Package 'SUE'

February 19, 2015

Type Package

Title Subsampling method

Version 1.0

Date 2013-01-28

Author Jim Yi

Maintainer Jim Yi <yijinpeng@live.cn>

Description This is a package for the subsampling method of robust estimation of linear regression models

Suggests MASS

License GPL-2

Repository CRAN

Date/Publication 2013-01-29 20:15:52

NeedsCompilation no

R topics documented:

SUE-package *Subsampling method*

Description

This is a package for the subsampling method of robust estimation of linear regression models

Details

Package:	SUE
Type:	Package
Version:	1.0
Date:	2013-01-28
License:	GPL-2

Author(s)

Jim Yi

Maintainer: Jim Yi <yijinpeng@live.cn>

`parameters` *Parameters Choice*

Description

This function computes the main parameters of the subsampling method. The values it returns are needed in following computing process of the function `SUE.lm`.

Usage

```
parameters(N, ns, a = 0.1, E = 0.99, p = 0.99, method = "default")
```

Arguments

`N`	The number of observations of a sample
`ns`	The subsample size
`a`	The proportion of contamination
`E`	The desired efficiency
`p`	The probability of having at least r* good subsamples
`method`	The options to generate parameters in different mechanisms

Value

`ns`	The subsamples size
`r`	The number of subsamples to be combined
`k`	The total number of subsamples to be gemerated

Author(s)

Jim Yi

Examples

```
##Generate parameters of the data with N=50 observations
parameters(50)

##Generate parameters using method="small.k"
parameters(50,a=0.2,method="small.k")

##Generate parameters using method="appro.k"
parameters(50,a=0.05,method="appro.k")
```

`SUE.fitted.values` *Fitted Values of SUE.lm Fits*

Description

Computes fitted values of a fit `SUE.lm` model.

Usage

```
SUE.fitted.values(fit)
```

Arguments

`fit`	An object of class `SUE.lm` , typically the result of a call to `SUE.lm`

Author(s)

Jim Yi

Examples

```
##See examples of the function SUE.lm
```

`SUE.lm` *The Subsampling Method for Linear Regression*

Description

This function computes the subsampling method estimators for linear regression.

Usage

```
SUE.lm(formula, data = list(), k, ns, r, constant = 0.25, consistency.check = TRUE)
```

Arguments

formula	it is an object that indicates the variables used in the regression. a formula object has the form y = x1 +x2 +...+xp, where y is the name of the dependent variable, and x1, ..., xp are the names of the explanatory variables.
data	This argument is used only if the variables belong to a data frame, in which case data is the name of the data frame.
k	It is the total number of subsamples to be generated.
ns	It is the subsample size.
r	It is the number of subsamples to be combined. The function parameters is especially designed to compute these three parameters of subsampling method.
constant	is a predetermined parameter which is used to control the distance between two estimated values. It only works under the condition that consistency,check = "TURE". The default value is 0.25. However. users can try different values to get better result. What has to be mentioned is that: if the value is set too small, the function will fail the consistency check easily which result in running the program for many more times, but if the value is too large, the result may be not reliable. It is user's job to balance these situations.
consistency.check	The argument decides if we conduct consistency check. The defualt value is TURE. We highly recommend to always check the consistency of the result after computing. It can sufficiently increase the reliability of subsampling method.

Value

Apart from the same output components as the object of class "lm", such as coefficients, residuals and fitted.values, the main components of the output are:

combined.sample	is the final combined sample generated by the subsampling method, It is supposed to be the fine data without outliers.
sample.size	is the sample size of the combined sample, which is convenient for user to compute the number of outliers.
mse	They are MSEs of the regressions of r chosen subsamples.
beta	They are coefficient parameters of the regressions of r chosen subsamples.
check	It is a logistic output which indicates whether the subsampling method fails the consistency check or not.

Author(s)

Jim Yi

Examples

```
##We analysis the well known stackloss data by using ordinary linear method and the subsampling method.
##We also try two m values, m = 2 and 4, which represent toughly 10% and 20% working
##proportion of outliers in the data. The subsample size is chosen to be the default size of ns = 11.
```

```
data(stackloss)
a1=lm(stack.loss~Air.Flow+Water.Temp+Acid.Conc.,data=stackloss)
a2=SUE.lm(stack.loss~Air.Flow+Water.Temp+Acid.Conc.,data=stackloss,k=57,ns=11,r=6,
consistency.check=TRUE,constant=0.25)
a3=SUE.lm(stack.loss~Air.Flow+Water.Temp+Acid.Conc.,data=stackloss,k=327,ns=11,r=5,
consistency.check=TRUE,constant=0.25)
par(mfrow=c(2,2))
plot(a1$fitted.values,a1$residuals,xlab="(a) fitted values",ylab="residuals",ylim=c(-12,12))
abline(h=0)
abline(h=9.7,lty=2)
abline(h=-9.7,lty=2)
plot(SUE.fitted.values(a2),SUE.residuals(a2),xlab="(b) fitted values",ylab="residuals",ylim=c(-12,12))
abline(h=0)
abline(h=9,lty=2)
abline(h=-9,lty=2)
plot(SUE.fitted.values(a3),SUE.residuals(a3),xlab="(c) fitted values",ylab="residuals",ylim=c(-12,12))
abline(h=0)
abline(h=3.75,lty=2)
abline(h=-3.75,lty=2)
```

`SUE.plot` *Plot for "SUE.lm" Objects*

Description

Fitted.values vs residuals plot for "SUE.lm" Objects

Usage

```
SUE.plot(fit)
```

Arguments

`fit`	An object of class `SUE.lm` , typically the result of a call to `SUE.lm`

Author(s)

Jim Yi

Examples

```
## We analysis the well-known stackloss data by using the subsampling method.

data(stackloss)
a=SUE.lm(stack.loss~Air.Flow+Water.Temp+Acid.Conc.,data=stackloss,k=57,ns=11,r=6,
consistency.check=TRUE,constant=0.25)
SUE.plot(a)
```

SUE.residuals *Residuals of SUE.lm Fits*

Description

Computes residuals of fit SUE.lm model.

Usage

```
SUE.residuals(fit)
```

Arguments

fit	An object of class SUE.lm , typically the result of a call to SUE.lm

Author(s)

Jim Yi

Examples

```
##See examples of the function SUE.lm
```

Index