Homework 6

Yi Chen

## R Markdown

### 1. data mannipulation

library(readxl)  
library(readr)  
library(dplyr)  
library(tidyr)  
library(smacof)  
data <- read\_csv("nations\_ALLDATA\_1920.csv")  
demographic <- read\_excel("nations demo data.xlsx")  
sex <- as.character(data[2,2:ncol(demographic)-1])  
birth <- as.character(data[1,2:ncol(demographic)-1])  
country <- unique(c(as.character(data$nation1),as.character(data$nation2)))  
data <- data %>% dplyr::select(2:21)  
  
## create the proximy   
dist\_list <- list()  
for (k in 1:length(sex)){  
 proximy <- matrix(NA,nrow = length(country),ncol = length(country))  
 rownames(proximy) <- country  
 colnames(proximy) <- country  
 for (i in rownames(proximy)){  
 for (j in colnames(proximy)){  
 if (i == j){  
 proximy[i,j] = 9  
 }else{  
 d <- as.numeric(data %>% filter(nation1==i & nation2==j) %>% dplyr::select(k+2))  
 proximy[j,i] = d  
 }  
 }  
 }  
 proximy\_dist <- as.dist(proximy)   
 dist\_list[[k]] <- sim2diss(proximy\_dist,method = 10,to.dist = TRUE)  
}  
  
dist\_list

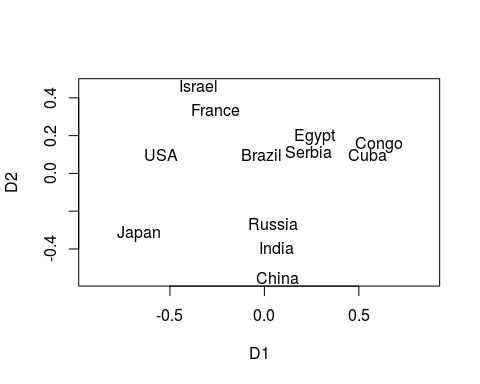
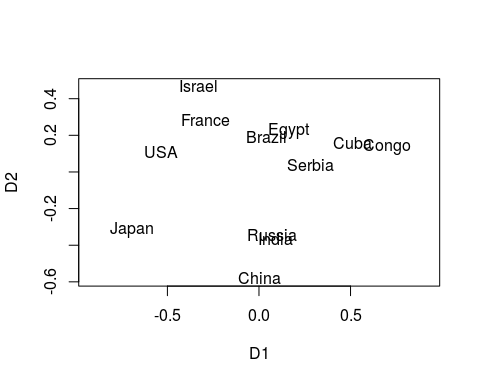
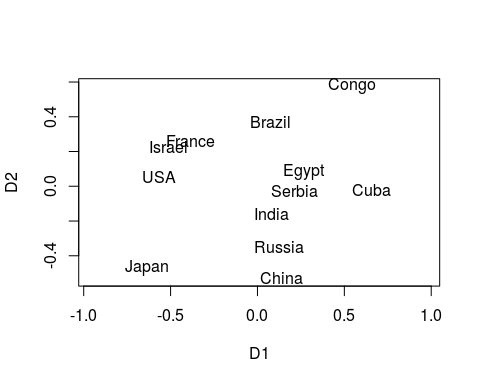
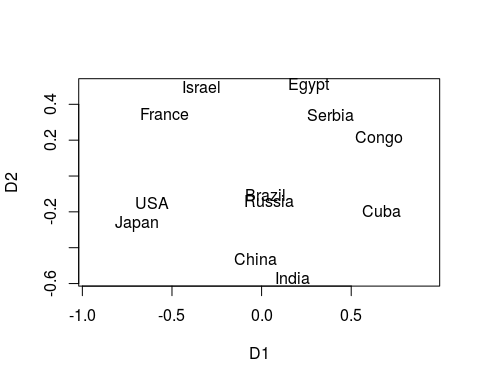
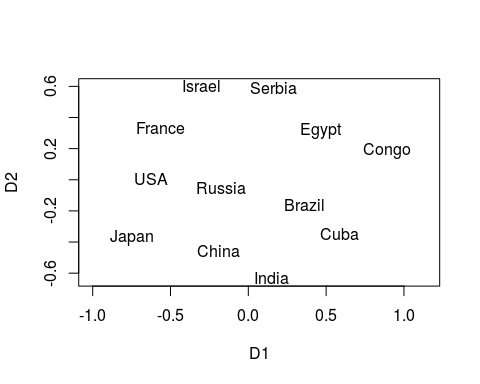
## [[1]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 7   
## Cuba 3 7   
## Egypt 6 2 5   
## France 5 7 7 6   
## India 4 7 6 7 9   
## Israel 5 9 8 6 2 6   
## Japan 7 9 8 7 4 4 4   
## China 4 8 3 6 5 3 9 3   
## Russia 5 7 4 6 6 5 7 4 4   
## USA 6 8 6 10 3 4 1 3 7 4   
## Serbia 8 7 7 4 3 7 8 9 8 3 8  
##   
## [[2]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 8   
## Cuba 5 5   
## Egypt 6 6 6   
## France 4 8 8 7   
## India 3 7 6 3 8   
## Israel 9 9 9 7 5 9   
## Japan 9 9 9 9 4 9 2   
## China 3 8 4 5 7 3 8 7   
## Russia 6 9 7 7 7 8 9 9 2   
## USA 6 8 7 7 3 8 7 5 7 6   
## Serbia 9 8 7 8 4 8 7 4 8 5 7  
##   
## [[3]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 8   
## Cuba 5 5   
## Egypt 6 6 6   
## France 4 8 8 7   
## India 3 7 6 3 8   
## Israel 9 9 9 7 5 9   
## Japan 9 9 9 9 4 9 2   
## China 3 8 4 5 7 3 8 7   
## Russia 6 9 7 7 7 8 9 9 2   
## USA 6 8 7 7 3 8 7 5 7 6   
## Serbia 9 8 7 8 4 8 7 4 8 5 7  
##   
## [[4]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 4   
## Cuba 5 7   
## Egypt 4 4 5   
## France 4 8 7 4   
## India 6 5 6 3 8   
## Israel 6 7 7 3 4 6   
## Japan 8 8 8 6 7 4 7   
## China 7 7 4 5 6 4 7 2   
## Russia 8 7 3 5 5 5 5 6 2   
## USA 6 8 6 4 3 6 4 5 7 3   
## Serbia 6 5 4 6 4 4 4 8 4 1 5  
##   
## [[5]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 5   
## Cuba 5 6   
## Egypt 6 4 4   
## France 6 7 9 8   
## India 4 6 3 2 6   
## Israel 8 7 8 8 3 6   
## Japan 7 8 8 7 7 7 6   
## China 8 7 7 3 4 3 7 4   
## Russia 7 7 6 6 3 5 4 7 2   
## USA 7 8 9 8 3 6 4 3 4 4   
## Serbia 6 5 4 4 7 5 5 8 8 7 9  
##   
## [[6]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 6   
## Cuba 3 6   
## Egypt 7 1 5   
## France 4 1 8 7   
## India 6 8 7 8 7   
## Israel 6 7 5 5 7 6   
## Japan 9 9 4 9 7 7 1   
## China 8 9 8 2 7 7 4 4   
## Russia 8 8 8 7 7 4 7 7 7   
## USA 3 3 8 7 8 4 7 6 3 7   
## Serbia 9 7 8 6 4 9 5 6 6 6 7  
##   
## [[7]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 7   
## Cuba 3 3   
## Egypt 3 2 5   
## France 7 9 8 7   
## India 2 5 6 3 7   
## Israel 5 8 7 5 7 7   
## Japan 7 9 9 9 4 6 8   
## China 4 8 6 7 7 4 6 4   
## Russia 3 7 4 6 4 3 7 6 2   
## USA 5 9 9 7 2 6 4 3 6 3   
## Serbia 5 5 5 4 6 7 6 8 5 3 9  
##   
## [[8]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 3   
## Cuba 2 6   
## Egypt 5 3 5   
## France 8 8 8 8   
## India 5 4 5 5 8   
## Israel 7 8 8 5 4 8   
## Japan 7 8 8 8 5 7 5   
## China 3 7 3 7 8 8 8 3   
## Russia 9 8 8 9 8 8 7 6 3   
## USA 4 8 8 5 3 7 2 5 9 8   
## Serbia 6 8 4 4 8 6 8 8 7 3 9  
##   
## [[9]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 5   
## Cuba 3 4   
## Egypt 5 3 3   
## France 5 8 6 4   
## India 4 3 5 2 6   
## Israel 4 7 3 2 5 6   
## Japan 4 9 5 7 4 6 4   
## China 3 6 5 6 5 3 6 3   
## Russia 6 7 3 6 3 5 5 4 2   
## USA 3 8 5 5 3 6 3 3 3 3   
## Serbia 6 6 4 4 4 6 4 7 7 3 5  
##   
## [[10]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 8   
## Cuba 4 8   
## Egypt 7 8 8   
## France 6 6 8 7   
## India 7 8 7 7 8   
## Israel 8 8 8 5 6 7   
## Japan 8 8 8 7 5 7 6   
## China 5 7 6 7 6 7 7 4   
## Russia 7 8 5 7 6 6 5 6 3   
## USA 5 8 8 7 5 6 5 6 6 7   
## Serbia 6 8 6 8 7 7 7 8 7 2 5  
##   
## [[11]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 4   
## Cuba 5 5   
## Egypt 6 4 4   
## France 6 4 4 7   
## India 4 8 5 6 6   
## Israel 7 7 7 3 4 6   
## Japan 7 7 7 8 4 4 5   
## China 3 4 4 5 5 3 5 2   
## Russia 5 5 5 4 4 5 5 5 2   
## USA 4 8 5 6 2 6 3 2 3 3   
## Serbia 7 6 7 5 5 5 5 6 6 5 7  
##   
## [[12]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 9   
## Cuba 5 9   
## Egypt 5 3 7   
## France 6 8 7 9   
## India 8 7 5 9 9   
## Israel 8 9 8 9 5 9   
## Japan 2 8 8 9 6 7 8   
## China 6 8 9 8 7 7 5 7   
## Russia 9 2 7 9 9 5 2 7 4   
## USA 4 9 8 9 5 3 8 5 6 6   
## Serbia 6 7 6 7 8 7 8 8 8 8 8  
##   
## [[13]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 1   
## Cuba 8 8   
## Egypt 7 9 6   
## France 9 4 8 4   
## India 9 7 5 3 6   
## Israel 9 8 9 2 5 2   
## Japan 8 9 8 3 5 6 4   
## China 8 8 3 2 5 2 3 1   
## Russia 8 9 8 3 8 4 7 2 2   
## USA 4 8 1 6 3 7 5 2 5 8   
## Serbia 5 7 5 3 5 3 2 3 7 6 5  
##   
## [[14]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 4   
## Cuba 4 4   
## Egypt 4 6 5   
## France 7 7 7 7   
## India 6 6 5 5 7   
## Israel 6 7 5 6 6 7   
## Japan 7 7 7 7 3 7 4   
## China 5 4 4 3 7 4 5 6   
## Russia 3 7 3 3 6 5 6 7 3   
## USA 6 7 6 6 4 5 4 3 6 6   
## Serbia 4 4 4 4 7 5 6 7 3 4 7  
##   
## [[15]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 4   
## Cuba 3 4   
## Egypt 7 3 4   
## France 5 7 5 7   
## India 5 8 7 4 7   
## Israel 6 7 7 2 5 7   
## Japan 8 9 6 8 4 4 5   
## China 5 7 4 8 8 3 8 2   
## Russia 5 7 4 8 4 6 7 7 4   
## USA 3 7 6 8 3 5 5 4 5 5   
## Serbia 7 8 6 6 4 6 6 8 5 2 7  
##   
## [[16]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 6   
## Cuba 6 6   
## Egypt 7 4 7   
## France 7 7 8 9   
## India 7 6 5 6 8   
## Israel 8 8 8 3 3 7   
## Japan 8 9 8 9 6 6 7   
## China 5 7 4 6 7 2 8 5   
## Russia 7 8 5 9 5 5 7 6 4   
## USA 4 7 9 8 3 7 3 5 8 8   
## Serbia 7 9 6 9 5 9 7 8 7 2 7  
##   
## [[17]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 6   
## Cuba 3 5   
## Egypt 6 7 5   
## France 8 7 7 8   
## India 6 7 5 7 7   
## Israel 8 8 5 6 7 8   
## Japan 8 9 7 6 4 4 8   
## China 9 8 7 7 6 4 8 3   
## Russia 7 9 7 7 4 8 7 5 3   
## USA 7 9 9 8 6 8 9 6 2 2   
## Serbia 6 6 7 6 6 6 6 7 8 2 10  
##   
## [[18]]  
## Brazil Congo Cuba Egypt France India Israel Japan China Russia USA  
## Congo 2   
## Cuba 3 7   
## Egypt 5 4 6   
## France 5 8 7 3   
## India 8 7 7 6 8   
## Israel 4 5 5 2 7 4   
## Japan 8 7 8 4 6 5 6   
## China 8 7 8 3 8 4 6 2   
## Russia 7 8 4 7 4 7 3 3 3   
## USA 3 6 2 5 3 10 6 7 8 3   
## Serbia 3 3 7 5 4 4 3 5 5 5 6

#### 2

###### 2a: individual difference MDS

There is an error when I try the r = 1. I just try r from 2 to 6, instead.

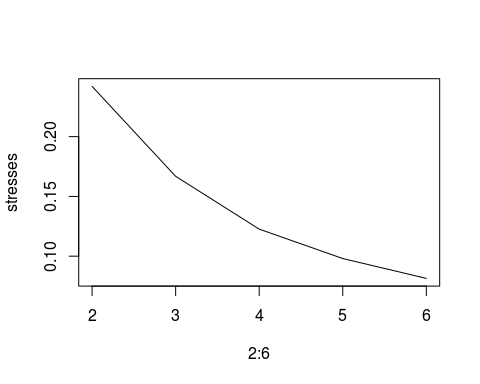
stresses <- c()  
for (r in 2:6){  
 country\_ind<- indscal(dist\_list,type="ordinal",init="torgerson",ndim = r)  
 stresses <- c(stresses,country\_ind$stress)  
 plot(country\_ind$gspace,asp=1,pch=' ')   
 text(country\_ind$gspace,colnames(proximy))  
}



stresses

## [1] 0.24206113 0.16679605 0.12259322 0.09797922 0.08137385

plot(2:6,stresses,type = 'l')



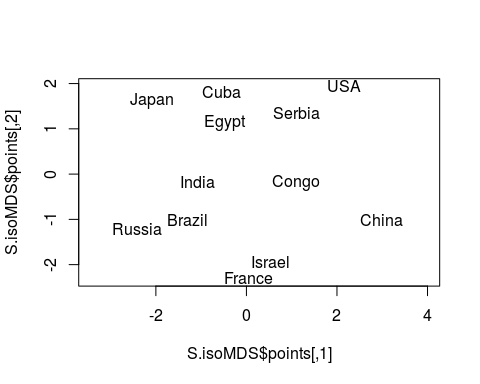
There are no big differences between the results from different dimension after r = 3. The overall configuration of the countries are similar. I personally feel the 4 dimension result is the best one: Russia and Japan are closer probably because of the deographic factors. USA and China are closer probably because of their social-economics status.

###### 2b average difference MDS

nations <- read\_csv("nations\_ALLDATA\_1920\_lowerhalf.csv")   
row.names(nations) <- nations %>% pull(X1)  
nations <- nations %>% dplyr::select(c(-1))  
nations <- nations %>% replace(is.na(.),0)   
S <- as.matrix(nations + t(nations))  
diag(S) <- 9  
S <- as.dist(S)  
S.isoMDS <- MASS::isoMDS(S,k=3)

## initial value 22.759334   
## iter 5 value 17.702845  
## iter 10 value 17.099188  
## final value 16.995915   
## converged

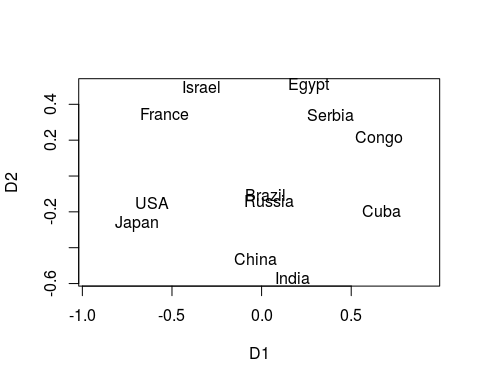
plot(S.isoMDS$points,asp = 1,pch=" ")   
text(S.isoMDS$points,row.names(nations))



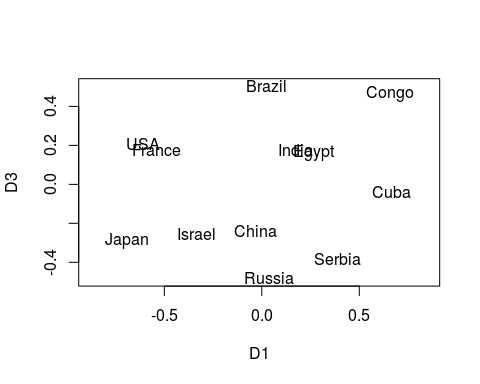
There are big difference in the result. In this configuration, China and USA are far away in the dimension 2. Same as Russia and Japan.

#### 3 INDSCAL Interpretation

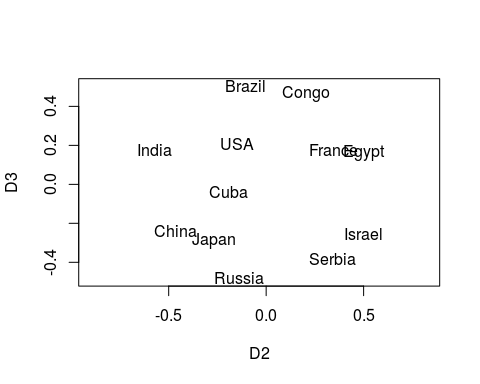
country\_ind<- indscal(dist\_list,type="ordinal",init="torgerson",ndim = 3)  
plot(country\_ind$gspace[,c(1,2)],asp=1,pch=' ')   
text(country\_ind$gspace[,c(1,2)],colnames(proximy))



plot(country\_ind$gspace[,c(1,3)],asp=1,pch=' ')   
text(country\_ind$gspace[,c(1,3)],colnames(proximy))



plot(country\_ind$gspace[,c(2,3)],asp=1,pch=' ')   
text(country\_ind$gspace[,c(2,3)],colnames(proximy))



For the first dimension:Gongo and Cuba are low while Israel, Janpan, USA, and France are low. Maybe this is related to the technology and education difference.

For the second dimension:Janpan, China are low, while Rgypt, Congo, and France are low. May be related to the cultural and demographic difference.

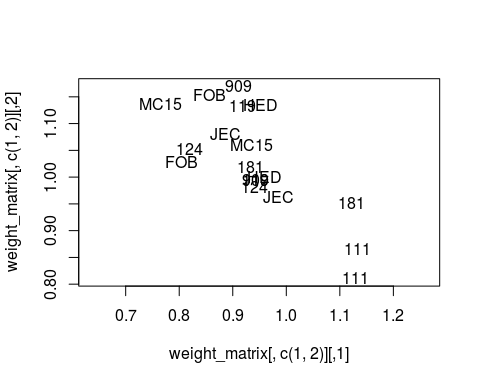
For the third dimension: Congo, Egypt and India are high, while Russia, China, and Japan are low. May be related to the economics difference.

#### 4

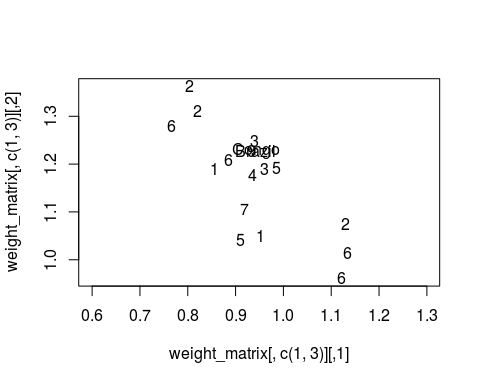
#country\_ind<- indscal(dist\_list,type="ordinal",init="torgerson",ndim = 4)  
weight\_matrix <- matrix(NA,nrow=length(sex),ncol = 3)  
for (i in 1:length(sex)){  
 weight\_matrix[i,] <- diag(country\_ind$cweights[[i]])  
}  
weight\_matrix

## [,1] [,2] [,3]  
## [1,] 0.9525080 1.1341427 1.0502796  
## [2,] 0.9417935 0.9946555 1.2270807  
## [3,] 0.9417935 0.9946555 1.2270807  
## [4,] 0.9401665 0.9806793 1.2479379  
## [5,] 1.1297687 0.8108202 1.0737825  
## [6,] 0.8040193 1.0276428 1.3624709  
## [7,] 1.1211207 0.9522580 0.9609639  
## [8,] 0.9861716 0.9636382 1.1927617  
## [9,] 0.9354330 1.0608883 1.1779747  
## [10,] 0.9596332 1.0003871 1.1896828  
## [11,] 0.9188522 1.1330881 1.1063076  
## [12,] 0.9113015 1.1694523 1.0417977  
## [13,] 0.8193659 1.0526524 1.3118363  
## [14,] 1.1332880 0.8648257 1.0133890  
## [15,] 0.8563618 1.1530766 1.1890325  
## [16,] 0.9329159 1.0192787 1.2277427  
## [17,] 0.8859778 1.0799848 1.2083233  
## [18,] 0.7654909 1.1370943 1.2804174

plot(weight\_matrix[,c(1,2)],asp=1,pch=' ')   
text(weight\_matrix[,c(1,2)],colnames(data)[4:12])



plot(weight\_matrix[,c(1,3)],asp=1,pch=' ')   
text(weight\_matrix[,c(1,3)], birth)



plot(weight\_matrix[,c(2,3)],asp=1,pch=' ')   
text(weight\_matrix[,c(2,3)], birth)

