

Elaborated Canonical Correspondence Analysis

Canonical Correspondence Analysis is a direct gradient analysis techniques and the ordination equivalent of multiple linear regression. For this procedure to be valid the assumptions of multiple regression, must be met. Screening of the data and evaluation of the assumptions is vital. Screening and assumption evaluation techniques are presented in prior labs. These techniques should be applied as appropriate for CCA. We will assume that the data has been screened and that the assumptions have met.

Setup

This dataset is part of an aquatic ecology dataset. The data is presented as two matrices. The first matrix includes the fish community data collected at a number of sites. The second matrix includes measures of the environment (chemical and physical) at the same sites. We want to reduce the number of variables (species) that we need consider and the new axes should represent the differences in the environment environments at the sites.

Set the working directory and load the datasets

```
setwd("C:\\Multi\\")  
  
DoubtsSp <- read.table("DoubtsSpe.txt", header=TRUE)  
  
DoubtsEnv <- read.table("DoubtsEnv.txt", header=TRUE)
```

Perform the CCA

```
library(vegan)  
  
DoubtsSp.cca <- cca(DoubtsSp ~ ., DoubtsEnv)  
  
DoubtsSp.cca  
  
summary(DoubtsSp.cca)
```

Plot of CCA Results

```
Scaling 1 – Triplot – Sites are emphasized  
windows(title="CCA triplot - scaling 1 - lc scores", 9, 9)  
  
plot(DoubtsSp.cca, scaling=1, display=c("sp", "lc", "cn"), main="Triplot CCA  
DoubtsSp ~ DoubtsEnv - scaling 1")
```

Scaling 1 – Biplot – Sites are emphasized

```
windows(title="CCA biplot - scaling 1", 9, 9)
```

```
plot(DoubsSp.cca, scaling=1, display=c("lc", "cn"), main="Biplot CCA DoubsSp  
~ DoubsEnv - scaling 1")
```

Scaling 2 – Species are emphasized

```
windows(title="CCA triplot - scaling 2 - lc scores", 9, 9)
```

```
plot(DoubsSp.cca, display=c("sp", "lc", "cn"), main="Triplot CCA DoubsSp ~  
DoubsEnv - scaling 2")
```

Scaling 2 – Biplot – Species are emphasized

```
windows(title="CCA biplot - scaling 2", 9, 9)
```

```
plot(DoubsSp.cca, scaling=2, display=c("sp", "cn"), main="Biplot CCA  
DoubsSp ~ DoubsEnv - scaling 2")
```

Permutation tests of CCA

Overall

```
anova(DoubsSp.cca, step=1000)
```

Each Axis

```
anova(DoubsSp.cca, by="axis", step=1000)
```

Do the Environmental Variables All Contribute to the Ordination

The above is similar to an entire model for multiple regression. Variables may or may not be contributing to the overall model. In this step, we are constructing a forward model of the ordination – only variables that contribute to the overall model are included.

```
DoubsSp.cca.step.forward <- ordistep(cca(DoubsSp ~ 1, data=DoubsEnv),  
scope=formula(DoubsSp.cca), direction="forward", pstep=1000)
```

Performing the Most Parsimonious CCA

```
DoubsSp.cca.pars <- cca(DoubsSp ~ das + alt + oxy + dbo, data=DoubsEnv)
```

Permutation Tests of the Parsimonious CCA

```
anova.cca(DoubsSp.cca.pars, step=1000)
```

```
anova.cca(DoubsSp.cca.pars, step=1000, by="axis")
```

```
vif.cca(DoubsSp.cca)
```

```
vif.cca(DoubsSp.cca.pars)
```

Since the vif (variance inflation factor) of das is higher than 10 we should probably rerun the procedure without das. Then rerun the permutation tests and diagnostics again. Plots should also be constructed for the final CCA.