Homework5

Homework 5

part 1

```
#install.packages("smacof")
library(readr)
library(smacof)
library(dplyr)
library(MASS)
nations <- read_csv("nations_ALLDATA_1920_lowerhalf.csv")
row.names(nations) <- nations %>% pull(X1)
(nations <- nations %>% dplyr::select(c(-1)))
```

```
## # A tibble: 12 x 12
##    Brazil Congo  Cuba Egypt France India Israel Japan China Russia   USA Ser
bia
##  *  <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl>  <dbl> <dbl> <dbl>  <dbl> <dbl> <lg
l>
##  1  NA    NA    NA    NA     NA    NA     NA    NA    NA     NA    NA    NA
##  2   4.61 NA    NA    NA     NA    NA     NA    NA    NA     NA    NA    NA
##  3   5.83  4.17 NA    NA     NA    NA     NA    NA    NA     NA    NA    NA
##  4   4.33  5.61  4.67 NA     NA    NA     NA    NA    NA     NA    NA    NA
##  5   4.11  3.22  2.78  3.39  NA    NA     NA    NA    NA     NA    NA    NA
##  6   4.61  3.56  4.39  5.06   2.61 NA     NA    NA    NA     NA    NA    NA
##  7   3.17  2.33  3     5.22   5     3.33  NA    NA    NA     NA    NA    NA
##  8   2.72  1.61  2.5   2.67   5.06  3.94   4.89 NA    NA     NA    NA    NA
##  9   4.61  2.89  4.83  4.72   3.61  5.89   3.44  6.17 NA     NA    NA    NA
## 10   3.56  2.67  4.56  3.56   4.44  4.33   3.94  4.11  7     NA    NA    NA
## 11   5.22  2.39  3.39  3.17   6.39  3.78   5.17  5.67  4.33   4.89 NA    NA
## 12   3.61  3.5   4.22  4.39   4.72  3.78   4.22  3.22  3.5    6     2.89 NA
```

part 2

```
nations <- nations %>% replace(is.na(.),0)
(S <- as.matrix(nations + t(nations)))
```

```
##        Brazil Congo Cuba Egypt France India Israel Japan China Russia  USA
## Brazil   0.00  4.61 5.83  4.33   4.11  4.61   3.17  2.72  4.61   3.56 5.22
## Congo    4.61  0.00 4.17  5.61   3.22  3.56   2.33  1.61  2.89   2.67 2.39
## Cuba     5.83  4.17 0.00  4.67   2.78  4.39   3.00  2.50  4.83   4.56 3.39
## Egypt    4.33  5.61 4.67  0.00   3.39  5.06   5.22  2.67  4.72   3.56 3.17
## France   4.11  3.22 2.78  3.39   0.00  2.61   5.00  5.06  3.61   4.44 6.39
## India    4.61  3.56 4.39  5.06   2.61  0.00   3.33  3.94  5.89   4.33 3.78
## Israel   3.17  2.33 3.00  5.22   5.00  3.33   0.00  4.89  3.44   3.94 5.17
## Japan    2.72  1.61 2.50  2.67   5.06  3.94   4.89  0.00  6.17   4.11 5.67
## China    4.61  2.89 4.83  4.72   3.61  5.89   3.44  6.17  0.00   7.00 4.33
## Russia   3.56  2.67 4.56  3.56   4.44  4.33   3.94  4.11  7.00   0.00 4.89
## USA      5.22  2.39 3.39  3.17   6.39  3.78   5.17  5.67  4.33   4.89 0.00
## Serbia   3.61  3.50 4.22  4.39   4.72  3.78   4.22  3.22  3.50   6.00 2.89
##        Serbia
## Brazil   3.61
## Congo    3.50
## Cuba     4.22
## Egypt    4.39
## France   4.72
## India    3.78
## Israel   4.22
## Japan    3.22
## China    3.50
## Russia   6.00
## USA      2.89
## Serbia   0.00
```

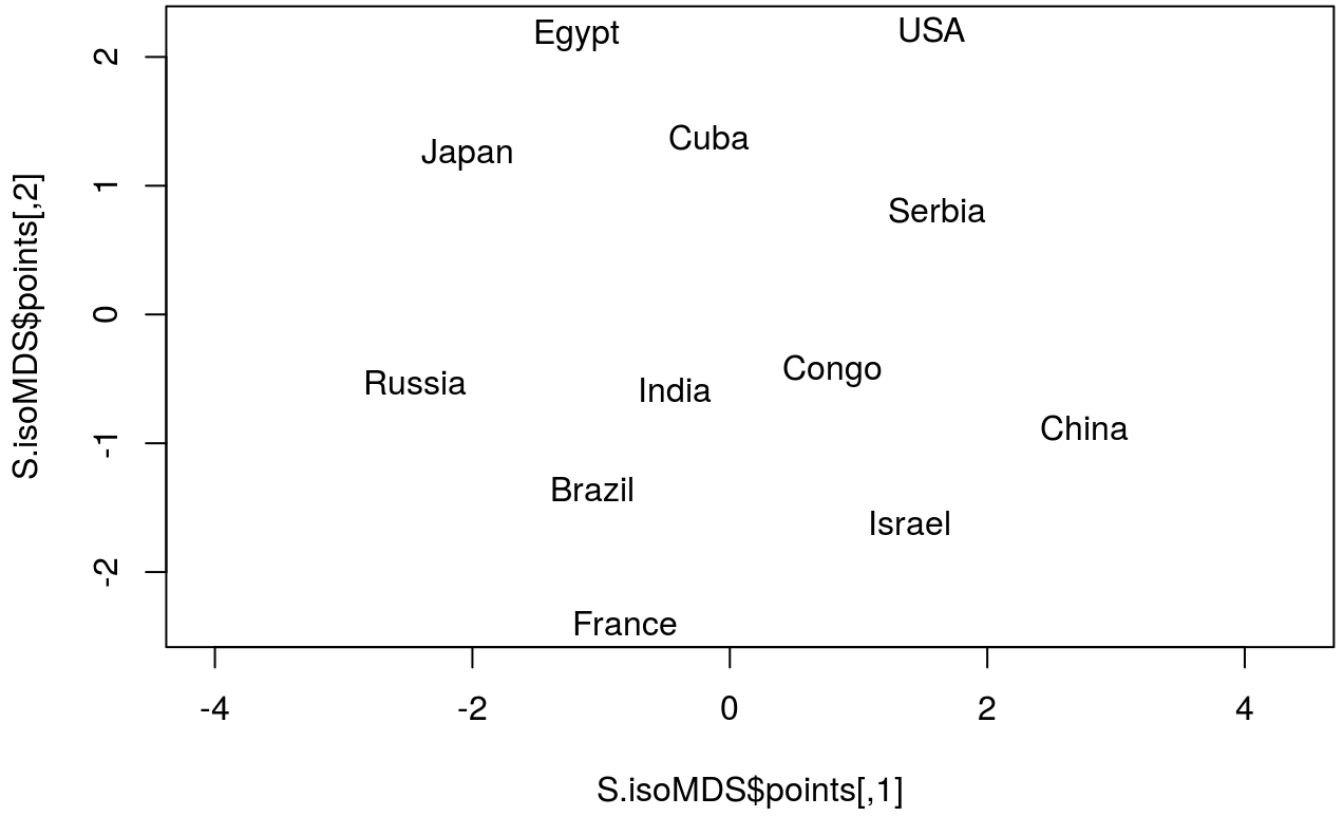
```
S.isoMDS <- MASS::isoMDS(S,k = 2,trace = TRUE)
```

```
## initial  value 31.825212
## iter   5 value 26.535255
## final  value 26.410546
## converged
```

```
S.isoMDS$stress
```

```
## [1] 26.41055
```

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



Close neighbors between Egypt and USA does not make sense.

part 3

```
(DD <- sim2diss(S,method = 10))
```

```
##         Brazil Congo  Cuba Egypt France India Israel Japan China Russia   USA
## Brazil  10.00  5.39  4.17  5.67   5.89  5.39   6.83  7.28  5.39   6.44  4.78
## Congo    5.39 10.00  5.83  4.39   6.78  6.44   7.67  8.39  7.11   7.33  7.61
## Cuba     4.17  5.83 10.00  5.33   7.22  5.61   7.00  7.50  5.17   5.44  6.61
## Egypt    5.67  4.39  5.33 10.00   6.61  4.94   4.78  7.33  5.28   6.44  6.83
## France   5.89  6.78  7.22  6.61  10.00  7.39   5.00  4.94  6.39   5.56  3.61
## India    5.39  6.44  5.61  4.94   7.39 10.00   6.67  6.06  4.11   5.67  6.22
## Israel   6.83  7.67  7.00  4.78   5.00  6.67  10.00  5.11  6.56   6.06  4.83
## Japan    7.28  8.39  7.50  7.33   4.94  6.06   5.11 10.00  3.83   5.89  4.33
## China    5.39  7.11  5.17  5.28   6.39  4.11   6.56  3.83 10.00   3.00  5.67
## Russia   6.44  7.33  5.44  6.44   5.56  5.67   6.06  5.89  3.00  10.00  5.11
## USA      4.78  7.61  6.61  6.83   3.61  6.22   4.83  4.33  5.67   5.11 10.00
## Serbia   6.39  6.50  5.78  5.61   5.28  6.22   5.78  6.78  6.50   4.00  7.11
##         Serbia
## Brazil   6.39
## Congo    6.50
## Cuba     5.78
## Egypt    5.61
## France   5.28
## India    6.22
## Israel   5.78
## Japan    6.78
## China    6.50
## Russia   4.00
## USA      7.11
## Serbia  10.00
```

```
(Dx <- sim2diss(S,method = 10,to.dist = TRUE))
```

```
##         Brazil Congo Cuba Egypt France India Israel Japan China Russia  USA
## Congo    5.39
## Cuba     4.17  5.83
## Egypt    5.67  4.39 5.33
## France   5.89  6.78 7.22  6.61
## India    5.39  6.44 5.61  4.94   7.39
## Israel   6.83  7.67 7.00  4.78   5.00  6.67
## Japan    7.28  8.39 7.50  7.33   4.94  6.06   5.11
## China    5.39  7.11 5.17  5.28   6.39  4.11   6.56  3.83
## Russia   6.44  7.33 5.44  6.44   5.56  5.67   6.06  5.89  3.00
## USA      4.78  7.61 6.61  6.83   3.61  6.22   4.83  4.33  5.67   5.11
## Serbia   6.39  6.50 5.78  5.61   5.28  6.22   5.78  6.78  6.50   4.00 7.11
```

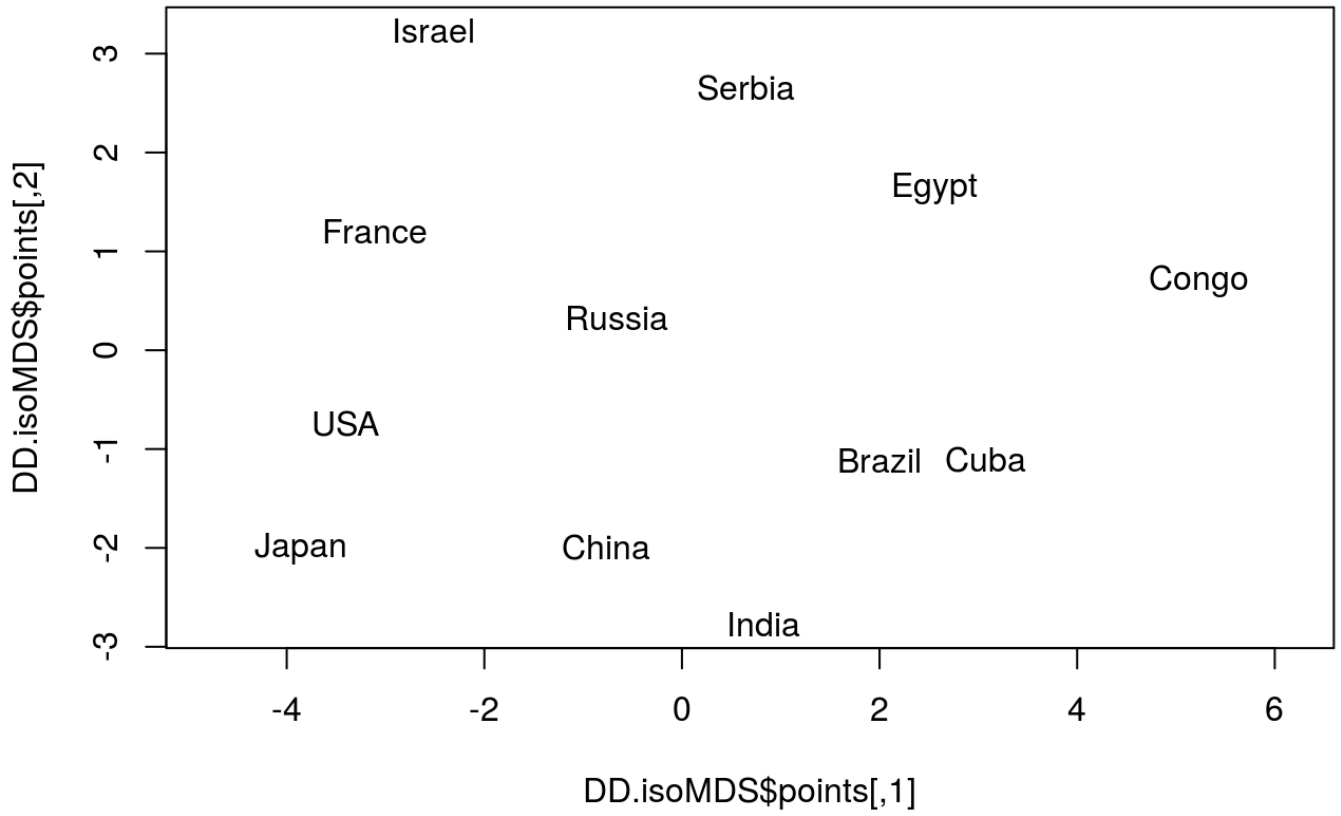
```
diag(DD) <- 0
DD.isoMDS <- MASS::isoMDS(DD,k = 2,trace = TRUE)
```

```
## initial  value 22.057320
## iter   5 value 17.421786
## iter  10 value 16.932740
## final  value 16.894084
## converged
```

```
DD.isoMDS$stress
```

```
## [1] 16.89408
```

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



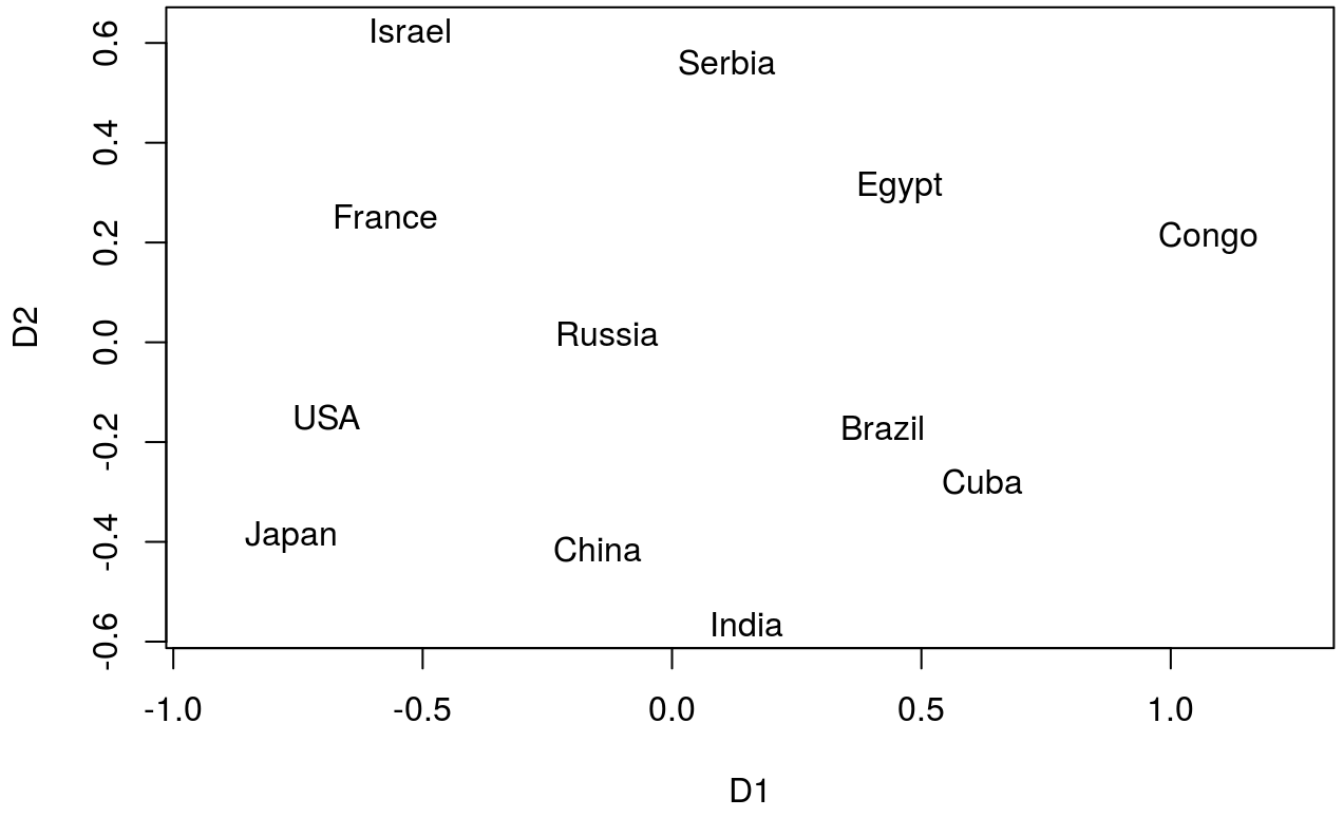
Firstly, the stress decrease a lot which means a big overall improvement of the model fit. Second the configuration of the new solution make much more sense. The developed countries like USA and Japan and France are group closer, while the developing countries like Cuba, Congo, and Egypt are closer.

part 4

```
(nn_sm <- smacof::smacofSym(Dx,ndim = 2,type = "ordinal"))
```

```
##
## Call:
## smacof::smacofSym(delta = Dx, ndim = 2, type = "ordinal")
##
## Model: Symmetric SMACOF
## Number of objects: 12
## Stress-1 value: 0.164
## Number of iterations: 65
```

```
plot(nn_sm$conf,asp=1,pch=' ')
text(nn_sm$conf,rownames(nations))
```

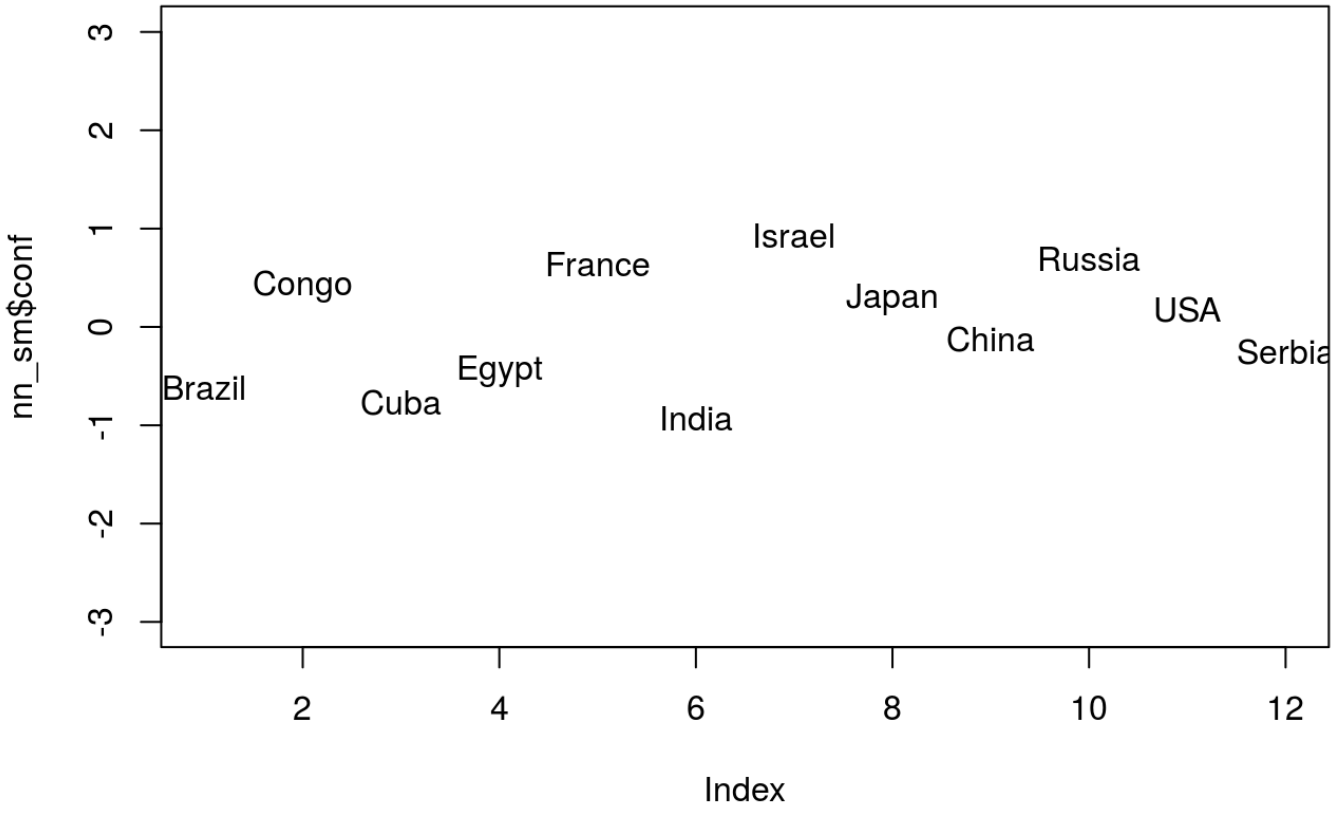


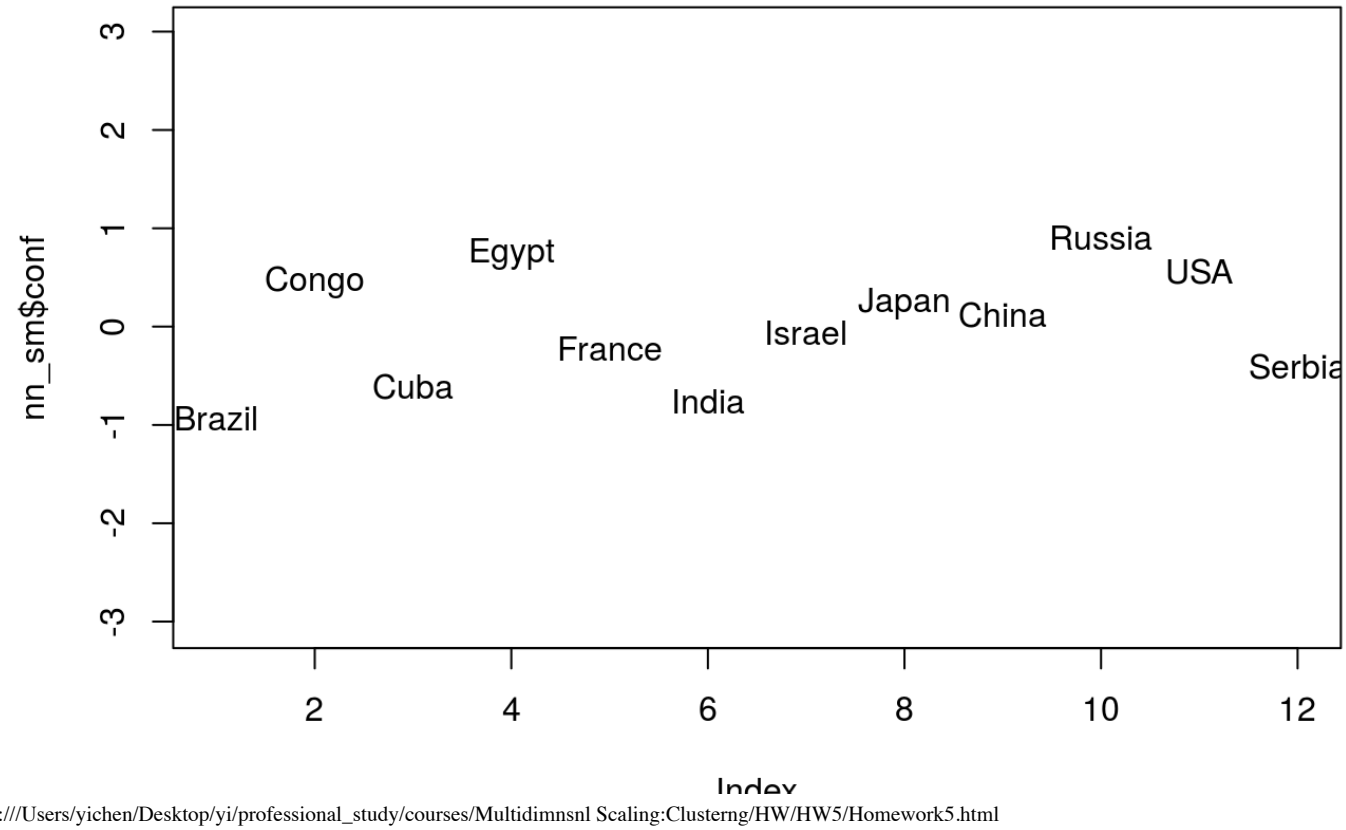
The results from isoMDS and smacof are very similar to each other. But smacof gives a much smaller stress. This means smacof have a better model fit. But the difference is limited in terms of interpreation and application of the results.

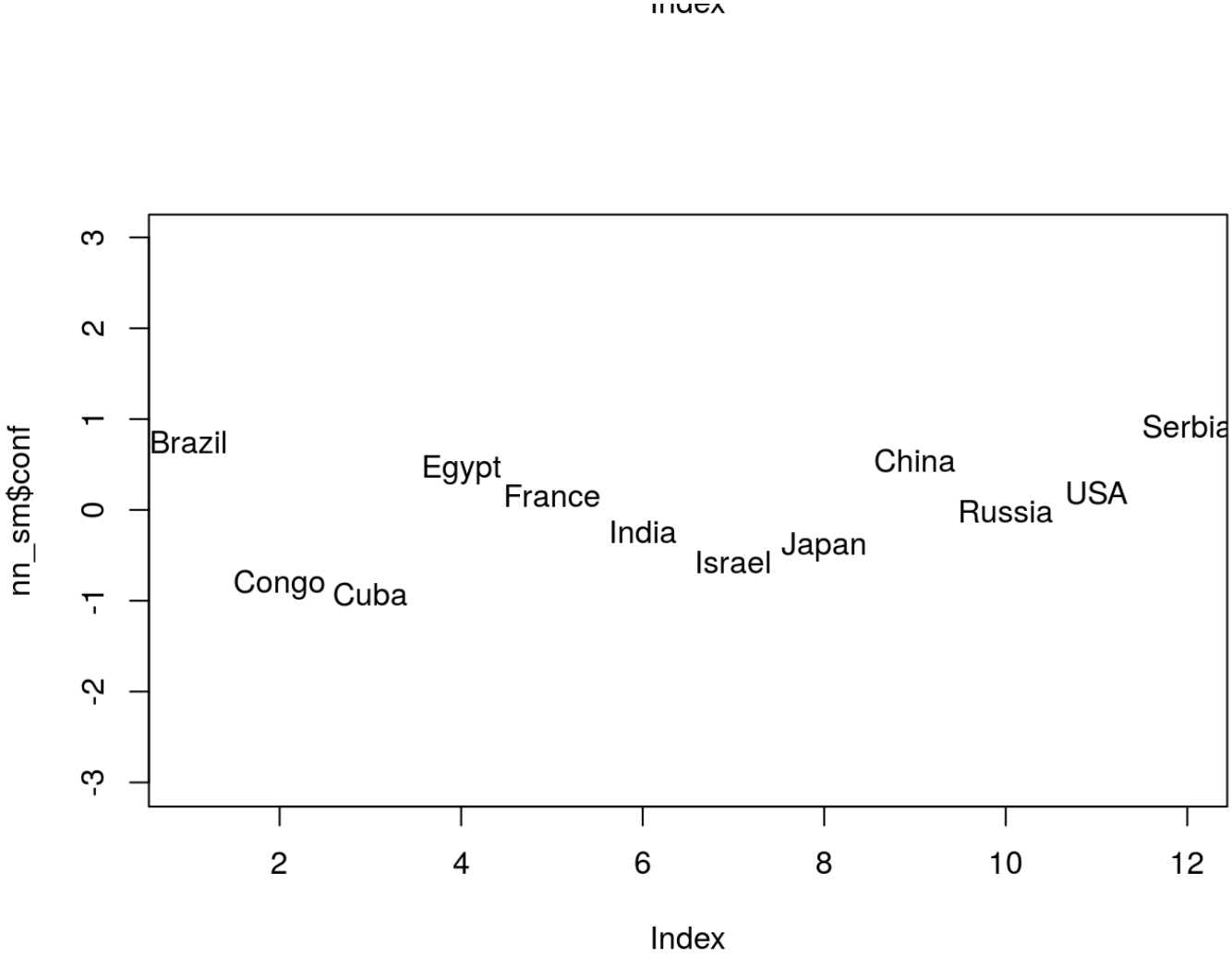
part 5

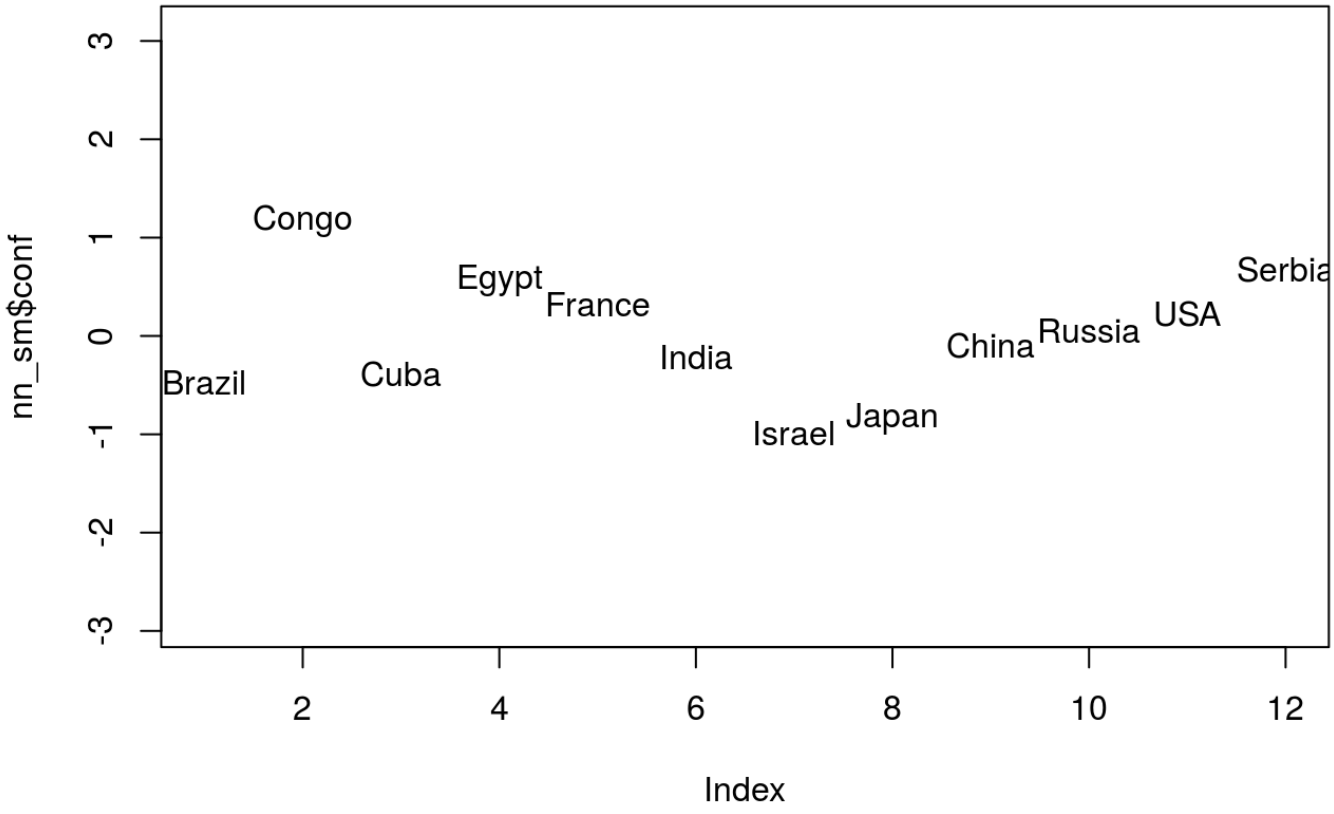
For one dimension

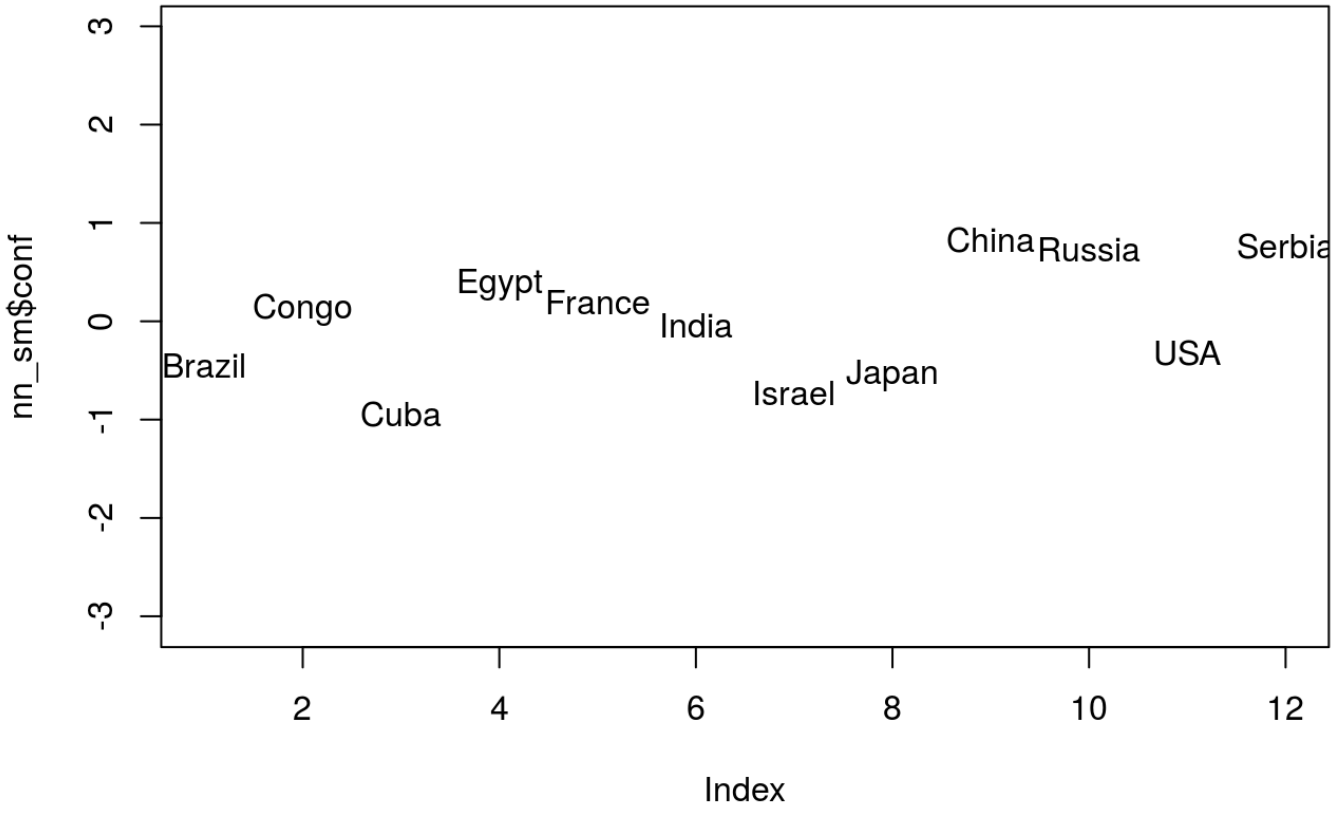
```
for (i in 1:10){
  nn_sm <- smacof::smacofSym(Dx,ndim = 1,type = "ordinal",init = "random")
  plot(nn_sm$conf,asp=1,pch=' ')
  text(nn_sm$conf,rownames(nations))
}
```

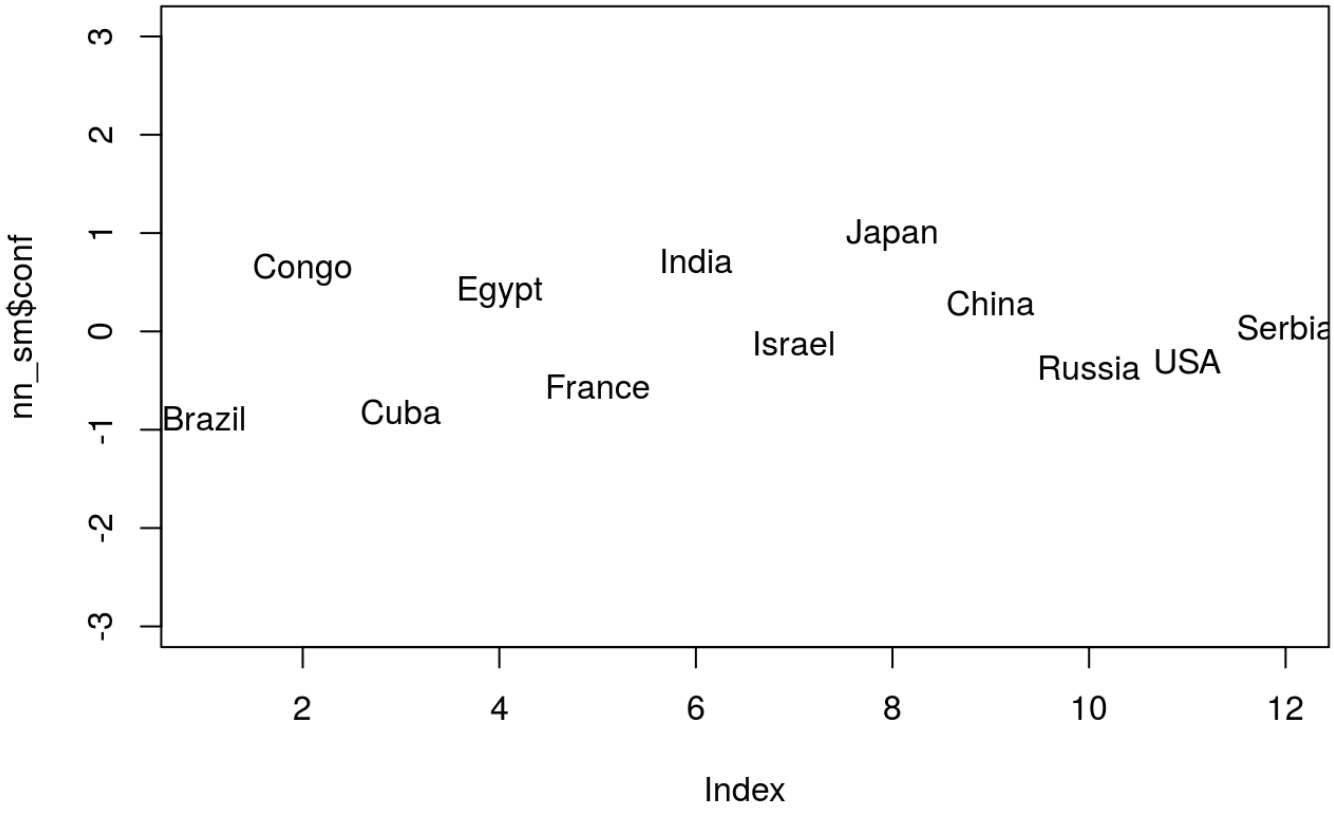

nn_sm$conf
3
2
1
0
-1
-2
-3
Israel
France
Russia
Congo
Japan
USA
China
Serbia
Egypt
Brazil
Cuba
India
2
4
6
8
10
12
Index

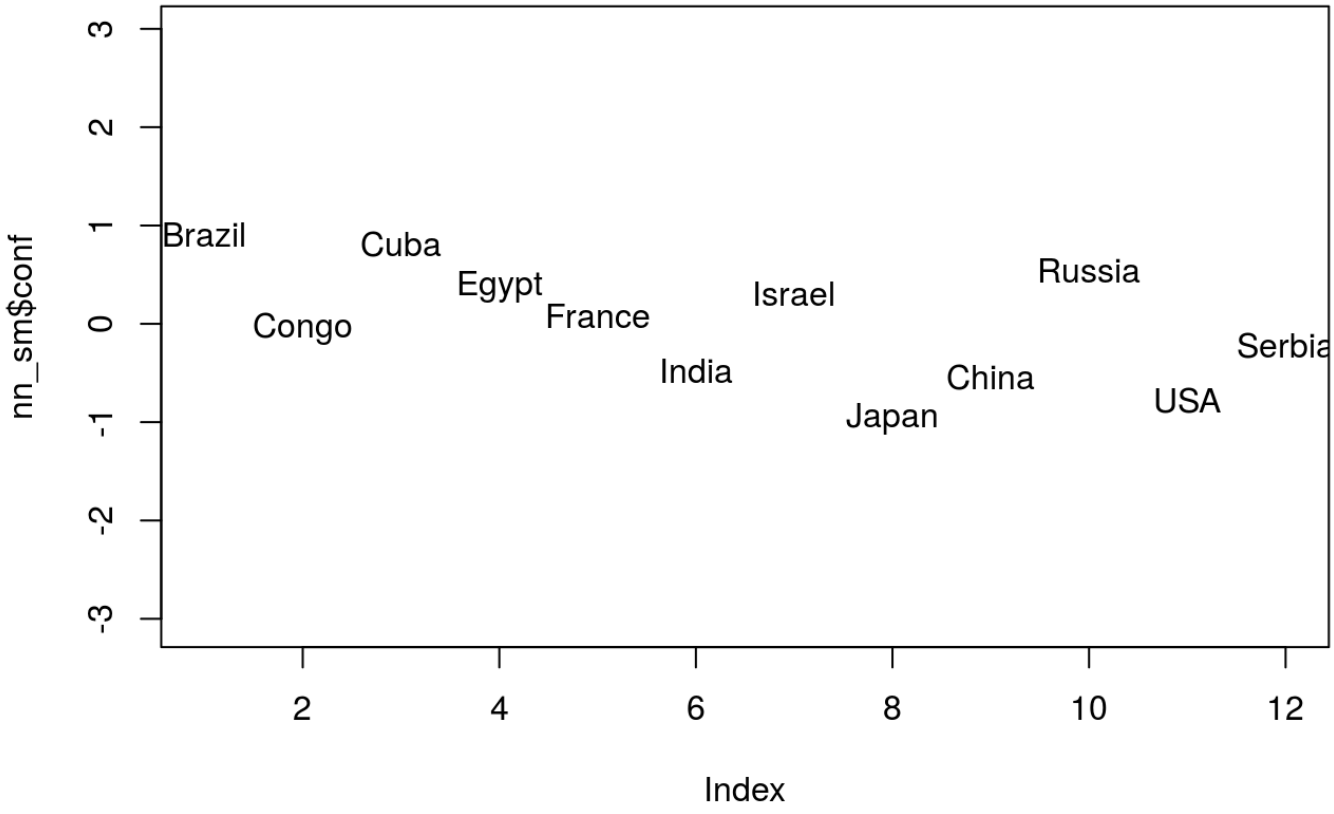

nn_sm$conf
3
2
1
0
-1
-2
-3
Russia
Egypt
Congo
USA
Japan
China
Israel
France
Serbia
Cuba
India
Brazil
2
4
6
8
10
12
Index

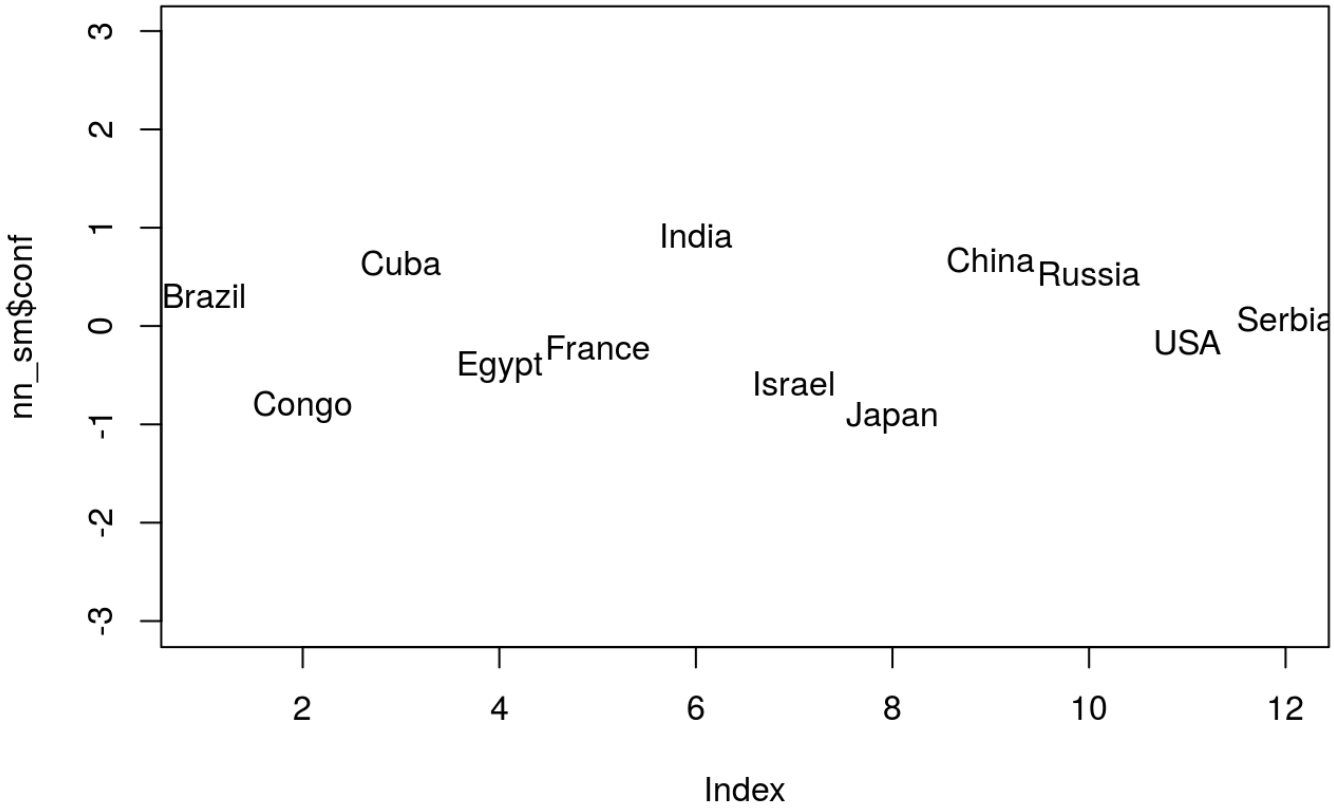

Index
nn_sm$conf
3
2
1
0
-1
-2
-3
Brazil
Egypt
France
India
Israel
Japan
China
Russia
USA
Serbia
Congo
Cuba
2
4
6
8
10
12
Index

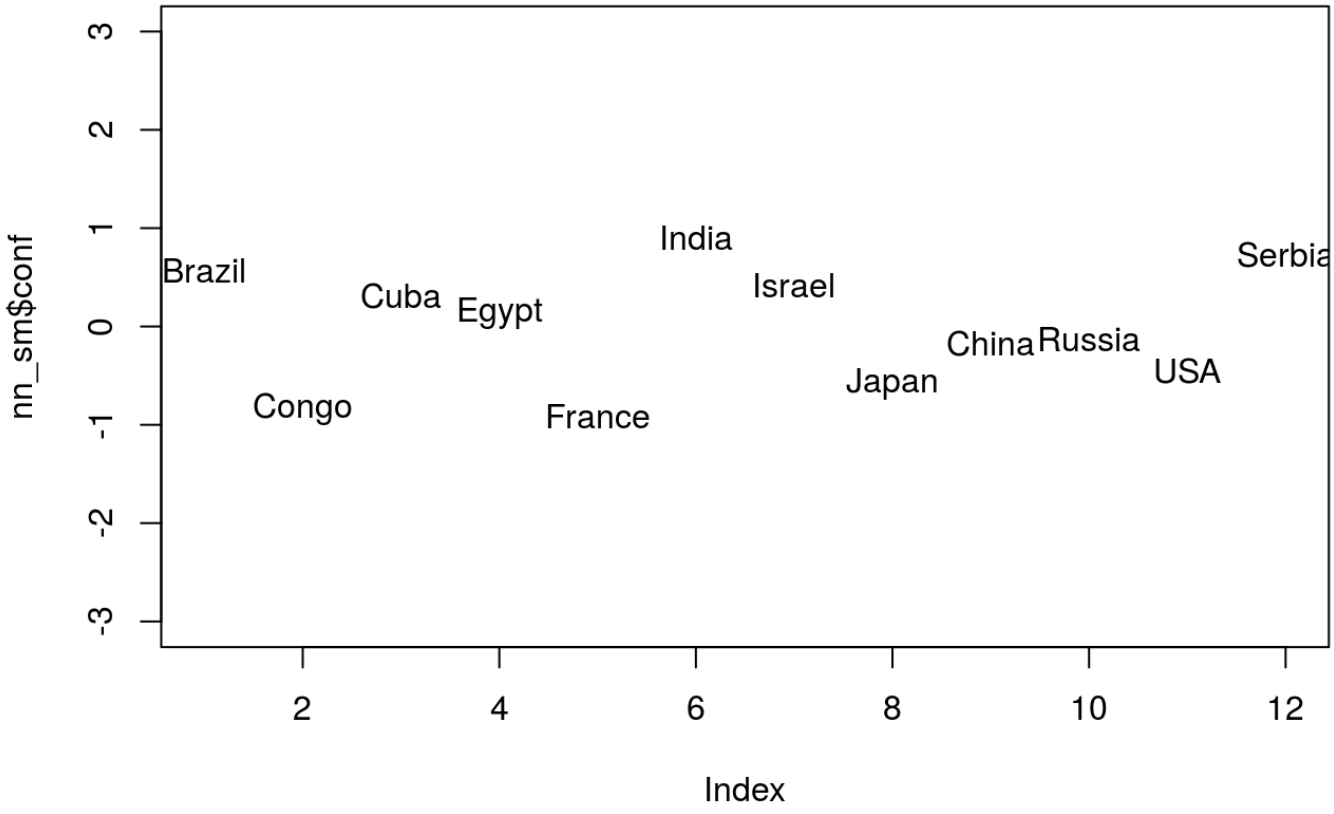

nn_sm$conf
3
2
1
0
-1
-2
-3
Congo
Egypt
France
Serbia
India
China
Russia
USA
Brazil
Cuba
Japan
Israel
2
4
6
8
10
12
Index

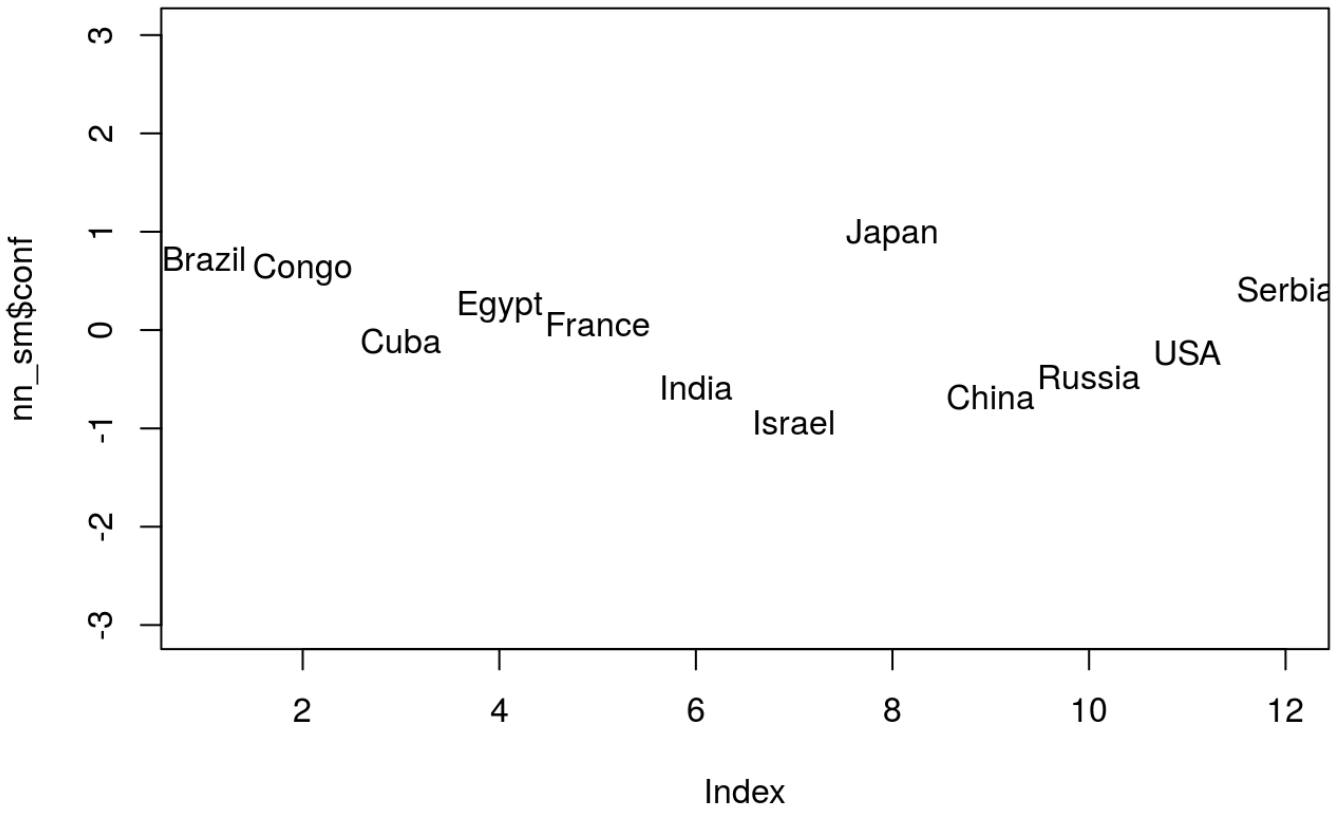

nn_sm$conf
3
2
1
0
-1
-2
-3
Brazil
Congo
Cuba
Egypt
France
India
Israel
Japan
China
Russia
Serbia
USA
2
4
6
8
10
12
Index


nn_sm$conf
3
2
1
0
-1
-2
-3
Congo
Egypt
India
Japan
China
Serbia
Israel
Russia
USA
France
Brazil
Cuba
2
4
6
8
10
12
Index


nn_sm$conf
3
2
1
0
-1
-2
-3
Brazil
Cuba
Egypt
France
Congo
India
Israel
Japan
China
Russia
USA
Serbia
2
4
6
8
10
12
Index

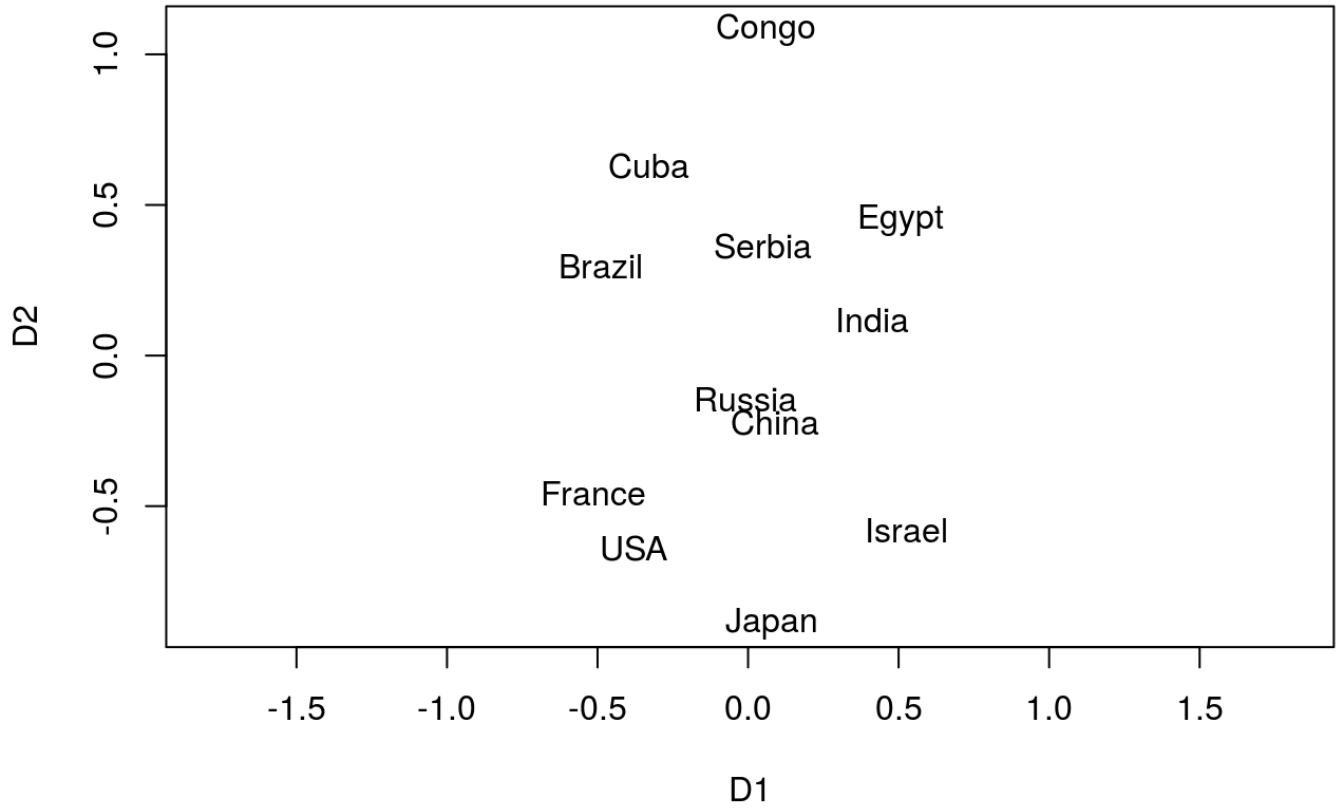

nn_sm$conf
3
2
1
0
-1
-2
-3
Brazil
Cuba
India
China
Russia
Serbia
USA
Egypt
France
Israel
Japan
Congo
2
4
6
8
10
12
Index

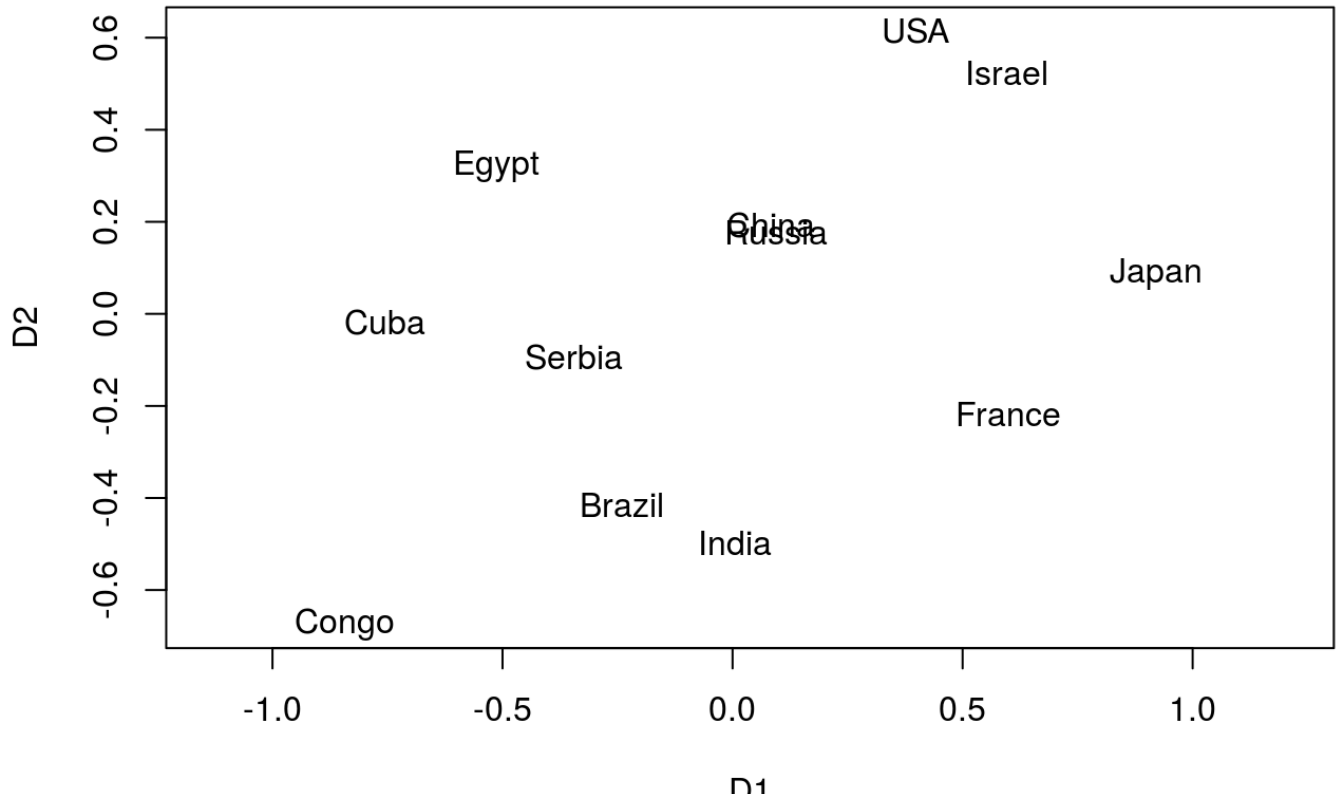

nn_sm$conf
3
2
1
0
-1
-2
-3
Brazil
Cuba
Egypt
India
Israel
Serbia
China
Russia
USA
Japan
Congo
France
2
4
6
8
10
12
Index

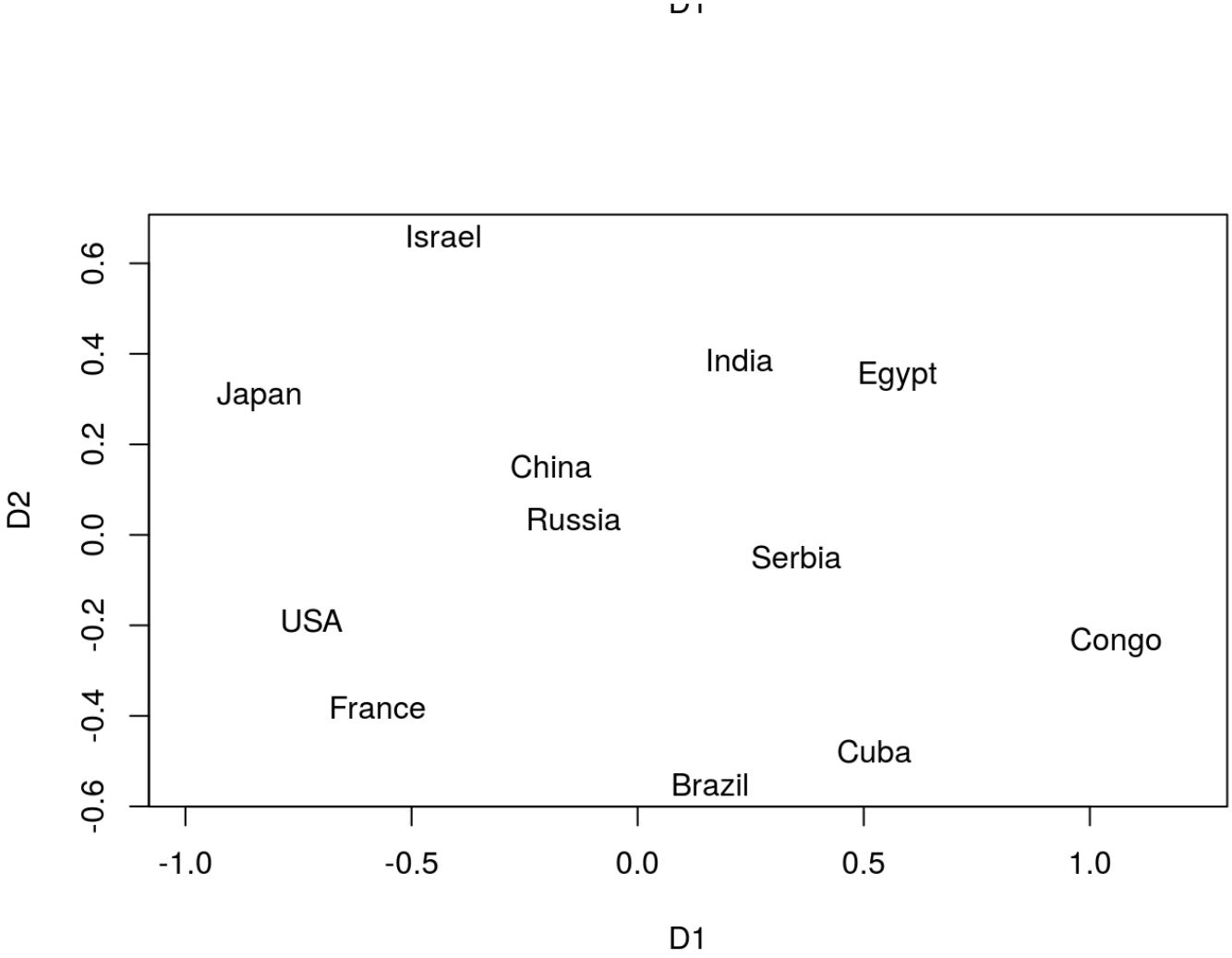


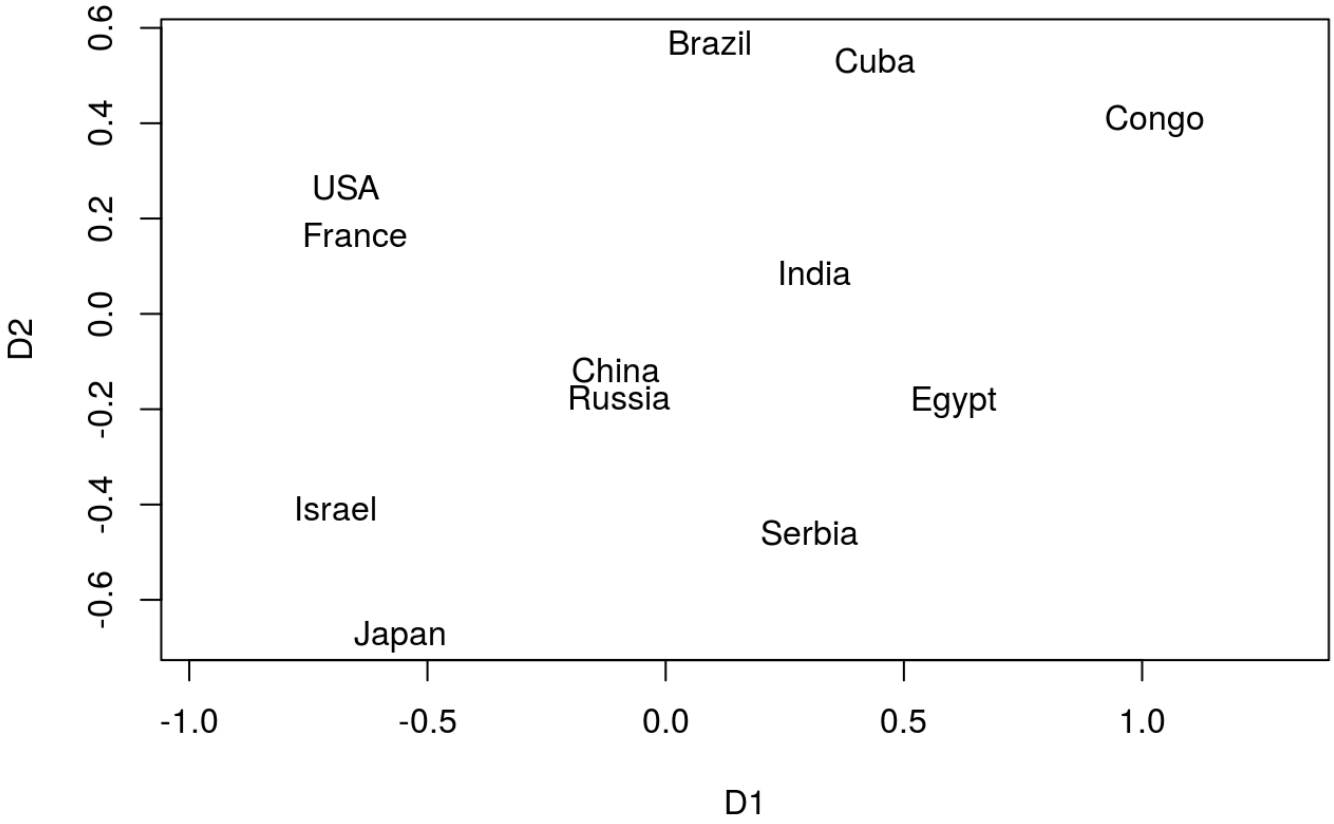
For two dimension

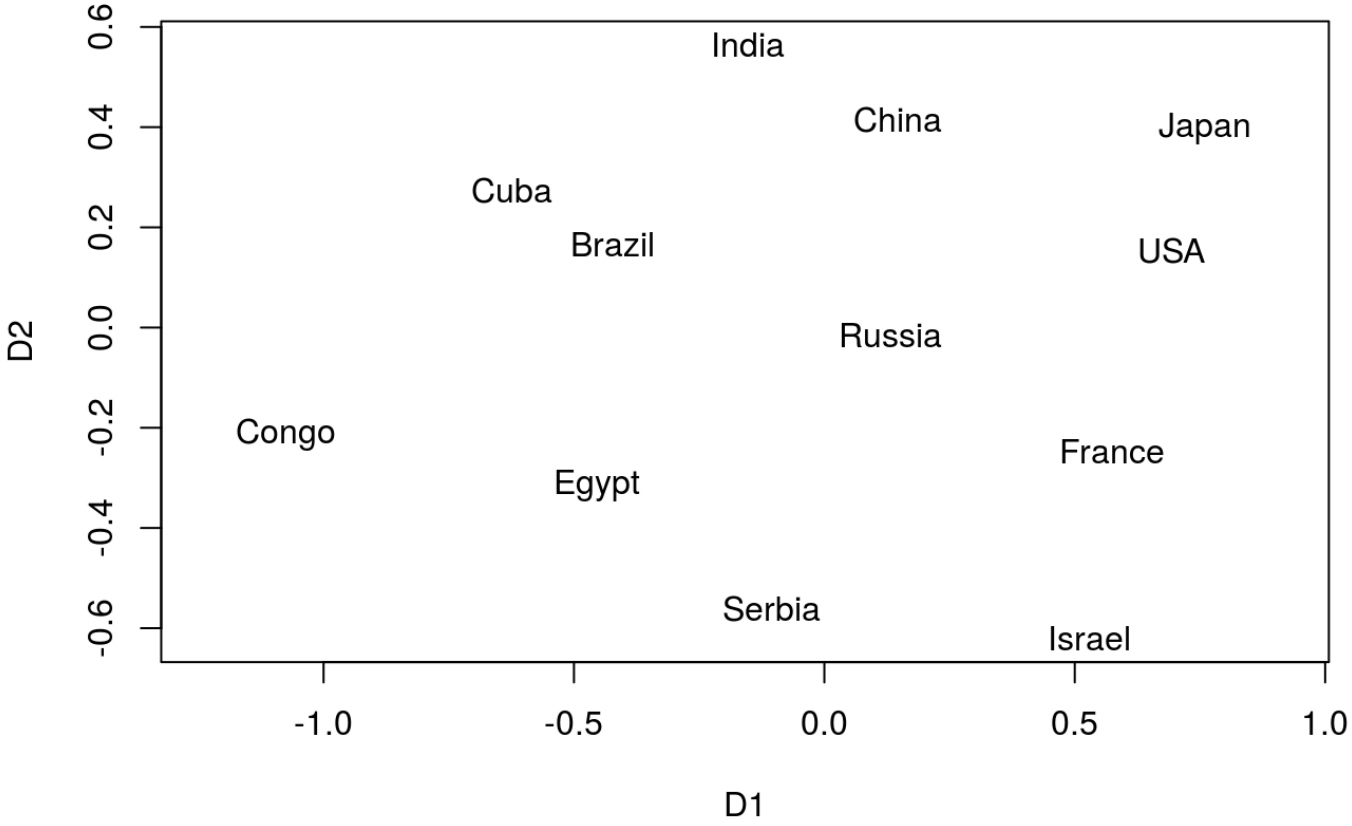
```
for (i in 1:10){
  nn_sm <- smacof::smacofSym(Dx,ndim = 2,type = "ordinal",init = "random")
  plot(nn_sm$conf,asp=1,pch=' ')
  text(nn_sm$conf,rownames(nations))
}
```

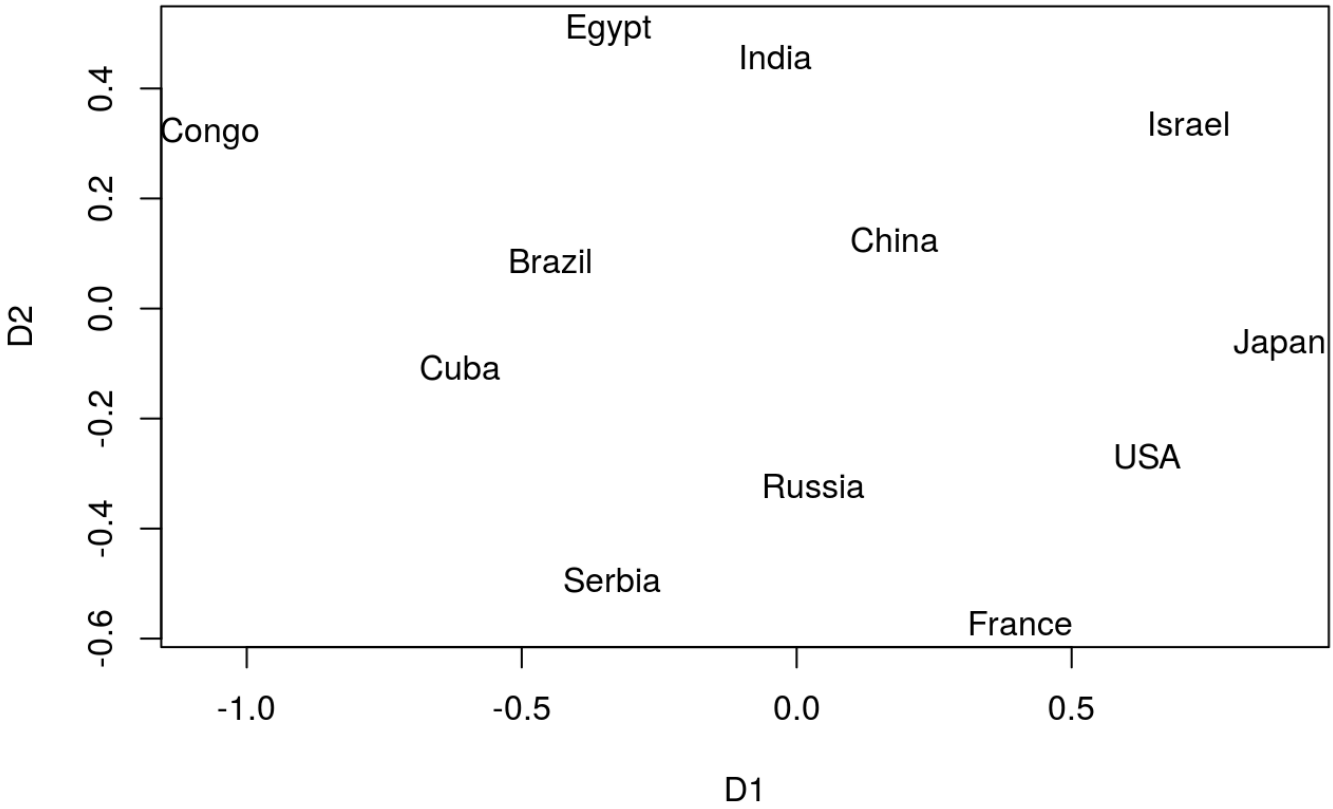

Congo
Cuba
Egypt
Serbia
Brazil
India
Russia
China
France
Israel
USA
Japan
D2
D1
1.0
0.5
0.0
-0.5
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5

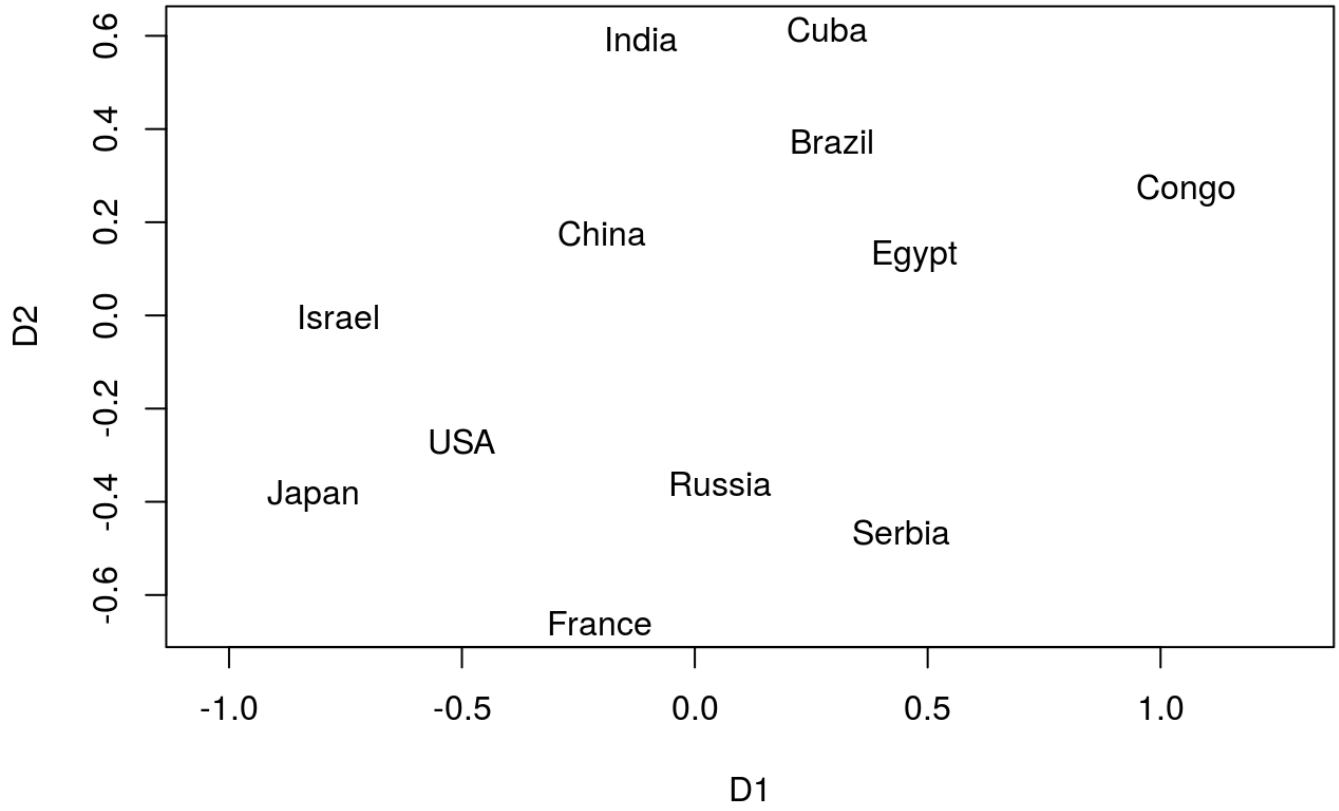

USA
Israel
Egypt
China
Russia
Japan
Cuba
Serbia
France
Brazil
India
Congo
D2
D1
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
1.0

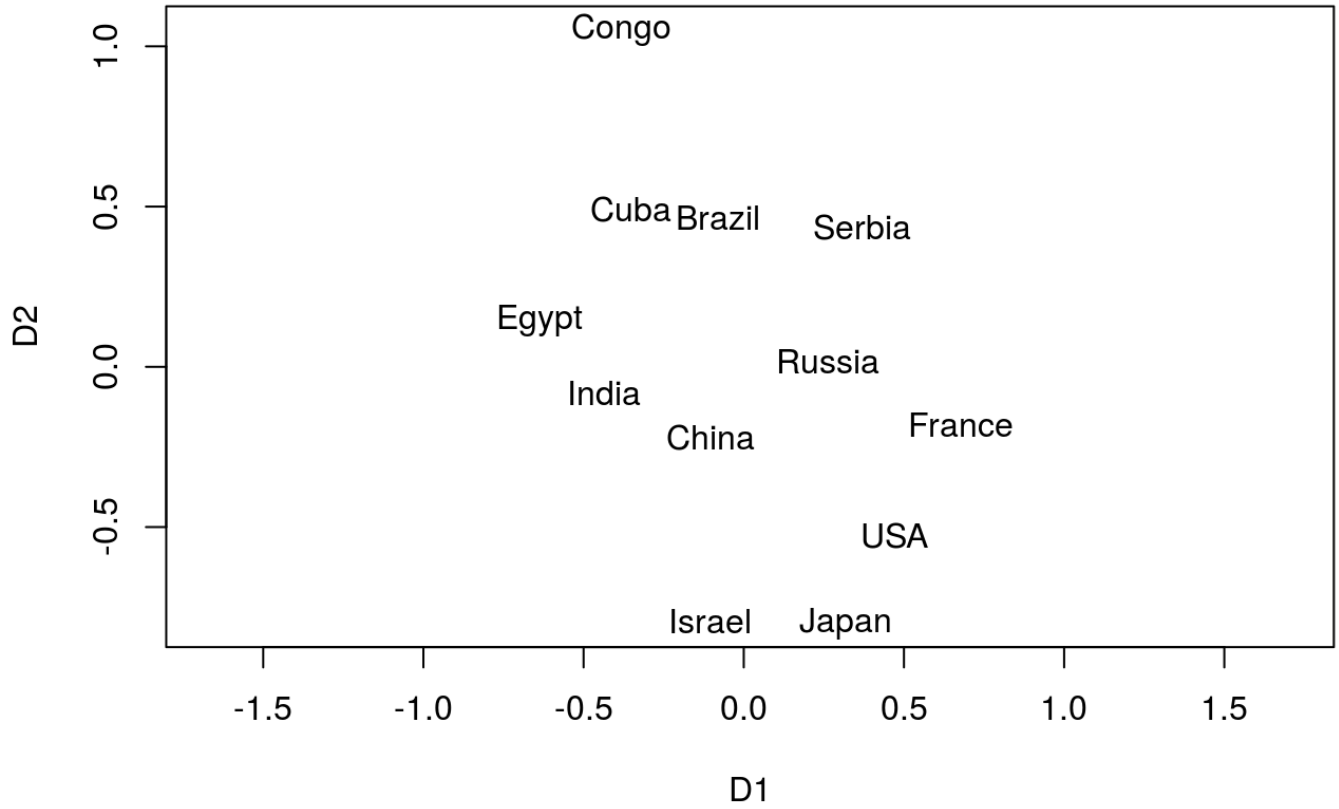

Israel
India
Egypt
Japan
China
Russia
Serbia
USA
Congo
France
Cuba
Brazil
D2
D1
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
1.0

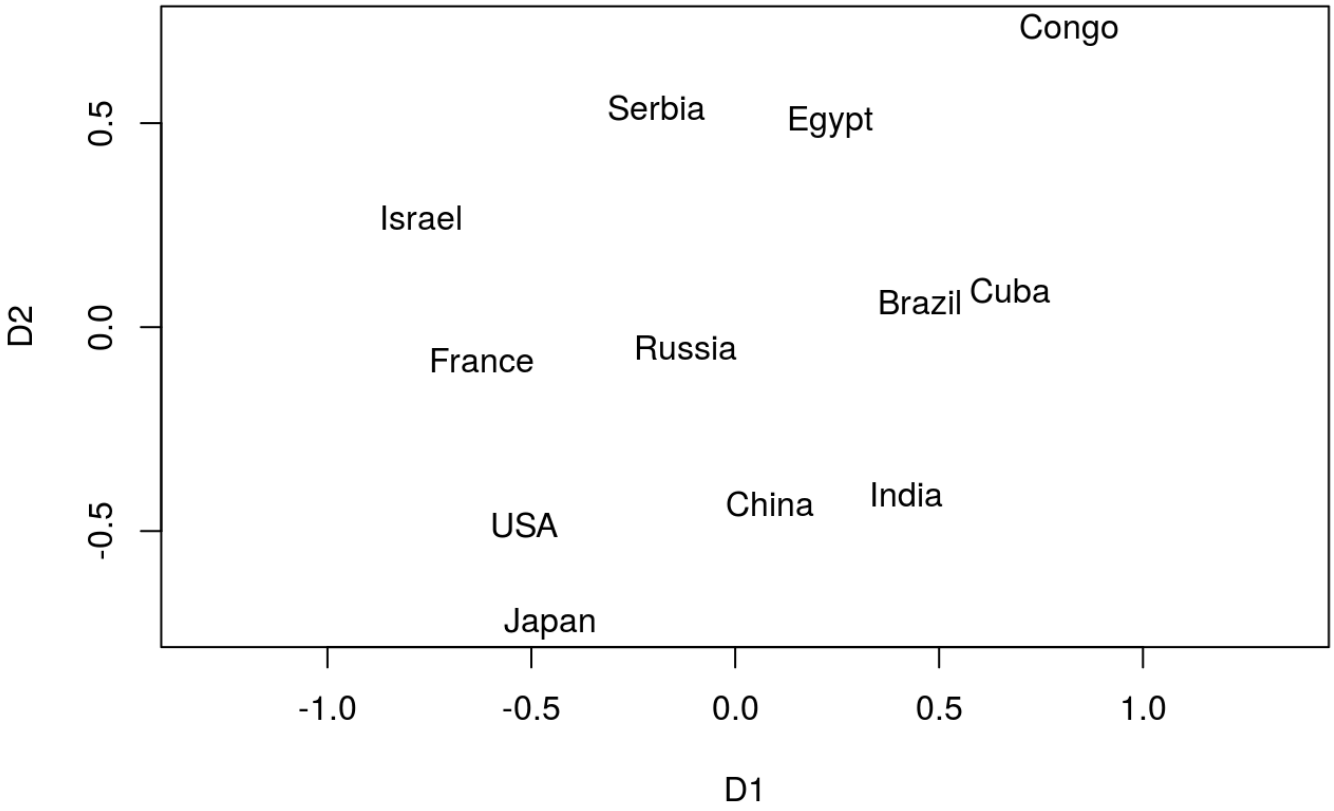

Brazil
Cuba
Congo
USA
France
India
China
Russia
Egypt
Israel
Serbia
Japan
D2
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
1.0
D1

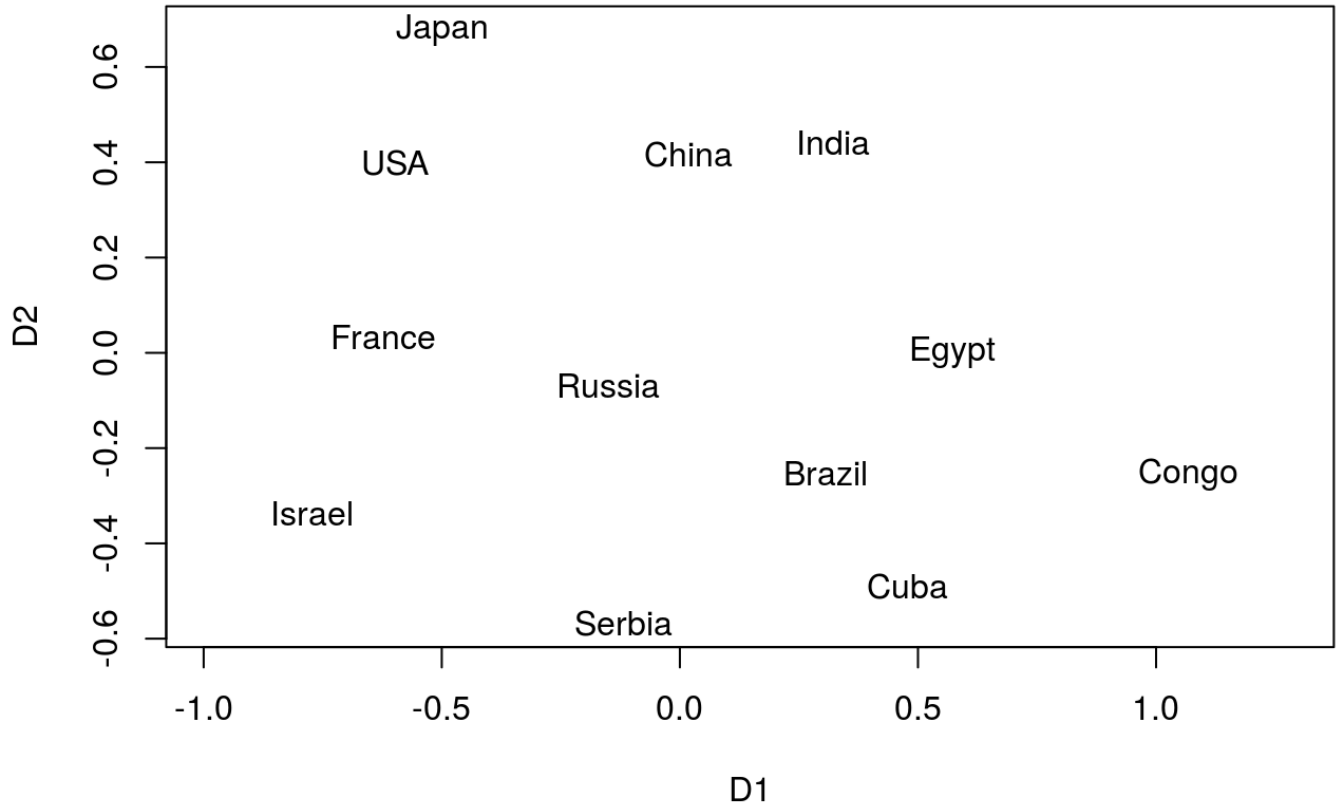

India
China
Japan
Cuba
Brazil
USA
Russia
Congo
France
Egypt
Serbia
Israel
D2
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
1.0
D1


Egypt
India
Congo
Israel
China
Brazil
Japan
Cuba
USA
Russia
Serbia
France
D2
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
D1


India
Cuba
Brazil
Congo
China
Egypt
Israel
USA
Japan
Russia
Serbia
France
D2
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.5
0.0
0.5
1.0
D1


Congo
Cuba
Brazil
Serbia
Egypt
Russia
India
China
France
USA
Israel
Japan
D2
1.0
0.5
0.0
-0.5
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
D1


Congo
Serbia
Egypt
Israel
Brazil
Cuba
France
Russia
China
India
USA
Japan
0.5
0.0
-0.5
D2
-1.0
-0.5
0.0
0.5
1.0
D1



There is almost no local minimum problem in the one dimension solution with random start. But there are obvious local mimimum problem when the solution space is two dimension.