BIOL 540 — Analysis of Ecological Communities

Monday 9:00 Lewis 407, Tuesday-Thursday 9:00-11:00, Lewis 407

Dave Roberts
Ecology Department
117 AJM Johnson Hall
droberts@montana.edu

Course Description

This course is designed to familiarize community ecologists with the analysis of ecological community data. Such data are generated by sampling multiple species at multiple sites, e.g. plant species abundances at sample plots across a range of environments, aquatic invertebrate species (or taxa) abundances from multiple benthic samples, or bird species abundances from pre-dawn surveys across a range of vegetation structures. In each case, there are observations of multiple species and observations of multiple attributes of the environment at each sample unit. Depending on class interest, we may also cover the analysis of individual species distributions. However, this is not a class in experimental design or time series, and is not appropriate for students with structured data requiring ANOVA-like analyses.

This is a practical, hands-on course emphasizing the interpretation of ecological analyses, and covers the majority of multivariate analyses in common use by community ecologists. The course is largely a lab course, and all grades are based on the quality of lab reports submitted; there are no exams. Lab reports are expected to be presented in a formal form with extensive annotated tabular and graphical output from the analyses. Submission of lab reports in HTML or PDF is strongly encouraged. Students are encouraged to use RMarkdown or other software like knitr or Sweave to write reports, but other software is certainly acceptable.

We will be using the R computer environment to conduct all analyses. Students will be expected to load their personal computers with a copy of the R package and all necessary material. No prior experience with R is necessary, but general familiarity with computer spreadsheets and data handling is helpful. Students are encouraged to bring their own datasets, but data will be made available to those who need it.

There will be a general lecture/literature review on Mondays, a lab introduction to the software and analyses on Tuesdays, and hands-on analysis Thursdays. There is no text, but readings from the primary literature will be made available.
Modules

- Introduction to the R language
  - obtaining and installing R
  - S language and syntax
  - loading datasets and saving results
  - functions and packages
- Graphical/tabular examination/presentation of multivariate data
  - scatterplots and boxplots
  - multi-layer plots, colors, and glyphs
- Species and community distribution models
  - generalized linear models (GLM)
  - generalized additive models (GAM)
  - classification and regression trees (CART/RANDOM FORESTS)
- Ordination
  - Dissimilarity and distance indices
  - Principal coordinates analysis (PCO)
  - Non-metric multidimensional scaling (NMDS)
  - Correspondence analysis and detrended correspondence analysis (CA and DCA)
  - Canonical correspondence analysis (CCA)
  - Fuzzy set ordination (FSO)
- Cluster analysis and discriminant analysis
  - hierarchical cluster analysis
  - PAM, optpart, and partana

Materials

There are no computer workstations in the lab. Rather, students will be expected to work on their own laptop. You will need to install R and several packages. In recent years many student have chosen to install Rstudio. I’m not a fan, but you’re welcome to work in that environment if you choose.

Work Flow

This a demanding course, and requires continual effort. It is critical that you stay up-to-date as new material is presented every week. You must make sure that your data are loaded fully and correctly as soon as possible. You must also be sure to save all your results when you first obtain them to avoid having to repeat analyses. R makes this very easy, but it requires a little discipline on your part.

Lab Reports

You will be graded on the quality of your lab reports. There are no exams. You will write one or two lab reports for each module, depending on the length of that module. I will want to see extensive results, both tabular and graphical, so you will want to make sure you know how to capture and save your results as you go.

Lab reports should be written in a nearly publishable format:
1) abstract (question, general approach, major findings)
2) introduction
   general ecological question of interest
   review of previous results if relevant
3) data
   description of data to be analyzed
4) methods
   description in statistical terms of approach
   description of software packages or functions employed
5) results
   extensive presentation of annotated tabular and graphical output
6) interpretation
   ecological perspective — did you learn anything about the species or communities
   methodological perspective — how well did the analyses work to answer your
   question, compared to previous methods where relevant.

General References

There is no text for this course. The material on the S language and the R program
is available on the web, and will be presented to you on CD-ROM. Statistics texts that
are likely to be helpful include:

Statistical Ecology

Blackwell


Chapman & Hall

Pielou, E.C. 1984. The interpretation of ecological data: a primer on classification and
ordination. Wiley-Interscience
Adler, J. 2012. R in a Nutshell. O'Reilly


Everitt, B. and T. Hothorn. An Introduction to Applied Multivariate Analysis with R.
Specific Readings

**R Syntax and Functions**
Spector. Chap 1. Data in R

Spector. Chap 2. Reading and Writing Data

Spector. Chap 5. Factors


Everitt and Hothorn. Chap 2. Looking at Multivariate Data Sets

**Gradient Response Models**


Tree Classifiers


Similarity, Dissimilarity, and Distance


Ordination


Borcard, Gillet and Legendre. Unconstrained Ordination. Chap 5

Borcard, Gillet, and Legendre. Canonical Ordination. Chap 6


**Cluster Analysis**


Kent, M. Numerical classification, cluster analysis, and phytosociology. Chap.


