Package ‘acca’

January 28, 2022

Version 0.2

Date 2022-01-28

Title A Canonical Correlation Analysis with Inferential Guaranties

Author Livio Finos

Maintainer livio finos <livio.finos@unipd.it>

Description
It performs Canonical Correlation Analysis and provides inferential guaranties on the correla-
tion components. The p-values are computed following the resampling method developed in Wink-

Encoding UTF-8

License GPL (>= 2)

RoxygenNote 7.1.2

Imports methods, stats, ggplot2, plyr

NeedsCompilation no

Repository CRAN

Date/Publication 2022-01-28 08:10:02 UTC

R topics documented:

  acca-package .................................................. 2
  acca-method .................................................. 2
  cc ............................................................... 3
  ccbiplot ...................................................... 4
  cc_inference .................................................. 5

Index 7
acca-package

A Canonical Correlation Analysis with Inferential Guaranties

Description


Author(s)

Livio Finos

Examples

```r
set.seed(1)
X=matrix(rnorm(500),100,5)
Y=matrix(rnorm(700),100,7)
Z=matrix(rnorm(200),100,2)
mod=cc(X,Y,Z)
mod
ccbiplot(mod)
mod=cc_inference(mod, B = 100, numb_cc = 3)
mod
```

acca-method

Methods for acca objects

Description

Methods for acca objects. The following are methods to extract and manipulate relevant information from a acca object.

Usage

```r
## S3 method for class 'acca'
print(x, ...)

## S3 method for class 'acca'
summary(object, ...)
```
### Description

Very similar `cancor()` of package `stats`, but 1) it allows for \( X \) and \( Y \) to be rank deficient, 2) it allows for categorical variables and 3) it allows for covariates, 4) more (see below).

### Usage

```r
cc(X, Y, Zx = NULL, Zy = Zx, numb_cc = NULL, fill.na = FALSE)
```

### Arguments

- **\( X \)**
  - See \texttt{x} in `cancor` for a proper documentation.
- **\( Y \)**
  - See \texttt{y} in `cancor` for a proper documentation.
- **\( Zx \)**
  - (=\texttt{NULL} by default) covariates (i.e. nuisance) of \( X \). If different from \texttt{NULL}, the \( X \) are residualized by \( Zx \) before entering `cc()`.
  - IMPORTANT: if \( Zx \) is not \texttt{NULL}, include the intercept (if appropriate!)
- **\( Zy \)**
  - (=\texttt{Zx} by default) covariates of \( Y \). Same use of \( Zx \).
- **\texttt{numb.cc}**
  - number of (pairs of) canonical correlations to be extracted
- **\texttt{fill.na}**
  - replace NA in \( X \) and \( Y \) with column mean before enter \texttt{cc}()

### Value

It returns an \texttt{acca} object. This object contains the same list as returned by \texttt{cc}, while it also contains \texttt{prop_expl_var} which is a list reporting the proportion of explained (total) variance of \( X \) and \( Y \) by each component (i.e. mode).

```r
L=matrix(rnorm(10),10,1) X=matrix(rnorm(50),10,5) Y=matrix(rnorm(70),10,7) Z=matrix(rnorm(20),10,2)
ccbiplot(mod)
```
The function adapts the `ggbiplot::ggbiplot()` function to an `acca`-object.

Usage

```r
ccbiplot(
  pcobj,  
  choices = 1:2,
  scale = 1,
  pc.biplot = TRUE,
  obs.scale = 1 - scale,
  var.scale = scale,
  groups = NULL,
  shapes = NULL,
  ellipse = FALSE,
  ellipse.prob = 0.68,
  labels = NULL,
  labels.size = 3,
  alpha = 1,
  var.axes = TRUE,
  circle = FALSE,
  circle.prob = 0.69,
  varname.size = 3,
  varname.adjust = 1.5,
  varname.abbrev = FALSE,
  arrows.color = NULL,
  ...
)
```

Arguments

- `pcobj` an object returned by `acca::cc()`
- `choices` which PCs to plot
- `scale` covariance biplot (scale = 1), form biplot (scale = 0). When scale = 1, the inner product between the variables approximates the covariance and the distance between the points approximates the Mahalanobis distance.
- `pc.biplot` for compatibility with `biplot.princomp()`
- `obs.scale` scale factor to apply to observations
- `var.scale` scale factor to apply to variables
- `groups` optional factor variable indicating the groups that the observations belong to. If provided the points will be colored according to groups
shapes which shapes to use when plotting observations
ellipse draw a normal data ellipse for each group?
ellipse.prob size of the ellipse in Normal probability
labels optional vector of labels for the observations
labels.size size of the text used for the labels
alpha alpha transparency value for the points (0 = transparent, 1 = opaque)
var.axes draw arrows for the variables?
circle draw a correlation circle?
circle.prob size of the circle in qchisq(circle.prob, df = 2)
var.name.size size of the text for variable names
var.name.adjust adjustment factor the placement of the variable names, >= 1 means farther from the arrow
var.name.abbrev whether or not to abbreviate the variable names
arrows.color color of the arrows
... not used yet

Value

It returns a \texttt{ggplot} object.

\begin{Verbatim}
\texttt{cc_inference}
\end{Verbatim}

\begin{Verbatim}
\texttt{cc_inference}
\end{Verbatim}

Description


Usage

\begin{Verbatim}
cc_inference(
  mod,  
  B = 100,  
  alpha_max = 0.5,  
  numb_cc = NULL,  
  resamp_type = "sign-flip",  
  light = FALSE
)
\end{Verbatim}
Arguments

- **mod**: an acca object.
- **B**: (100 by default) number of random sampling.
- **alpha_max**: stop if p-value > alpha_max (alpha_max=0.5 by default).
- **numb_cc**: stop after computing p-values for the first numb_cc are computed(numb_cc=NULL - the default - means compute all).
- **resamp_type**: "sign-flip" (by default) or "permutation".
- **light**: If TRUE the orthogonalization of the residuals of the projection on Z is not performed. For large sample size the two methods tend to overlap.

Value

It returns an acca object (see cc) with p-values for each pair of the numb_cc components.

Examples

```r
set.seed(1)
X=matrix(rnorm(500),100,5)
Y=matrix(rnorm(700),100,7)
Z=matrix(rnorm(200),100,2)
mod=cc(X,Y,Z)
mod

ccbiplot(mod)

mod=cc_inference(mod, B = 100, numb_cc = 3)
mod
```
Index

acca-method, 2
acca-package, 2

cancor, 3
cc, 3, 3, 6
cc_inference, 5
ccbiplot, 4

ggplot, 5

print.acca(acca-method), 2

summary.acca(acca-method), 2