

Canonical correspondence analysis

Description

Demonstration function for canonical correspondence analysis (ter Braak 1986, 1987), following the algorithm described in Subsection 11.2.1 of Legendre and Legendre (1998). The results are identical to those obtained with CANOCO. No tests are performed; use `vegan`'s function `cca()` to compute the global test of significance and tests of individual canonical axes.

Usage

```
CCA(Y, X, use.svd=TRUE)
# Write basic statistics of eigen-decomposition and VIF to the R window
x or print(x, ...)
# Draw a triplot
triplot(x, xax=1, yax=2, scaling=1, graph.type="Z", aspect=1, cex=1,
color.sites="black", color.sp="red", color.env="blue", ...)
```

Arguments

<code>Y</code>	Response data matrix
<code>X</code>	Explanatory data matrix
<code>use.svd</code>	TRUE: the decomposition is done by svd (default). FALSE: the decomposition is done by eigen . The signs of the coefficients along any one axis may differ between the two methods.
<code>x</code>	Name of the output object of function <code>CCA()</code> .
<code>xax, yax</code>	Axes that will be used to draw the triplot. Default: <code>xax=1, yax=2</code> .
<code>scaling</code>	<code>scaling = {1, 2, 3}</code> to obtain a triplot with the corresponding scaling. See details.
<code>graph.type</code>	Either "Z" or "F"; "Z" = fitted site scores (default), "F" = site scores.
<code>aspect</code>	Aspect ratio of the plot; <code>aspect=1</code> makes the scales the same along the two axes. Use <code>aspect=NA</code> to remove the effect of parameter <code>aspect</code> in the triplot.
<code>cex</code>	A numerical value giving the amount by which plotting text and symbols should be magnified relative to the default, <code>cex=1</code> .
<code>color</code>	Color of the site, species, and explanatory variable symbols and labels in triplots. Defaults: <code>color.sites="black", color.sp="red", color.env="blue"</code> .
<code>...</code>	Other parameters passed to the <code>print</code> or <code>triplot</code> functions.

Details

Canonical correspondence analysis (CCA) of a table of frequencies producing scaling 1, 2, or 3 triplots. All variables in the response matrix `Y` must be frequency-like. Negative values are not allowed in the response matrix in CCA.

Scaling type 1 triplot uses V for species, Z1 or F for sites, and biplot.sc1 for explanatory variables (notation as in Legendre and Legendre 1998, Section 9.4). The sites are at the centroids (barycentres) of the species. This projection preserves the chi-square distance among the sites

Scaling type 2 triplot uses Fhat for species, Z2 or Vhat for sites, and biplot.sc2 for explanatory variables. The species are at the centroids (barycentres) of the sites. This projection preserves the chi-square distance among the species.

Scaling type 3 triplot is a compromise between scalings 1 and 2. This scaling, called “symmetric” in CANOCO, does not preserve the chi-square distances among the species nor among the sites. It uses spec3 for species, Z3 or site3 for sites, and biplot.sc3 for explanatory variables.

Matrices Z1, Z2 and Z3 contain the *site scores which are linear combinations of explanatory variables* (the CCA model fitted values) whereas matrices F, Vhat and site3 contain projections of matrix Qbar onto the canonical axes.

Algorithmic notes – The data matrix is transformed into matrix Qbar of the contributions to chi-square following equation 9.32 of Legendre and Legendre (1998). Then Qbar is regressed on standardized X, with weights, and the matrix of fitted values is decomposed by [svd](#) (default) or by [eigen](#).

Value

Function CCA returns a list containing the following results and matrices:

x\$general

\$inertia	Total inertia in matrix Qbar.
\$values	CA eigenvalues.
\$rel.values	Relative eigenvalues.
\$cum.rel	Cumulative sum of the relative eigenvalues.
\$r.spec.env	Correlations between the corresponding F and Z axes.
\$cor.env.axes	Inter-set correlations of explanatory variables with axes.
\$vif	Variance inflation factors of the explanatory variables.
\$k	Rank of the canonical solution.

x\$scaling1

\$species, \$sites.Z Matrices required to produce the scaling=1 triplot: V for species, Z1 and F for sites, and biplot.sc1 for the scores of the explanatory variables.

x\$scaling2

\$species, \$sites.Z Matrices required to produce the scaling=2 triplot: Fhat for species, Z2 and \$sites.F, \$biplot Vhat for sites, and biplot.sc2 for the scores of the explanatory variables.

x\$scaling3

\$species, \$sites.Z Matrices required to produce the scaling=3 triplot: spec3 for species, Z3 and \$sites.F, \$biplot site3 for sites, and biplot.sc3 for the scores of the explanatory variables.

<u>x\$other</u>	
\$U, \$Uhat	Matrices produced by SVD, used to compute the matrices needed for triplots.
\$X.stand	Matrix of explanatory variables X standardized as described in Legendre and Legendre (1998, Subsection 11.2.1).
\$Qbar	Contains the Y data transformed into contributions to chi-square following equation 9.32 of Legendre and Legendre (1998, Section 9.4).
\$site.names	Vector of site names.
\$sp.names	Vector of species names.
\$env.names	Vector of explanatory variable names.

In triplots, the arrows representing the scores of the explanatory variables may be longer than given by the coordinates of the corresponding variables in the \$biplot matrix. The elongation is an integer multiple factor of their lengths. That factor is computed by the *triplot.CCA()* function.

References

Aart, P. J. M. (van der) and N. Smeenk-Enserink. 1975. Correlations between distributions of hunting spiders (Lycosidae, Ctenidae) and environmental characteristics in a dune area. *Neth. J. Zool.* 25: 1-45.

Legendre, P. and Legendre, L. 1998. *Numerical Ecology*. 2nd English ed. Elsevier, Amsterdam.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.

ter Braak, C. J. F. 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* 69: 69-77.

Author

Pierre Legendre, Département de Sciences Biologiques, Université de Montréal.

Examples

```
# Example: data from Table 11.3 of Legendre and Legendre (1998)
reef = matrix(c(1,0,0,11,11,9,9,7,7,5,0,0,1,4,5,6,7,8,9,10,0,0,0,0,17,0,13,0,10,0,
0,0,0,0,7,0,10,0,13,0,0,0,0,8,0,6,0,4,0,2,0,0,0,1,0,2,0,3,0,4,2,5,0,6,6,10,4,
6,6,0,4,6,2,2,6,1,5,6,2,1,4,1,3,0,2,4,4,4,0,3,1,2,3,4,5,6,7,8,9,10,0,0,0,0,1,
0,1,0,1,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,1,0,1),10,13)
rownames(reef) = c("Site1", "Site2", "Site3", "Site4", "Site5", "Site6", "Site7", "Site8", "Site9",
"Site10")
colnames(reef) = c("Sp.1", "Sp.2", "Sp.3", "Sp.4", "Sp.5", "Sp.6", "Sp.7", "Sp.8", "Sp.9",
"Depth", "Coral", "Sand", "Other")
res = CCA(reef[,1:9], reef[,10:13])
res                                # Print out a summary of the results
triplot(res) or biplot.CCA(res)    # Draw a triplot (default: scaling=1)
```

```

summary(res)                # Print the structure of the output object
res$scaling1                # Print out all matrices used for the scaling 1 triplot
res$scaling2$species        # Print out the matrix of species scores used in scaling 2

# Example: the spider data of Aart and Smeenk-Enserink (1975), available in library mvpart.
# The spider data frame has 28 rows and 18 columns. The first 12 columns are abundances of
# different species of spiders and the next 6 are explanatory data (environmental variables).
library(mvpart)
data(spider)                # Note: this data file only has numbers for row names
rownames(spider) <- paste("Site.",1:nrow(spider),sep="")
res = CCA(spider[,1:12], spider[,13:18])
res                          # Print out a summary of the results
triplot(res, scaling=2, color.sites="blue") # Draw a triplot (scaling=2) with axes 1 and 2 (default)
triplot.CCA(res, scaling=2, color.sites="blue") # Draw a triplot (scaling=2) with axes 1 and 2
triplot(res, xax=1, yax=3)    # Draw a triplot (scaling=1) with axes 1 and 3

# Draw two triplots side by side in a graphic window
par(mfrow=c(1,2))
triplot(res, scaling=1)
triplot(res, scaling=2)

```