The roblm Package

June 10, 2006

Type Package

Title MM-regression estimators

Version 0.6

Date 2006-03-29

Author Matias Salibian-Barrera <matias@stat.ubc.ca>

Maintainer Matias Salibian-Barrera <matias@stat.ubc.ca>

Description Robust regression estimators

License GPL 2


R topics documented:

  coleman.dat .......................................................... 2
  plot.roblm .......................................................... 2
  print.roblm.null .................................................... 3
  print.roblm ........................................................ 3
  print.summary.roblm ................................................. 4
  roblm.Chi.prime.2 ................................................ 5
  roblm.Chi.prime ................................................... 5
  roblm.Chi .......................................................... 6
  roblm.control ....................................................... 6
  roblm.fit.MM ....................................................... 7
  roblm.MM ........................................................... 8
  roblm-package ...................................................... 9
  roblm.Psi.prime ................................................... 9
  roblm.Psi .......................................................... 10
  roblm .............................................................. 11
  roblm.S ............................................................ 12
  stackloss.dat ...................................................... 13
  summary.roblm ..................................................... 14

Index 15
coleman.dat

Coleman data

Description
A good description of these data is coming in future releases

Usage
data(coleman.dat)

Format
A data set

Source
Coleman

References
Coleman

plot.roblm
plot method for roblm

Description
Diagnostic plots for elements of class roblm

Usage
## S3 method for class 'roblm':
plot(x,
which = 1:5, caption = c("Standardized residuals vs. Robust Distances",
"Normal Q-Q vs. Residuals", "Response vs. Fitted Values",
"Residuals vs. Fitted Values",
"Sqrt of abs(Residuals) vs. Fitted Vaules"),
panel = points, sub.caption = deparse(x$call), main = "",
ask = prod(par("mfcol")) < length(which) && dev.interactive(),
..., p=0.025)

Arguments
x an object as created by roblm
which integer number between 1 and 5 to specify which plot is desired
caption Caption for the different plots
panel Panel
main Main title
References

Robust diagnostic plots as in Rousseuw and van Zomeren (1990)

See Also

plot.lm

print.roblm.null  

print method for objects of class roblm

Description

Print method for elements of class roblm

Usage

## S3 method for class 'roblm.null':
print(x, digits = max(3,getOption("digits") - 3), ...)

Arguments

x  an object as created by roblm
digits  number of digits to use
...  optional arguments to be passed to print

See Also

print.lm

print.roblm  

print method for objects of class roblm

Description

Print method for elements of class roblm

Usage

## S3 method for class 'roblm':
print(x, digits = max(3,getOption("digits") - 3), ...)

Arguments

x  an object as created by roblm
digits  number of digits to use
...  optional arguments to be passed to print
print.summary.roblm

Arguments

x an object as created by roblm
digits number of digits to display
...

See Also

print.lm

print.summary.roblm

print method for objects of class summary.roblm

Description

Print method for elements of class summary.roblm

Usage

## S3 method for class 'summary.roblm':
print(x, digits = max(3, getOption("digits") - 3),
symbolic.cor =
  x$symbolic.cor, signif.stars = getOption("show.signif.stars"),
  ...)

Arguments

x an object as created by summary.roblm
digits number of digits to display
symbolic.cor logical variable indicating whether to use symbols to display the estimated correlation matrix of the coefficient estimators
signif.stars logical variable indicating whether to use stars to display different levels of significance in the individual t-tests
...

See Also

print.summary.lm
roblm.Chi.prime.2  Second derivative of Tukey's bi-square loss function

Description

Usage

roblm.Chi.prime.2(x, cc)

Arguments

x  argument
cc tuning constant

Details

Value

Author(s)

Matias Salibian-Barrera

roblm.Chi.prime  Derivative of Tukey's bi-square loss function

Description

Usage

roblm.Chi.prime(x, cc)

Arguments

x  argument
cc tuning constant

Details

Value
Author(s)

Matias Salibian-Barrera

Description

Tukey’s bi-square loss function

Usage

roblm.Chi(x, cc)

Arguments

x  argument
cc  tuning constant

Details

Value

Author(s)

Matias Salibian-Barrera

Description

Tuning parameters for the MM-regression estimator and the associated S-estimator

Usage

roblm.control(seed = 99, Nres = 500,
              tuning.chi = 1.54764, bb = 0.5, tuning.psi = 4.685061,
              groups = 5, n.group = 400, k.fast.s = 1,
              max.it = 50,
              compute.rd = TRUE
              )
Arguments

Nres  Number of re-sampling candidates to be used to find the initial S-estimator. This parameter is currently set to 500, which works well in most situations (see References below). User-choice capability will be added in future releases.

seed  Random seed for the re-samples used in obtaining candidates for the initial S-estimator.

tuning.chi  Tuning constant for the S-estimator. The choice 1.54764 yields a 50% breakdown estimator.

bb  Expected value under the normal model of the "chi" function with tuning constant equal to tuning.chi. This is used to compute the S-estimator.

max.it  Maximum number of IRWLS iterations.

tuning.psi  Tuning constant for the re-descending M-estimator. The choice 4.685061 yields an estimator with asymptotic efficiency of 95% for normal errors.

compute.rd  If TRUE robust distances (based on the MCD robust covariance matrix) are computed for the robust diagnostic plots. This may take some time to finish, specially for large data sets.

groups  This parameter is for the fast-S algorithm. Number of random subsets to use when the data set is large.

n.group  This parameter is for the fast-S algorithm. Size of each of the groups above.

k.fast.s  This parameter is for the fast-S algorithm. Number of local improvement steps for each re-sampling candidate.

Author(s)

Matias Salibian-Barrera

See Also

roblm

Description

Fit function for MM-regression estimators

Usage

roblm.fit.MM()

Arguments

Details

This function is used by roblm and not intended to be used on its own.
Value

A list that contains regression coefficient estimators and the robustly estimated error standard deviation

See Also

roblm

---

**roblm.MM**

*Function to compute MM-regression estimators*

**Description**

This function performs RWLS iterations to find an MM-regression estimator

**Usage**

`roblm.MM(x, y, beta.initial, scale, control)`

**Arguments**

- `x` Design matrix
- `y` Response vector
- `beta.initial` Vector of initial estimators. Usually an S-regression estimator
- `scale` A robust residual scale estimate. Usually an S-scale estimator.
- `control` A list of control parameters as returned by `roblm.control`

**Details**

This function is used by `roblm.fit.MM` and not intended to be used on its own.

**Value**

A list with the following elements:

- `coef` The MM-regression estimator
- `cov` The covariance matrix of the MM-regression estimator
- `control` The control list used
- `scale` The residual scale estimate
- `seed` The random number generator seed
- `converged` TRUE if the RWLS iterations converged, FALSE otherwise

**Author(s)**

Matias Salibian-Barrera

**References**

Yohai, 1987
**roblm-package**  
*MM-regression estimators*

**Description**

MM-regression estimators

**Details**

- **Package:** roblm
- **Type:** Package
- **Version:** 0.6
- **Date:** 2006-03-28
- **License:** GPL 2.0

MM-regression estimators

**Author(s)**

Matias Salibian-Barrera

Maintainer: Matias Salibian-Barrera <matias@stat.ubc.ca>

**References**


---

**roblm.Psi.prime**  
*Derivative of Tukey’s bi-square score function*

**Description**

**Usage**

roblm.Psi.prime(x, cc)
Arguments

- `x` argument
- `cc` tuning constant

Details

Value

Author(s)

Matias Salibian-Barrera

---

roblm.Psi  

*Tukey's bi-square score function*

Description

Usage

`roblm.Psi(x, cc)`

Arguments

- `x` argument
- `cc` tuning constant

Details

Value

Author(s)

Matias Salibian-Barrera
**roblm**

**MM-regression estimators**

**Description**

MM-regression estimators

**Usage**

```r
roblm(formula, data = list(), weights, na.action,
      model = TRUE, x = FALSE, y = FALSE, singular.ok = TRUE,
      contrasts = NULL, offset = NULL, control = roblm.control())
```

**Arguments**

- `formula`: formula
- `data`: data
- `weights`: weights
- `na.action`: na.action
- `model`: model
- `x`: x
- `y`: y
- `singular.ok`: singular.ok
- `contrasts`: contrasts
- `offset`: offset
- `control`: control

**Details**

This function computes and MM-regression estimator as described in Yohai (1987). It uses a bi-square re-descending score function, and by default it returns a highly robust and highly efficient estimator (with 50% breakdown point and 95% asymptotic efficiency for normal errors). It uses an S-estimator (Rousseeuw and Yohai, 1984) for the errors which is also computed with a bi-square score function. The S-estimator is computed using the Fast-S algorithm of Salibian-Barrera and Yohai (2006). Standard errors are computed using the formulas for homoscedastic and independent errors of Croux, Dhaene and Hoorelbeke (2003).

**Value**

An object of class `roblm`. A list that includes the following components:

- `coef`: The MM-regression estimator
- `scale`: The S-scale estimator
- `s`: The auxiliary S-regression estimator
- `cov`: The estimated covariance matrix of the regression coefficients
- `residuals`: Residuals associated with the MM-estimator
- `fitted.values`: Fitted values associated with the MM-estimator
- `converged`: TRUE if the IRWLS iterations converged
roblm.S

Note

Author(s)

Matias Salibian-Barrera

References


See Also

roblm.control

Examples

data(coleman.dat, package='roblm')
summary( roblm(y~., data=coleman.dat) )

---

roblm.S  S-regression estimators

Description

This function computes the S-regression estimator

Usage

roblm.S(x, y, control)

Arguments

x  Design matrix
y  Response vector
control  list as retruned by roblm.control

Details

This function is used by roblm.fit.MM and not intended to be used on its own
Value

A list containing the following elements:

- coef: Vector of S-regression coefficient estimates
- cov: Covariance matrix of the coefficient estimators
- control: The control list as given to it
- scale: The S-scale residual estimate
- seed: The value of seed

Author(s)

Matias Salibian-Barrera

References


Description

A good description of these data is coming in future releases

Usage

data(stackloss.dat)

Format

A data set

Source

Source

References

References
**Description**

Summary method for elements of class `roblm`

**Usage**

```r
## S3 method for class 'roblm':
summary(object, correlation = FALSE,
         symbolic.cor = FALSE, ...)
```

**Arguments**

- `object`: an object as created by `roblm`
- `correlation`: logical variable indicating whether to compute the correlation matrix of the coefficient estimators
- `symbolic.cor`: logical variable indicating whether to use symbols to display the correlation matrix of the coefficient estimators
- `...`: Extra arguments

**See Also**

`summary.lm`
Index

*Topic datasets
  coleman.dat, 1
  stackloss.dat, 13

*Topic package
  roblm-package, 8

*Topic regression
  plot.roblm, 2
  print.roblm, 3
  print.roblm.null, 3
  print.summary.roblm, 4
  roblm, 10
  roblm.control, 6
  roblm.fit.MM, 7
  roblm.MM, 7
  roblm.S, 12
  summary.roblm, 13

*Topic robust
  plot.roblm, 2
  print.roblm, 3
  print.roblm.null, 3
  print.summary.roblm, 4
  roblm, 10
  roblm.Chi, 5
  roblm.Chi.prime, 5
  roblm.Chi.prime.2, 4
  roblm.control, 6
  roblm.fit.MM, 7
  roblm.MM, 7
  roblm.Psi, 10
  roblm.Psi.prime, 9
  roblm.S, 12
  summary.roblm, 13

coleman.dat, 1
plot.roblm, 2
print.roblm, 3
print.roblm.null, 3
print.summary.roblm, 4
roblm, 7, 10
roblm-package, 8
roblm.Chi, 5
roblm.Chi.prime, 5
roblm.Chi.prime.2, 4
roblm.control, 6, 12
roblm.fit.MM, 7
roblm.MM, 7
roblm.Psi, 10
roblm.Psi.prime, 9
roblm.S, 12
stackloss.dat, 13
summary.roblm, 13