HW5

Mitch Oshima

February 10, 2016

Load in data from vegan package.

```
library(vegan)
```

```
## Warning: package 'vegan' was built under R version 3.1.3
```

```
## Loading required package: permute
```

```
## Warning: package 'permute' was built under R version 3.1.3
```

```
## Loading required package: lattice
```

```
## Warning: package 'lattice' was built under R version 3.1.3
```

```
## This is vegan 2.3-3
```

```
library(dummies)
```

```
## dummies-1.5.6 provided by Decision Patterns
```

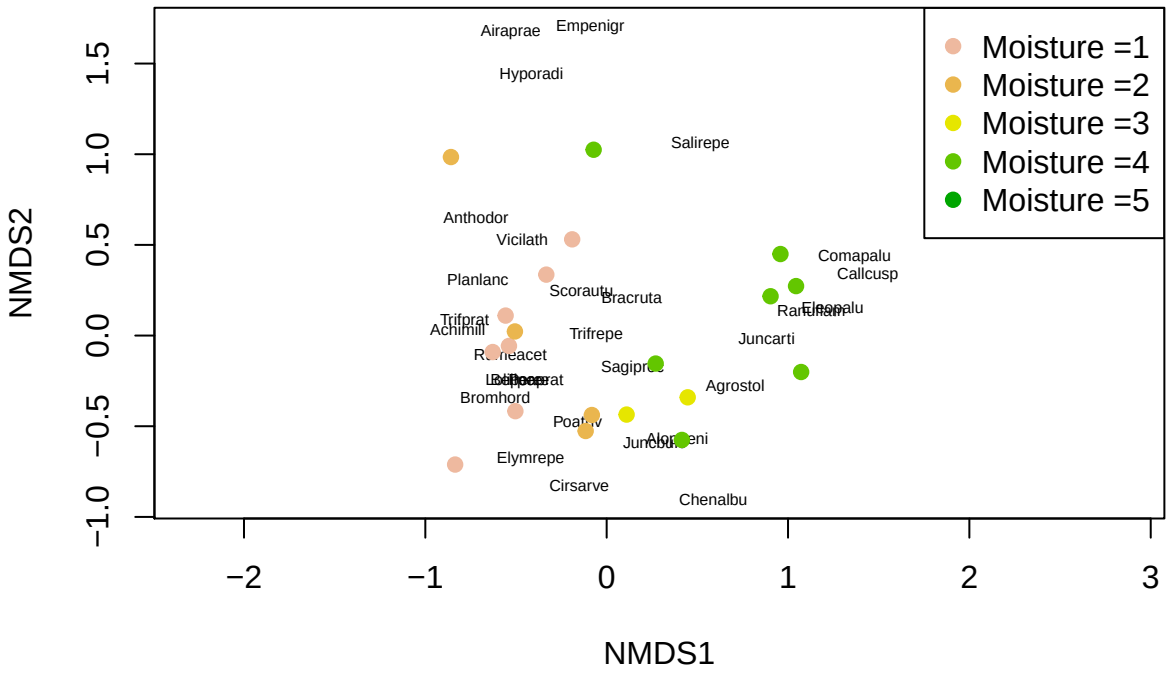
```
data(dune)
data(dune.env)
?dune
```

1.) This type of analysis allows us to view different variables by which datasets may be clustered or correlated. After running the NMDS analysis and plotting the results we can see that the species seem to be loosely clustered by levels of moisture.

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

```
## Run 0 stress 0.1192678
## Run 1 stress 0.1192679
## ... procrustes: rmse 0.0001269712  max resid 0.000383958
## *** Solution reached
```

```
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)
```

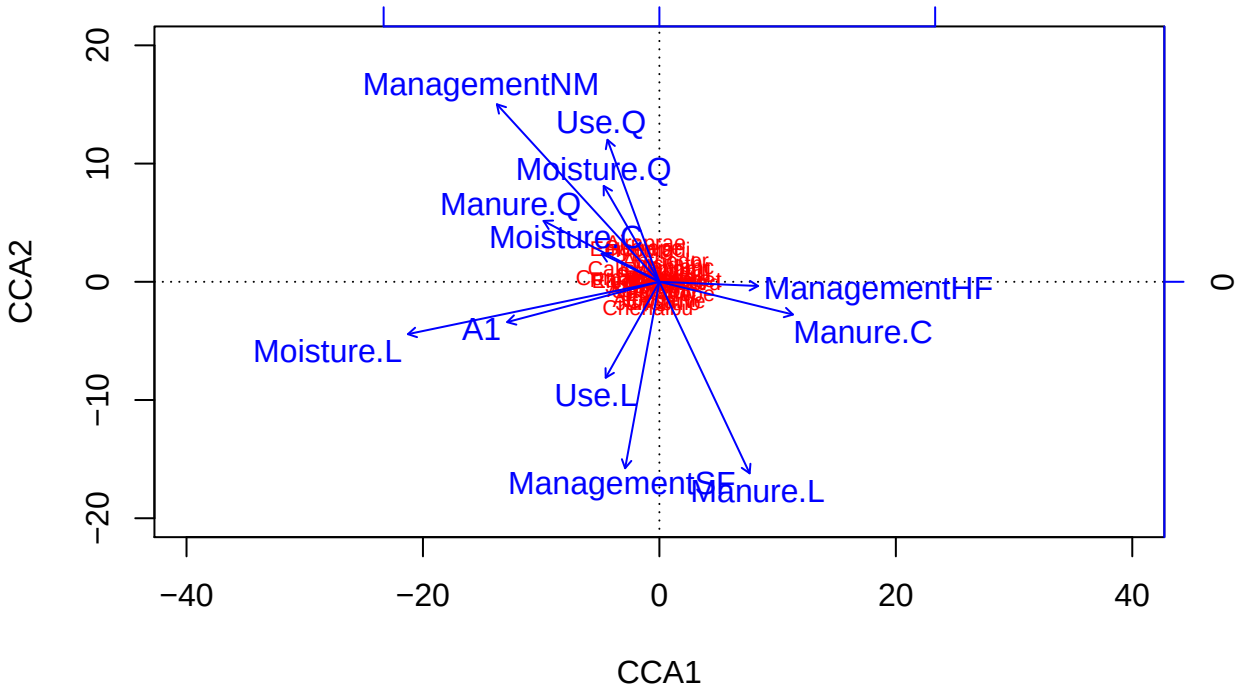


2.) Using the cca function from the vegan package I performed a canonical correspondence analysis of the dune data as explained by all of the environmental data. The cca test resulted in a total inertia of 2.1153. The constrained axes acounted for about 1.5 of the total while the unconstrained axes added .61. The plot for this analysis did not show a particularly strong relationship within the variables. Using an ANOVA test I found that moisture and manure have the highest

```
dune_ca1 = cca(dune ~ A1 + Moisture + Management + Use + Manure, dune.env)
dune_ca1
```

```
## Call: cca(formula = dune ~ A1 + Moisture + Management + Use +
## Manure, data = dune.env)
##
##               Inertia Proportion Rank
## Total          2.1153     1.0000
## Constrained    1.5032     0.7106   12
## Unconstrained  0.6121     0.2894    7
## Inertia is mean squared contingency coefficient
## Some constraints were aliased because they were collinear (redundant)
##
## Eigenvalues for constrained axes:
##   CCA1   CCA2   CCA3   CCA4   CCA5   CCA6   CCA7   CCA8   CCA9  CCA10
## 0.4671 0.3410 0.1761 0.1532 0.0953 0.0703 0.0589 0.0499 0.0318 0.0260
##  CCA11  CCA12
## 0.0228 0.0108
##
## Eigenvalues for unconstrained axes:
##     CA1     CA2     CA3     CA4     CA5     CA6     CA7
## 0.27237 0.10876 0.08975 0.06305 0.03489 0.02529 0.01798
```

```
#creates a plot showing each axis
plotdune1 = plot(dune_ca1, xlim=c(-10,10), ylim=c(-20, 20), display=c('sp','bp'), scaling=1)
```




```
anova(dune_ca1, by='margin')
```

```
## 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, data = dune.env)
##            Df ChiSquare      F Pr(>F)
## A1          1   0.11070 1.2660  0.218
## Moisture    3   0.31587 1.2041  0.231
## Management  2   0.15882 0.9081  0.576
## Use         2   0.13010 0.7439  0.786
## Manure      3   0.25490 0.9717  0.515
## Residual    7   0.61210
```

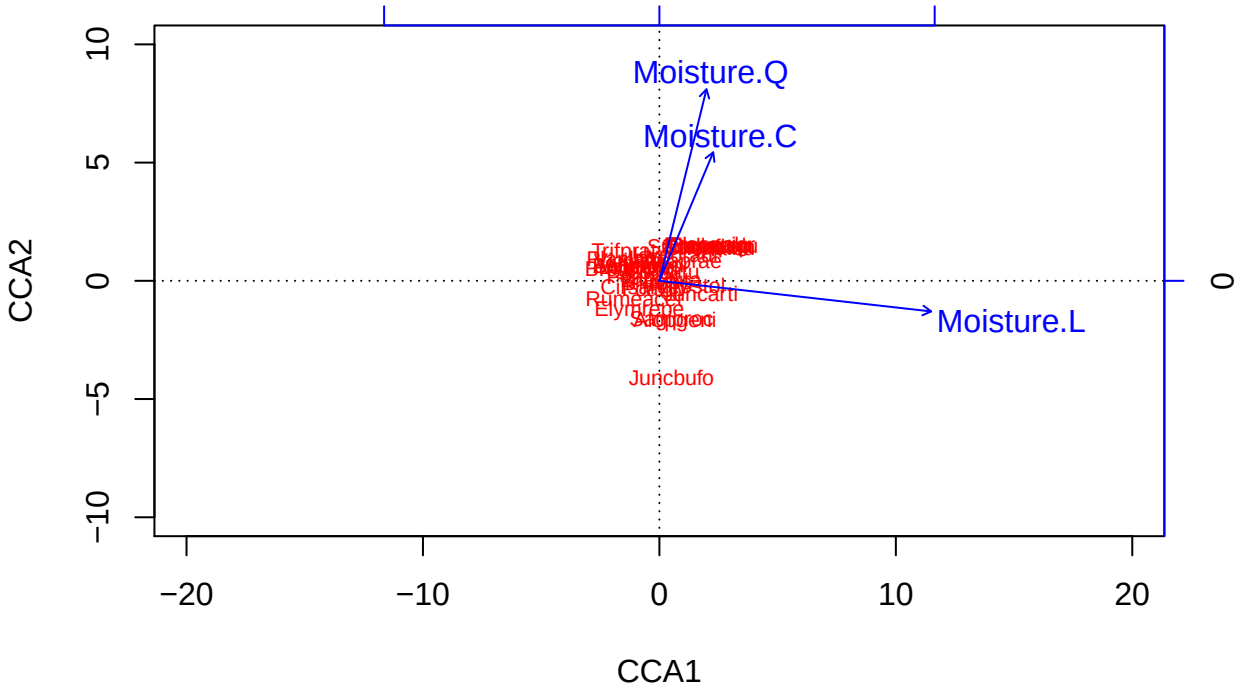
After looking at every variable in the dune.env dataset, moisture seems to have the strongest relationship with the data. I then only looked at the moisture variable from the dune environment data. The total inertia did not chage. Howver, the constrained axes accounted for .62 while the unconstrained axes accounted for 1.48. This makes sense since there are less explanatory variables being included in the model.

```
dune_ca2 = cca(dune ~ Moisture, dune.env)
dune_ca2
```

```
## Call: cca(formula = dune ~ Moisture, data = dune.env)
##
##               Inertia Proportion Rank
## Total          2.1153     1.0000
## Constrained    0.6283     0.2970    3
## Unconstrained  1.4870     0.7030   16
```

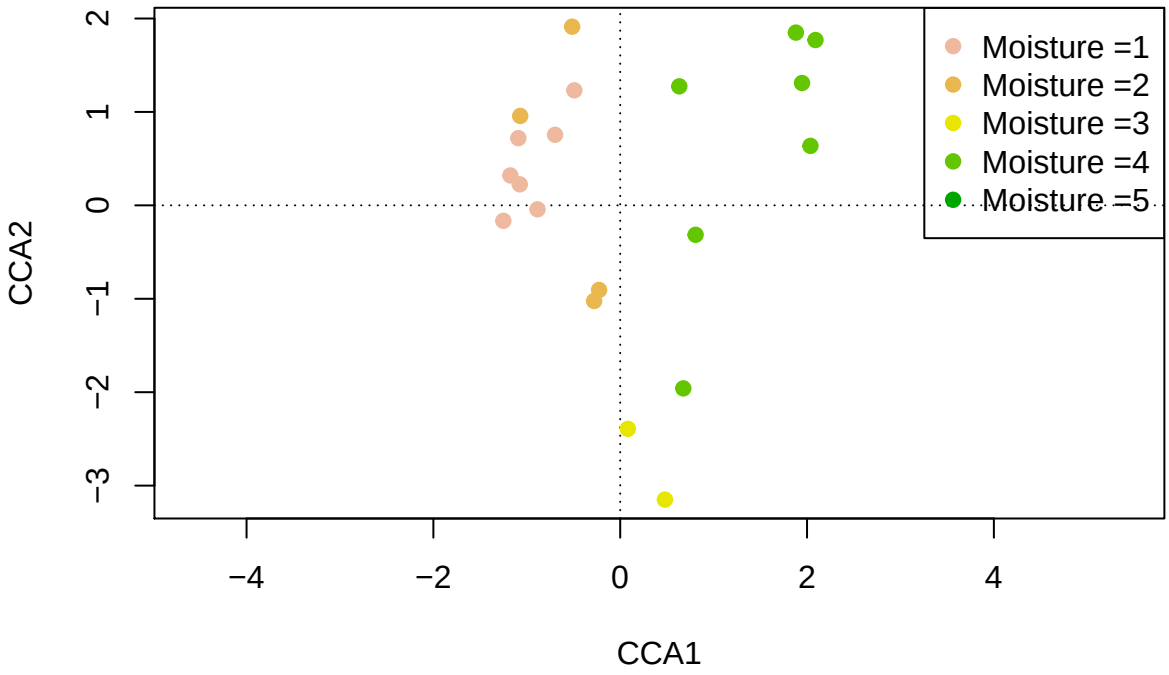
```
## Inertia is mean squared contingency coefficient
##
## Eigenvalues for constrained axes:
##   CCA1   CCA2   CCA3
## 0.4187 0.1330 0.0766
##
## Eigenvalues for unconstrained axes:
##    CA1    CA2    CA3    CA4    CA5    CA6    CA7    CA8    CA9   CA10
## 0.4098 0.2259 0.1761 0.1234 0.1082 0.0908 0.0859 0.0609 0.0566 0.0467
##   CA11   CA12   CA13   CA14   CA15   CA16
## 0.0419 0.0201 0.0143 0.0099 0.0085 0.0080
```

```
plotdune2 = plot(dune_ca2, xlim=c(-10,10), ylim=c(-10, 10), display=c('sp','bp'), scaling=1)
```



Graphically, the plot of the CCA test is not very helpful. To create a more helpful visualization I created an empty plot and then added the points of the species. These points were then color coded by moisture level to better display any clustering that appears.

```
#create new plot
plot(dune_ca2, type ='n')
color_vect = rev(terrain.colors(6))[-1]
points(dune_ca2, 'sites', pch=19, col=color_vect[dune.env$Moisture])
legend('topright', paste("Moisture =", 1:5, sep=''), col=color_vect, pch=19)
```




```
anova(dune_ca2, by='margin')
```

```
## Permutation test for cca under NA model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ Moisture, data = dune.env)
##          Df ChiSquare      F Pr(>F)
## Moisture  3   0.62831 2.2536  0.003 **
## Residual 16   1.48695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.) At first glance, the two test performed on the data set does not seem to agree. In the first analysis there was a visible striation within the data based on moisture while in the CCA, there is no clear relationship between species and moisture. However, when we change the plot of the CCA test we can see that there is in fact a similar amount of clustering that occurs based on moisture levels.