Package ‘rrr’
December 9, 2016

Title Reduced-Rank Regression
Version 1.0.0

URL http://github.com/chrisaddy/rrr

Description Reduced-rank regression, diagnostics and graphics.

Depends R (>= 3.2.0)
Imports Rcpp, MASS, magrittr, dplyr, ggplot2, plotly, GGally
Suggests testthat, knitr, rmarkdown, readr

rrr.R data.R

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 5.0.1

VignetteBuilder knitr

NeedsCompilation no

Author Chris Addy [aut, cre]

Maintainer Chris Addy <chris.william.addy@gmail.com>

Repository CRAN

Date/Publication 2016-12-09 15:15:55

R topics documented:

COMBO17 ......................................................... 2
pairwise_plot .................................................... 2
pendigits ........................................................ 3
rank_trace ...................................................... 4
residuals ......................................................... 5
rrr ............................................................... 6
scores .......................................................... 8
threewise_plot ............................................... 9
tobacco ........................................................ 10
Description

COMBO-17 galaxy photometric catalogue, 216, 219, 235

Usage

COMBO17

Format

A data frame with 3462 observations on 65 numeric variables.

References


Description

Pairwise Plots

Usage

pairwise_plot(x, y, type = "pca", pair_x = 1, pair_y = 2, rank = "full", k = 0, interactive = FALSE, point_size = 2.5)

Arguments

x  
data frame or matrix of predictor variables

y  
data frame or matrix of response variables

type  
type of reduced-rank regression model to fit. type = "identity", the default, uses $\Gamma = I$ to fit a reduced-rank regression. type = "pca" fits a principal component analysis model as a special case of reduced-rank regression. type = "cva" fits a canonical variate analysis model as a special case of reduced-rank regression. type = "lda" fits a linear discriminant analysis model as a special case of reduced-rank regression.
pair_x variable to be plotted on the X-axis
pair_y variable to be plotted on the Y-axis
rank rank of coefficient matrix.
k small constant added to diagonal of covariance matrices to make inversion easier.
interactive logical. If interactive = FALSE, the default, plots a static pairwise plot. If interactive = FALSE plots an interactive pairwise plot.
point_size size of points in scatter plot.

Value

ggplot2 object if interactive = FALSE; plotly object if interactive = TRUE.

References


Examples

data(pendigits)
digits_features <- pendigits[,1:34]
digits_class <- pendigits[,35]
pairwise_plot(digits_features, digits_class, type = "pca", pair_x = 1, pair_y = 3)

library(dplyr)
data(COMBO17)
galaxy <- as_data_frame(COMBO17)
galaxy <- select(galaxy, -starts_with("e"), -Nr, -IFS:-IFD)
galaxy <- na.omit(galaxy)
galaxy_x <- select(galaxy, -Rmag:-chi2red)
galaxy_y <- select(galaxy, Rmag:chi2red)
pairwise_plot(galaxy_x, galaxy_y, type = "cva")

data(iris)
iris_x <- iris[,1:4]
iris_y <- iris[5]
pairwise_plot(iris_x, iris_y, type = "lda")

Description

pen-based handwritten digit recognition, 211, 234, 274, 348, 391, 631
Usage

pendigits

Format

a data frame with 10992 observations on 36 unnamed variables

Source


References


---

**rank_trace**

**Rank Trace Plot**

Description

rank_trace is a plot used to determine the effective dimensionality, i.e., \( t = \text{rank}(C) \), of the reduced-rank regression equation.

Usage

```r
rank_trace(x, y, type = "identity", k = 0, plot = TRUE, interactive = FALSE)
```

Arguments

- `x` data frame or matrix of predictor variables
- `y` data frame or matrix of response variables
- `type` type of reduced-rank regression model to fit. `type = "identity"`, the default, uses \( \Gamma = I \) to fit a reduced-rank regression. `type = "pca"` fits a principal component analysis model as a special case of reduced-rank regression. `type = "cva"` fits a canonical variate analysis model as a special case of reduced-rank regression. `type = "lda"` fits a linear discriminant analysis model as a special case of reduced-rank regression.
- `k` small constant added to diagonal of covariance matrices to make inversion easier.
- `plot` if FALSE, returns data frame of rank trace coordinates.
- `interactive` if TRUE, creates an interactive plotly graphic.

Value

plot of rank trace coordinates if `plot = TRUE`, the default, or data frame of rank trace coordinates if `plot = FALSE`. 
residuals

References


Examples

```r
data(tobacco)
tobacco_x <- tobacco[,4:9]
tobacco_y <- tobacco[,1:3]
gamma <- diag(1, dim(tobacco_y)[2])
rank_trace(tobacco_x, tobacco_y)
rank_trace(tobacco_x, tobacco_y, plot = FALSE)
rank_trace(tobacco_x, tobacco_y, type = "cva")
data(pendigits)
digits_features <- pendigits[, -35:-36]
rank_trace(digits_features, digits_features, type = "pca")
library(dplyr)
data(COMBO17)
galaxy <- as_data_frame(COMBO17)
galaxy <- select(galaxy, starts_with("e."), -Nr, -UFS:-IFD)
galaxy <- na.omit(galaxy)
galaxy_x <- select(galaxy, -Rmag:-chi2red)
galaxy_y <- select(galaxy, Rmag:chi2red)
rank_trace(galaxy_x, galaxy_y, type = "cva")
```

**residuals**  
*Reduced-Rank Regression Residuals*

Description

residuals calculates the regression residuals for reduced-rank regression and canonical variate analysis.

Usage

```r
residuals(x, y, type = "identity", rank = "full", k = 0, plot = TRUE)
```

Arguments

- **x**: data frame or matrix of predictor variables
- **y**: data frame or matrix of response variables
- **type**: type of reduced-rank regression model to fit. type = "identity", the default, uses $\Gamma = I$ to fit a reduced-rank regression. type = "pca" fits a principal component analysis model as a special case of reduced-rank regression. type = "cva" fits a canonical variate analysis model as a special case of reduced-rank regression. type = "lda" fits a linear discriminant analysis model as a special case of reduced-rank regression.
Fit Reduced-Rank Regression Model

**Description**

`rrr` fits a reduced-rank regression model.

**Usage**

```r
rrr(x, y, type = "identity", rank = "full", k = 0)
```

**Value**

scatterplot matrix of residuals if `plot` = TRUE, the default, or a data frame of residuals if `plot` = FALSE.

**References**

**Arguments**

- `x`: data frame or matrix of predictor variables
- `y`: data frame or matrix of response variables
- `type`: type of reduced-rank regression model to fit. `type = "identity"`, the default, uses $\Gamma = I$ to fit a reduced-rank regression. `type = "pca"` fits a principal component analysis model as a special case of reduced-rank regression. `type = "cva"` fits a canonical variate analysis model as a special case of reduced-rank regression. `type = "lda"` fits a linear discriminant analysis model as a special case of reduced-rank regression.
- `rank`: rank of coefficient matrix.
- `k`: small constant added to diagonal of covariance matrices to make inversion easier.

**Value**

A list containing estimates of coefficients and means, and eigenvalue-based diagnostics.

**References**


**Examples**

```r
data(tobacco)
tobacco_x <- tobacco[,4:9]
tobacco_y <- tobacco[,1:3]
rrr(tobacco_x, tobacco_y, rank = 1)

data(pendigits)
digits_features <- pendigits[, -35:-36]
rrr(digits_features, digits_features, type = "pca", rank = 3)

library(dplyr)
data(COMBO17)
galaxy <- as_data_frame(COMBO17)
galaxy <- select(galaxy, -starts_with("e"), -Nr, -UFS:-IFD)
galaxy <- na.omit(galaxy)
galaxy_x <- select(galaxy, -Rmag:-chi2red)
galaxy_y <- select(galaxy, Rmag:chi2red)
rrr(galaxy_x, galaxy_y, type = "cva", rank = 2)

data(iris)
iris_x <- iris[,1:4]
iris_y <- iris[5]
rrr(iris_x, iris_y, type = "lda")
```
scores

Compute Latent Variable Scores

Description

Compute Latent Variable Scores

Usage

scores(x, y, type = "pca", rank = "full", k = 0)

Arguments

- x: data frame or matrix of predictor variables
- y: data frame or matrix of response variables
- type: type of reduced-rank regression model to fit. type = "identity", the default, uses \( \Gamma = I \) to fit a reduced-rank regression. type = "pca" fits a principal component analysis model as a special case of reduced-rank regression. type = "cva" fits a canonical variate analysis model as a special case of reduced-rank regression. type = "lda" fits a linear discriminant analysis model as a special case of reduced-rank regression.
- rank: rank of coefficient matrix.
- k: small constant added to diagonal of covariance matrices to make inversion easier.

References


Examples

data(pendigits)
digits_features <- pendigits[, -35:-36]
scores(digits_features, digits_features, type = "pca", rank = 3)

library(dplyr)
data(COMBO17)
galaxy <- as_data_frame(COMBO17)
galaxy <- select(galaxy, -starts_with("e."), -Nr, -UFS:-IFD)
galaxy <- na.omit(galaxy)
galaxy_x <- select(galaxy, -Rmag:-chi2red)
galaxy_y <- select(galaxy, Rmag:chi2red)
scores(galaxy_x, galaxy_y, type = "cva", rank = 4)

data(iris)
iris_x <- iris[,1:4]
iris_y <- iris[5]
scores(iris_x, iris_y, type = "lda")
threewise_plot

3-D Reduced Rank Regression Plots

Description

Create three-dimensional, interactive plotly graphics for exploration and diagnostics.

Usage

threewise.plot(x, y, type = "pca", pair_x = 1, pair_y = 2, pair_z = 3,
               rank = "full", k = 0, point.size = 2.5)

Arguments

x       data frame or matrix of predictor variables
y       data frame or matrix of response variables
type    type of reduced-rank regression model to fit. type = "identity", the default, uses \( \Gamma = I \) to fit a reduced-rank regression. type = "pca" fits a principal component analysis model as a special case of reduced-rank regression. type = "cva" fits a canonical variate analysis model as a special case of reduced-rank regression. type = "lda" fits a linear discriminant analysis model as a special case of reduced-rank regression.
pair_x  variable to be plotted on the X-axis
pair_y  variable to be plotted on the Y-axis
pair_z  variable to be plotted on the Z-axis
rank    rank of coefficient matrix.
k       small constant added to diagonal of covariance matrices to make inversion easier.
point.size size of points in scatter plot.

Value

three-dimensional plot. If type = "pca" returns three principal components scores - defaulted to the first three - against each other. If type = "cva" returns three-dimensional plot of residuals. If type = "lda" returns three-dimensional plot of three linear discriminant scores plotted against each other.

Examples

## Not run:
data(pendigits)
digits_features <- pendigits[, -35:-36]
threewise.plot(digits_features, digits_class, type = "pca", k = 0.0001)

library(dplyr)
tobacco

**MMST TOBACCO DATA**

**Description**

chemical composition of tobacco, 183, 187

**Usage**

tobacco

**Format**

a data frame with 25 observations on the following 9 variables.

- ‘Y1.BurnRate’ a numeric vector
- ‘Y2.PercentSugar’ a numeric vector
- ‘Y3.PercentNicotine’ a numeric vector
- ‘X1.PercentNitrogen’ a numeric vector
- ‘X2.PercentChlorine’ a numeric vector
- ‘X3.PercentPotassium’ a numeric vector
- ‘X4.PercentPhosphorus’ a numeric vector
- ‘X5.PercentCalcium’ a numeric vector
- ‘X6.PercentMagnesium’ a numeric vector

**References**


Index

+Topic datasets
  COMBO17, 2
  pendigits, 3
  tobacco, 10

COMBO17, 2

pairwise_plot, 2
pendigits, 3

rank_trace, 4
residuals, 5
rrr, 6

scores, 8

threewise_plot, 9
tobacco, 10