Part 6 Multidimensional Scaling (MDS)

Natalia Levshina © 2015

Part of the course taught at the University of Mainz, Germany
26-28 May 2015

eWAVE data

http://ewave-atlas.org/

```
> library(Rling)
> data(eWAVE) #represents 76 varieties
of English and 20 different linguistic
features, as well as the region, type
and geographic coordinates
```

Recoding the data

```
> eWAVE1 <- eWAVE # create a copy of the original data set
> eWAVE1[eWAVE1 == "?"] <- NA
> eWAVE1[eWAVE1 == "X"] <- NA
> eWAVE2 <- lapply(eWAVE1[, 1:20], function(x) ordered(x, levels = c("D", "C", "B", "A")))
> eWAVE2 <- data.frame(eWAVE2)
```

Computing the Gower distances

```
> library(cluster)#install first!
> ling.dist <- daisy(eWAVE2) # equivalent to daisy(eWAVE2, metric = "gower")
> summary(ling.dist)
2850 dissimilarities, summarized :
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
0.00000 0.28070 0.35556 0.37555 0.45000 1.00000       8
Metric :  mixed ;  Types = O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O, O
Number of objects : 76
```

Replacing the missing values and 0

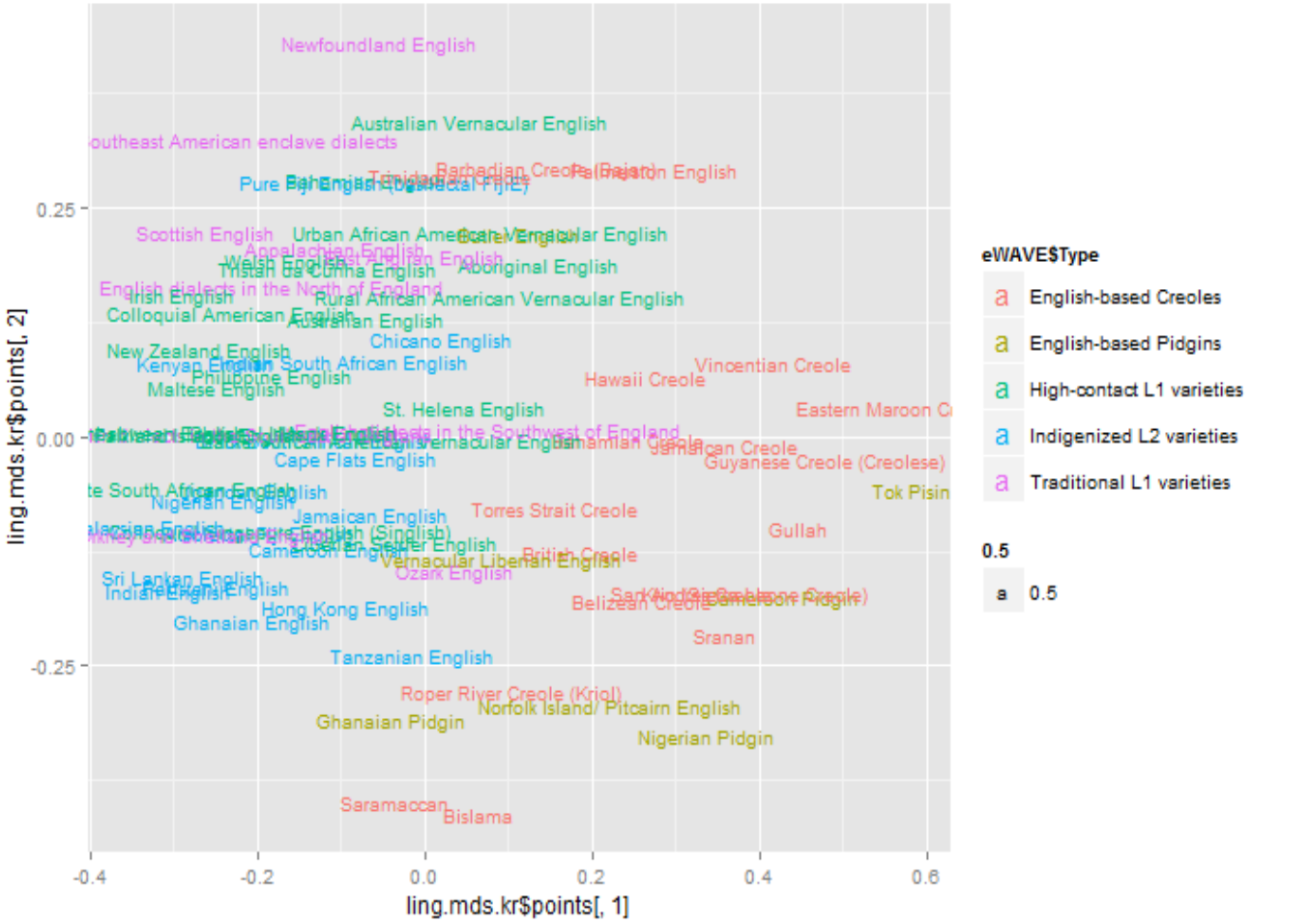
```
> ling.dist1 <- ling.dist
> ling.dist1[is.na(ling.dist)] <- mean(ling.dist, na.rm = TRUE)# replace the missing values with average distances
> ling.dist2 <-ling.dist1
> ling.dist2[ling.dist2 == 0] <- 0.001 # replace 0 with very small distances, necessary for isoMDS()
```

Performing non-metric MDS

```
> library(MASS)
> ling.mds.kr <- isoMDS(ling.dist2)
#equivalent to isoMDS(ling.dist2, k = 2)
initial  value 32.470515
iter   5 value 24.586167
final  value 24.416368
converged
```

Creating the MDS map

```
> library(ggplot2) #install first!
> qplot(ling.mds.kr$points[, 1],
ling.mds.kr$points[, 2], label =
rownames(eWAVE), cex = 0.5, col =
eWAVE$Type, geom = "text")
```


Newfoundland English
Australian Vernacular English
Scottish English
Urban African American Vernacular English
Aboriginal English
Rural African American Vernacular English
Colloquial American English
Chicano English
New Zealand English
Vincentian Creole
Hawaii Creole
Maltese English
St. Helena English
Cape Flats English
Jamaican English
Torres Strait Creole
Gullah
Guyanese Creole (Creolese)
Tok Pisin
Nigerian English
Sri Lankan English
Hong Kong English
Ghanaian English
Tanzanian English
Sranan
Roper River Creole (Kriol)
Norfolk Island/ Pitcairn English
Ghanaian Pidgin
Nigerian Pidgin
Saramaccan
Bislama
0.25
0.00
-0.25
-0.4
-0.2
0.0
0.2
0.4
0.6
ling.mds.kr$points[, 2]
ling.mds.kr$points[, 1]
eWAVE$Type
a English-based Creoles
a English-based Pidgins
a High-contact L1 varieties
a Indigenized L2 varieties
a Traditional L1 varieties
0.5
a 0.5