CptS 591: Elements of Network Science - Assignment 2

Hitesh Bahar

2/25/2019

```
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 3.5.2
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```
require(scales)
```

```
## Loading required package: scales
```

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.5.2
```

```
setwd("D:/NetworkScience/assignment2")
```

Question-1

The fourth real world network I have chosen here is *Astrophysics collaborations*, which contains the collaboration network of scientists posting preprints on the astrophysics archive at www.arxiv.org, 1995-1999, as compiled by M. Newman. It is a weighted network.

The URL for the data is *http://www-personal.umich.edu/~mejn/netdata/astro-ph.zip*

In this questoin we will find the node importance using centrality measure of the following:

1. Degree
2. Eccentricity
3. Closeness
4. Betweenness
5. Katz index
6. PageRank
7. Kleinberg's Authority score
8. Kleinberg's Hub score

```
politcalNetwork <- read_graph("polblogs.gml", format = "gml")
internetNetwork <- read_graph("internet.gml",format = "gml")
neuralNetwork <- read_graph("neuralnetwork.gml", format = "gml")
astroPhysics <- read_graph("astro-ph.gml", format = "gml")
```

```
dfGraph  <- data.frame(Network = character(0),
                 Degree = numeric(0),
                 Eccentricity = numeric(0),
                 Closeness = numeric(0),
                 Betweenness = numeric(0),
                 Katz_Index = numeric(0),
                 PageRank = numeric(0),
                 Kleinbergs_Authority_score = numeric(0),
                 Kleinbergs_Hub_score = numeric(0),
                 stringsAsFactors = FALSE)
#Political netowk
maxDegree <- which.max(degree(politcalNetwork))
maxEccentricity <- which.max(eccentricity(politcalNetwork))
maxCloseness <- which.max(closeness(politcalNetwork))
maxBetweeness <- which.max(betweenness(politcalNetwork))
KatzIndex <- NA
maxPageRank <- which.max(page_rank(politcalNetwork)$vector)
maxKAuthScore <-  which.max(authority_score(politcalNetwork)$vector)
maxKHubScore <- which.max(hub_score(politcalNetwork)$vector)

dfGraph[nrow(dfGraph)+1,] <- c("Political Netowrk",maxDegree,maxEccentricity,maxCloseness,maxBetweeness


#neuralNetwork
maxDegree <- which.max(degree(neuralNetwork))
maxEccentricity <- which.max(eccentricity(neuralNetwork))
maxCloseness <- which.max(closeness(neuralNetwork))
maxBetweeness <- which.max(betweenness(neuralNetwork))
KatzIndex <- NA
maxPageRank <- which.max(page_rank(neuralNetwork)$vector)
maxKAuthScore <-  which.max(authority_score(neuralNetwork)$vector)
maxKHubScore <- which.max(hub_score(neuralNetwork)$vector)

dfGraph[nrow(dfGraph)+1,] <- c("neural Network",maxDegree,maxEccentricity,maxCloseness,maxBetweeness,Ka

#Internet Network
maxDegree <- which.max(degree(internetNetwork))
maxEccentricity <- which.max(eccentricity(internetNetwork))
maxCloseness <- which.max(closeness(internetNetwork))
maxBetweeness <- which.max(betweenness(internetNetwork))
KatzIndex <- NA
maxPageRank <- which.max(page_rank(internetNetwork)$vector)
maxKAuthScore <-  which.max(authority_score(internetNetwork)$vector)
maxKHubScore <- which.max(hub_score(internetNetwork)$vector)

dfGraph[nrow(dfGraph)+1,] <- c("Internet Network",maxDegree,maxEccentricity,maxCloseness,maxBetweeness,

#astroPhysics
maxDegree <- which.max(degree(astroPhysics))
maxEccentricity <- which.max(eccentricity(astroPhysics))
maxCloseness <- which.max(closeness(astroPhysics))
maxBetweeness <- which.max(betweenness(astroPhysics))
KatzIndex <- NA
```

```
maxPageRank <- which.max(page_rank(astroPhysics)$vector)
maxKAuthScore <-  which.max(authority_score(astroPhysics)$vector)
maxKHubScore <- which.max(hub_score(astroPhysics)$vector)

dfGraph[nrow(dfGraph)+1,] <- c("Astro Physics",maxDegree,maxEccentricity,maxCloseness,maxBetweeness,Katz

kable(dfGraph, format = "markdown")
```

Network	Degree	Eccentricity	Closeness	Betweenness	Katz_Index	PageRank	Kleinbergs_Authority_score	Kleinbergs_Hub_score
Political Netowrk	855	794	293	855	NA	155	155	512
neural Network	45	82	260	178	NA	45	45	126
Internet Network	4	9200	23	4	NA	4	4	4
Astro Physics	5503	1613	5503	913	NA	1232	5503	5503

Observation:

1. *Political Network:* In this network, nodes 855 and 155 are important an node 855 has maximum degree as well as betweenness and node 155 has highest page rank and Kleinberg's authority score
2. *Internet Network:* In this network, node 4 is important as it has highest degree, betweenness, pagerank, Kleinberg's authority score and Kleinberg's hub score
3. *Neural Network:* Here node 45 is an important node as it has highest degree,pagerank and Kleinberg's authority score
4. *AstroPhysics Network:* Here node 55-3 is important as it has the highesst degree, closeness, Kleinberg's authority score and Kleinberg's hub score

Question-2(2.1, 2.2)

```
er20 <- erdos.renyi.game(20,80,type ="gnm")
er40 <- erdos.renyi.game(40,120,type="gnm")
bb20 <- barabasi.game(n=20, m=5, directed=FALSE)
bb40 <- barabasi.game(n= 40, m= 3, directed = FALSE)

df <- data.frame(Graph = character(0),
                 n = numeric(0),
                 m = numeric(0),
                 d_min = numeric(0),
                 d_max = numeric(0),
                 l = numeric(0),
                 D = numeric(0),
                 ccg = numeric(0),
                 lambda_2 = numeric(0),
                 lambda_N = numeric(0),
                 stringsAsFactors = FALSE)

df[nrow(df)+1,] <- c("Erdos-Renyi 20",
                     vcount(er20),
```

```r
                     ecount(er20),
                     min(degree(er20)),
                     max(degree(er20)),
                     mean_distance(er20),
                     diameter(er20),
                     transitivity(er20, type= "globalundirected"),
                     sort(eigen(laplacian_matrix(er20))$values,decreasing = FALSE)[2],
                     sort(eigen(laplacian_matrix(er20))$values,decreasing = TRUE)[1])

df[nrow(df)+1,] <- c("Erdos-Renyi 40",
                     vcount(er40),
                     ecount(er40),
                     min(degree(er40)),
                     max(degree(er40)),
                     mean_distance(er40),
                     diameter(er40),
                     transitivity(er40, type= "globalundirected"),
                     sort(eigen(laplacian_matrix(er40))$values,decreasing = FALSE)[2],
                     sort(eigen(laplacian_matrix(er40))$values,decreasing = TRUE)[1])

df[nrow(df)+1,] <- c("Barabasi 20",
                     vcount(bb20),
                     ecount(bb20),
                     min(degree(bb20)),
                     max(degree(bb20)),
                     mean_distance(bb20),
                     diameter(bb20),
                     transitivity(bb20, type= "globalundirected"),
                     sort(eigen(laplacian_matrix(bb20))$values,decreasing = FALSE)[2],
                     sort(eigen(laplacian_matrix(bb20))$values,decreasing = TRUE)[1])

df[nrow(df)+1,] <- c("Barabasi 40",
                     vcount(bb40),
                     ecount(bb40),
                     min(degree(bb40)),
                     max(degree(bb40)),
                     mean_distance(bb40),
                     diameter(bb40),
                     transitivity(bb40, type= "globalundirected"),
                     sort(eigen(laplacian_matrix(bb40))$values,decreasing = FALSE)[2],
                     sort(eigen(laplacian_matrix(bb40))$values,decreasing = TRUE)[1])

kable(df, format = "markdown")
```

Graph	n	m	d_min	d_max	l	D	ccg	lambda_2	lambda_N
Erdos-Renyi 20	20	80	4	12	1.58947368421053	3	0.414473684210526	3.16357460938392	14.90348295467666
Erdos-Renyi 40	40	120	3	13	2.18205128205128	4	0.160237388724036	1.36244124459626	14.94568290212081
Barabasi 20	20	85	5	13	1.56842105263158	3	0.509749303621170	3.76132471856702	15.0896315592716

Graph	n	m	d_min	d_max	l	D	ccg	lambda_2	lambda_N
Barabasi 40	40	114	3	14	2.2025641025641	4	0.200524246395801	0.682992814912 8	15.58751092964952

Observation:

- It can be observed above that the number of edges for both Barabasi and Erdos_Renyi Graphs are almost same.
- Highest Maximum and minimum degrees(17 and 5 respectively) are seen in Barabasi Graph 20, which means that there are more edges from a vertex in Barabasi graph when compared with Erdos-Renyi.
- In Both the graphs the average path lengths are almost equal. Here, for Erods-Renyi 20 and Barabasi 20, the average number of steps along the shortest path is ~1.5, while for Erods-Renyi 40 and Barabasi 40, the average number of steps along the shortest path is ~2.2
- From the Global Clustering Coefficent(ccg), Erdos Renyi 40, has the lowest vale, which means there are less nodes in the graph which cluster together.
- From the second smallest eigen value, Barabasi 40 has the less algebraic connectivity in the graph when compared with other graphs.
- From the largest eigen value, Barabasi 20 has the maximum spectral radius in the graph when compared with other graphs.

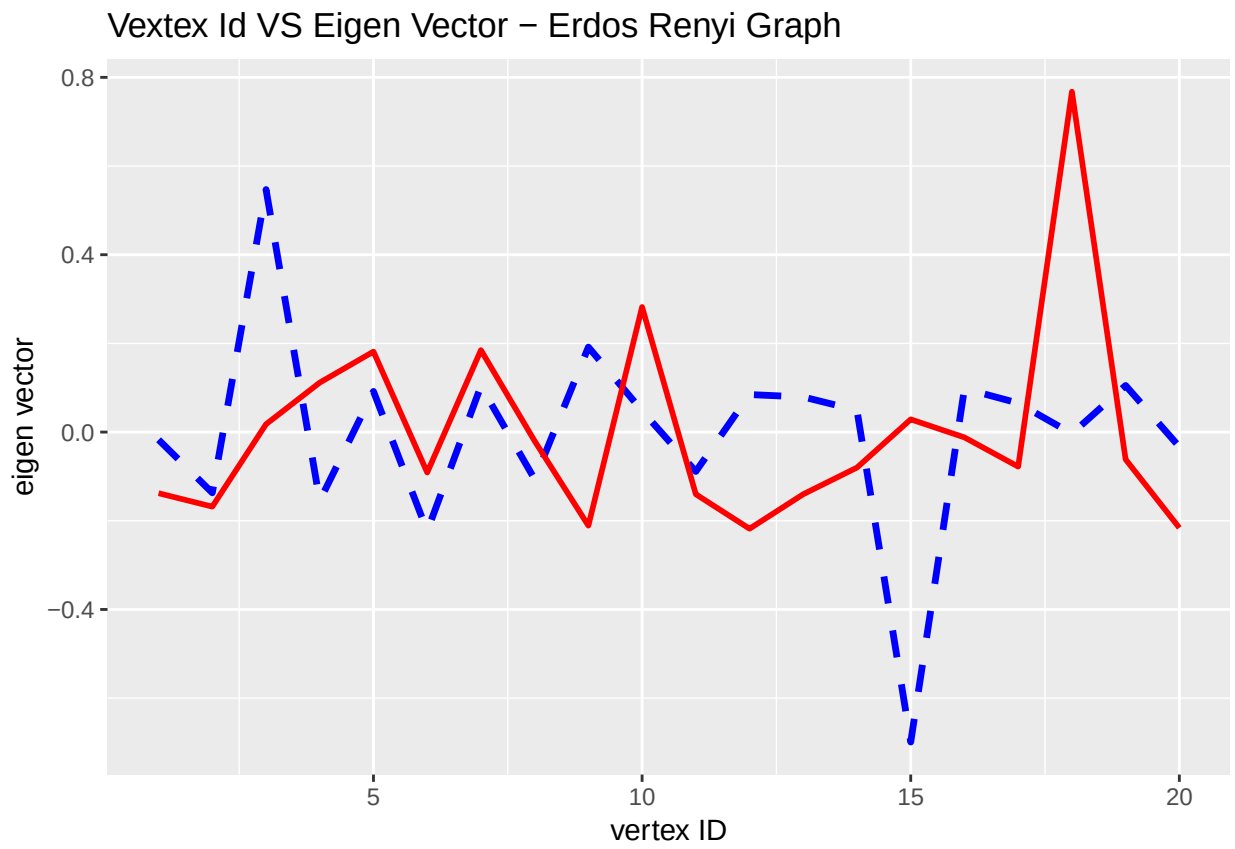
Question-2(2.3)

Plotting the eigenvector corresponding to the second-smallest eigenvalue (??2) and the eigenvector corresponding to the largest eigenvalue (??n)

```
x_axis <- c(1:20)
lambda2 <- eigen(graph.laplacian(er20))[["vectors"]][,19]
lambdaN <- eigen(graph.laplacian(er20))[["vectors"]][,which.max(eigen(graph.laplacian(er20))$values)]

plotdfER <- data.frame(x_axis, lambda2, lambdaN)

ggplot(plotdfER, aes(x = plotdfER$x_axis)) +
  geom_line(aes(y = plotdfER$lambda2), colour="blue", size = 1.2, linetype="dashed") +
  geom_line(aes(y = plotdfER$lambdaN), colour = "red", size = 1) +
  scale_colour_manual("",breaks = c("lambda2", "lambdaN"),values = c("red", "blue")) +
  ylab(label="eigen vector") +
  xlab("vertex ID") +
  labs(title = "Vextex Id VS Eigen Vector - Erdos Renyi Graph")
```

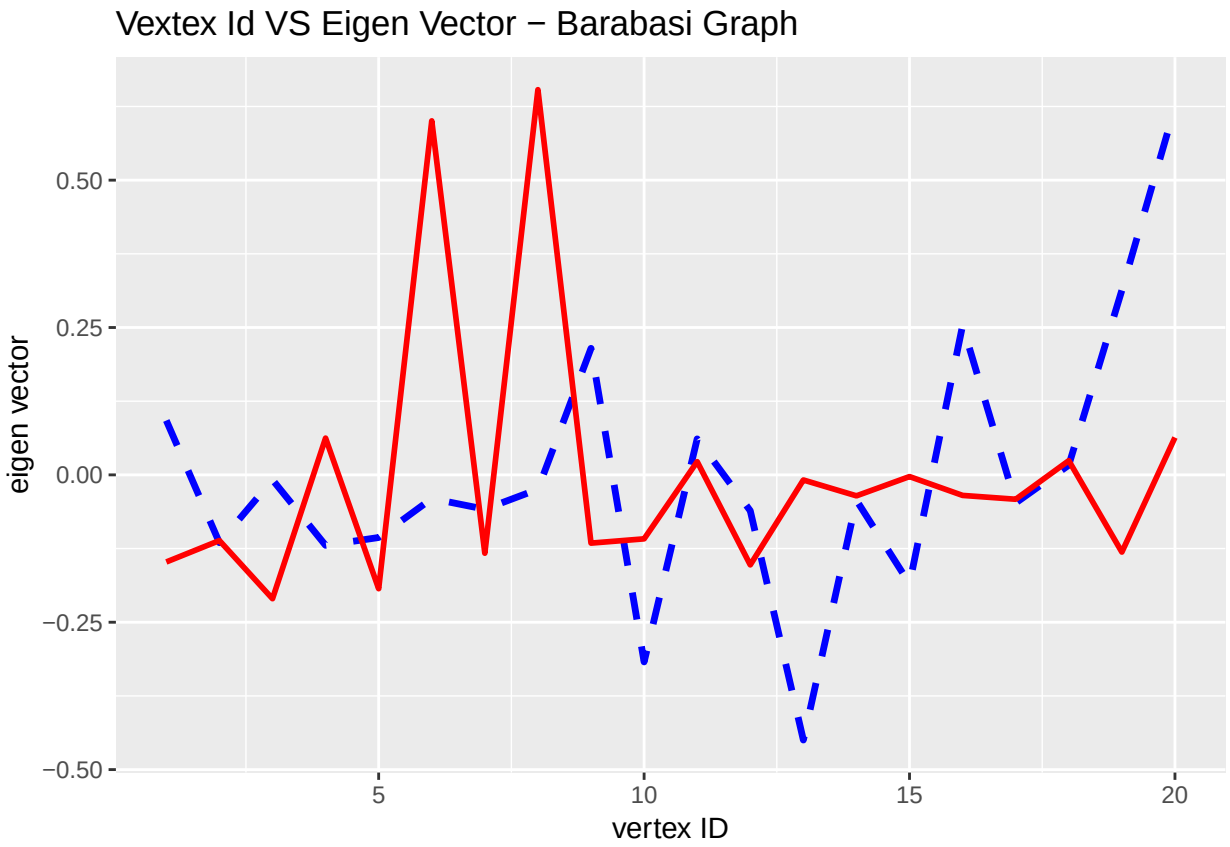



```
lambda2 <- eigen(graph.laplacian(bb20))[["vectors"]][,19]
lambdaN <- eigen(graph.laplacian(bb20))[["vectors"]][,which.max(eigen(graph.laplacian(er20))$values)]

plotdfBB <- data.frame(x_axis, lambda2, lambdaN)

ggplot(data = plotdfER, aes(x = plotdfBB$x_axis)) +
  geom_line(aes(y = plotdfBB$lambda2), colour="blue", size = 1.2, linetype="dashed") +
  geom_line(aes(y = plotdfBB$lambdaN), colour = "red", size = 1) +
  scale_colour_manual("",breaks = c("lambda2", "lambdaN"),values = c("red", "blue")) +
  ylab(label="eigen vector") +
  xlab("vertex ID") +
  labs(title = "Vextex Id VS Eigen Vector - Barabasi Graph") +
  scale_colour_manual("", breaks = c("lambda 2", "lambda N"))
```

```
## Scale for 'colour' is already present. Adding another scale for
## 'colour', which will replace the existing scale.
```



Observation:

It can be observed that both the Barabasi and Erdos Renyi Graphs are irregular by nature. However, The eigen values of both the second lowest lambda(lambda2) and highest lambda(lambda N) of Barabasi graph are found close to zero. It can be seen that there is a deviation of eigen values in the eigen vectors for both the graphs.

Question-3