Assignment 5: MDS of the new nations data using isoMDS & smacof

As part of a classroom demo, we collected similarity ratings of 12 nations, a semi-replication of Wish's nations data collected in the 1970's. It is not a perfect replication, because at least one of the countries Wish used as timuli (Yugoslavia) no longer exists. The file "nations_ALLDATA_1920_lowerhalf.csv" contains the averaged similarity ratings from this class and last's year's section, as a lowerhalf matrix.

1) read the new nations data into R, and convert it to type matrix (using the first colum as the rownames:

```
nnations <-
read.table("C:/Users/corter/Desktop/HUDM5124/nations_ALLDATA_1920_lowerhalf.csv",sep=",",heade
r=TRUE)
# use first column of nations as rownames
m<-ncol(nnations)
nn <- as.matrix(nnations[,2:m])
m <- m-1
rownames(nn) <- nnations[,1]
```

2) First, let's analyze the data incorrectly (using isoMDS), treating the similiarities as dissimilarities, to see what effect this has on the derived configuration and stress. Plug 0 for the missing-data value "NA", and symmetrize the matrix, as swe did in previous examples, calling the resulting symmetric similarities matrix "S". Then run isoMDS on this similarities matrix using the Euclidean distance model and ndim=2. Print the derived configuration. Comment on this configuration -- do you see any close neighbors that don't make sense? Record the value of stress-1 and comment on it (NOTE: isoMDS reports stress-1, the square-root version, multiplied by a factor of 100. See demo in file "investigate stress in isoMDS.R").

3) Now, we will re-analyze the data, correctly converting it to dissimilarities first. To convert S to dissimilarities, we candefine R code to subtract all the entries from a large constant (as in previous assignments), or we can use the smacof package utility "sim2diss" -- see the posted R manual, or try "help(sim2diss)" once you have loaded and attached to the smacof library. Try using sim2diss with and without the "to.dist=TRUE" option:

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

The commands above define a full matrix DD containing dissimilarities (subtracitng every element in S from the constant 10), and define matrix Dx as a lowerhalf matrix of dissimilarities, defining it as type *distance* (a known type of object in R).
Run isoMDS on DD or Dx. Compare the 2-D solution to the one you obtained in step 2), both in terms of the configuration and in terms of stress.

4) Now we will try using the smacof routine on the matrix Dx:

```
nn_sm<-smacofSym(Dx,ndim=2, type=ordinal)
nn_sm
# plot the 2-D solution and label the points
plot(nn_sm$conf,asp=1,pch=' ')
text(nn_sm$conf,rownames(nn))
```

Print out the 2-dim configuration, and note the stress-1 value. Compare these to the solution you obtained in step 3 using isoMDS. Comment.

5) [Optional] Perform 10 runs of smacof (ordinal MDS) on the new nations data with a RANDOM start, using the *init="random"* parameter. Do this with 1-dimensional and 2-dimensional solutions. Do you see any evidence of a local minimum problem in any of the runs?