Package 'PPtreeViz'

October 12, 2022

Title Projection Pursuit Classification Tree Visualization

Description Tools for exploring projection pursuit classification tree using various projection pursuit indexes.

Version 2.0.4

Author Eun-Kyung Lee

Maintainer Eun-Kyung Lee <lee.eunk@gmail.com>

Date 2019-12-6

Depends R (>= 2.10), gridExtra, ggplot2, partykit, Rcpp

Suggests MASS (>= 3.1-20), stats

LinkingTo Rcpp, RcppArmadillo

License GPL (>= 2)

RoxygenNote 7.0.2

NeedsCompilation yes

Repository CRAN

Date/Publication 2019-12-06 08:40:02 UTC

R topics documented:

ENTROPYindex1D   *ENTROPY PPindex*

Description

ENTROPY projection pursuit index

Usage

```
ENTROPYindex1D(origclass,origdata,proj)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
proj	projection matrix

Details

Calculate ENTROPY projection pursuit index value

Examples

```
data(iris)
ENTROPYindex1D(iris[,5],as.matrix(iris[,1,drop=FALSE]))
```

GINIindex1D   *GINI PPindex*

Description

GINI projection pursuit index

Usage

```
GINIindex1D(origclass,origdata,proj)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
proj	projection matrix

Details

Calculate GINI projection pursuit index value

Examples

```
data(iris)
GINIindex1D(iris[,5],as.matrix(iris[,1,drop=FALSE]))
```

Huberplot	*Huber plot*

Description

Huber plot for 2D data

Usage

```
Huberplot(origdata2D,origclass,PPmethod="LDA",weight=TRUE,r=1,
          lambda=0.5,opt.proj=TRUE,UserDefFtn=NULL,...)
```

Arguments

origdata2D	2-dimensional numerical data for Huber plot
origclass	class information vector of data
PPmethod	method for projection pursuit; "LDA", "PDA", "Lr", "GINI", "ENTROPY", and "UserDef"
weight	weight flag in LDA, PDA and Lr index
r	r in Lr index
lambda	lambda in PDA index
opt.proj	flag to show the best projection in the plot
UserDefFtn	User defined index function when PPmethod="UserDef"
...	arguments to be passed to methods

Details

Draw Huber plot for 2-dimensional data with various PP indices and the histogram of the projected data onto the optimal projection to explore the behavior of the projection prsuit indices

References

Lee, EK., Cook, D., Klinke, S., and Lumley, T.(2005) Projection Pursuit for Exploratory Supervised Classification, Journal of Computational and Graphical Statistics, 14(4):831-846.

Examples

```
data(iris)
Huberplot(iris[,1:2],iris[,5],PPmethod="LDA")
```

LDAindex   *LDA PPindex*

Description

LDA projection pursuit index

Usage

```
LDAindex(origclass,origdata,proj,weight=TRUE)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
proj	projection matrix
weight	weight flag in LDA

Details

Calculate LDA projection pursuit index value

References

Lee, EK., Cook, D., Klinke, S., and Lumley, T.(2005) Projection Pursuit for exploratory supervised classification, Journal of Computational and Graphical statistics, 14(4):831-846.

Examples

```
data(iris)
LDAindex(iris[,5],as.matrix(iris[,1:4]))
```

LDAopt   *PP optimization using LDA index*

Description

PP optimization using LDA index

Usage

```
LDAopt(origclass,origdata,q=1,weight=TRUE,...)
```

Arguments

origclass	class information vector of data
origdata	data matrix without class information
q	dimension of projection vector
weight	weight flag in LDA index
...	arguments to be passed to methods

Details

Find the q-dimensional optimal projection using LDA projectin pursuit index

Value

indexbest maximum LDA index value

projbest optimal q-dimensional projection matrix

origclass original class information vector

origdata original data matrix without class information

References

Lee, EK., Cook, D., Klinke, S., and Lumley, T.(2005) Projection Pursuit for Exploratory Supervised Classification, Journal of Computational and Graphical Statistics, 14(4):831-846.

Examples

```
data(iris)
LDA.proj.result <- LDAopt(iris[,5],iris[,1:4])
LDA.proj.result$indexbest
LDA.proj.result$projbest
```

Lrindex	*Lr PPindex*

Description

Lr projection pursuit index

Usage

```
Lrindex(origclass,origdata,proj,weight=TRUE,r=1)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
proj	projection matrix
weight	weight flag in Lr index
r	r in Lr index

Details

Calculate Lr projection pursuit index value

References

Lee, EK., Cook, D., Klinke, S., and Lumley, T.(2005) Projection Pursuit for exploratory supervised classification, Journal of Computational and Graphical statistics, 14(4):831-846.

Examples

```
data(iris)
Lrindex(iris[,5],as.matrix(iris[,1:4]),r=1)
```

PDAindex	*PDA PPindex*

Description

PDA projection pursuit index

Usage

```
PDAindex(origclass,origdata,proj,weight=TRUE,lambda=0.1)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
proj	projection matrix
weight	weight flag in PDA
lambda	lambda in PDA index

Details

Calculate PDA projection pursuit index value

References

Lee, EK., Cook, D.(2010) A projection pursuit index for large p small n data, Statistics and Computing, 20:381-392.

Examples

```
data(iris)
PDAindex(iris[,5],as.matrix(iris[,1:4]),lambda=0.2)
```

PDAopt   *PP optimization using PDA index*

Description

PP optimization using PDA index

Usage

```
PDAopt(origclass,origdata,q=1,weight=TRUE,lambda=0.1,...)
```

Arguments

origclass	class information vector of data
origdata	data matrix without class information
q	dimension of projection vector
weight	weight flag in PDA index
lambda	lambda in PDA index
...	arguments to be passed to methods

Details

Find the q-dimensional optimal projection using PDA projectin pursuit index

Value

indexbest maximum PDA index value

projbest optimal q-dimensional projection matrix

origclass original class information vector

origdata original data matrix without class information

References

Lee, EK, Cook, D.(2010) A Projection Pursuit Index for Large p Small n Data, Statistics and Computing, 20:381-392.

Examples

```
data(iris)
PDA.proj.result <- PDAopt(iris[,5],iris[,1:4],weight=TRUE,q=2,lambda=0.1)
PDA.proj.result$indexbest
PDA.proj.result$projbest
```

---

plot.PPtreeclass   *PPtree plot*

Description

projection pursuit classification tree plot

Usage

```
## S3 method for class 'PPtreeclass'
plot(
  x,
  font.size = 17,
  width.size = 1,
  main = "Projection Pursuit Classification Tree",
  sub = NULL,
  ...
)
```

Arguments

x	PPtreeclass object
font.size	font size of plot
width.size	size of eclipse in each node.
main	main title
sub	sub title
...	arguments to be passed to methods

Details

Draw projection pursuit classification tree with tree structure. It is modified from a function in party library.

References

Lee, EK(2017) PPtreeViz: An R Package for Visualizing Projection Pursuit Classification Trees, Journal of Statistical Software <doi:10.18637/jss.v083.i08>

Examples

```
data(iris)
Tree.result <- PPTreeclass(Species~., data = iris,"LDA")
Tree.result
plot(Tree.result)
```

PPclassify   *predict PPtree*

Description

predict projection pursuit classification tree

Usage

```
PPclassify(Tree.result,test.data,Rule,true.class=NULL,...)
```

Arguments

Tree.result	PPtreeclass object
test.data	the test dataset
Rule	split rule 1: mean of two group means 2: weighted mean of two group means - weight with group size 3: weighted mean of two group means - weight with group sd 4: weighted mean of two group means - weight with group se 5: mean of two group medians 6: weighted mean of two group medians - weight with group size 7: weighted mean of two group median - weight with group IQR 8: weighted mean of two group median - weight with group IQR and size
true.class	true class of test dataset if available
...	arguments to be passed to methods

Details

Predict class for the test set with the fitted projection pursuit classification tree and calculate prediction error.

Value

predict.class predicted class

predict.error number of the prediction errors

References

Lee, YD, Cook, D., Park JW, and Lee, EK(2013) PPtree: Projection Pursuit Classification Tree, Electronic Journal of Statistics, 7:1369-1386.

Examples

```
data(iris)
n <- nrow(iris)
tot <- c(1:n)
n.train <- round(n*0.9)
train <- sample(tot,n.train)
test <- tot[-train]
Tree.result <- PPTreeclass(Species~.,data=iris[train,],"LDA")
PPclassify(Tree.result,iris[test,1:4],1,iris[test,5])
```

PPclassNodeViz   *PPtree node visualization*

Description

Visualization tools to explore each node of PPtree

Usage

```
PPclassNodeViz(PPclassOBJ,node.id,Rule,legend,std,image,diff.prop)
```

Arguments

PPclassOBJ	PPregclass object
node.id	node ID
Rule	cutoff rule
legend	flag to represent legend in the plot. Default value is TRUE
std	flag to standardize data before drawing plot
image	flag to draw image plot of correlation matrix
diff.prop	percentage of number of variables with significant differences and shown in red in the bar chart style means

Details

For the inner node, four plots are provided - the bar chart style plot with projection pursuit coefficients of each variable, the histogram of the projected data, the bar chart style plots with means of each variables for the left and the right group, and the image plot of correlations.

References

Lee, YD, Cook, D., Park JW, and Lee, EK(2013) PPtree: Projection Pursuit Classification Tree, Electronic Journal of Statistics, 7:1369-1386.

Examples

```
data(iris)
Tree.result <- PPTreeclass(Species~., data = iris,"LDA")
Tree.result
PPclassNodeViz(Tree.result,1,1)
```

PPopt	*Projection pursuit optimization*

Description

PP optimization using various projection pursuit indices

Usage

```
PPopt(origclass,origdata,q=1,PPmethod="LDA",weight=TRUE,r=1,
             lambda=0.1,energy=0,cooling=0.999,TOL=0.0001,maxiter = 50000)
```

Arguments

origclass	class information vector
origdata	data matrix without class information
q	dimension of projection matrix
PPmethod	method for projection pursuit; "LDA", "PDA", "Lr", "GINI", and "ENTROPY"
weight	weight flag in LDA, PDA and Lr index
r	r in Lr index
lambda	lambda in PDA index
energy	energy parameter
cooling	cooling parameter
TOL	tolerance
maxiter	number of maximum iteration

Details

Find the q-dim optimal projection using various projectin pursuit indices with class information

Value

indexbest maximum LDA index value

projbest optimal q-dim projection matrix

origclass original class information vector

origdata original data matrix without class information

References

Lee, EK., Cook, D., Klinke, S., and Lumley, T.(2005) Projection Pursuit for exploratory supervised classification, Journal of Computational and Graphical statistics, 14(4):831-846.

Examples

```
data(iris)
PP.proj.result <- PPopt(iris[,5],as.matrix(iris[,1:4]))
PP.proj.result
```

PPoptViz   *PPopt visualization*

Description

Visualize PPopt result

Usage

```
PPoptViz(PPoptOBJ)
```

Arguments

PPoptOBJ   PPoptim object. result from LDAopt, PDAopt, and PPopt

Details

Visualize the result of projection pursuit optimization

References

Lee, EK(2017) PPtreeViz: An R Package for Visualizing Projection Pursuit Classification Trees, Journal of Statistical Software <doi:10.18637/jss.v083.i08>

Examples

```
data(iris)
PPoptViz(LDAopt(iris[,5],iris[,1:4],q=1))
PPoptViz(LDAopt(iris[,5],iris[,1:4],q=2))
```

PPTreeclass   *Projection pursuit classification tree*

Description

Construct the projection pursuit classification tree

Usage

```
PPTreeclass(formula,data, PPmethod="LDA",weight=TRUE,
                   r=1,lambda=0.1,energy=0,maxiter=50000,...)
```

Arguments

formula	an object of class "formula"
data	data frame
PPmethod	method for projection pursuit; "LDA", "PDA", "Lr", "GINI", and "ENTROPY"
weight	weight flag in LDA, PDA and Lr index
r	r in Lr index
lambda	lambda in PDA index
energy	parameter for the probability to take new projection
maxiter	maximum iteration number
...	arguments to be passed to methods

Details

Find tree structure using various projection pursuit indices of classification in each split.

Value

Tree.Struct tree structure of projection pursuit classification tree

projbest.node 1 dimensional optimal projections of each node split

splitCutoff.node cutoff values of each node split

origclass original class

origdata original data

References

Lee, YD, Cook, D., Park JW, and Lee, EK(2013) PPtree: Projection Pursuit Classification Tree, Electronic Journal of Statistics, 7:1369-1386.

Examples

```
data(iris)
Tree.result <- PPTreeclass(Species~.,data = iris,"LDA")
Tree.result
```

---

predict.PPtreeclass   *predict PPtree*

Description

predict projection pursuit classification tree

Usage

```
## S3 method for class 'PPtreeclass'
predict(object, newdata = NULL, Rule = 1, ...)
```

Arguments

object	a fitted object of class inheriting from "PP.Tree.class"
newdata	the test dataset
Rule	split rule 1: mean of two group means 2: weighted mean of two group means - weight with group size 3: weighted mean of two group means - weight with group sd 4: weighted mean of two group means - weight with group se 5: mean of two group medians 6: weighted mean of two group medians - weight with group size 7: weighted mean of two group median - weight with group IQR 8: weighted mean of two group median - weight with group IQR and size
...	arguments to be passed to methods

Details

Predict class for the test set with the fitted projection pursuit classification tree and calculate prediction error.

References

Lee, YD, Cook, D., Park JW, and Lee, EK(2013) PPtree: Projection Pursuit Classification Tree, Electronic Journal of Statistics, 7:1369-1386.

Examples

```
data(iris)
n <- nrow(iris)
tot <- c(1:n)
n.train <- round(n*0.9)
train <- sample(tot,n.train)
test <- tot[-train]
Tree.result <- PPTreeclass(Species~., data=iris[train,],"LDA")
predict(Tree.result)
```

---

print.PPtreeclass	*Print PP.Tree.class result*

Description

Print PP.Tree.class result

Usage

```
## S3 method for class 'PPtreeclass'
print(x, coef.print = FALSE, cutoff.print = FALSE, verbose = TRUE, ...)
```

Arguments

x	PPtreeclass object
coef.print	print projection coefficients in each node ifTRUE
cutoff.print	print cutoff values in each node if TRUE
verbose	print if TRUE, no output if FALSE
...	arguments to be passed to methods

Details

Print the projection pursuit classification tree result

References

Lee, EK(2017) PPtreeViz: An R Package for Visualizing Projection Pursuit Classification Trees, Journal of Statistical Software <doi:10.18637/jss.v083.i08>

Examples

```
data(iris)
Tree.result <- PPTreeclass(Species~.,data = iris,"LDA")
Tree.result
print(Tree.result,coef.print=TRUE,cutoff.print=TRUE)
```

Index