

Package ‘FRCC’

January 20, 2025

Title Fast Regularized Canonical Correlation Analysis

Version 1.1.0

Description Contains the core functions associated with Fast Regularized Canonical Correlation Analysis. Please see the following for details: Raul Cruz-Cano, Mei-Ling Ting Lee, Fast regularized canonical correlation analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473 <[doi:10.1016/j.csda.2013.09.020](https://doi.org/10.1016/j.csda.2013.09.020)>.

License GPL (>= 2)

Encoding UTF-8

RoxygenNote 7.2.1

Imports calibrate, CCP, corpcor, graphics, MASS, stats, utils

Depends R (>= 2.10)

LazyData true

Suggests grDevices

NeedsCompilation no

Author Raul Cruz-Cano [aut, cre]

Maintainer Raul Cruz-Cano <raulcruz@iu.edu>

Repository CRAN

Date/Publication 2022-10-29 09:10:42 UTC

Contents

custom.draw.circle	2
frcc	2
generate_multivariate_normal_sample	4
microRNA	5
off.diagonal.lambda	6
plot_units	6
plot_variables	7
rearrange.frcc	8

Index**10**

`custom.draw.circle` *Draws a circle*

Description

Given a center, radius and color, this function draws a circle.

Usage

```
custom.draw.circle(x, y, r, col)
```

Arguments

<code>x</code>	X coordinate of the center
<code>y</code>	Y coordinate of the center
<code>r</code>	Radius of the circle
<code>col</code>	Color of the circle

Value

This function does not return a value, it just draws a circle.

Author(s)

Michael Bedward

References

<http://www.r-bloggers.com/circle-packing-with-r/>

`frcc` *This function implements the Fast Regularized Canonical Correlation Analysis*

Description

This function implements the Fast Regularized Canonical Correlation algorithm described in [Cruz-Cano et al., 2014].

The main idea of the algorithm is using the minimum risk estimators of the correlation matrices described in [Schafer and Strimmer, 2008] during the calculation of the Canonical correlation Structure.

It can be considered an extension of the work for two set of variables (blocks) mentioned in [Tenenhaus and Tenenhaus, 2011].

Usage

```
frcc(X, Y)
```

Arguments

X numeric matrix (n by p) which contains the observations on the X variables.
 Y numeric matrix (n by q) which contains the observations on the Y variables.

Value

A list with the following components of the Canonical Structure:

cor Canonical correlations.
 p_values The corresponding p-values for the each of the canonical correlations.
 canonical_weights_X The canonical weights for the variables of the dataset X.
 canonical_weights_Y The canonical weights for the variables of the dataset Y.
 canonical_factor_loadings_X The inter-set canonical factor loadings for the variables of the dataset X.
 canonical_factor_loadings_Y The inter-set canonical factor loadings for the variables of the dataset Y.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.
 Schafer, J; Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Statistical Applications in Genetics and Molecular Biology 4:14, Article 32.
 Tenenhaus, A.; Tenenhaus, M. (2011). Regularized Generalized Canonical Correlation Analysis. Psychometrika 76:2, DOI: 10.1007/S11336-011-9206-8.

Examples

```
# Example # 1 Multivariate Normal Data
p<-10
q<-10
n<-50
res<-generate_multivariate_normal_sample(p,q,n)
X<-res$X
Y<-res$Y
rownames(X)<-c(1:n)
colnames(X)<-c(1:p)
```

```

colnames(Y)<- c(1:q)
my_res<-frcc(X,Y)
print(my_res)
#Example #2 Soil Specification Data
data(soilspec)
list_of_units_to_be_used<-sample(1:nrow(soilspec),14)
X<- soilspec[list_of_units_to_be_used,2:9]
Y<- soilspec[list_of_units_to_be_used,10:15]
colnames(X)<-c("H. pubescens", "P. bertolonii", "T. pretense",
              "P. sanguisorba", "R. squarrosus", "H. pilosella", "B. media", "T. drucei")
colnames(Y)<- c("d","P","K","d x P", "d x K","P x K")
my_res<-frcc(X,Y)
grDevices::dev.new()
plot_variables(my_res,1,2)
#Example #3 NCI-60 micrRNA Data
data("Topoisomerase_II_Inhibitors")
data("microRNA")
my_res <- frcc(t(microRNA),-1*t(Topoisomerase_II_Inhibitors))
for( i in 1:dim(microRNA)[2])
{
  colnames(microRNA)[i]<-substr(colnames(microRNA)[i], 1, 2)
}#end for i

```

```
generate_multivariate_normal_sample
```

It generates a sample from a multivariate normal distribution function

Description

It generates a sample from a multivariate normal distribution function with the cross-covariance matrix described in [Cruz-Cano et al. 2012].

Usage

```
generate_multivariate_normal_sample(p, q, n)
```

Arguments

p	Number of desired variables in the dataset X.
q	Number of desired variables in the dataset Y.
n	sample size desired.

Value

A list of n sample units with the values for the variables of the data sets X and Y.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.

Examples

```
p<-10
q<-10
n<-50
res<-generate_multivariate_normal_sample(p,q,n)
X<-res$X
Y<-res$Y
rownames(X)<-c(1:n)
colnames(X)<-c(1:p)
colnames(Y)<- c(1:q)
my_res<-frcc(X,Y)
```

microRNA

Example data sets

Description

These are data sets used to demonstrate the functions in this package.

Usage

microRNA

soilspec

Topoisomerase_II_Inhibitors

Format

Each is a dataframe.

Examples

```
microRNA
soilspec
Topoisomerase_II_Inhibitors
```

`off.diagonal.lambda` *Calculates the value of the shrinkage coefficient for the off-diagonal matrices*

Description

Calculates the value of the shrinkage coefficient for the off-diagonal matrices as described in [Cruz-Cano et al., 2012]

Usage

```
off.diagonal.lambda(xs, p, q)
```

Arguments

`xs` Matrix with the values for the datasets X and Y.
`p` Number of variables in the dataset X.
`q` Number of variables in the dataset Y.

Value

Shrinkage coefficient for the off-diagonal matrices used to calculate the FRCC canonical structure.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.

`plot_units` *Plots the experimental units in the Canonical Variates Space*

Description

This function plots the experimental units used in the FRCCA as points in a two-dimensional plane in which the axis are the canonical variates selected by the user

Usage

```
plot_units(X, Y, res.mrcc, i, text_size = 0.8, point_size = 2)
```

Arguments

X	numeric matrix (n by p) which contains the observations on the X variables.
Y	numeric matrix (n by p) which contains the observations on the Y variables.
res.mrcc	List containing a canonical structure provided by the function frcc for the dataset X and Y.
i	Canonical Variate which will be used for the axes (X for horizontal and Y for vertical).
text_size	Character expansion factor for the labels of the experimental units.
point_size	Character expansion factor for the point representing the experimental units.

Value

This function just creates the units plot. It does not return a value.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.

Examples

```
#Example: NCI-60 microRNA Data
data("Topoisomerase_II_Inhibitors")
data("microRNA")
my_res <- frcc(t(microRNA), -1*t(Topoisomerase_II_Inhibitors))
for( i in 1:dim(microRNA)[2])
{
  colnames(microRNA)[i]<-substr(colnames(microRNA)[i], 1, 2)
}#end for i
grDevices::dev.new()
plot_units(t(microRNA), -1*t(Topoisomerase_II_Inhibitors), my_res, 1, 1, text_size=0.01)
```

plot_variables

Plot variables in the Canonical Factor Loadings Space

Description

This function plots the variables used in the FRCCA as points in a two-dimensional plane in which the axis are the canonical factor loadings selected by the user.

Usage

```
plot_variables(res.mrcc, i, j, inner_circle_radius = 0.5, text_size = 0.8)
```

Arguments

<code>res.mrcc</code>	List containing a canonical structure provided by the function <code>frcc</code> .
<code>i</code>	Canonical Factor Loadings which will be used as the horizontal axis.
<code>j</code>	Canonical Factor Loadings which will be used as the vertical axis.
<code>inner_circle_radius</code>	Radius of the circle which is used to determine which variables are significant. Only the significant variables will be labeled.
<code>text_size</code>	Character expansion factor for the labels of the variables.

Value

This function just creates the variables plot. It does not return a value.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.

Examples

```
# Example: Multivariate Normal Data
p<-10
q<-10
n<-50
res<-generate_multivariate_normal_sample(p,q,n)
X<-res$X
Y<-res$Y
rownames(X)<-c(1:n)
colnames(X)<-c(1:p)
colnames(Y)<- c(1:q)
my_res<-frcc(X,Y)
grDevices::dev.new()
plot_variables(my_res,1,2,text_size=1.0)
```

<code>rearrange.frcc</code>	<i>Rearranges the canonical structure according to the canonical correlations</i>
-----------------------------	---

Description

By using the minimum risk estimators of the correlation matrices instead of the sample correlation matrices the FRCC algorithm might disrupt the order of the canonical correlations and hence of the canonical structure. This is unacceptable for the algorithm used to calculate the p-values which requires the canonical correlations to be ordered in a descending order. This function rearranges the canonical structure according to the canonical correlations from largest to smallest.

Usage

```
rearrange.frcc(res.frcc)
```

Arguments

`res.frcc` List containing a canonical structure produced by the function `frcc`.

Value

A list containing the sorted canonical structure.

Author(s)

Raul Cruz-Cano

References

Cruz-Cano, R.; Lee, M.L.T.; Fast Regularized Canonical Correlation Analysis, Computational Statistics & Data Analysis, Volume 70, 2014, Pages 88-100, ISSN 0167-9473, <https://doi.org/10.1016/j.csda.2013.09.020>.

Index

* **Internal**

- custom.draw.circle, 2
- off.diagonal.lambda, 6
- rearrange.frcc, 8

* **datasets**

- microRNA, 5

custom.draw.circle, 2

frcc, 2

generate_multivariate_normal_sample, 4

microRNA, 5

off.diagonal.lambda, 6

plot_units, 6

plot_variables, 7

rearrange.frcc, 8

soilspec (microRNA), 5

Topoisomerase_II_Inhibitors (microRNA),
5