

10. Análisis multivariante (II): Técnicas de ordenación en ecología.

Ecología Metodológica y Cuantitativa (5C1)
Departamento de Ecología e Hidrología

Curso 2008–09

Índice

1. Datos de gradientes	1
2. Análisis en componentes principales	1
2.1. ACP	2
2.2. Representación gráfica en ACP	2
2.3. ACP en un gradiente largo	3
3. Análisis de correspondencias	3
3.1. AC	3
3.2. Representación gráfica del AC	3
3.3. ACB en un gradiente corto	4
4. Variables pasivas	4
5. Análisis de correspondencias canónicas	4
6. Ejercicios adicionales	5

Antes de empezar

Iniciar R. Cargar el archivo de funciones de la asignatura, `funciones.R`:
`source("http://www.um.es/docencia/emc/datos/funciones.R")`

1. Datos de gradientes

En la figuras 1 y 2 se refleja la abundancia de distintas especies en dos situaciones, respectivamente, un gradiente corto y un gradiente largo. En el primer caso se presentan cuatro especies (A , B , C y D) y el en segundo tres (A , B y C).

Para realizar un análisis multivariante necesitamos una matriz de datos de muestras por especies. Se ha construido las matrices de datos considerando diez que sobre el gradiente en cada uno de los puntos, localidades, numerados de 1 a 10 se ha obtenido una unidad de muestreo.

2. Análisis en componentes principales

En lo esencial, un análisis en componentes principales proporciona una transformación de la matriz de datos considerando nuevas variables o ejes (también llamadas componentes).

Los resultados del análisis reflejan:

- valores propios de la matriz diagonalizada;

- porcentaje del total de inercia que supone cada autovalor;
- coordenadas de las variables en los nuevos ejes; y
- coordenadas de los individuos en los nuevos ejes.

2.1. ACP

Ejercicios. Bloque 1:

1. Leer el archivo `gradc.dat`, crear la variable `gradc`, y representar gráficamente la matriz de datos, excluyendo la columna 5 que representa la variable ambiental:

```
read.table("http://www.um.es/docencia/emc/datos/gradc.dat")->gradc
gradec<-gradc[,1:4]
matplot(gradec,type="l")
```

2. Realizar un análisis en componentes principales de `gradec` utilizando matriz de correlaciones:

```
acp(scale(gradec))->gradec.acp; gradec.acp
¿qué significado tiene $inercia?
```

3. Calcular y comparar la matriz de distancias de las muestras con datos estandarizados y la nueva matriz de coordenadas:

```
dist(scale(gradec))
dist(gradec.acp$cfilas)
```

¿Proporcionan los mismos resultados? ¿Por qué?

4. Calcular las correlaciones entre las nuevas variables y las originales:

```
cor(gradec.acp$cf,gradec)
```

¿Qué relación hay entre estas correlaciones y la matriz `gradec.acp$ccolumnas`?

5. Calcular las correlaciones entre las nuevas variables y la variable ambiental original (`gradc[,5]`):

```
cor(gradec.acp$cf,gradc[,5])->cva;cva
```

¿Qué relación hay entre estas correlaciones y la matriz `gradec.acp$ccolumnas`?

2.2. Representación gráfica en ACP

Ejercicios. Bloque 2:

1. Estudiar todos los planos del análisis para representar las unidades de muestreo.

```
pairs(gradec.acp$cf,cex=1+(2/(10:1)),asp=1,col=grey(seq(0,1,0.1)),pch=16)
```

¿Cuáles son, en términos de información, los mejores ejes? ¿Qué relación, en cuanto a varianza, hay entre los mejor y los peor de los ejes? ¿Se puede trasladar esta idea a la matriz de especies?

2. Utilizar la función `dibujaacp()` para representar conjuntamente casos y especies (biplot). Esta función necesita un objeto con los resultados del análisis, los ejes que se han de representar en abscisas y ordenadas, y opcionalmente los colores para objetos y variables.

```
dibujaacp(gradec.acp,1,2)
```

3. Ordenar las filas y las columnas de la matriz según las coordenadas del primer eje del análisis.

```
ordenamatrix(gradec,gradec.acp)
```

4. Representar la variable ambiental en el plano obtenido a partir de la correlación de esta con los ejes:

```
arrows(0,0,cva[1],cva[2],col=2)
```

2.3. ACP en un gradiente largo

Ejercicios. Bloque 3:

1. Leer el archivo `gradl.dat`, crear la variable `gradl`, y representar gráficamente la matriz de datos, excluyendo la columna 4 que representa la variable ambiental:

```
read.table("http://www.um.es/docencia/emc/datos/gradl.dat")->gradl
gradel<-gradl[,1:3]
matplot(gradel,type="l")
```

2. Realizar un análisis de componentes principales a la matriz `gradel` siguiendo los pasos del apartado anterior.

¿Se obtiene una representación adecuada del gradiente en este caso? ¿Cómo se justifica?

3. Análisis de correspondencias

El análisis de correspondencias (binarias) es similar al de componentes principales pero existen diferencias importantes. En primer lugar utiliza como entrada una matriz de frecuencias —una gran tabla de contingencia para dos variables, de ahí el nombre de binarias—; además considera que los objetos (filas y columnas) son semejantes si son proporcionales, a diferencia del análisis en componentes principales que considera semejantes a los de magnitudes (tamaño) parecidas.

Como en el caso anterior obtendremos un vector de autovalores y dos matrices: las coordenadas de las filas y las de las columnas.

Existe un caso especialmente interesante: algunos de los autovalores (los primeros) presentan valores próximos a 1. Esta situación indica que nos encontramos ante una matriz de datos disjunta, es decir, hay submatrices dentro de la matriz que no comparten valores.

3.1. AC

Ejercicios. Bloque 4:

Para realizar un análisis de correspondencias binarias ha de utilizarse la función `acb()`; que proporcionan unos resultados análogos a la función `acp()`. Como en el caso anterior: procederemos a analizar la matriz `gradel`.

1. Realizar un análisis de correspondencias utilizando

```
acb(gradel)->gradel.acb; gradel.acb
```

¿qué significado tiene `$inercia`? ¿cuantos ejes son necesarios para representar los datos con una calidad suficiente?

3.2. Representación gráfica del AC

Ejercicios. Bloque 5:

1. Representar los resultados del análisis de correspondencias mediante la función `dibujaacb()` análoga a `dibujaacp()` utilizada anteriormente.

```
dibujaacb(gradel.acb,1,2)
```

¿Cuál es la interpretación de la posición relativa de las columnas (especies) y las filas (muestras)?

2. ¿Puede hablarse de ventajas de una técnica de ordenación sobre la otra? ¿Son estas ventajas una consecuencia de la naturaleza de la matriz de datos?

3. Ordenar las filas y las columnas de la matriz según las coordenadas del primer eje del análisis.

```
ordenamatrix (grad, grad.acb)
```

3.3. ACB en un gradiente corto

Ejercicios. Bloque 6:

1. Utilizando la matriz de datos de especies correspondientes al gradiente corto, `gradec`, realizar un análisis de correspondencias, tal como se hizo en el apartado anterior.
¿Se obtiene una representación adecuada del gradiente en este casos? ¿Cómo se justifica?
-
-

4. Variables pasivas

Ejercicios. Bloque 7:

Utilizando los datos del muestreo de aves en sistemas forestales de la Región de Murcia, realizar un análisis de correspondencias con las siguientes especies: `sylmel` (curruca cabecinegra), `sylcan` (curruca carrasqueña), `sylund` (curruca rabilarga), `parcri` (herrerillo capuchino), `parate` (carbonero garrapinos), `parmaj` (carbonero común), `carchl` (verderón común), `carcar` (jilguero) y `carcan` (pardillo común).

```
read.table("http://www.um.es/docencia/emc/datos/avesca.dat") ->avesca  
acb(avesca) ->avesca.acb
```

1. Representar el diagrama de ordenación (ejes 1 y 2), etiquetando los puntos correspondientes a las especies:

```
dibujaacb(avesca.acb)
```

2. Para interpretar los ejes como gradientes ambientales leer el fichero `amb.dat`, que contiene la información ambiental (altitud, temperatura, precipitación y cobertura arbórea) de cada una de las unidades de muestreo.

```
read.table("http://www.um.es/docencia/emc/datos/amb.dat") ->amb
```

3. Seguidamente calcular las correlaciones entre estas variables ambientales y las coordenadas de las unidades de muestreo (las filas) en los ejes 1 y 2 del análisis de correspondencias.

```
cor(amb, avesca.acb$cf[,1:2])
```

5. Análisis de correspondencias canónicas

Ejercicios. Bloque 8:

Realizaremos un CCA de los datos de la matriz `avesca` y `amb`. Cargar la librería de funciones `vegan`:

```
library(vegan)
```

1. Realizar el análisis mediante la función `cca()`:

```
cca(avesca, amb) ->aves.cca
```

¿Qué resultados proporciona el análisis?

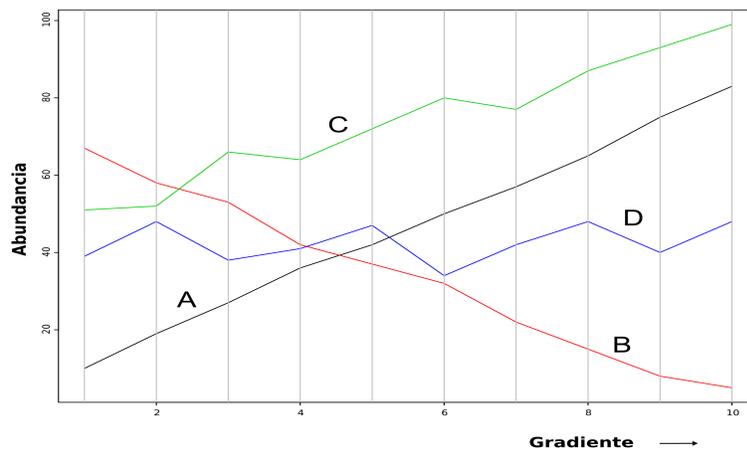


Figura 1: Abundancia de las especies *A*, *B*, *C* y *D* en un gradiente ambiental corto.

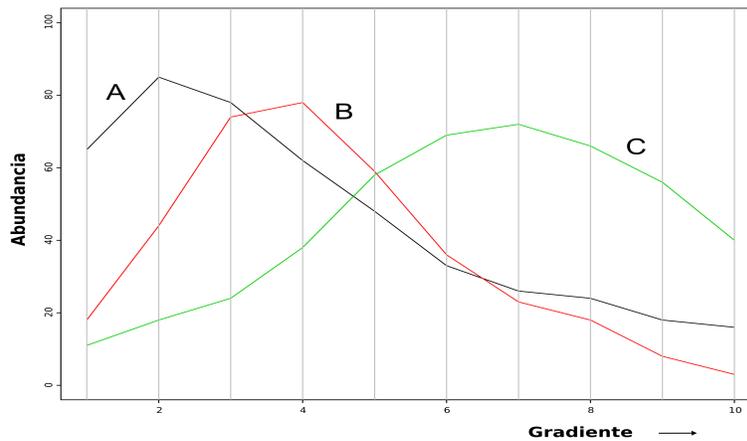


Figura 2: Abundancia de las especies *A*, *B* y *C* en un gradiente ambiental largo.

2. Representar gráficamente el análisis:

```
plot.cca(aves.cca, type="text")
```

¿Cómo puede interpretarse el resultado obtenido? ¿Se asemeja el comportamiento de las variables ambientales en CCA y el obtenido como variables pasivas en CA? ¿Qué consecuencia puede sacarse de ello?

6. Ejercicios adicionales

1. Realizar un análisis de correspondencias de los datos del archivo `marmenor.dat`, utilizando sólo las variables correspondientes a las especies de aves: `marmenor[, 1:3]`
2. Realizar un análisis de componentes principales de los datos del archivo `marmenor.dat`, utilizando sólo las variables correspondientes a las especies de aves: `marmenor[, 1:3]`

Representar como variables pasivas las restantes variables.