Package 'ChemoSpec'

October 16, 2009

Type Package

Title Exploratory Chemometrics for Spectroscopy

Version 1.2

Date 2009-10-15

Author Bryan A. Hanson DePauw University, Greencastle Indiana USA

Maintainer Bryan A. Hanson <hanson@depauw.edu>

Description A collection of functions for plotting spectra (NMR, IR etc) and carrying out various forms of exploratory data analysis, such as HCA and PCA. The design allows comparison of data from samples which fall into groups such as treatment vs. control. Robust methods appropriate for this type of high-dimensional data are available. ChemoSpec is designed to be very user friendly and suitable for people with limited background in R.

License GPL-3

LazyLoad yes

Depends chemometrics, robustbase, RColorBrewer, plyr, pcaPP, mvtnorm, mvoutlier, pls, lattice, grid

URL http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

R topics documented:

`ChemoSpec-package` *Exploratory Chemometrics for Spectroscopy*

Description

A collection of functions for plotting spectra (NMR, IR etc) and carrying out various forms of exploratory data analysis, such as HCA and PCA. The design permits comparison of data from samples which fall into groups such as treatment vs. control. Robust methods appropriate for this type of high-dimensional data are available. ChemoSpec is designed to be very user friendly and suitable for people with limited background in R. A vignette illustrating typical operations is available.

Details

Package:	ChemoSpec
Type:	Package
Version:	1.1
Date:	2009-09-23
License:	GPL-3
LazyLoad:	yes

Author(s)

Bryan A. Hanson, DePauw University, Greencastle Indiana USA

Maintainer: Bryan A. Hanson ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

chkSpectra *Verify the Integrity of a Spectra Object*

Description

Utility function to verify that the structure of a "Spectra" object (an instance of an S3 object) is internally consistent. Rather than directly manipulating a "Spectra" object, one should manipulate it via trimNbin or removeSample. Should not see much direct use by users.

Usage

```
chkSpectra(spectra, confirm = FALSE)
```

Arguments

spectra	An object of S3 class "Spectra" to be checked.
confirm	Logical indicating whether or not to write the results to screen, as would be desirable for interactive use.

Details

This function is similar in spirit to validObject in the S4 world. When used at the console, and the object is OK, no message is written unless confirm = TRUE. At the console, if there is a problem, messages are issued regardless of the value of confirm. When used in a function, this function is silent (assuming confirm = FALSE) unless there is a problem.

Value

None; messages will be printed at the console if there is a problem.

Called by

A majority of ChemoSpec functions.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

Examples

```
data(CuticleIR)
chkSpectra(CuticleIR, confirm = TRUE) # OK
# What's next works, but is the wrong way to manipulate a this-is-escaped-code{ object.
# Use \code{\LinkA{trimNbin}{trimNbin}} or \code{\LinkA{removeSample}{removeSample}} inst
remove <- c(20:40)
CuticleIR$freq <- CuticleIR$freq[-remove]
chkSpectra(CuticleIR, confirm = TRUE) # not OK, you didn't listen to me!
```

`classPCA` *Classical PCA of Spectra Objects*

Description

A wrapper which carries out classical PCA analysis on a `"Spectra"` object. The data are row- and column-centered, and the user can select various options for scaling.

Usage

```
classPCA(spectra, choice = "noscale")
```

Arguments

`spectra`	An object of S3 class `"Spectra"`
`choice`	A character string indicating the choice of scaling. One of `c("noscale", "autoscale", "Pareto")`.

Details

The scale choice `autoscale` scales the columns by their standard deviation. `Pareto` scales by the square root of the standard deviation.

Value

An object of class `prcomp`, modified to include a list element called `$method`, a character string describing the pre-processing carried out and the type of PCA performed (it appears on plots which you might make).

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html
K. Varmuza and P. Filzmoser *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, 2009.

See Also

`prcomp` for the underlying function, `robPCA` for analogous robust PCA calculations.
For displaying the results, `plotScree`, `plotScores`, `plotLoadings`, `plot2Loadings`.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plotScores(CuticleIR, title = "Cuticle IR Spectra", results, pcs = c(1,2), ellipse = "rob
```

colLeaf *Color Dendrogram of a Spectra Object*

Description

Utility function called by HCA and ScoreHCA via dendrapply. Not intended for end-users.

Usage

```
colLeaf(n, spectra)
```

Arguments

n	The leaf number (an integer).
spectra	An object of S3 class "Spectra".

Value

None. The leaf colors of the relevant dendrogram object are modified to correspond to those in spectra

Called by

HCA, ScoreHCA

Note

The basic idea was found in the help archives. I can't write this kind of stuff!

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

CuticleIR *IR Ppectra of Plant Cuticles by Genotype and Treatment*

Description

IR spectra obtained by ATR sampling on leaves of *Portulaca oleracea* (common purslane). There are 157 spectra, divided into four groups from a G x E experiment. Two genotypes were studied, golden (G), and tall green (T). Two temperature regimes were employed, experimental (E, 35C) and control (C, 22C). Sample name GC10 means golden phenotype, control treatment, plant no. 11.

Usage

```
data(CuticleIR)
```

Format

List of 7
$ freq : num [1:1242] 501 503 505 507 509 ...
$ data : num [1:157, 1:1242] 0.205 0.247 0.219 0.203 0.234 ...
$ names : chr [1:157] "GC10" "GC11" "GC12" "GC14" ...
$ groups: Factor w/ 4 levels "GC","GE","TC",..: 1 1 1 1 1 1 1 1 1 1 ...
$ colors: chr [1:157] "brown2" "brown2" "brown2" "brown2" ...
$ unit : chr [1:3] "Wavenumbers" "Absorbance" "TRUE"
$ desc : chr "Kelly's Complete IR Data Set, Summer 2009"
- attr(*, "class")= chr "Spectra"

Details

Noisy regions at the extremes of the frequency range have been removed. The region from 1800 - 2500 wavenumbers was also removed as it is uninformative.

Source

Data obtained by Kelly Summers at DePauw University, Summer 2009.

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

Examples

```
data(CuticleIR)
chkSpectra(CuticleIR)
sumSpectra(CuticleIR) # this also runs chkSpectra() before doing the summary
```

---

getManyCsv *Merge CSV Files in a Directory into a Spectra Object*

Description

This function will read all .csv files in a directory, and use the file names to construct group membership and assign colors. All the data is placed into an object of S3 class "Spectra".

Usage

```
getManyCsv(gr.crit = NULL, gr.cols = c("auto"), freq.unit = "no frequency unit p
format = "original", int.unit = "no intensity unit provided", descrip = "no desc
```

Arguments

gr.crit	Group Criteria. A vector of character strings which will be searched for among the file names in order to assign an individual spectrum/sample to group membership.

gr.cols	Group Colors. Either the word "auto", in which case colors will be automatically assigned, or a vector of acceptable color names with the same length as gr.crit. In the latter case, colors will be assigned one for one, so the first element of gr.crit is assigned the first element of gr.col and so forth. See details below for some other things to consider.
freq.unit	A character string giving the units of the x-axis (frequency or wavelength).
NMR	A logical indicating if the type of spectra is traditionally plotted with higher numbers on the left (e.g. NMR or IR).
format	A character string giving the format of the csv files to be processed. Currently set to "original" but not used; intended for future flexibility.
int.unit	A character string giving the units of the y-axis (some sort of intensity).
descrip	A character string describing the data set that will be stored. This string is used in some plots so it is recommended that it's length be less than about 40 characters.
out.file	A file name acceptable to the save function. The completed object of S3 class "Spectra" will be written to this file.

Details

The linking of groups with colors is handled by groupNcolor. The user may specify any color name known to R. However, if you plan to use rggobi and GGobi to view the data later, keep in mind that ggobi only uses certain color schemes (although there are many options). In the case of ChemoSpec, two particular options have been hard-coded into the function plotScoresG for simplicity. If you plan to use plotScoresG, you should use choose from one of the two color schemes now if you want all your graphics to use the same scheme. Keep in mind that these colors must be used in order (though you can use the order of argument gr.crit to associate a particular group with a particular color:

primary scheme: `c("red3", "dodgerblue4", "forestgreen", "purple4", "orangered", "yellow", "orangered4", "violetred2")`

pastel scheme: `c("seagreen", "brown2", "skyblue2", "hotpink3", "chartreuse3", "darkgoldenrod2", "lightsalmon3", "gray48")`

If you want to see what these colors look like, use `display.brewer.pal(8, "Set1)` or `display.brewer.pal(8, "Set2)`. Finally, the difficulty to bear in mind here is that R plots are generally on a white background, so pale colors should be avoided, while GGobi plots on a black background, so dark colors should be avoided!

Value

An object of S3 class "Spectra" and name "spectra" will be written to out.file. You can change it's name later by loading it and assigning it to a new name.

Calls

groupNcolor

Called by

Top level function, called by user.

Warning

Files whose names are not matched using gr.crit are still incorporated into the "Spectra" object, but they are not assigned a group or color and therefore don't plot, though they do take up space in a plot! I will fix this eventually.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

---

groupNcolor *Assign Group Membership and Colors for a Spectra Object*

Description

A utility function which looks for gr.crit in the file names of .csv files and assigns group membership. Also assigns a color to each group. Not intended for users.

Usage

```
groupNcolor(spectra, gr.crit = NULL, gr.cols = c("auto"))
```

Arguments

spectra	An object of S3 class "Spectra". Until this function acts on "Spectra" it is not quite complete.
gr.crit	As per getManyCsv
gr.cols	As per getManyCsv

Value

A *complete* object of S3 class "Spectra". Until this function has done its job, an object of class "Spectra" will not pass checks as the assembl is not complete (see chkSpectra).)

Called by

getManyCsv

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

getManyCsv for details.

HCA *Plot HCA Results of a Spectra Object*

Description

A wrapper which carries out HCA and plots a dendrogram colored by the information in a "Spectra" object. All methods for computing the cluster distances are available.

Usage

```
HCA(spectra, title = "no title provided", method = "complete", ...)
```

Arguments

spectra	An object of S3 class "Spectra".
title	A character string for the plot title.
method	A character string acceptable as a method in hclust.
...	Other parameters to be passed to the plotting functions.

Value

None. The side effect is a plot.

Calls

chkSpectra, colLeaf

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

hclust for the underlying function. ScoreHCA for similar analysis of PCA scores from a "Spectra" object.

Examples

```
data(CuticleIR)
HCA(CuticleIR, title = "Cuticle IR Spectra", method = "complete")
```

labelExtremes *Label Extreme Values in a Data Set*

Description

A utility function which plots the sample names next to the sample points. The number of samples labeled can be specified by passing it from the calling function. Never called by the user.

Usage

```
labelExtremes(data, names, tol)
```

Arguments

data	A matrix containing the x values of the points/samples in the first column, and the y values in the second.
names	A character vector of sample names. Length must match the number of rows in x.
tol	A number describing the fraction of points to be labeled. `tol = 1.0` labels all the points; `tol = 0.05` labels the most extreme 5 percent.

Value

None. Annotates the plot with labels.

Called by

plotScoresDecoration

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

pcaBoot *Cross-Validation of Classical PCA Results for a Spectra Object*

Description

This function carries out classical PCA on the data in a "Spectra" object using a cross-validation method. Nothing more than a wrapper to Peter Filzmoser's pcaCV method with some small plotting changes.

Usage

```
pcaBoot(spectra, pcs, choice = "noscale", repl = 50, segments = 4, segment.type
```

Arguments

spectra	An object of S3 class "Spectra".
choice	A character string indicating the choice of scaling. One of c("noscale", "autoscale", "Pareto").
pcs	As per pcaCV where it is called amax; an integer giving the number of PC scores to include.
repl	As per pcaCV; the number of replicates to perform.
segments	As per pcaCV
segment.type	As per pcaCV
length.seg	As per pcaCV
trace	As per pcaCV
...	Parameters to be passed to the plotting routines

Value

None (????). Side effect is a plot.

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html
K. Varmuza and P. Filzmoser *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, 2009.

See Also

pcaCV for the underlying function.

Examples

```
data(CuticleIR)
results <- pcaBoot(CuticleIR, pcs = 5, choice = "noscale")
```

pcaDiag *Outlier Diagnostic Plots for PCA of a Spectra Object*

Description

A function to carry diagnostics on the PCA results for a "Spectra" object. Basically a wrapper to Filzmoser's pcaDiagplot which colors everything according to the scheme stored in the "Spectra" object. Works with PCA results of either class "prcomp" or class "princomp". Works with either classical or robust PCA results.

Usage

```
pcaDiag(spectra, pca, pcs = 3, quantile = 0.975, plot = c("OD", "SD"), ...)
```

Arguments

spectra	An object of S3 class "Spectra".
pca	An object of class prcomp or prcomp, modified to include a character string ($method) describing the pre-processing carried out and the type of PCA performed.
pcs	As per pcaDiagplot. The number of principal components to include.
quantile	As per pcaDiagplot. The significance criteria to use as a cutoff.
plot	A character string, indicating whether to plot the score distances or orthogonal distances, or both. Options are c("OD", "SD").
...	Additional parameters to be passed to the plotting functions.

Details

If both plots are desired, the output should be directed to a file rather than the screen. Otherwise, the 2nd plot overwrites the 1st in the active graphics window. Alternatively, just call the function twice, once specifying OD and once specifying SD.

Value

A list is returned as described in pcaDiagplot, so the result must be assigned or it will appear at the console. Side effect is a plot.

Calls

q2rPCA, labelExtremes

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html
K. Varmuza and P. Filzmoser *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, 2009.

See Also

pcaDiagplot in package chemometrics for the underlying function.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
temp <- pcaDiag(CuticleIR, results, pcs = 2, plot = "OD")
```

---

plot2Loadings *Plot PCA Loadings from a Spectra object Against Each Other*

Description

Plots two PCA loadings specified by the user, and labels selected (extreme) points. Typically used to determine which variables (frequencies) are co-varying, although in spectroscopy most peaks are represented by several variables and hence there is a lot of co-varying going on. Also useful to determine which variables are contributing the most to the clustering on a score plot.

Usage

```
plot2Loadings(spectra, pca, title = "no title provided", loads = c(1, 2), tol =
```

Arguments

spectra	An object of S3 class "Spectra".
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
loads	A vector of two integers specifying which loading vectors to plot.
tol	A number describing the fraction of points to be labeled. tol = 1.0 labels all the points; tol = 0.05 labels the most extreme 5 percent.
...	Other parameters to be passed to the plotting routines.

Value

None. Side effect is a plot.

Calls

labelExtremes

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

See plotLoadings to plot one loading against the original variable (frequency) axis.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plot2Loadings(CuticleIR, title = "Cuticle IR Spectra", results, loads = c(1,2), tol = 0.0
```

---

plotLoadings *Plot PCA Loadings for a Spectra Object*

Description

Creates a multi-panel plot of loadings along with a reference spectrum.

Usage

```
plotLoadings(spectra, pca, title = "no title provided", loads = c(1), ref = 1, .
```

Arguments

spectra	An object of S3 class "Spectra"
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
loads	An integer vector giving the loadings to plot. More than 3 loadings creates a useless plot using the default graphics window.
ref	An integer specifying the reference spectrum to plot, which appears at the bottom of the plot.
...	Additional parameters to be passed to plotting functions.

Value

None. Side effect is a plot.

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

See plot2Loadings to plot two loadings against each other.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plotLoadings(CuticleIR, results, title = "Cuticle IR Spectra", loads = c(1), ref = 1)
```

---

plotScores *Plot PCA Scores of a Spectra Object*

Description

Plots the requested PCA scores using the color scheme derived from a "Spectra" object. Options are provided to add confidence ellipses for each group in the object. The ellipses may be robust or classical. Option to label the extreme points provided.

Usage

```
plotScores(spectra, pca, title = "no title provided", pcs = c(1, 2), ellipse = "
```

Arguments

spectra	An object of S3 class "Spectra"
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
pcs	A vector of two integers specifying the PCA scores to plot.
ellipse	A character vector specifying the type of ellipses to be plotted. One of c("both", "none", "cls", "rob"). cls specifies classical confidence ellipses, rob specifies robust confidence ellipses.
tol	A number describing the fraction of points to be labeled. tol = 1.0 labels all the points; tol = 0.05 labels the most extreme 5 percent.
...	Additional parameters to be passed to the plotting functions.

Value

None. Side effect is a plot.

Calls

plotScoresDecoration, plotScoresCor, chkSpectra

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

For other ways of displaying the results, plotScree, plotLoadings, plot2Loadings. For a 3D plot of the scores, see plotScores3D.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plotScores(CuticleIR, title = "Cuticle IR Spectra", results, pcs = c(1,2), ellipse = "bot
```

---

plotScores3D *3D PCA Score Plot for a Spectra Object*

Description

Creates a basic 3D plot of PCA scores from the analysis of a "Spectra" object, color coded according the to scheme stored in the object.

Usage

```
plotScores3D(spectra, pca, pcs = c(1:3), title = "no title provided", view = lis
```

Arguments

spectra	An object of S3 class "Spectra"
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
pcs	A vector of three integers specifying the PCA scores to plot.

`view` A list of viewing transformations to be applied to the data. May contain values for x, y and z axes; keep in mind that the order of the transformations is important. For example, specifying `view = list(x = 45, y = 10)` produces a different view than `view = list(y = 10, x = 45)`. The list may be as along as you like - the series of transformations representing an accumulation of tweaks to achieve the desired view.

Value

None. Side effect is a plot.

Calls

`chkSpectra`

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

For a 2D plot of the scores, see `plotScores`. For more sophisticated 3D plots, use ggobi.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plotScores3D(CuticleIR, results, title = "Cuticle IR Spectra")
```

---

plotScoresCor *Compute Confidence Ellipses*

Description

A utility function which when given a x,y data set computes both classical and robust confidence ellipses. Never called by the user.

Usage

```
plotScoresCor(x, quan = 1/2, alpha = 0.025)
```

Arguments

x	As per cor.plot.
quan	As per cor.plot.
alpha	As per cor.plot.

Value

A list with the following elements (a simpler version of that in the original function cor.plot):

x.cls	The x values for the classical ellipse.
y.cls	The y values for the classical ellipse.
c	The correlation value for the classical ellipse.
x.rob	The x values for the robust ellipse.
y.rob	The y values for the robust ellipse.
r	The correlation value for the robust ellipse.

Called by

plotScores

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

See function cor.plot in package mvoutlier on which this function is based.

---

plotScoresDecoration

Decorate PCA Score Plot of a Spectra Object

Description

Utility function to carry out misc. labeling functions on the PCA score plot of a "Spectra" object. Never called by the user.

Usage

```
plotScoresDecoration(spectra, pca, pcs = c(1, 2), tol = "none")
```

Arguments

spectra	An object of S3 class "Spectra"
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
pcs	A vector of two integers specifying the PCA scores to plot.
tol	A number describing the fraction of points to be labeled. tol = 1.0 labels all the points; tol = 0.05 labels the most extreme 5 percent.

Value

None. The score plot is decorated.

Calls

labelExtremes

Called by

plotScores

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

---

plotScoresG *Plot PCA Scores of a Spectra Object in GGobi*

Description

Plots the specified PCs from a PCA analysis of a Spectra object using GGobi.

Usage

```
plotScoresG(spectra, pca, pcs = c(1:3), scheme = "primary")
```

Arguments

spectra	An object of S3 class "Spectra".
pca	An object of class prcomp.
pcs	A vector of integers specifying the scores to plot. GGobi handles multivariate data, and so it is not necessary to use only 3 dimensions.
scheme	The color scheme to use, one of c("primary", "pastel"). See details for more on color choices.

Details

In general, users many specify any valid R color for use in ChemoSpec. However, the color schemes available to GGobi are limited. If you want all your graphics to use the same color scheme, you need to specify a color scheme useable with GGbi at the time the Spectra object is created using the function getManyCsv. Two particular color schemes have been hard-coded into the function described here; the choices are:

primary: `c("red3", "dodgerblue4", "forestgreen", "purple4", "orangered", "yellow", "orangered4", "violetred2")`

pastel: `c("seagreen", "brown2", "skyblue2", "hotpink3", "chartreuse3", "darkgoldenrod2", "lightsalmon3", "gray48")`

These are very close matches to the color schemes used in GGobi, which in turn come from RColorBrewer. If you want to see what these colors look like, use `display.brewer.pal(8, "Set1)` or `display.brewer.pal(8, "Set2)`.

The behavior of this function is to first check to see if a GGobi-compatible color scheme already exists in spectra. If it does, it is re-used with the same mapping of colors to groups as originally specified and stored in spectra. If the color scheme is not GGobi-compatible, color assignments for GGobi window are generated and the mapping of old colors to new colors is reported at the console.

Finally, a difficulty to bear in mind is that R plots are generally on a white background, so pale colors should be avoided, while GGobi plots on a black background (interactively), so dark colors should be avoided! Plus, ggplot2 plots on a grey background by default! So it is not necessarily possible to specify a single color scheme that works with all possible plotting systems in R.

Value

None. Side effect is a plot in an X11 window created by GGobi. The mapping of colors from the Spectra object to the GGobi color schemes is printed at the console, since the original color choices may not be acceptable to GGobi.

Calls

Requires that GGobi be installed and X11 running.

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html
http://www.ggobi.org/
D. Cook & D. F. Swayne, Interactive and Dynamic Graphics for Data Analysis with R and GGobi, Springer 2007

See Also

For plotting only 3 scores with rudimentary control over the view, see plotScores3D. For plotting 2 scores, see plotScores.

Examples

```
data(CuticleIR)
require(ggobi)
results <- classPCA(CuticleIR, choice = "noscale")
## Not run: plotScoresG(CuticleIR, results) # GGobi runs interactively using X11
```

---

plotScree *Scree Plot of PCA Results for a Spectra Object*

Description

Function to draw a scree plot illustrating the importance of the components in a PCA analysis of a "Spectra" object.

Usage

```
plotScree(pca, title = "no title provided", ...)
```

Arguments

pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
...	Additional parameters to be passed to plotting functions.

Details

Technically, if you add $method to the PCA results, this will plot a scree plot for any PCA results, not just those from "Spectra" objects.

Value

None. Side effect is a plot.

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
plotScree(results, title = "Cuticle IR Spectra")
```

`plotSpectra` *Plot Spectra Object*

Description

Plots the spectra stored in a `"Spectra"` object. One may choose which spectra to plot, and the x range to plot. Spectra may be plotted offset or stacked. The vertical scale is controlled by a combination of several parameters.

Usage

```
plotSpectra(spectra, title = "no title provided", which = c(1), xrange = range(s
yrange = c(0, max(spectra$data)), offset = 0, amplify = 1, lab.pos = mean(spectr
```

Arguments

`spectra`	An object of S3 class `"Spectra"`
`title`	A character string for the plot title.
`which`	An integer vector specifying which spectra to plot, and the order.
`xrange`	A vector giving the limits of the x axis desired, for instance `c(200, 1000)`, expressed in the units of the stored spectra.
`yrange`	A vector giving the limits of the y axis desired, for instance `c(0, 15)`. This parameter depends upon the range of values in the stored spectra and defaults to the height of the largest peak in the data set. Interacts with the next two arguments, as well as the number of spectra to be plotted as given in `which`. Trial and error is used to adjust all these arguments to produce the desired plot.
`offset`	A number specifying the vertical offset between spectra if more than one is plotted. Set to 0.0 for a stacked plot.
`amplify`	A number specifying an amplification factor to be applied to all spectra. Useful for magnifying spectra so small features show up (though large peaks will then be clipped, unless you zoom on the x axis).
`lab.pos`	A number giving the location for the identifying label. Generally, pick an area that is clear in all spectra plotted. If no label is desired, give `lab.pos` outside the plotted x range.
`...`	Additional parameters to be passed to plotting functions.

Value

None. Side effect is a plot.

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

Examples

```
data(CuticleIR)
plotSpectra(CuticleIR, "Cuticle IR Spectra",
        which = c(10:11, 40:41, 100:101, 140:141, 150:151),
        xrange = c(3600, 500), yrange = c(0,10),
        offset = 0.8, amplify = 1.0, lab.pos = 2000)
```

---

q2rPCA *Conversion Between PCA Classes*

Description

Utility to convert objects of S3 class "prcomp" (Q-mode PCA) to objects of S3 class "princomp" (R-mode PCA) or *vice-versa*. Not likely to be called by most users.

Usage

```
q2rPCA(x)
r2qPCA(x)
```

Arguments

x	An object of either class "prcomp" or class "princomp". It will be converted to a form that can be used by functions expecting either class.

Details

In the conversion, the necessary list elements are added; the "old" elements are not deleted (and user added list elements are not affected). To indicate this, the class attribute is updated to include class "conPCA". The new object can then be used by functions expecting either class prcomp or princomp. For details of the structure of prcomp or princomp, see their respective help pages.

Value

A list of class "conPCA". Note that the order of the elements will vary depending upon the direction of conversion.

loadings	The loadings from "princomp", or a copy of the rotations from "prcomp".
scores	The scores from "princomp", or a copy of the x values from "prcomp".
call	The call. Objects of class "prcomp" do not store the original call, so a place holder is used. Otherwise the unchanged call from "princomp".
n.obs	The number of observations from "princomp", or computed from the 1st dimension of x in "prcomp".
class	"conPCA" is pre-pended to the existing class.
sdev	Unchanged from original.
center	Unchanged from original.
scale	Unchanged from original.

Called by

robPCA

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

prcomp, princomp

---

removeSample *Remove Samples from a Spectra Object*

Description

Removes specified samples from a "Spectra" object.

Usage

```
removeSample(spectra, rem.sam)
```

Arguments

spectra	An object of S3 class "Spectra"
rem.sam	Either an integer vector specifying the samples to be removed, or a character vector giving the sample names to be removed.

Details

If rem.sam is a character vector, the sample names are grepped for the corresponding values. Remember that the grepping process is greedy, i.e. grepping for "XY" find not only "XY" but also "XYZ".

Value

A modified object of S3 class "Spectra"

Calls

chkSpectra

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

trimNbin to remove selected frequencies from a "Spectra" object.

Examples

```
data(CuticleIR)
new1 <- removeSample(CuticleIR, rem.sam = 20) # removes the 20th spectrum/sample
new2 <- removeSample(CuticleIR, rem.sam = "GE") # removes all samples whose name contains
new3 <- removeSample(CuticleIR, rem.sam = "GE10") # removes one spectrum/sample with this
```

---

robPCA *Robust PCA of a Spectra Object*

Description

A wrapper which carries out robust PCA analysis on a "Spectra" object. The data are row- and column-centered, and the user can select various options for scaling.

Usage

```
robPCA(spectra, choice = "noscale")
```

Arguments

spectra	An object of S3 class "Spectra"
choice	A character vector describing the type of scaling to be carried out. One of c("noscale", "mad").

Value

An object of classes "conPCA" and "princomp" (see q2rPCA). It includes a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on plots which you might make).

Calls

chkSpectra, r2qPCA

Called by

Top level function, called by user.

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html
K. Varmuza and P. Filzmoser *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, 2009.

See Also

See PCAgrid on which this function is based. For the classical version, see classPCA.
For displaying the results, plotScree, plotScores, plotLoadings,

Examples

```
data(CuticleIR)
results <- robPCA(CuticleIR, choice = "mad")
plotScores(CuticleIR, title = "Cuticle IR Spectra", results, pcs = c(1,2), ellipse = "rob
```

ScoreHCA *HCA on PCA scores from a Spectra Object*

Description

A wrapper which performs HCA on the scores from a PCA of a "Spectra" object, color-coding the results as specified in the object.

Usage

```
ScoreHCA(spectra, pca, title = "no title provided", scores = c(1:5), method = "c
```

Arguments

spectra	An object of S3 class "Spectra"
pca	An object of class prcomp, modified to include a list element called $method, a character string describing the pre-processing carried out and the type of PCA performed (it appears on the plot). This is automatically provided if ChemoSpec functions classPCA or robPCA were used to create pca.
title	A character string for the plot title.
scores	A vector of integers specifying which scores to use for the HCA.
method	A character string acceptable as a method in hclust.
...	Additional parameters to be passed to the plotting functions.

Value

None. Side effect is a plot.

Calls

colLeaf

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

hclust for the underlying function. See HCA for HCA of the entire data set stored in the "Spectra" object.

Examples

```
data(CuticleIR)
results <- classPCA(CuticleIR, choice = "noscale")
ScoreHCA(CuticleIR,  results, scores = c(1:5), title = "Cuticle IR Spectra",  method = "c
```

---

specSurvey *Plot Std Dev of Frequencies in a Spectra Object*

Description

Computes the standard deviation of each frequency value for a "Spectra" object and plots it against the frequency. Useful for identifying uninformative spectral regions.

Usage

```
specSurvey(spectra, title = "No title provided", ...)
```

Arguments

spectra	An object of S3 class "Spectra"
title	A character string for the plot title.
...	Additional parameters to be passed to the plotting functions.

Value

None. Side effect is a plot.

Calls

chkSpectra

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

Examples

```
data(CuticleIR)
specSurvey(CuticleIR, title = "Cuticle IR Spectra")
```

Spectra *Spectra Objects*

Description

In ChemoSpec, spectral data sets are stored in an S3 class called Spectra, which contains a variety of information in addition to the spectra themselves. Spectra objects are created by getManyCsv or similar functions (no others currently exist).

Structure

The structure of a Spectra object is a list of 7 elements and an attribute as follows:

element	*type*	*description*
$freq	num	A common frequency (or wavelength) axis for all the spectra.
$data	num	The intensities for the spectra. A matrix of dimension no. samples x no. frequency points.
$names	chr	The sample names for the spectra; length must be no. samples.
$groups	Factor	The group classification of the samples; length must be no. samples.
$colors	chr	The colors for each sample; length must be no. samples. Groups and colors correspond.
$unit	chr	Three entries, the first giving the x axis unit, the second the y axis unit, and the third, a logical i
$desc	chr	A character string describing the data set. This appears on plots and therefore should probably l
- attr	chr "Spectra"	The S3 class designation.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html

See Also

sumSpectra to summarize a Spectra object. chkSpectra to verify the integrity of a Spectra object.

---

sumSpectra *Summarize a Spectra Object*

Description

Provides a summary of a "Spectra" object, essentially a more spectroscopist-friendly version of str().

Usage

```
sumSpectra(spectra)
```

Arguments

`spectra`	An object of S3 class `"Spectra"`

Details

Prior to summarizing, `chkSpectra` is run with confirm = FALSE. If there are problems, warnings are issued to the console and the summary is not done.

Value

None.

Warning

The computation of frequency resolution does not currently take into account any regions removed from the middle of the range and hence will be in error.

Calls

`chkSpectra`

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

`http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html`

Examples

```
data(CuticleIR)
sumSpectra(CuticleIR)
```

`trimNbin` *Edit and Bin Frequencies in a Spectra Object*

Description

Function to remove specified frequencies from a `"Spectra"` object, for instance, noisy or uninformative regions. The bin function is currently not implemented.

Usage

```
trimNbin(spectra, rem.freq)
```

Arguments

`spectra`	An object of S3 class `"Spectra"`
`rem.freq`	An integer vector of frequencies to be removed. See details.

Details

Frequencies should be given as integers. A certain amount of rounding will occur. A specification such as `rem.freq = c("min", 1500)` will remove all values from the minimum up to 1500. See the examples for further valid specifications. If `"min"` or `"max"` are used, the action is to create a range of frequencies to remove with the adjacent numerical value.

Value

A modified object of S3 class `"Spectra"`

Calls

`chkSpectra`

Called by

Top level function, called by user.

Author(s)

Bryan A. Hanson, DePauw University. ⟨hanson@depauw.edu⟩

References

`http://academic.depauw.edu/~hanson/ChemoSpec/ChemoSpec.html`

See Also

`removeSample` to remove one or more samples from a `"Spectra"` object.

Examples

```
data(CuticleIR)
new1 <- trimNbin(CuticleIR, rem.freq = c(750:1000)) # removes all frequencies in the rang
new2 <- trimNbin(CuticleIR, rem.freq = c("min", 1000)) # removes all frequencies from the
new3 <- trimNbin(CuticleIR, rem.freq = c("min", 1000, 3000, "max")) # frequencies between
new4 <- trimNbin(CuticleIR, rem.freq = c("min", 1000, 2000:2500, 3000, "max")) # frequenc
## Not run: new5 <- trimNbin(CuticleIR, rem.freq = c("min", 1000 "max")) # silliness, rem
```

Index