} v_p$$

Hence as $n \to \infty$, $\dfrac{\lambda_i}{\lambda_1^n} \to 0$ for $i \in \{2, \cdots, p\}$.

Then clearly we have that

$$\lim_{n\to\infty} \frac{A^n v^*}{\lambda_1^n} = \alpha_1 v_1$$

where v_1 is the eigenvector associated with the greatest eigenvalue. Hence, v_1 is the dominant eigenvector.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Question 6

I used the following data set to construct the weighted adjacency matrix or boundary length between New England states: `http://users.econ.umn.edu/~holmes/data/BORDLIST.html`
Degree centrality:

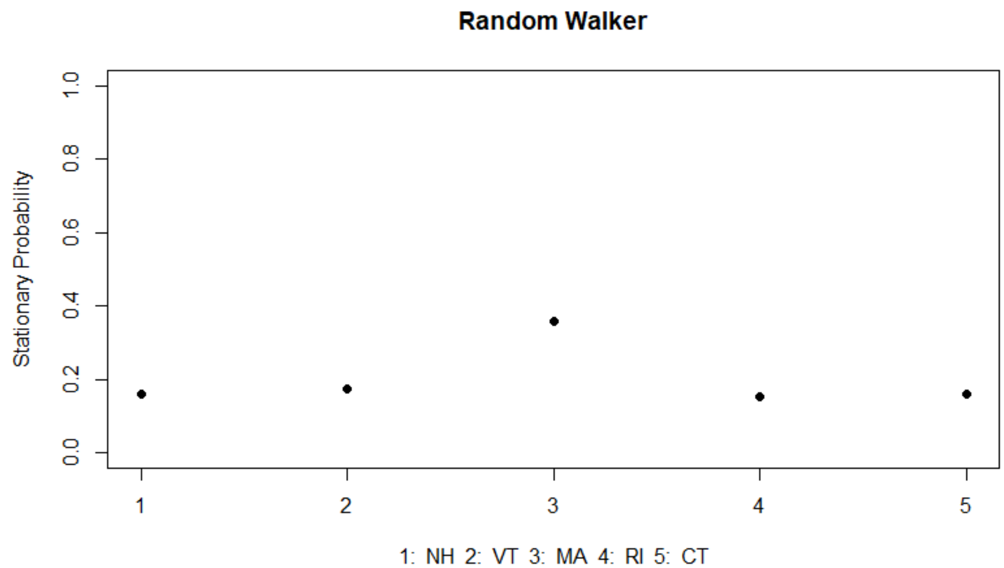
```
randomWalker <- function(A, t = 1e3){
  p <- ncol(A)
  v <- sample.int(p, 1)
  pi <- rep(0, p)
  pi[v] <- 1
  for(time in 2:t){
    v <- sample(which(A[,v] != 0), 1)
    pi[v] <- (1 + time*pi[v])/(time + 1)
    pi[-v] <- (time*pi[-v])/(time + 1)
    states <- colnames(A)
    plot(pi, pch = 19, xlim = c(1, p),
         ylim = c(0, 1), xlab = paste("1: ", states[1],
                                 " 2: ", states[2],
                                 " 3: ", states[3],
                                 " 4: ", states[4],
                                 " 5: ", states[5]),
         ylab = "Stationary Probability",
         main = "Random Walker")
    Sys.sleep(0.05)
  }
  d <- colSums(A)
  P <- sweep(A,2,d,"/")
  return(P)
}

NE.states <- c("NH", "VT", "MA", "RI", "CT")
border.length <- matrix(c(0, 180, 86.7, 0, 0,
                          180, 0, 40.9, 0, 0,
                          86.7, 40.9, 0, 62.6, 86.5,
                          0, 0, 62.6, 0, 40.3,
                          0, 0, 86.5, 40.3, 0),
                        nrow = 5,
                        ncol = 5,
                        byrow = TRUE)
rownames(border.length) <- NE.states
colnames(border.length) <- NE.states

randomWalker(border.length)
```

The image below shows MA as the most central state in New England by degree centrality.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu



Eigenvector centrality:
We compute the dominant eigenvector v of the adjacency matrix and rank the vertices in the graph according to their corresponding values in v with greater values indicating greater centrality.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

```
B <- matrix(c(0, 1, 1, 0, 0,
              1, 0, 1, 0, 0,
              1, 1, 0, 1, 1,
              0, 0, 1, 0, 1,
              0, 0, 1, 1, 0),
            nrow = 5,
            ncol = 5,
            byrow = TRUE)

e <- eigen(B)
vals <- e$values
vecs <- e$vectors
dom.vec <- vecs[,1]
dom.vec <- as.matrix(abs(dom.vec))
rownames(dom.vec) <- NE.states
dom.vec
The output:
> dom.vec
        [,1]
NH 0.3941027
VT 0.3941027
MA 0.6154122
RI 0.3941027
CT 0.3941027
```

Hence, the ranking is as follows: MA, NH, VT, CT, RI.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

Question 7

Prove that P_α is a stochastic matrix, i.e. that the column sums of P_α are equal to one.

We have that

$$P_\alpha = (1-\alpha)P + \alpha J$$

where J is the matrix of all ones times the reciprocal of the number of vertices in the graph. Recall how we defined P initially then

$$P_\alpha = (1-\alpha)P + \alpha J$$

$$= (1-\alpha)\begin{bmatrix} \frac{A[1,1]}{d_1} & \frac{A[1,2]}{d_2} & \cdots & \frac{A[1,p]}{d_p} \\ \frac{A[2,1]}{d_1} & \frac{A[2,2]}{d_2} & \cdots & \frac{A[2,p]}{d_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A[p,1]}{d_1} & \frac{A[p,2]}{d_2} & \cdots & \frac{A[p,p]}{d_p} \end{bmatrix} + \alpha \begin{bmatrix} \frac{1}{p} & \frac{1}{p} & \cdots & \frac{1}{p} \\ \frac{1}{p} & \frac{1}{p} & \cdots & \frac{1}{p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{p} & \frac{1}{p} & \cdots & \frac{1}{p} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{(1-\alpha)A[1,1]}{d_1} & \frac{(1-\alpha)A[1,2]}{d_2} & \cdots & \frac{(1-\alpha)A[1,p]}{d_p} \\ \frac{(1-\alpha)A[2,1]}{d_1} & \frac{(1-\alpha)A[2,2]}{d_2} & \cdots & \frac{(1-\alpha)A[2,p]}{d_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{(1-\alpha)A[p,1]}{d_1} & \frac{(1-\alpha)A[p,2]}{d_2} & \cdots & \frac{(1-\alpha)A[p,p]}{d_p} \end{bmatrix} + \begin{bmatrix} \frac{\alpha}{p} & \frac{\alpha}{p} & \cdots & \frac{\alpha}{p} \\ \frac{\alpha}{p} & \frac{\alpha}{p} & \cdots & \frac{\alpha}{p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha}{p} & \frac{\alpha}{p} & \cdots & \frac{\alpha}{p} \end{bmatrix}$$

Please consider just the sum of first column, the rest will follow:

$$= \begin{bmatrix} \frac{(1-\alpha)A[1,1]}{d_1} \\ \frac{(1-\alpha)A[2,1]}{d_1} \\ \vdots \\ \frac{(1-\alpha)A[p,1]}{d_1} \end{bmatrix} + \begin{bmatrix} \frac{\alpha}{p} \\ \frac{\alpha}{p} \\ \frac{\alpha}{p} \\ \vdots \\ \frac{\alpha}{p} \end{bmatrix} = \begin{bmatrix} \frac{(1-\alpha)A[1,1]p + d_1\alpha}{d_1 p} \\ \frac{(1-\alpha)A[2,1]p + d_1\alpha}{d_1 p} \\ \vdots \\ \frac{(1-\alpha)A[p,1]p + d_1\alpha}{d_1 p} \end{bmatrix}$$

Then

$$\frac{(1-\alpha)A[1,1]p + d_1\alpha}{d_1 p} + \frac{(1-\alpha)A[2,1]p + d_1\alpha}{d_1 p} + \cdots + \frac{(1-\alpha)A[p,1]p + d_1\alpha}{d_1 p}$$

$$= \frac{(1-\alpha)(A[1,1]p + (1-\alpha)A[2,1]p + \cdots + (1-\alpha)A[p,1]p) + \alpha p d_1}{d_1 p}$$

$$= \frac{(1-\alpha)p(A[1,1] + A[2,1] + \cdots + A[p,1]) + \alpha p d_1}{d_1 p}$$

Angyalka Valcsics
Angyalka.Valcsics@maine.edu

$$= \frac{(1-\alpha)pd_1 + \alpha pd_1}{d_1 p} = (1-\alpha) + \alpha = 1$$

Hence with the other column sums computed similarly to the first column sum, we conclude that P_α is a column stochastic matrix.

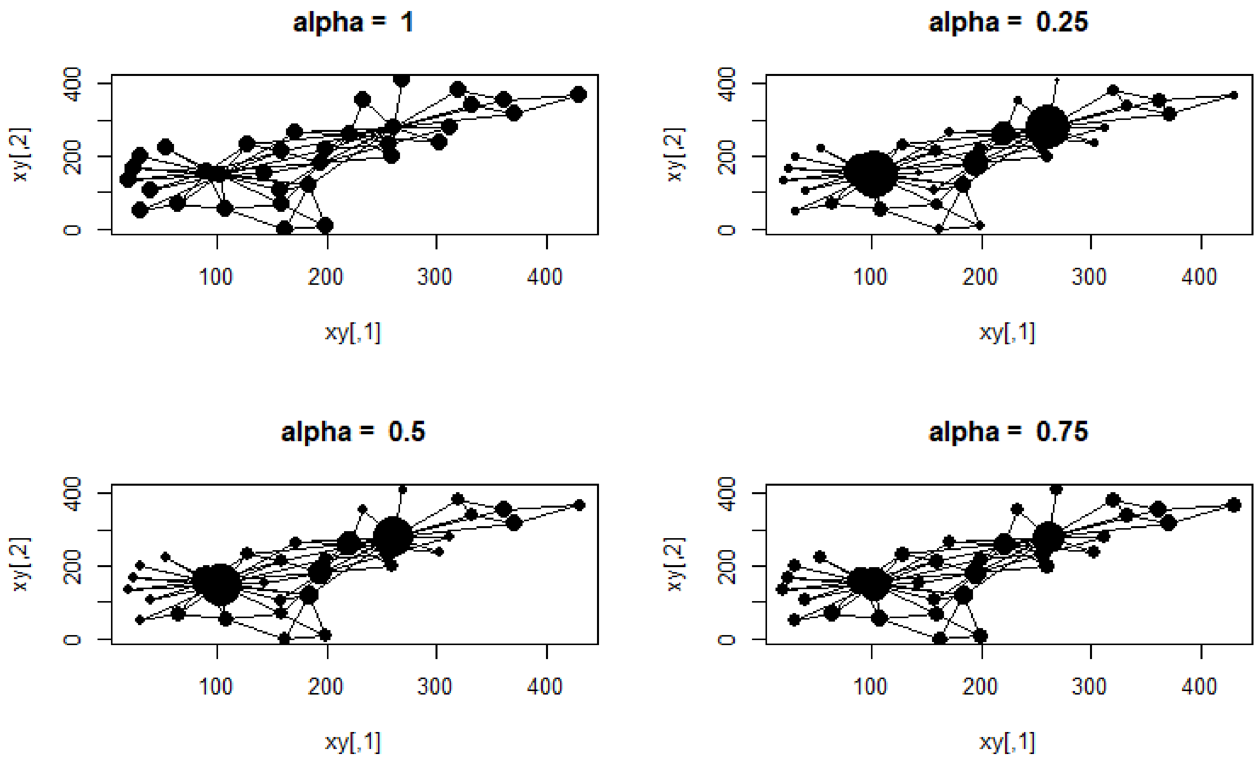
Now to demonstrate the effect of α on the centrality of vertices in the Karate Club network.

```
# w was given to us,
# it is unimportant for our purposes as to how it was loaded.
W <- W+t(W)
A <- 1*(W>0)
g <- graph.adjacency(A,mode="undirected")
xy <- tk_coords(tkplot(g))
RW <- function(A, xy, alpha){
  p <- ncol(A)
  E <- matrix(0, sum(A)/2, 2)
  h <- 0
  for(i in 1:(p-1)){
    for(j in (i+1):p){
      if(A[i,j] == 1){
        h <- h+1
        E[h,] <- c(i, j)}}}
  d <- colSums(A)
  P <- sweep(A,2,d,"/")
  J <- matrix(1, nrow = p, ncol = p)
  J <- (1/p)*J
  P.alpha <- (1-alpha)*P + alpha*J
  f <- eigen(P.alpha)
  pi <- abs(f$vectors[,1])
  plot(xy, cex = 10*pi, pch = 19, main = paste("alpha = ", alpha))
  for (h in 1:nrow(E)) {
    segments(xy[E[h,1], 1],
             xy[E[h,1], 2],
             xy[E[h,2], 1],
             xy[E[h,2], 2])
  }}
```

```
par(mfrow = c(2,2))
RW(A, xy, alpha = 1)
RW(A, xy, alpha = 0.25)
RW(A, xy, alpha = 0.5)
RW(A, xy, alpha = 0.75)
```

Then we have the following plots.

Angyalka Valcsics
Angyalka.Valcsics@maine.edu



We can see that when $\alpha = 1$, the vertices all have the same size. When α is close to 0, the senseis vertices have a greater importance relative to the other members. As $\alpha \to 1$, the importance of the members evens out.