Package 'pass'

February 20, 2015

Type Package

Title Prediction and Stability Selection of Tuning Parameters

Version 1.0

Date 2012-12-21

Author Yixin Fang, Wei Sun, Junhui Wang

Maintainer Yixin Fang <yixin.fang@nyumc.org>

Description To implement two methods, Kappa and PASS, for selecting tuning parameters in regularized procedures such as LASSO, SCAD, adaptive LASSO, aiming for variable selection in regularized regression

License GPL-2

LazyLoad yes

Depends R (>= 2.10.0), MASS, lars, ncvreg

Repository CRAN

Date/Publication 2013-01-12 15:01:38

NeedsCompilation no

R topics documented:

agree.twosets *Agreement of Two Subsets*

Description

To calculate Cohen's Kappa coefficients of two subsets

Usage

```
agree.twosets(aset1, aset2, p.tot)
```

Arguments

aset1	The first subsets
aset2	The second subsets
p.tot	The total number of variables

Value

ratio	The Kappa coefficient of the input two subsets

References

Cohen (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20: 37-46.

cv.twosets *Two-fold Cross-Validation*

Description

To calcuate two-fold cross-validation based on a random splitting

Usage

```
cv.twosets(data1, beta1.hat, data2, beta2.hat)
```

Arguments

data1	The first subsample
beta1.hat	The sparse solution obtained from the first subsample
data2	The second subsample
beta2.hat	The sparse solution obtained from the second subsample

Value

cv.value	The two-fold cross-validation value.

pass *Prediction and Stability Selection of Tuning Parameters*

Description

To perform two methods, Kappa and PASS, for selecting tuning parameters in regularized procedures such as LASSO, SCAD, and adaptive LASSO

Usage

```
pass(data, base = "LASSO", lambda.grid=NULL, num.grid=20, num.split = 20, alpha = 0.1)

## S3 method for class 'pass'
print(x, ...)
## S3 method for class 'pass'
plot(x, ...)
```

Arguments

data	It is an n by (p+1) matrix, where the first p columns form the design matrix and the last column is response vector.
base	It is the base procedure used for variable selection. Three choices of base are "LASSO", "SCAD", and "aLASSO".
lambda.grid	It is a vector consisting of the values of tuning parameter lambda to be evaluated. If lambda.grid=NULL, a grid of lambda's will be decided automatically, with specified number of lambda's to be considered.
num.grid	It is the number of lambda's to be considered, where a grid of lambda's is decided manually or automatically. The default value is 20.
num.split	It is the number of random half-half splittings. The default value is 20.
alpha	It is the threshold only used for the Kappa selection method. It is not a tuning parameter. The default value is 0.1.
x	This is the output object of class "pass" to be used in print.pass and plot.pass.
...	Not used.

Details

Because the data matrix will be centerized so that the column means are zero, there is no need an intercept column in the data matrix. Function print.lass(x) prints the two estimated optimal values of tuning parameter lambda and function plot.lass(x) plots the two tuning parameter selection processess, where x is the output of function pass.

Value

`pass.values`	The values evaluated over lambda.grid using the PASS criterion. A curve based on these values can be drawn using function `plot.pass`. The maximum point is selected as the estimated optimal value for the tuning parameter lamda.
`kappa.values`	The values evaluated over lambda.grid using the Kappa criterion. A curve based on these values can be drawn using function `plot.pass`. The maximum point (adjusted for the threshold `alpha`) is selected as the estimated optimal value for the tuning parameter lamda.
`lambda.pass`	The estimated optimal value for the tuning parameter lambda using the PASS criterion
`lambda.kappa`	The estimated optimal value for the tuning parameter lambda using the Kappa criterion (adjusted for the threshold `alpha`)
`beta.pass`	The estimated coefficients using selected lambda by the PASS criterion
`beta.kappa`	The estimated coefficients using selected lambda by the Kappa criterion (adjusted for the threshold `alpha`)
`subset.pass`	The selected submodel by the PASS criterion
`subset.kappa`	The selected submodel by the Kappa criterion (adjusted for the threshold `alpha`)

Author(s)

Yixin Fang, Wei Sun, Junhui Wang

References

(1) Sun, Wang, and Fang (2012+) Consistent selection of tuning parameters via variable selection stability. Revision Submmitted. Available at arXiv.

(2) Fang, Wang, and Sun (2012+) A PASS for tuning parameter selection in regularized regression. Submmitted. Available at arXiv.

Examples

```
library(MASS)
library(lars)
library(ncvreg)

beta=c(3,1.5,0,0,2,0,0,0)
p=8
n=100
sigma=1
rho=0.5

set.seed(100)
x=matrix(0, n, p)
x[,1]<-rnorm(n, 0, 1)
for (i in 2:p) x[,i]<-rho*x[,i-1]+sqrt(1-rho^2)*rnorm(n, 0, 1)
y=x%*%beta+sigma*rnorm(n, 0, 1)
data<-cbind(x,y)
```

```
lambda.grid=10^seq(-2,2,length=20)
results<-pass(data=data, base="LASSO", lambda.grid=lambda.grid, num.grid=20, num.split=20)
print(results)
plot(results)
```

Index