Package 'morepls'

May 29, 2025

Type Package

Title Interpretation Tools for Partial Least Squares Regression

Version 0.2.1

Imports pls, ggplot2, ggrepel, rlang, descriptio

Suggests plsVarSel, ggforce

Description Various kinds of plots (observations, variables, correlations, weights, regression coefficients and Variable Importance in the Projection) and aids to interpretation (coefficients, Q2, correlations, redundancies) for partial least squares regressions computed with the 'pls' package, following Tenenhaus (1998, ISBN:2-7108-0735-1).

License GPL (>= 2)

Encoding UTF-8

RoxygenNote 7.3.2

URL https://framagit.org/nicolas-robette/morepls, https://nicolas-robette.frama.io/morepls/

NeedsCompilation no

Author Nicolas Robette [aut, cre]

Maintainer Nicolas Robette <nicolas.robette@uvsq.fr>

Repository CRAN

Date/Publication 2025-05-29 10:10:02 UTC

Contents

`get_coef` *Standardized and raw coefficients*

Description

Computes the standardized and raw coefficients of a PLS regression, with p-values and confidence intervals from a jackknife procedure.

Usage

```
get_coef(object, y = NULL, ncomp = NULL,
 ci = 0.95, raw = FALSE)
```

Arguments

`object`	an object of class `mvr` from `pls` package. It must be cross-validated with `jackknife = TRUE`
`y`	the name of the response variable whose coefficients are plotted. If `NULL` (default), the first response variable is used.
`ncomp`	the number of components to use for computing coefficients
`ci`	the confidence level of the confidence interval. Default is 0.95.
`raw`	logical. If `FALSE` (default), standardized coefficients are computed. If `TRUE`, raw coefficients are computed.

Value

A data frame with coefficients, standard deviation, t-values, p-values and confidence intervals.

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

`plo_coef`

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls",
           jackknife = TRUE)
res <- get_coef(pls)
head(res)
```

`get_cor` *Correlations between variables and scores*

Description

Computes correlations between variables and scores from a PLS regression.

Usage

```
get_cor(object)
```

Arguments

`object`	an object of class `mvr` from `pls` package.

Value

A list with the following elements :

`Xt`	correlations between X variables and X scores
`Yt`	correlations between Y variables and X scores
`Xu`	correlations between X variables and Y scores
`Yu`	correlations between Y variables and Y scores
`XY`	correlations between X variables and Y variables
`tu`	correlations between X scores and Y scores

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

get_red, plo_cor

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
res <- get_cor(pls)
str(res)
```

---

get_Q2 *Q2 and cumulative Q2 indexes*

Description

Computes Q2 and cumulative Q2 indexes from a PLS regression.

Usage

```
get_Q2(object)
```

Arguments

object	an object of class mvr from pls package. It has to be cross-validated

Value

A list with the following elements :

Q2kh	Q2 index by X variable and number of components
Q2h	Q2 index by number of components
Q2cumkh	cumulative Q2 index by X variable and number of components
Q2cumh	cumulative Q2 index by number of components

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
res <- get_Q2(pls)
str(res)
```

`get_red` *R2 and redundancies*

Description

Computes R2 and redundancies between variables and scores from a PLS regression.

Usage

```
get_red(object)
```

Arguments

`object`	an object of class `mvr` from `pls` package.

Value

A list with the following elements :

`Xt`	R2 and redundancies between X variables and X scores
`Yt`	R2 and redundancies between Y variables and X scores
`Xu`	R2 and redundancies between X variables and Y scores
`Yu`	R2 and redundancies between Y variables and Y scores
`Xtcum`	cumulative R2 and redundancies between X variables and X scores
`Ytcum`	cumulative R2 and redundancies between Y variables and X scores
`Xucum`	cumulative R2 and redundancies between X variables and Y scores
`Yucum`	cumulative R2 and redundancies between Y variables and Y scores

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

get_cor

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
res <- get_red(pls)
str(res)
```

---

plo_coef *Plot of coefficients*

Description

Plots the coefficients from a PLS regression.

Usage

```
plo_coef(object, y = NULL, ncomp = NULL, sort = FALSE,
                     col = "darkgreen", repel = FALSE,
                     max.pval = NULL, whiskers = FALSE, ci = 0.95)
```

Arguments

object	an object of class mvr from pls package
y	the name of the response variable whose coefficients are plotted. If NULL (default), the first response variable is used.
ncomp	the number of components to use for computing coefficients
sort	logical. If TRUE, bars are sorted by decreasing coefficients. Default is FALSE.
col	color of the bars
repel	logical. If TRUE, the names of the variables are repelled with geom_text_repel. Default is FALSE
max.pval	coefficients with jack-knife p-values higher than max.pval have a more transparent color bar. If NULL (default), all bars have the same opacity. If not NULL, object must be cross-validated with jackknife = TRUE.
whiskers	logical. If TRUE, whiskers are added to represent the confidence interval of the coefficients. Default is FALSE. If TRUE, object must be cross-validated with jackknife = TRUE.
ci	the confidence level of the confidence interval. Only used if whiskers is TRUE. Default is 0.95.

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) *Multivariate calibration*. Chichester: Wiley.

Tenenhaus, M. (1998) *La Regression PLS. Theorie et Pratique*. Editions TECHNIP, Paris.

See Also

plo_ctr, plo_vip, jack.test,

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls",
           jackknife = TRUE)
plo_coef(pls)
plo_coef(pls, max.pval = 0.05)
plo_coef(pls, whiskers = TRUE)
```

---

plo_cor *Plot of correlations*

Description

Plots the correlations between (X and Y) variables and the components (X scores) of a PLS regression.

Usage

```
plo_cor(object, comps = 1:2, which = "both", min.cor = NULL,
        lim = NULL, circles = NULL, col = NULL, size = 3.88)
```

Arguments

object	an object of class mvr from pls package
comps	the components to use. Default is c(1,2).
which	character string. If "both" (default), X and Y variables are plotted. If "X", only X variables are plotted. If "Y", only Y variables are plotted.
min.cor	numerical value. The minimal correlation with one or the other component for a variable to be plotted. If NULL (default), all the variables are plotted.
lim	numerical value. The limit of the scale (in absolute value). If NULL (default), the limits are automatically determined from the range of tha data.
circles	vector of numeric values. Circles are added to the plot at radiuses specified in circles. If NULL (default), no circle is plotted.
col	colors for the names of the variables. Only one value should be provided if which is "X" or "Y", a vector of two if which is "both". If NULL (default), colors are set automatically.
size	numerical value. The size of the names of the variables.

Value

a ggplot2 object

Note

This is what Tenenhaus calls the univariate interpretation of the PLS components, as opposed to the multivariate interpretation (see plo_var).

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

get_cor, plo_var

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
```

```
plo_cor(pls)
# plot with circles corresponding to
# correlations of 0.5 and 1
plo_cor(pls, lim = 1, circles = c(0.5, 1), col = c("pink", "purple"))
```

---

plo_ctr *Plot of weights*

Description

Plots the weights of X variables from a PLS regression.

Usage

```
plo_ctr(object, comp = 1, sort = FALSE, col = "tomato4", repel = FALSE)
```

Arguments

object	an object of class mvr from pls package
comp	the component to use. Default is 1.
sort	logical. If TRUE, bars are sorted by decreasing VIPs. Default is FALSE.
col	color of the bars
repel	logical. If TRUE, the names of the variables are repelled with geom_text_repel. Default is FALSE

Details

According to Tenenhaus, the contribution of a variable to the construction of a component is measured by the squared loading weight. For a given component, the sum af the squared loading weights is equal to 1. This plot represents the loading weights, which keeps the information about their sign. Dashed lines are added at +/- sqrt(1/p), with p the number of X variables, which corresponds to the average contribution to the construction of the component.

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

plo_coef, plo_vip

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
plo_ctr(pls)
```

---

plo_inter *Plot of Interactions*

Description

Plots the interaction between two categorical supplementary variables for a PLS regression.

Usage

```
plo_inter(object, var1, var2, excl1 = NULL, excl2 = NULL,
  comps = c(1,2), shapesize = 1, textsize = 3,
  force = 1, max.overlaps = Inf,
  lines = TRUE, dashes = TRUE)
```

Arguments

object	an object of class mvr from pls package
var1	factor. The first categorical supplementary variable.
var2	factor. The second categorical supplementary variable.
excl1	character vector of categories from the var1 to exclude from the plot. If NULL (default), all the supplementary categories are plotted.
excl2	character vector of categories from the var2 to exclude from the plot. If NULL (default), all the supplementary categories are plotted.
comps	the components to use. Default is c(1,2).
shapesize	Size of the shapes. Default is 1.
textsize	Size of the labels of categories. Default is 3.
force	Force of repulsion between overlapping text labels. Defaults to 1. If 0, labels are not repelled at all.
max.overlaps	Exclude text labels that overlap too many things. Defaults to Inf, which means no labels are excluded.

lines	logical. Whether to add colored lines between the points of the categories of v1. Default is TRUE.
dashes	logical. Whether to add gray dashed lines between the points of the categories of v2. Default is TRUE.

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

plo_sup, plo_part

Examples

```
library(pls)
data(mpg, package = "ggplot2")
pls <- mvr(displ ~ cty + hwy + cyl,
           ncomp = 3,
           data = mpg,
           validation = "CV",
           method = "oscorespls")
supvars <- data.frame(class = factor(mpg$class), trans = factor(mpg$trans))
plo_inter(pls, factor(mpg$class), factor(mpg$trans), lines = FALSE, dashes = FALSE)
```

---

plo_obs *Plot of scores*

Description

Plots the scores of the observations of a PLS regression.

Usage

```
plo_obs(object, comps = 1:2, col = "black", size = 1.5)
```

Arguments

object	an object of class mvr from pls package
comps	the components to use. Default is c(1,2).
col	the color of the points.
size	numerical value. The size of the points.

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
plo_obs(pls)
```

---

plo_part *Plot of Main and Partial Effect*

Description

Plots the the main and partial effects of a supplementary variable for a PLS regression, with one or more supplementary partialled out.

Usage

```
plo_part(object, var, controls, excl = NULL,
  comps = c(1,2), shapesize = 1.5, col = "black",
  textsize = 4, force = 1, max.overlaps = Inf,
  lines = TRUE, dashes = TRUE, alpha = 0.3, legend = "right")
```

Arguments

object	an object of class mvr from pls package
var	factor. The categorical supplementary variable.
controls	data frame of supplementary variables to be partialled out (i.e. control variables)
excl	character vector of categories from the var to exclude from the plot. If NULL (default), all the supplementary categories are plotted.
comps	the components to use. Default is c(1,2).
shapesize	Size of the shapes. Default is 1.5.
col	the color for the labels and lines. Default is "black".
textsize	Size of the labels of categories. Default is 4.
force	Force of repulsion between overlapping text labels. Defaults to 1. If 0, labels are not repelled at all.
max.overlaps	Exclude text labels that overlap too many things. Defaults to Inf, which means no labels are excluded.
lines	logical. Whether to add colored lines between the points of the categories of v1. Default is TRUE.
dashes	logical. Whether to add gray dashed lines between the points of the categories of v2. Default is TRUE.
alpha	Numerical value. Transparency of the partial effects. Default is 0.3.
legend	the position of legends ("none", "left", "right", "bottom", "top", or two-element numeric vector). Default is right.

Value

a ggplot2 object

Note

The partial effects of the supplementary variable are computed with the Average Marginal Effects of a linear regression, with individual coordinates as dependent variable, and the supplementary and control variables as independent variables.

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

plo_sup, plo_inter

Examples

```
library(pls)
data(mpg, package = "ggplot2")
pls <- mvr(displ ~ cty + hwy + cyl,
           ncomp = 3,
           data = mpg,
           validation = "CV",
           method = "oscorespls")
plo_part(pls, factor(mpg$class), factor(mpg$trans), lines = FALSE, dashes = FALSE)
```

---

plo_sup *Plot of Supplementary Variables*

Description

Plots the categories of supplementary variables for a PLS regression.

Usage

```
plo_sup(object, vars, excl = NULL, comps = c(1,2),
 shapesize = 2, textsize = 3, vlab = TRUE, force = 1,
 max.overlaps = Inf, dashes = TRUE)
```

Arguments

object	an object of class mvr from pls package
vars	A data frame of categorical supplementary variables. All these variables should be factors.
excl	character vector of supplementary categories to exclude from the plot, specified in the form "namevariable.namecategory" (for instance "Gender.Men"). If NULL (default), all the supplementary categories are plotted.
comps	the components to use. Default is c(1,2).
shapesize	Size of the shapes. Default is 2.
textsize	Size of the labels of categories. Default is 3.
vlab	Logical. If TRUE (default), the variable name is added as a prefix for the labels of the categories.
force	Force of repulsion between overlapping text labels. Defaults to 1. If 0, labels are not repelled at all.
max.overlaps	Exclude text labels that overlap too many things. Defaults to Inf, which means no labels are excluded.
dashes	Logical. Should one add lines between categories ? Default is TRUE.

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

plo_var

Examples

```
library(pls)
data(mpg, package = "ggplot2")
pls <- mvr(displ ~ cty + hwy + cyl,
           ncomp = 3,
           data = mpg,
           validation = "CV",
           method = "oscorespls")
supvars <- data.frame(class = factor(mpg$class), trans = factor(mpg$trans))
plo_sup(pls, supvars, dashes = FALSE)
```

---

plo_var *Plot of loadings*

Description

Plots the loadings of the variables of a PLS regression.

Usage

```
plo_var(object, comps = 1:2, which = "both", col = NULL,
 size = 3.88, Yline = TRUE, col.Yline = "firebrick3")
```

Arguments

object	an object of class mvr from pls package
comps	the components to use. Default is c(1,2).
which	character string. If "both" (default), X and Y variables are plotted. If "X", only X variables are plotted. If "Y", only Y variables are plotted.
col	colors for the names of the variables. Only one value should be provided if which is "X" or "Y", a vector of two if which is "both". If NULL (default), colors are set automatically.
size	numerical value. The size of the names of the variables.

`Yline`	logical. If `TRUE` (default), a line is drawn through the origin and the coordinates of the response variable, and a second line orthogonal to the first one. This is aimed at facilitating the interpretation.
`col.Yline`	the color of the lines drawn if `Yline` is `TRUE`. Default is "firebrick3".

Value

a `ggplot2` object

Note

This is what Tenenhaus calls the multivariate interpretation of the PLS components, as opposed to the univariate interpretation provided by the correlations (see `plo_cor`). This superposes Y loadings (vectors from the C matrix) and projections, i.e. modified weights (vectors of the W* matrix).

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

`plo_cor`

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
plo_var(pls)
```

---

plo_vip *Plot of VIPs*

Description

Plots the Variable Importance in Projections (VIP) indexes of a PLS regression.

Usage

```
plo_vip(object, ncomp = NULL, sort = FALSE,
col = "steelblue4", repel = FALSE)
```

Arguments

object	an object of class mvr from pls package
ncomp	the number of components to use for computing VIPs
sort	logical. If TRUE, bars are sorted by decreasing VIPs. Default is FALSE.
col	color of the bars
repel	logical. If TRUE, the names of the variables are repelled with geom_text_repel. Default is FALSE

Value

a ggplot2 object

Author(s)

Nicolas Robette

References

Martens, H., Næs, T. (1989) Multivariate calibration. Chichester: Wiley.

Tenenhaus, M. (1998) La Regression PLS. Theorie et Pratique. Editions TECHNIP, Paris.

See Also

VIP

Examples

```
library(pls)
data(yarn)
pls <- mvr(density ~ NIR,
           ncomp = 5,
           data = yarn,
           validation = "CV",
           method = "oscorespls")
plo_vip(pls)
```

Index