International Trade Network

The size and structure of international trade flows varies significantly over time. This exercise is based in part on

```
Luca De Benedictis and Lucia Tajoli. (2011). 'The World Trade
Network.' *The World Economy*, 34:8, pp.1417-1454.
```

The trade data are from Katherine Barbieri and Omar

```
Keshk. (2012). *Correlates of War Project Trade Data Set*,
Version 3.0. available at
[http://correlatesofwar.org](http://correlatesofwar.org).
```

The volume of goods traded between countries has grown rapidly over the past century, as technological advances lowered the cost of shipping and countries adopted more liberal trade policies. At times, however, trade flows have decreased due to disruptive events such as major wars and the adoption of protectionist trade policies. In this exercise, we will explore some of these changes by examining the network of international trade over several time periods. The data file `trade.csv` contains the value of exports from one country to another in a given year. The names and descriptions of variables in this data set are:

Name	Description
`country1`	Country name of exporter
`country2`	Country name of importer
`year`	Year
`exports`	Total value of exports (in tens of millions of dollars)

The data are given for years 1900, 1920, 1940, 1955, 1980, 2000, and 2009.

Question 1

We begin by analyzing international trade as an unweighted, directed network. For every year in the data set, create an adjacency matrix whose entry (i, j) equals 1 if country i exports to country j. If this export is zero, then the entry equals 0. We assume that missing data, indicated by `NA`, represents zero trade. Plot the 'network density', which is defined over time as follows,

$$\text{network density} = \frac{\text{number of edges}}{\text{number of potential edges}}$$

The `graph.density` function can compute this measure given an adjacency matrix. Interpret the result.

Answer 1

```
par(cex = 1.5)
trade <- read.csv("data/trade.csv", stringsAsFactors = FALSE)

library(igraph)
```

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

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

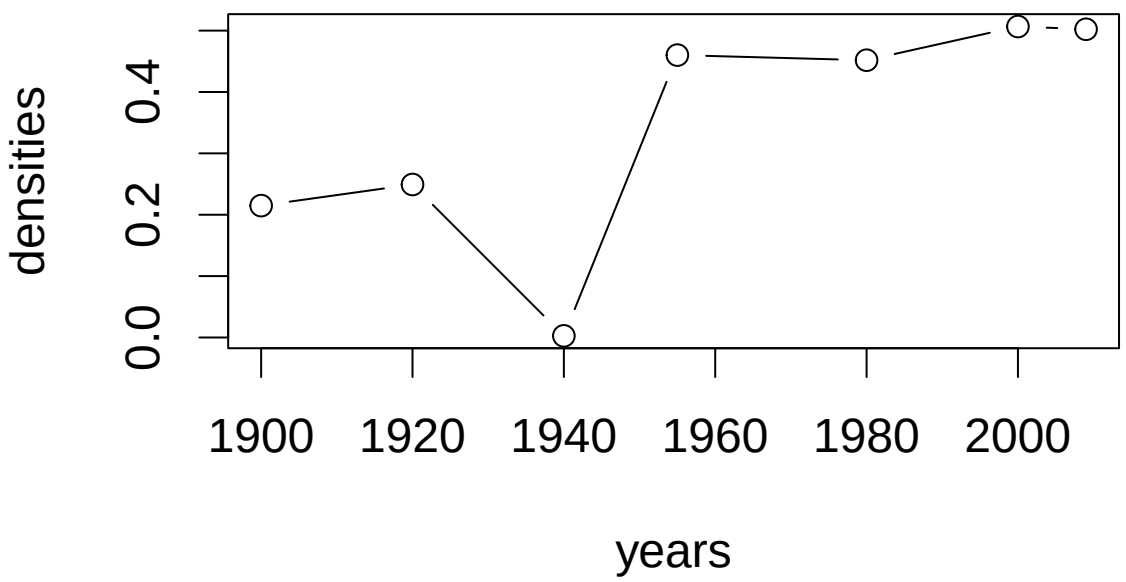
```
## The following object is masked from 'package:base':
##
##     union
```

```
## change NA to 0
trade$exports[is.na(trade$exports)] <- 0
## Create a loop over three years in the dataset
years <- unique(trade$year)
densities <- rep(NA, length(years))
adj.mat <- list() # create an empty list
for(i in 1:length(years)){
    ## Subset the dataset
    trade.year <- trade[trade$year == years[i],]
    ## Create a vector of all unique countries in this year
    countries <- unique(c(trade.year$country1,
                          trade.year$country2))
    ## Construct an unweighted, directed adjacency matrix
    adj.year <- matrix(0, nrow = length(countries),
                       ncol = length(countries))
    rownames(adj.year) <- colnames(adj.year) <- countries

     for (j in 1:nrow(trade.year)){
        exporter <- trade.year$country1[j]
        importer <- trade.year$country2[j]
        adj.year[exporter, importer] <- trade.year$exports[j]
    }
    ## unweighted network
    adj.year <- ifelse(adj.year > 0, 1, 0)
    ## Create network object and calculate graph density
    net <- graph.adjacency(adj.year, diag = FALSE)
    densities[i] <- graph.density(net)
    adj.mat[[i]] <- net # save network for each year
}
dim(adj.mat)
```

```
## NULL
```

```
names(adj.mat) <- years
## Plot network density over time
plot(years, densities, type = "b",
     main = "Density of International Trade Network")
```



The trade network has generally grown more dense from the period from 1900 to 2009. This means that over time, countries tend to gain more trade partners, exporting to a higher proportion of potential trade partners. However, there is a notable decrease in density in the year 1940, most likely reflecting protectionist policies adopted in the inter-war period.

Visualizaiton

Let's plot a random subset of the network in the year of 2009, using igraph.

```
library(igraph)
set.seed(02138)

#Slice year 2009
trade09 <-  subset(trade, year==2009)

#Countries
countries <- unique(c(trade09$country1,
                      trade09$country2))
adj09 <- matrix(NA, ncol=length(countries),nrow=length(countries))
colnames(adj09) <- countries
rownames(adj09) <- countries
for (j in 1:nrow(trade09)){
  exporter = trade09$country1[j]
  importer = trade09$country2[j]
  adj09[exporter, importer] <- trade09$exports[j]
}

## unweighted network
adj09 <- ifelse(adj09 > 0, 1, 0)

#Get rid of disconnected nodes
adj_comp = adj09[-which(rowSums(adj09,na.rm=T)==0,),
```

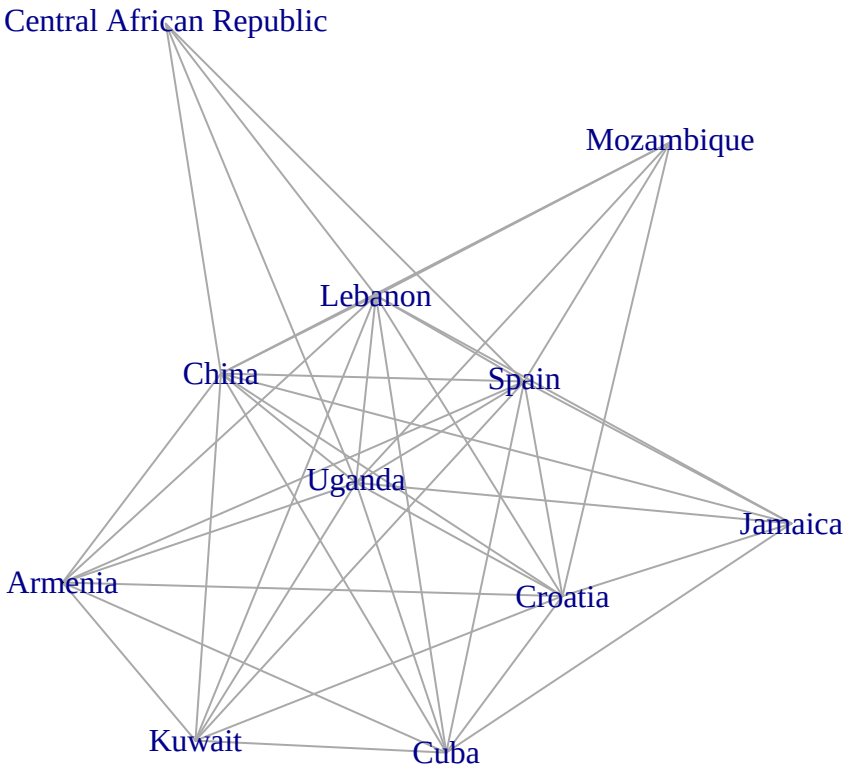
```
                -which(rowSums(adj09,na.rm=T)==0,)]

#Sample 50
set.seed(100)
ss <- sample(1:nrow(adj_comp),11)


adj_ss  <- adj_comp[ss,ss]
#Plot the graph
network <- graph_from_adjacency_matrix(adj_ss, mode = c("undirected"), weighted = NULL, diag = FALSE)

V(network)$label <-row.names(adj_ss)

par(mar=c(1,1,1,1))
plot(network, edge.arrow.size=0,vertex.size=0.001)
```



Question 2

For the years 1900, 1955, and 2009, compute the measures of centrality based on degree, betweenness, and closeness (based on total degree) for each year. For each year, list the five countries that have the largest values of these centrality measures. How do the countries on the lists change over time? Briefly comment on the results.

Answer 2

```
## Centrality measures for 1900, 1955, 2009
degree.1900 <- degree(adj.mat[["1900"]])
between.1900 <- betweenness(adj.mat[["1900"]])
```

```
close.1900 <- closeness(adj.mat[["1900"]], mode = "all")
degree.1955 <- degree(adj.mat[["1955"]])
between.1955 <- betweenness(adj.mat[["1955"]])
close.1955 <- closeness(adj.mat[["1955"]], mode = "all")
degree.2009 <- degree(adj.mat[["2009"]])
between.2009 <- betweenness(adj.mat[["2009"]])
close.2009 <- closeness(adj.mat[["2009"]], mode = "all")
## Identify 5 countries that rank highest on each
sort(degree.1900, decreasing = TRUE)[1:5]
```

```
##           United Kingdom United States of America                    France
##                       65                       54                       47
##                  Germany                  Belgium
##                       47                       42
```

```
sort(between.1900, decreasing = TRUE)[1:5]
```

```
##           United Kingdom United States of America                     Japan
##                419.45519                133.10400                 85.20877
##          Austria-Hungary                  Germany
##                 77.91829                 63.79488
```

```
sort(close.1900, decreasing = TRUE)[1:5]
```

```
##           United Kingdom United States of America                  Germany
##               0.02702703               0.02380952               0.02222222
##                   France                  Belgium
##               0.02173913               0.02040816
```

```
sort(degree.1955, decreasing = TRUE)[1:5]
```

```
##          United Kingdom German Federal Republic              Netherlands
##                     151                     147                     146
##                   Italy                  France
##                     146                     145
```

```
sort(between.1955, decreasing = TRUE)[1:5]
```

```
##           United Kingdom                   Italy United States of America
##                 296.9171                244.8158                 220.8367
##  German Federal Republic                  France
##                 210.4708                153.9190
```

```
sort(close.1955, decreasing = TRUE)[1:5]
```

```
##          United Kingdom             Netherlands German Federal Republic
##              0.01234568              0.01219512              0.01219512
##                   Italy                  France
##              0.01219512              0.01204819
```

```
sort(degree.2009, decreasing = TRUE)[1:5]
```

```
##           United Kingdom                   China                   India
##                      363                     363                     362
## United States of America                  France
##                      361                     361
```

```
sort(between.2009, decreasing = TRUE)[1:5]
```

```
##                   Taiwan United States of America                    Canada
##                 742.0159                  521.2458                  485.6058
##                    Japan                     India
##                 473.5176                  462.6919
```

```
sort(close.2009, decreasing = TRUE)[1:5]
```

```
##                    China                    India United States of America
##              0.005235602              0.005235602              0.005208333
##           United Kingdom                   France
##              0.005208333              0.005208333
```

In earlier periods, Western countries (e.g., the United Kingdom, Germany, the United States) tended to dominate the international trade network in terms of both closeness and betweenness centrality. By 2009, however, the top 5 countries had grown significantly more diverse. For example, at that time, China and India were the two top-ranking countries in closeness centrality, and Taiwan was the top-ranking country in betweenness centrality.

Question 3

We now analyze the international trade network as a weighted, directed network in which each edge has a non-negative weight proportional to its corresponding trade volume. Create an adjacency matrix for such network data. For the years 1900, 1955, and 2009, compute the centrality measures from above for the weighted trade network. Instead of degree, however, compute the *graph strength*, which in this case equals the sum of imports and exports with all adjacent nodes. The `graph.strength` function can be used to compute this weighted version of degree. For betweenness and closeness, we use the same function as before, i.e., `closeness` and `betweenness`, which can handle weighted graphs appropriately. Do the results differ from those of the unweighted network? Examine the top five countries. Can you think of another way to calculate centrality in this network that accounts for the value of exports from each country? Briefly discuss.

Answer 3

```
## the following code is similar to what we had earlier
wadj.mat <- list() # create an empty list
years <- c(1900, 1955, 2009)
for(i in 1:length(years)){
    ## Subset the dataset
    trade.year <- trade[trade$year == years[i],]
    ## Create a vector of all unique countries in this year
    countries <- unique(c(trade.year$country1,
                          trade.year$country2))
    ## Construct an unweighted, directed adjacency matrix
    adj.year <- matrix(0, nrow = length(countries),
                       ncol = length(countries))
    rownames(adj.year) <- colnames(adj.year) <- countries
    for (j in 1:nrow(trade.year)) {
        exporter <- trade.year$country1[j]
        importer <- trade.year$country2[j]
        adj.year[exporter, importer] <- trade.year$exports[j]
    }
    ## Create network object and calculate graph density
    wadj.mat[[i]] <- graph.adjacency(adj.year, weighted = TRUE,
                                     diag = FALSE)
}
names(wadj.mat) <- years
```

```r
## Centrality measures for 1900, 1955, 2009
strength.1900 <- graph.strength(wadj.mat[["1900"]])
between.1900 <- betweenness(wadj.mat[["1900"]])
close.1900 <- closeness(wadj.mat[["1900"]], mode = "all")
strength.1955 <- graph.strength(wadj.mat[["1955"]])
between.1955 <- betweenness(wadj.mat[["1955"]])
close.1955 <- closeness(wadj.mat[["1955"]], mode = "all")
strength.2009 <- graph.strength(wadj.mat[["2009"]])
between.2009 <- betweenness(wadj.mat[["2009"]])
```

```
## Warning in betweenness(wadj.mat[["2009"]]): At core/centrality/
## betweenness.c:111 : Some weights are smaller than epsilon, calculations may
## suffer from numerical precision.
```

```r
close.2009 <- closeness(wadj.mat[["2009"]], mode = "all")
## Identify 5 countries that rank highest on each
sort(strength.1900, decreasing = TRUE)[1:5]
```

```
##           United Kingdom                  Germany United States of America
##                2836.4000                2081.9900                1857.0103
##                   France                  Belgium
##                1242.0400                 750.0987
```

```r
sort(between.1900, decreasing = TRUE)[1:5]
```

```
##           Peru          Japan          Italy United Kingdom        Bolivia
##            636            485            347            311            302
```

```r
sort(close.1900, decreasing = TRUE)[1:5]
```

```
##                    Japan                 Thailand                  Belgium
##               0.04457737               0.04414445               0.04395043
##                  Romania United States of America
##               0.04390550               0.04388837
```

```r
sort(strength.1955, decreasing = TRUE)[1:5]
```

```
## United States of America           United Kingdom  German Federal Republic
##                26285.523                15852.263                11387.990
##                   Canada                   France
##                 9845.702                 6418.747
```

```r
sort(between.1955, decreasing = TRUE)[1:5]
```

```
## Luxembourg   Paraguay   Thailand    Iceland   Bulgaria
##  3270.0000   552.7667   509.1389   486.5333   479.3333
```

```r
sort(close.1955, decreasing = TRUE)[1:5]
```

```
## Luxembourg    Iceland     Panama   Honduras   Paraguay
## 0.09600851 0.09360189 0.09169178 0.09121014 0.08925065
```

```r
sort(strength.2009, decreasing = TRUE)[1:5]
```

```
## United States of America                    China                    Japan
##                2550432.8                2441328.4                1182584.8
##                   France           United Kingdom
##                 867601.3                 733534.2
```

```
sort(between.2009, decreasing = TRUE)[1:5]
```

```
##      Oman   Albania     Qatar Swaziland     Nauru
##  6585.996  5744.364  4991.218  4804.891  3287.748
```

```
sort(close.2009, decreasing = TRUE)[1:5]
```

```
##        Qatar      Namibia Turkmenistan   Tajikistan         Laos
## 0.0001489483 0.0001489483 0.0001489483 0.0001489483 0.0001489483
```

The betweenness and closeness centrality measures for the weighted networks produce results substantially different from those of the unweighted networks. In particular, they assign high betweenness and closeness centrality scores to countries that do not have large economies, such as Qatar and Albania. These scores suggest that while these countries may not trade in large quantities themselves, they link to other large trading partners that would otherwise be distantly connected. A weighted version of degree centrality shows that the United States and other Western countries have played a central role in the world trade network.

Question 4

Apply the PageRank algorithm to the weighted trade network separately for each year. For each year, identify the 5 most influential countries according to this algorithm. In addition, examine how the ranking of PageRank values has changed over time for each of the following five countries – US, United Kingdom, Russia, Japan, and China. Briefly comment on the patterns you observe.

Answer 4

```
?page.rank
## PageRank measures
pr.scores <- matrix(NA, nrow = length(years), ncol = 5)
colnames(pr.scores) <- c("United States of America", "United Kingdom",
                         "Russia", "Japan", "China")
rownames(pr.scores) <- years
for (i in 1:length(years)) {
    pr <- page.rank(wadj.mat[[i]])$vector
    cat("year", years[i], ":\n")
    print(sort(pr, decreasing = TRUE)[1:5])
    pr.order <- names(pr)[order(pr, decreasing = TRUE)]
    pr.scores[i, ] <- match(colnames(pr.scores), pr.order)
}
```

```
## year 1900 :
##           United Kingdom                  Germany United States of America
##               0.20958376               0.12107723               0.10116542
##                   France                  Belgium
##               0.07196933               0.04630940
## year 1955 :
## United States of America           United Kingdom  German Federal Republic
##               0.11837325               0.09744337               0.06151723
##                   Russia                   Canada
##               0.04329905               0.04110264
## year 2009 :
## United States of America                    China                   France
##               0.11591121               0.08227763               0.04025726
##                    Japan           United Kingdom
##               0.03784552               0.03541415
```

```
pr.scores
```

```
##      United States of America United Kingdom Russia Japan China
## 1900                        3              1      8    10    16
## 1955                        1              2      4     8    14
## 2009                        1              5     16     4     2
```

The United States has had the highest PageRank since 1940. Prior to then, United Kingdom was the most influential according to the PageRank algorithm. In recent years, the rankings of Japan and then China have risen according to PageRank values while the rankings of countries like the United Kingdom and Russia have fallen. Western European countries like Germany and France remain among the most influential countries.