Package ‘labeltodendro’

February 20, 2015

Version 1.3
Date 2013-01-18
Title Convert labels or tables to a dendrogram
Author Vahid PARTOVI NIA, Anthony Davison, and Arpit Chaudhary
Maintainer Vahid PARTOVI NIA <vahid.partovin@polymtl.ca>
Depends R (>= 2.15.2)
Description The package offers a dendrogram representation of series
of labels, this is specially needed in Markov chain Monte Carlo
clustering. If you have a dendrogram in your mind you can
easily put series of meaningful labels in a matrix and heights
in a vector, then convert them to a dendrogram abject.
License GPL (>= 2)
URL http://labeltodendro.r-forge.r-project.org/
Repository CRAN
Repository/R-Forge/Project labeltodendro
Repository/R-Forge/Revision 16
Repository/R-Forge/DateTimeStamp 2013-01-18 08:24:42
Date/Publication 2013-01-27 08:13:24
NeedsCompilation yes

R topics documented:

  colorplot ................................................................. 2
  is.aggvec .............................................................. 3
  labclust ................................................................. 3
  labeltodendro ............................................................ 4
  labeltotable .............................................................. 5
  plot.labclust ............................................................ 6
  relabel ................................................................. 7
  relabel.matrix .......................................................... 8
  selectlabels ............................................................. 8
  tabletodendro ............................................................ 10
Index

\begin{itemize}
\item \textbf{colorplot} \hspace{1cm} \textit{colored dendrogram plot of a labclus object}
\end{itemize}

\textbf{Description}

The dendrogram is cut at the appropriate value, defined in a labclus object, and leaves are coloured respectively.

\textbf{Usage}

\begin{verbatim}
colorplot(x,h,horiz=FALSE,...)
\end{verbatim}

\textbf{Arguments}

\begin{itemize}
\item \textbf{x} \hspace{1cm} A labclus object to be plotted.
\item \textbf{h} \hspace{1cm} The height that the tree should be cut, automatically extracted from a labclus object.
\item \textbf{horiz} \hspace{1cm} Clarifies to plot the dendrogram horizontally or vertically.
\item \textbf{...} \hspace{1cm} Some of dendrogram plot options, but not all.
\end{itemize}

\textbf{Details}

The colorplot function is a modified version of HeatPlus package to cut a dendrogram and colour its leaves. Leaves with single element won’t be coloured.

\textbf{See Also}

\begin{verbatim}
plot.labclus
\end{verbatim}

\textbf{Examples}

\begin{verbatim}
x<-matrix(c(  c(1,1,1,1,1,1,1,1,1),  c(2,2,2,2,2,2,2,2,2), #exchangeable labels are allowed like this  c(1,1,1,1,1,1,1,1,1),  c(2,2,1,2,1,2,3,3,3),  c(1,1,2,1,1,3,3,3,3), #or like this  c(1,1,2,1,1,3,3,3,3),  c(1,1,2,3,4,5,5,5,5),  c(1,1,2,3,4,5,5,5,5),  c(1,2,3,4,5,6,7,8,9)), ncol=9, byrow=TRUE) #each iteration in a row

y<-labeltodendro(x) mydendro<-asNdendrogram(y) colorplot(y,horiz=TRUE,lwd=3,cluscol=c("red","green","blue")) #leaves with single subjects won’t get your specified color
plot(labeltodendro(x),horiz=TRUE) # changes orientation and dendrogram height starts from zero
\end{verbatim}
is.aggvec  
checks whether or not one vector is agglomeration of the another

Description

The function is useful to remove labels which are not agglomeration or division of each other.

Usage

is.aggvec(labagg, labdiv)

Arguments

labagg  A vector of integers.
labdiv  Another vector of integers.

Details

The function checks if labagg is agglomeration of labdiv. Exchangeable labels are allowed.

Examples

lab1 <- c(1, 2, 1, 2, 1, 2, 1, 2, 1)
lab2 <- c(6, 3, 6, 4, 6, 5, 6, 2, 6)  # exchangeable labels are allowed, 1 have the same role as 6
is.aggvec(lab1, lab2)  # should be TRUE
is.aggvec(lab1, lab2)  # should be FALSE

labclust  is an object

Description

The labclust object is very similar to the hclust object which additionally contains information about where the tree should be cut. You can convert any labclust object to a dendrogram by as.dendrogram

Examples

x <- matrix(c(
  c(1, 1, 1, 1, 1, 1, 1, 1, 1),
  c(2, 2, 2, 1, 2, 2, 2, 2),
  c(1, 1, 1, 1, 2, 2, 2, 2),
  c(2, 2, 1, 2, 3, 3, 3, 3),
  c(1, 1, 2, 1, 3, 3, 3, 3),
  c(1, 1, 2, 1, 3, 3, 3, 3),
  c(1, 1, 2, 3, 4, 5, 5, 5),
  c(1, 1, 2, 3, 4, 5, 5, 5),
  c(1, 1, 2, 3, 4, 5, 5, 5),
  c(1, 1, 2, 3, 4, 5, 5, 5),
), nrow = 10, byrow = TRUE)


c(1,1,2,3,4,5,5,5,5),
c(1,2,3,4,5,6,7,8,9)), ncol=9, byrow=TRUE)

y<-labeltodendro(x)
class(y) # can be regarded as hclust too

labeltodendro

converts series of labels to a labclust object

Description
The function calls other internal functions to count labels, make a table of agglomerations and divisions, and gives a labclust object.

Usage

labeltodendro(x,labels=NULL)

Arguments

x A numeric matrix of integers with iterations in rows and labels in columns. Each row of x corresponds to a grouping.

labels A vector of stings concering labels for each column of x; The names of associated to the clustered subjects.

Details
The most frequent label will be used to search for agglomerations and divisions and will be used for plotting and coloring the resulting dendrogram.

Value

merge Returns a matrix of merges according to hclust format.

height Gives heights of each branch of merge matrix.

hcut Shows where the tree should be cut.

See Also
tabledodendro, labeltotable, colorplot, labelst
Examples

```r
x<-matrix(c(
  c(1L,1L,1L,1L,1L,1L,1L,1L,1L), #exchangeable labels are allowed like this
  c(2L,2L,2L,2L,2L,2L,2L,2L,2L),
  c(1L,1L,1L,1L,1L,1L,1L,1L,1L),
  c(2L,2L,2L,2L,2L,2L,2L,2L,2L),
  c(1L,2L,1L,3L,3L,3L,3L,3L,3L), #or like this
  c(1L,2L,1L,3L,3L,3L,3L,3L,3L),
  c(1L,2L,3L,4L,5L,5L,5L,5L,5L),
  c(1L,2L,3L,4L,5L,5L,5L,5L,5L),
  c(1L,2L,3L,4L,5L,6L,7L,8L,9L), ncol=9, byrow=TRUE) #each iteration in a row
# this creates a matrix of labels

y<-labeltodendro(x)
mydendro<-as.dendrogram(y)
colorplot(y) #uses labclust plot
plot(mydendro) #uses dendrogram plot
plot(mydendro,center=TRUE)#you may like to center the dendrogram leaves
```

Description

The function is useful to make a summary table of Markov chain Monte Carlo labels for dendrogram representation. Exchangeable labels are allowed.

Usage

```r
labeltotable(x)
```

Arguments

- `x` a numeric matrix of integers with iterations in rows and labels in columns; each row of `x` corresponds to a grouping

Details

The function may not be computationally efficient for large number of iterations or large number of clustered subjects. This function counts labels and then orders the resulting table based on agglomeration and divisions of the most frequent label.
Value

labmat  A matrix of labels ordered in agglomerative order.
freq    The vector of frequencies corresponding to each row of labmat.

See Also

tabletodendro, labeltodendro, selectlabels

Examples

x<-matrix(c(
  c(1,1,1,1,1,1,1,1,1),  # exchangeable labels are allowed
  c(1,1,1,1,1,2,2,2,2),
  c(2,2,1,2,3,3,3,3),
  c(1,1,2,1,3,3,3,3),
  c(1,1,2,1,3,3,3,3),
  c(1,1,2,3,4,5,5,5),
  c(1,1,2,3,4,5,5,5),
  c(1,2,3,4,5,6,7,8,9)), ncol=9, byrow=TRUE)  # each iteration in a row
# this creates a matrix of labels
labtotab<-labeltotable(x)
# counts labels and summarises x

labtotab$lab
# gives labels

labtotab$freq
# gives the corresponding frequencies to the labels

plot.labclust  

Dendrogram plot of a labclust object

Description

The resulting plot does not show where the tree should be cut, instead it is very flexible.

Arguments

x          A labclust object to be plotted.
...        All dendrogram plot options

See Also

colorplot
Examples

x<-matrix(c(
  c(1,1,1,1,1,1,1,1,1),  # exchangeable labels are allowed like this
  c(1,1,1,1,2,2,2,2),
  c(2,2,1,2,2,3,3,3),
  c(1,1,2,1,3,3,3,3),  # or like this
  c(1,1,2,1,3,3,3,3),
  c(1,1,2,3,4,5,5,5),
  c(1,1,2,3,4,5,5,5),
  c(1,2,3,4,5,6,7,8,9)), ncol=9, byrow=TRUE)  # each iteration in a row

y<-labeltodendro(x)
plot(y)  # S3 method, uses plot.labclust
plot(y, center=TRUE)  # you may like to center the dendrogram
abline(h=y$hcut)  # or show where the dendrogram should be cut

relabel

relabels a vector, insisting that the labels appear in increasing order

Description

If you need to relabel a huge matrix, use relabel.matrix.

Usage

relabel(x)

Arguments

x  a vector of integers

Details

This calls a C function, hence it is relatively fast.

See Also

relabel.matrix, selectlabels

Examples

x<-c(2,2,2,1,2,2,2)  # labels do not appear in increasing order
relabel(x)  # gives the relabelled vector
**relabelNmatrix**  
_relabels a matrix label insisting that the labels of each row appear in increasing order_

### Description

We strongly recommend to use this function before tabletodendro or selectlabels, if you defined the label matrix by yourself.

### Usage

`relabelNmatrix(x)`

#### Arguments

- **x**  
  A numeric matrix of integers.

#### See Also

`tabletodendro`, `selectlabels`

#### Examples

```r
labmat<-matrix(c(
  c(1L,1L,1L,1L,1L,1L,1L,1L,1L),
  c(2L,2L,2L,2L,1L,2L,2L,2L,2L), # labels do not appear in increasing order
  c(2L,2L,1L,2L,3L,3L,3L,3L,3L), # here too
  c(1L,1L,2L,2L,2L,2L,3L,3L,3L),
  c(1L,2L,3L,4L,5L,6L,7L,8L,9L), ncol=9, byrow=TRUE)

relabelNmatrix(labmat)  # gives the relabelled matrix
```

**selectlabels**  
_it is useful to sort labels in agglomerative order_

### Description

The labels which are not agglomerations or divisions of the most frequent label, will be removed. The function does not allow exchangeable labels.

### Usage

`selectlabels(labmat, freq)`
selectlabels

Arguments

labmat A matrix of labels, each row corresponding to a labelling.
freq A vector of frequencies, each element corresponding to frequency of a labmat row.

Details

The function is useful if you have already counted your labels but you are not sure that they are in agglomerative order. The functions assumes your labels appear in increasing order, if your lables are like that, you should use relabel.matrix before.

Value

labmat Series of labels, sorted in agglomerative order.
freq A vector of frequencies corresponding to the sorted labels.

See Also
relabel.matrix

Examples

labmat<-matrix(c(
c(1,1,1,1,1,1,1,1,1),
c(2,2,2,1,2,2,2,2), #labels do not appear in increasing order
c(2,2,1,2,2,3,3,3), # here too, this is the most frequent label according to freq
c(1,1,2,2,2,2,3,3), # this is not agglomeration nor division of the most frequent label
c(1,2,3,4,5,6,7,8,9)), ncol=9, byrow=TRUE)
freq<-c(1,2,3,2,1)
labmat<-relabel.matrix(labmat)
labmat
# now labels of labmat appear in increasing order
y<-selectlabels(labmat,freq)
y$labmat
#now y is in agglomerative order
y$freq
# frequencies corresponding to ordered labels

plot(tabletodendro(y$labmat,y$freq))
# see the resulting dendrogram
**tabletodendro**

*converts a label matrix and a frequency vector to a labclust object*

**Description**

The function assumes that labels appear in increasing order and also each row sorted in agglomerative order.

**Usage**

```r
tabletodendro(labmat, freq, labels=NULL)
```

**Arguments**

- `labmat`: A label matrix, each row sorted in agglomerative order. Labels should appear in increasing order.
- `freq`: A vector of frequencies corresponding to each row of `labmat` matrix.
- `labels`: A vector of frequencies corresponding to each row of `labmat` matrix.

**Details**

If your label matrix does not have labels in increasing order use `relabel.matrix`. If the label matrix is not sorted in agglomerative order use `selectlabels`.

**See Also**

`labclust`, `relabel.matrix`, `selectlabels`

**Examples**

```r
labmat <- matrix(c(
  c(1,1,1,1,1,1,1,1,1),
  c(1,1,1,1,1,1,1,1,1),
  c(1,1,2,2,1,1,1,1,1),
  c(1,1,2,3,3,4,4,4,4)), ncol=9, byrow=TRUE)
freq <- c(5,10,15)

plot(tabletodendro(labmat, freq, labels=paste("subject",1:9))) # plots the resulting dendrogram
plot(tabletodendro(labmat, freq, center=TRUE)) # you may like to center the dendrogram
```
Index

colorplot, 2, 4, 6
is.aggvec, 3
labclust, 3, 4, 10
labeltodendro, 4, 6
labeltotable, 4, 5
plot(plot.labclust), 6
plot.labclust, 2, 6
relabel, 7
relabel.matrix, 7, 8, 9, 10
selectlabels, 6–8, 8, 10
tabletodendro, 4, 6, 8, 10