homework_6

Jasmine Williamson

2024-11-06

Homework 6

Constrained Ordination

Question 1: Constrained ordination is a type of ordination whereby the axes are constrained by environmental factors. For your dataset, which are the "response" variables and which are the "predictor" or "environmental" variables? Alternatively, do you think a symmetric analysis is more appropriate for your data (as opposed to an asymmetric or hypothesis-testing approach), and if so, why?

The response variables are the salamander data, and the predictor variables are my environmental matrix, i.e. downed wood, etc. I think symmetric could be more useful for my data for exploring structure and relationships.

Question 2: Perform a Redundancy Analysis (RDA) on your dataset (you can use a simple, transformationbased,

or distance-based approach, but justify your choice!). Interpret the results by discussing the relationship between the response and predictor variables. What percentage of variance in the response data is explained by the predictor variables? Provide R output and visualizations (e.g., RDA triplots) to support your analysis.

I'm using tbRDA because it works well with data that has a lot of zeros.

```
sal.hel <- decostand(sals, "hellinger")

tbRDA <- rda(sal.hel ~ ., env_cont)
```

```
##
## Some constraints or conditions were aliased because they were redundant. This
## can happen if terms are linearly dependent (collinear): 'logs'
```

```
summary(tbRDA)
```

```
##
## Call:
## rda(formula = sal.hel ~ jul_date + elev + temp + hum + canopy_cov +    veg_cov + dwd_cov + fwd_cov
##
## Partitioning of variance:
##                 Inertia Proportion
```

```
## Total            0.3394     1.0000
## Constrained      0.0677     0.1995
## Unconstrained  0.2717     0.8005
##
## Eigenvalues, and their contribution to the variance
##
## Importance of components:
##                          RDA1     RDA2     PC1      PC2
## Eigenvalue            0.05114 0.01656 0.1775 0.09423
## Proportion Explained  0.15067 0.04879 0.5229 0.27764
## Cumulative Proportion 0.15067 0.19946 0.7224 1.00000
##
## Accumulated constrained eigenvalues
## Importance of components:
##                          RDA1     RDA2
## Eigenvalue            0.05114 0.01656
## Proportion Explained  0.75540 0.24460
## Cumulative Proportion 0.75540 1.00000
```

```r
anova(tbRDA, permutations = how(nperm = 999))
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = sal.hel ~ jul_date + elev + temp + hum + canopy_cov + veg_cov + dwd_cov + fwd_c
##           Df Variance      F Pr(>F)
## Model     15 0.067699 1.8438  0.011 *
## Residual 111 0.271709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(tbRDA, by = "axis", permutations = how(nperm = 999))
```

```
## Permutation test for rda under reduced model
## Forward tests for axes
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = sal.hel ~ jul_date + elev + temp + hum + canopy_cov + veg_cov + dwd_cov + fwd_c
##           Df Variance       F Pr(>F)
## RDA1       1 0.051140 23.3387  0.002 **
## RDA2       1 0.016559  7.5572  0.988
## Residual 124 0.271709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(tbRDA, by = "terms", permutations = how(nperm = 999))
```

```
## Permutation test for rda under reduced model
## Terms added sequentially (first to last)
## Permutation: free
```

```
## Number of permutations: 999
##
## Model: rda(formula = sal.hel ~ jul_date + elev + temp + hum + canopy_cov + veg_cov + dwd_cov + fwd_c
##             Df Variance       F Pr(>F)
## jul_date    1 0.000028  0.0115  0.990
## elev        1 0.004956  2.0247  0.133
## temp        1 0.000859  0.3510  0.672
## hum         1 0.003031  1.2383  0.294
## canopy_cov  1 0.032173 13.1435  0.001 ***
## veg_cov     1 0.003876  1.5834  0.199
## dwd_cov     1 0.002125  0.8681  0.422
## fwd_cov     1 0.004299  1.7563  0.187
## soil_moist  1 0.002106  0.8603  0.421
## dwd_count   1 0.001529  0.6245  0.520
## stumps      1 0.001473  0.6016  0.525
## size_cl     1 0.000373  0.1523  0.867
## decay_cl    1 0.004677  1.9105  0.145
## char_cl     1 0.000705  0.2880  0.745
## length_cl   1 0.005490  2.2429  0.115
## Residual  111 0.271709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
RsquareAdj(tbRDA)
```

```
## $r.squared
## [1] 0.1994623
##
## $adj.r.squared
## [1] 0.09128155
```

```r
goodness(tbRDA)
```

```
##           RDA1      RDA2
## oss  0.05935405 0.1259271
## enes 0.27107790 0.2964183
```

```r
vif.cca(tbRDA)
```

```
##   jul_date      elev      temp       hum canopy_cov   veg_cov   dwd_cov
##   2.193660  2.400873  3.780248  3.421559   3.521891  1.931301  2.245364
##    fwd_cov soil_moist dwd_count    stumps       logs   size_cl  decay_cl
##   2.059265  1.938509  7.228652 14.209641         NA  1.151702  2.359334
##    char_cl  length_cl
##   1.930887   7.079383
```

```r
groups <- levels(factor(env$trt))
pt_col <- viridis(length(groups))

site.sc <- summary(tbRDA)[2]
spe.sc <- summary(tbRDA)[1]
```

```r
env.sc <- summary(tbRDA)[4]

plot(tbRDA,
    choices=c(1,2),
    type="n",
    scaling=2,
    main="tbRDA of Hellinger-transformed Sal Count",
    xlab="RDA 1",
    ylab="RDA 2")
for (i in 1:length(groups)){
    dim_choice <- site.sc$sites[env$SMU==groups[i],]
    points(dim_choice[,1], dim_choice[,2],
    pch=19,
    cex=1.4,
    col=pt_col[i])
    }
text(env.sc$biplot*2,
    row.names(env.sc$biplot),
    col="darkgray")
arrows(0, 0, env.sc$biplot[,1]*1.7, env.sc$biplot[,2]*1.7,
    lwd=2,
    length=0.1,
    col="darkgray")
text(spe.sc$species*2,
    row.names(spe.sc$species))
arrows(0, 0, spe.sc$species[,1]*1.5, spe.sc$species[,2]*1.5,
    lwd=2,
    length=0.1)
legend(x="bottomright",
    legend=groups,
    col=pt_col,
    pch=19)
```

tbRDA of Hellinger−transformed Sal Count
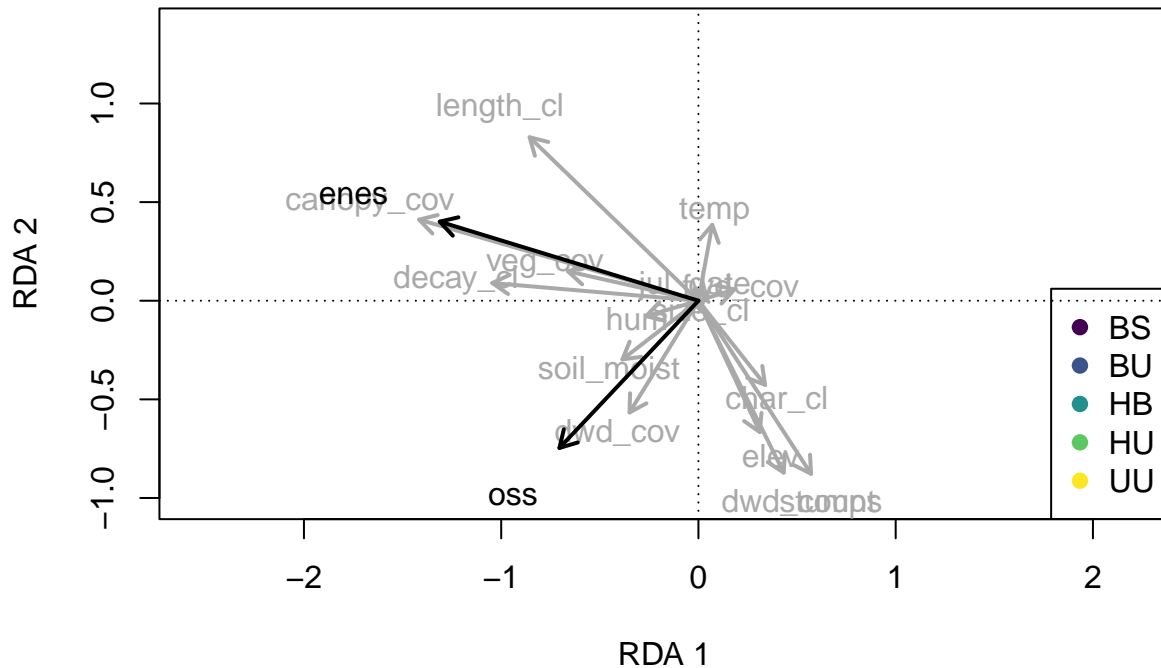
The total constrained inertia from my predictors is about 20%, which isnt great. Of that, RDA1 accounts for about 91% and RDA2 accounts for about 8%.

Question 3: Next, perform a Canonical Correspondence Analysis (CCA) on your dataset. Compare and contrast the results of the RDA with the CCA. Discuss the assumptions made in RDA and CCA and whether they are met by your dataset. How do the results differ in terms of species-environment (or other response-predictor) relationships?

I cant get this to work with my data, as it doesnt want any zeros in the community data matrix.

```
#CCA <- cca(sals ~ ., env_cont)
         #summary(CCA)

#Error in cca.default(d$X, d$Y, d$Z) :
#  all row sums must be >0 in the community data matrix
```

Question 4: Conduct a Co-inertia Analysis (CoIA), even if your "Y1" and "Y2" matrices are better suited for a directional hypothesis. How do the two datasets relate to one another? Use the RV coefficient to assess the strength of the shared structure between the datasets.

```
dudi.sal <- dudi.pca(sal.hel, scale=FALSE, scannf=FALSE)
    dudi.env <- dudi.pca(env_cont, scale=TRUE, scannf = FALSE)
```

```r
coia <- coinertia(dudi.sal, dudi.env, scannf = FALSE, nf=2)
    summary(coia)
```

```
## Coinertia analysis
##
## Class: coinertia dudi
## Call: coinertia(dudiX = dudi.sal, dudiY = dudi.env, scannf = FALSE,
##     nf = 2)
##
## Total inertia: 0.1152
##
## Eigenvalues:
##     Ax1     Ax2
## 0.10116 0.01403
##
## Projected inertia (%):
##     Ax1     Ax2
##   87.82   12.18
##
## Cumulative projected inertia (%):
##     Ax1   Ax1:2
##   87.82  100.00
##
## Eigenvalues decomposition:
##         eig     covar       sdX       sdY      corr
## 1 0.10115803 0.3180535 0.3778140 1.797890 0.4682298
## 2 0.01403038 0.1184499 0.4404449 1.240336 0.2168223
##
## Inertia & coinertia X (dudi.sal):
##      inertia       max     ratio
## 1  0.1427434 0.1939982 0.7357975
## 12 0.3367351 0.3367351 1.0000000
##
## Inertia & coinertia Y (dudi.env):
##     inertia      max     ratio
## 1  3.232408 3.838730 0.8420516
## 12 4.770841 6.341899 0.7522733
##
## RV:
##   0.08229505
```

```r
randtest(coia, nrepet = 999)
```

```
## Monte-Carlo test
## Call: randtest.coinertia(xtest = coia, nrepet = 999)
##
## Observation: 0.08229505
##
## Based on 999 replicates
## Simulated p-value: 0.001
## Alternative hypothesis: greater
##
```

```
##       Std.Obs   Expectation      Variance
## 4.9586735439 0.0302232434 0.0001102743
```

```
plot(coia)
```
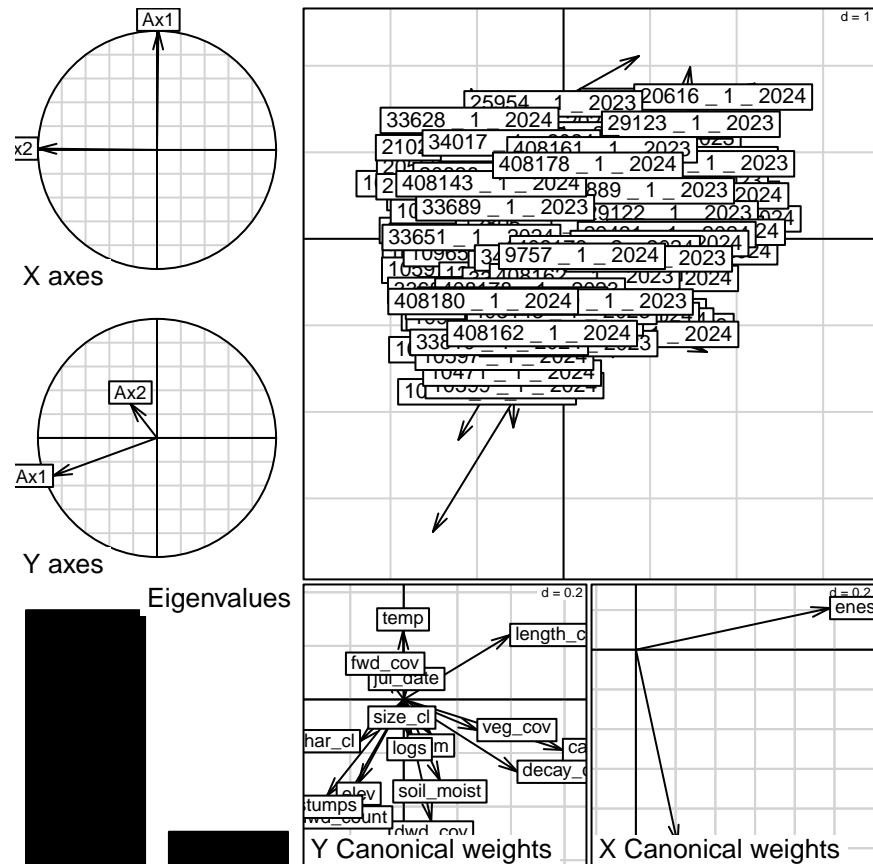

The total inertia is 0.11, which is pretty low, and means that the variance in the two datasets is not explained by their common structure. The RV is 0.08, which again is very low. This reinforces that there is not a strong relationship between the environmental variables and the salamander data. I dont know why and I dont like it.

Question 5: Based on the analyses you conducted (RDA, CCA, CoIA), summarize the main ecological insights gained from your data. What limitations did you encounter in your analysis and how might they affect your interpretations? Which method do you feel is most appropriate for your data and why?

These all felt like a bit of a bust for my data. Not sure if its user error, but none of them reported notable relationships between the matrices, and I couldnt get the CCA to work at all. Maybe it just means that my environmental data actually doesnt have any strong relationships with my salamander data, which would probably mean I made mistakes in my study design somehow, because we know that environmental variables should be related to salamander occupancy and abundance.

The tbRDA plot does associate OSS with soil moisture and downed wood cover, which tracks what we know about the species, and ENES with canopy cover, which means its occurring in our less disturbed sites.