Multivariate Models

Morgan Will

2/8/2022

For this assignment will be analyzing data on the Vegetation and Environment in Dutch Dune Meadows.

To import the data and read the metadata run the following:

```
library(vegan)

## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.5-7

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

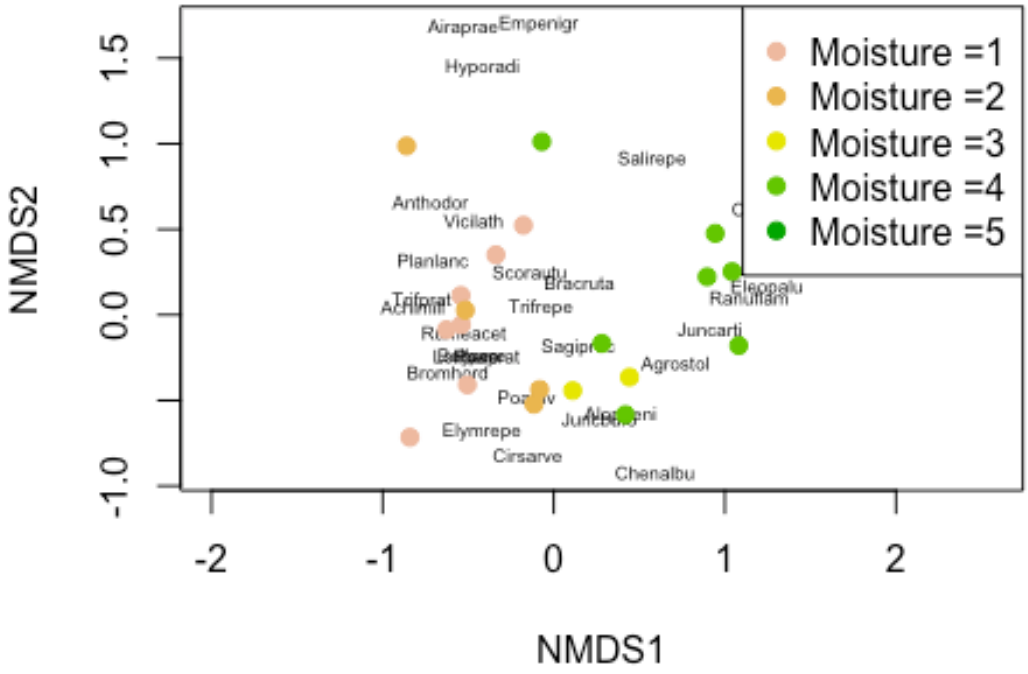
1. Conduct an indirect ordination on the dune plant community. Specifically, visually examine a NMDS plot using the bray-curtis distance metric. Below is some code to help you develop a potential plot that emphasizes the role of the environmental variable "Moisture".

```
dune_mds = metaMDS(dune)

## Run 0 stress 0.1192678
## Run 1 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 0.0202693  max resid 0.06495736
## Run 2 stress 0.1192679
## Run 3 stress 0.1886532
## Run 4 stress 0.1183186
## ... Procrustes: rmse 2.060329e-05  max resid 6.744392e-05
## ... Similar to previous best
## Run 5 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 4.133276e-06  max resid 1.413106e-05
## ... Similar to previous best
## Run 6 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 3.753197e-06  max resid 1.202564e-05
## ... Similar to previous best
## Run 7 stress 0.1183186
## ... Procrustes: rmse 8.207059e-06  max resid 2.711828e-05
## ... Similar to previous best
```

```
## Run 8 stress 0.1183186
## ... Procrustes: rmse 2.659901e-06  max resid 7.159281e-06
## ... Similar to previous best
## Run 9 stress 0.1183186
## ... Procrustes: rmse 6.035697e-06  max resid 1.689058e-05
## ... Similar to previous best
## Run 10 stress 0.1205054
## Run 11 stress 0.1183186
## ... Procrustes: rmse 2.948913e-06  max resid 1.008585e-05
## ... Similar to previous best
## Run 12 stress 0.1183186
## ... Procrustes: rmse 3.481541e-06  max resid 1.211545e-05
## ... Similar to previous best
## Run 13 stress 0.1192678
## Run 14 stress 0.1192679
## Run 15 stress 0.1192678
## Run 16 stress 0.1192678
## Run 17 stress 0.1192678
## Run 18 stress 0.1809577
## Run 19 stress 0.1183186
## ... Procrustes: rmse 2.923253e-06  max resid 8.873056e-06
## ... Similar to previous best
## Run 20 stress 0.1808911
## *** Solution reached

plot(dune_mds, type='n')
text(dune_mds, 'sp', cex=0.5)
# 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 NMDS plot displays a point for each site being evaluated. The farther apart any two sites are, the larger the variation in plant diversity is between the two sites. These sites were then colored to show the moisture value associated with each. The coloration shows that sites with higher moisture values tend to clump together (have relatively similar plant diversity), while sites with lower moisture values are also loosely grouped. Though some pattern is visible, the groupings are rather lose and have overlap, indicating that the effect of moisture on plant diversity may not be very strong, or some other variable is affecting the spread.

What is the goal of creating such a plot?

NMDS plots are indirect ordinations, so the purpose is to look for potential patterns in the data. In the above NMDS plot, the spread of the points is determined by variability in plant diversity between sites (the environmental variables have no effect on the spread). These points were then colored to show the moisture values associated with each, to see if this variable may help explain the spread displayed. Overall, this plot is meant to explore the spread of the diversity data, and look for potential variables that may help explain the variability in plant diversity between sites.

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

This plot suggests that moisture may explain some of the variability in plant diversity seen between sites. Further investigation is needed to verify this effect.

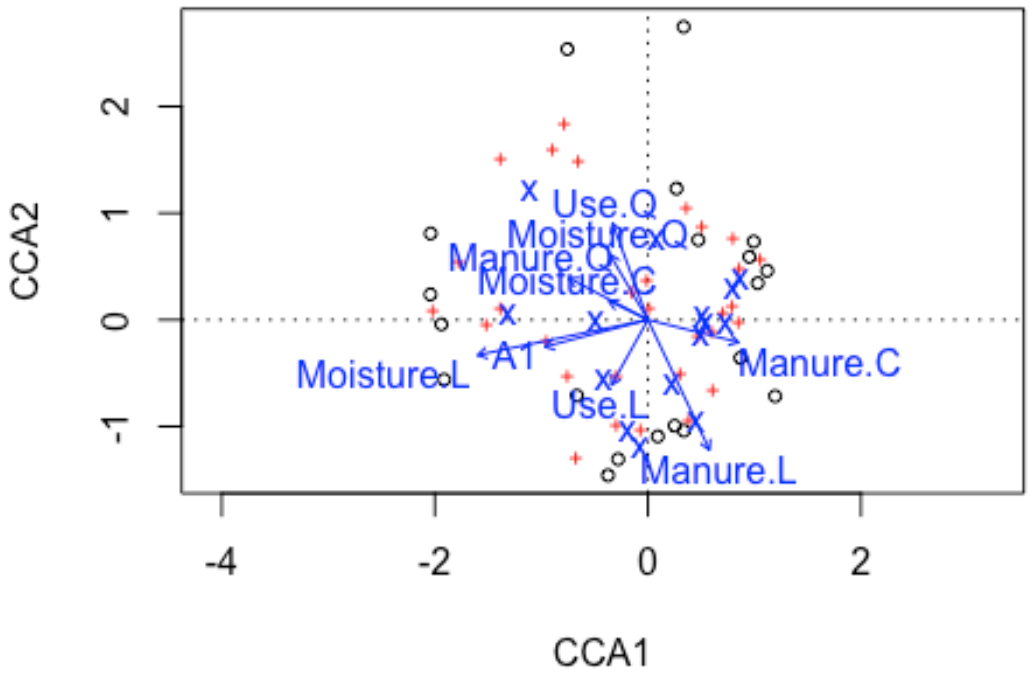
2. Carry out a direct ordination using CCA in order to test any potential hypotheses that you developed after examining the MDS plot. Specifically, carry out a test of the entire model (i.e., including all constrained axes) and also carry out tests at the scale of individual explanatory variables you included in your model if you included more than one variable. Plot your results.

```
library(car)

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.0.5

#Create and plot model
dune_cca = cca(dune ~ ., data=dune.env)
ordiplot(dune_cca)
```




```
#Test of entire model (all constrained axes)
anova(dune_cca, permutations = 999)
```

```
## Permutation test for cca under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ A1 + Moisture + Management + Use + Manure, dat
a = dune.env)
##          Df ChiSquare      F Pr(>F)
## Model    12    1.5032 1.4325  0.034 *
## Residual  7    0.6121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Test of individual explanatory variables
anova(dune_cca, by='margin', permutations = 999)
```

```
## Permutation test for cca under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ A1 + Moisture + Management + Use + Manure, dat
a = dune.env)
##            Df ChiSquare      F Pr(>F)
## A1          1   0.11070 1.2660  0.230
## Moisture    3   0.31587 1.2041  0.229
## Management  2   0.15882 0.9081  0.566
## Use         2   0.13010 0.7439  0.781
## Manure      3   0.25490 0.9717  0.520
## Residual    7   0.61210
```

Interpretation:

CCAs are designed to explain the spread of data along individual axes, each representing a potential explanatory variable. The above CCA shows the spread of the data (the variability in plant diversity between sites), with explanatory variables shown as blue arrows. The length of the arrow represents the strength of the variable in terms of explaining differences in site variability. Though many variables are represented on the plot, none are overly evident as driving variables behind the variability in diversity (all are of similar length, arranged in different directions). This is supported by the anova, with none of the variables having significant values ("best descriptor" of variability was A1 (thickness of soil A1 horizon), at p=0.245).

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? Which analysis do you find to be more useful?**

These analyses both show that moisture MAY be affecting variability in plant diversity. The NMDS plot shows a stronger relationship between moisture and diversity than the CCA, which may be due to covariability between moisture and A1 (NMDS only shows moisture, while CCA takes both into account). Overall, these analyses show that there may be a relationship between moisture and plant diversity, but more analysis needs to be done to tease out other potential factors involved.