Homework 4: Multivariate Models

Caroline Oliver

2/13/2018

```r
library(vegan)
```

```
## Warning: package 'vegan' was built under R version 3.2.5
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.4-0
```

```r
data(dune)
data(dune.env)
?dune
```

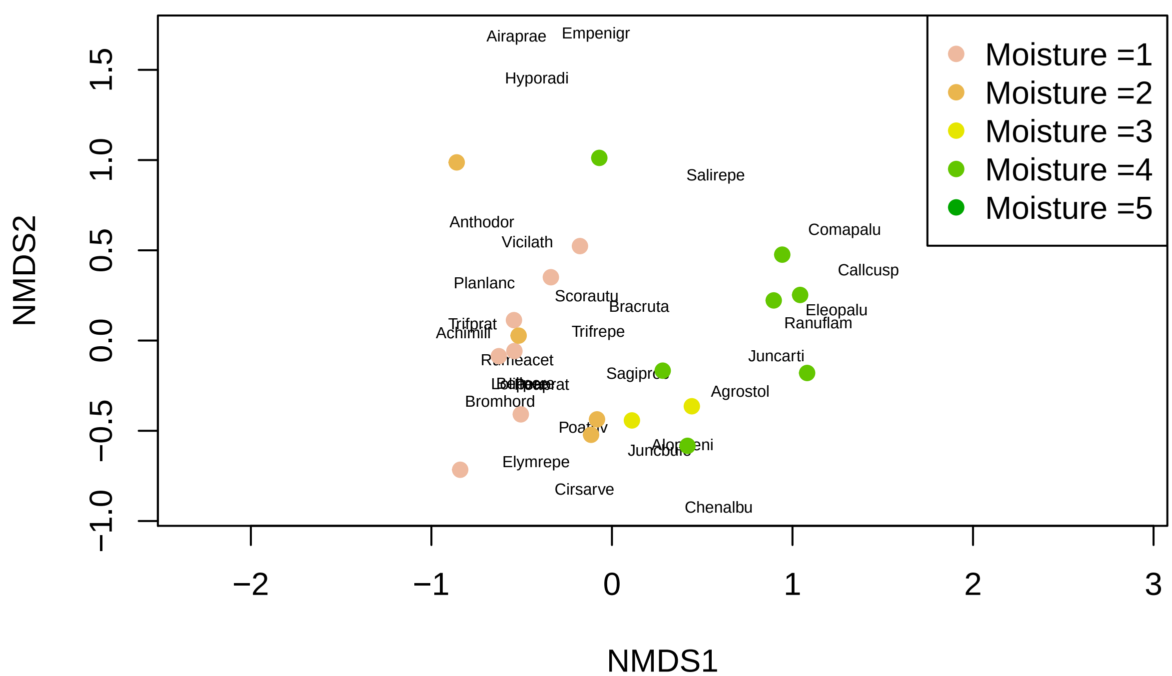
Question 1

```r
dune_mds = metaMDS(dune, distance = "bray")
```

```
## Run 0 stress 0.1192678
## Run 1 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 0.02026829  max resid 0.06494856
## Run 2 stress 0.1192678
## Run 3 stress 0.1889648
## Run 4 stress 0.1192683
## Run 5 stress 0.1886532
## Run 6 stress 0.1192685
## Run 7 stress 0.1183186
## ... Procrustes: rmse 2.306422e-05  max resid 7.289111e-05
## ... Similar to previous best
## Run 8 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 5.924691e-06  max resid 1.232806e-05
## ... Similar to previous best
## Run 9 stress 0.1192678
## Run 10 stress 0.1192688
## Run 11 stress 0.1808911
## Run 12 stress 0.1183186
## ... Procrustes: rmse 3.925527e-05  max resid 0.0001254755
## ... Similar to previous best
## Run 13 stress 0.2035424
## Run 14 stress 0.1192681
## Run 15 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 9.368676e-06  max resid 3.232005e-05
## ... Similar to previous best
## Run 16 stress 0.1183186
```

```
## ... Procrustes: rmse 5.092769e-05  max resid 0.0001679102
## ... Similar to previous best
## Run 17 stress 0.1192687
## Run 18 stress 0.2086875
## Run 19 stress 0.1192679
## Run 20 stress 0.1192678
## *** Solution reached
```

```r
plot(dune_mds, type='n')
text(dune_mds, 'sp', cex=.5)
# generate vector of colors
color_vect = rev(terrain.colors(6))[-1]
points(dune_mds, 'sites', pch=19, col=color_vect[dune.env$Moisture])
legend('topright', paste("Moisture =", 1:5, sep=''), col=color_vect, pch=19)
```



Describe how you interpret the graphic

The graph represents the various moisture levels and the plants that rank near each of the various moisture levels. If a plant species is near a moisture marker or if a group of species is located within or near a group of similar/same moisture markers than it is logical to assume that those specie(s) require that level of moisture to survive in the dunes.

What is the goal of creating such a plot?

The MDS plot helps to visualize the distribution of the moisture and which species the different moisture values correlate to. This helps us to know what the best growing conditions are for a certain species and which species possibly will work well together in the same space/grow together in nature.

Does this analysis suggest any interesting findings with respect to the dune vegetation?

The plot shows mostly clumps or groups of species near various moisture measures, however, 3 species (Airaprae, Empenigr, and Hyporadi) are located at the top, center of the plot not near any moisture level markers. Salirepe also has a greater distance between the species marker and a moisture marker. These 4 species have the most distance between species markers and moisture markers, while most other species are in closer proximity to a moisture marker. This could mean that these plants require little to no moisture compared to the other species in the dataset.

Question 2

Set Up

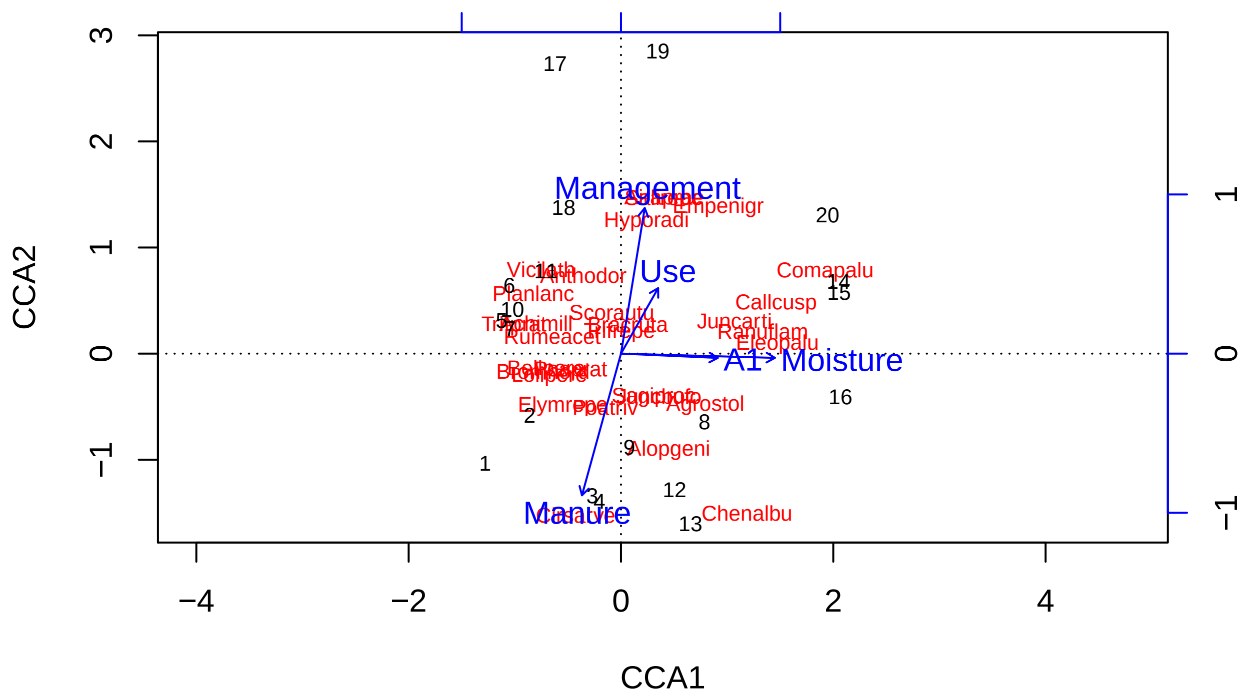
```r
dune.env$Moisture = as.numeric(dune.env$Moisture)
dune.env$Manure = as.numeric(dune.env$Manure)

# SF = 1, BF = 2, HF = 3, NM = 4
dune.env$Management = as.numeric(factor(dune.env$Management , levels=c("SF", "BF", "HF", "NM")))

# Haypastu = 1, Hayfield = 2, Pasture = 3
dune.env$Use = as.numeric(factor(dune.env$Use , levels=c("Haypastu", "Hayfield", "Pasture")))
```

```r
dune_cca = cca(dune, dune.env)
dune_cca
```

```
## Call: cca(X = dune, Y = dune.env)
##
##               Inertia Proportion Rank
## Total          2.1153     1.0000
## Constrained    0.9200     0.4349    5
## Unconstrained  1.1953     0.5651   14
## Inertia is mean squared contingency coefficient
##
## Eigenvalues for constrained axes:
##   CCA1   CCA2   CCA3   CCA4   CCA5
## 0.4305 0.2607 0.1350 0.0597 0.0340
##
```

```
## Eigenvalues for unconstrained axes:
##     CA1     CA2     CA3     CA4     CA5     CA6     CA7     CA8     CA9
## 0.30147 0.18494 0.14033 0.13423 0.08590 0.07420 0.07120 0.06259 0.04629
##    CA10    CA11    CA12    CA13    CA14
## 0.03898 0.02276 0.01715 0.01068 0.00456
```

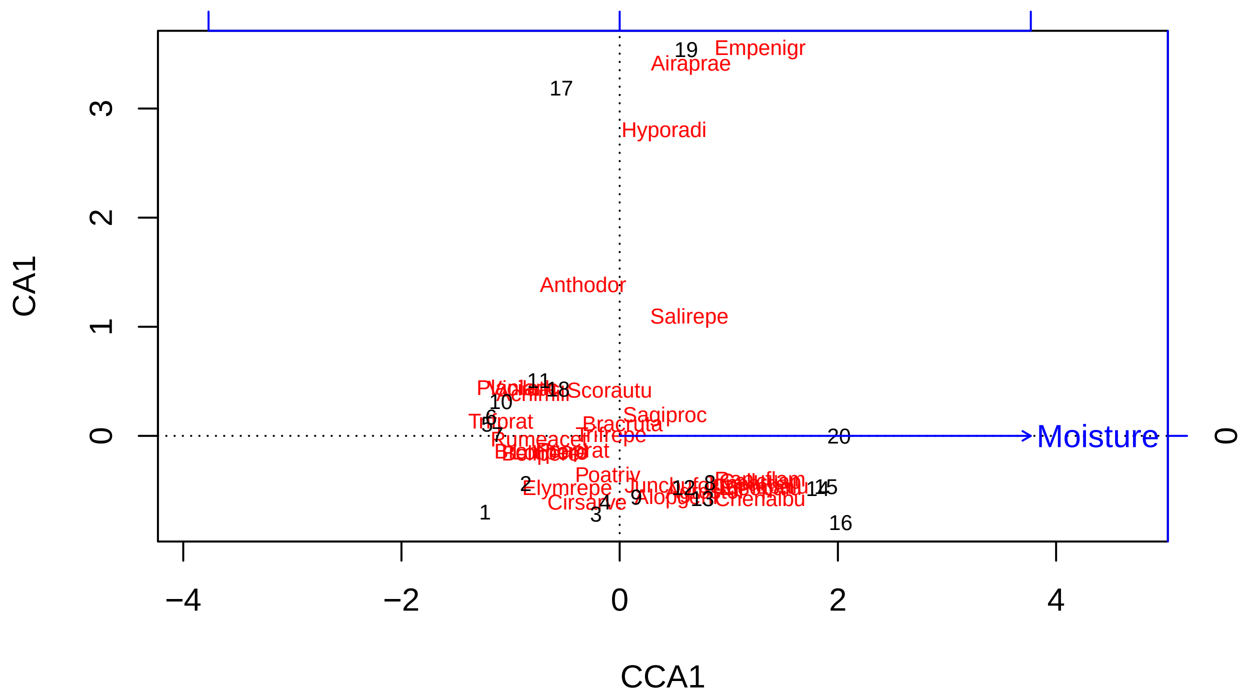
```
plot(dune_cca)
```




```
dune_cca2 = cca(dune ~ Moisture, dune.env)
dune_cca2
```

```
## Call: cca(formula = dune ~ Moisture, data = dune.env)
##
##               Inertia Proportion Rank
## Total          2.1153     1.0000
## Constrained    0.4108     0.1942    1
## Unconstrained  1.7045     0.8058   18
## Inertia is mean squared contingency coefficient
##
## Eigenvalues for constrained axes:
##   CCA1
## 0.4108
##
## Eigenvalues for unconstrained axes:
##    CA1    CA2    CA3    CA4    CA5    CA6    CA7    CA8
## 0.4180 0.3148 0.1835 0.1510 0.1082 0.0925 0.0878 0.0737
## (Showed only 8 of all 18 unconstrained eigenvalues)
```

```r
plot(dune_cca2)
```



For the second model that I did I only used Moisture as a predictor so that I could compare it better to the graph in the first question. To note, similar species that were farther away from the moisture markers in the first plot are stand out species in this plot as well, along with another, Anthodor.

Question 3

Do your two analyses agree with one another or complement one another or do these two analyses seem to be suggesting different take home messages?

For the initial model in Question 2 where all predictors are used, it shows that (obviously) moisture is not the only thing that effects the various species nor is it the most prominent deciding variable. The 3 species that seemed to be farthest away from the moisture markers in the first graph are more integrated with the species markers in this graph most likely because the other variables were used. However, like I stated above in question 2, the cca model only using moisture as a predictor has a very similar interpretation as the graph from question 1.

Which analysis do you find to be more useful?

Personally I found the NMDS model more useful. It has a clearner looking graph and handles more varaition than the CCA model.