Package ‘MTest’

October 6, 2023

Type Package

Title A Procedure for Multicollinearity Testing using Bootstrap

Version 1.0.2

Date 2023-10-06

Maintainer Víctor Morales-Oñate <victor.morales@uv.cl>

Description
Functions for detecting multicollinearity. This test gives statistical support to two of the most famous methods for detecting multicollinearity in applied work: Klein’s rule and Variance Inflation Factor (VIF). See the URL for the papers associated with this package, as for instance, Morales-Oñate and Morales-Oñate (2015) <doi:10.33333/rp.vol51n2.05>.

Depends R (>= 4.1.0)

License GPL (>= 3)

Encoding UTF-8

Imports car, ggplot2,plotly

Repository CRAN

URL https://github.com/vmoprojs/MTest

BugReports https://github.com/vmoprojs/MTest/issues

LazyData true

NeedsCompilation no

Author Víctor Morales-Oñate [aut, cre]
(<https://orcid.org/0000-0003-1922-6571>),
Bolívar Morales-Oñate [aut] (<https://orcid.org/0000-0003-4980-8759>)

Date/Publication 2023-10-06 13:10:02 UTC

R topics documented:

MTest *MTest*

Description

MTest is a nonparametric test based on bootstrap for detecting multicollinearity. This test gives statistical support to two of the most famous methods for detecting multicollinearity in applied work: Klein's rule and Variance Inflation Factor (VIF for essential multicollinearity).

Usage

```
MTest(object, nboot = 100,
                  nsam = NULL,trace = FALSE,seed = NULL,
                  valor_vif = 0.9)
```

Arguments

`object`	an object representing a model of an appropriate class (mainly "lm"). This is used as the model in MTest.
`nboot`	Numeric; number of bootstrap iterations to obtain the probability distribution of R squared (global and auxiliar).
`nsam`	Numeric; sample size for bootstrap samples.
`trace`	Logical; prints iteration process.
`seed`	Numeric; seed value for the bootstrap in nboot parameter.
`valor_vif`	Numeric; value to be compared in kleins rule.

Details

MTest generates a bootstrap distribution for the coefficient of determination which lets the researcher assess multicollinearity by setting a statistical significance α, or more precisely, an achieved significance level (ASL) for a given threshold.

Consider the regression model

$Y_i = \beta_0 X_{0i} + \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + u_i$

where $i = 1, ..., n$, $X_{j,i}$ are the predictors with $j = 1, ..., p$, $X_0 = 1$ for all i and u_i is the gaussian error term.

In order to describe Klein's rule and VIF methods, we need to define *auxiliary regressions* associated to model. An example of an auxiliary regressions is:

$X_{2i} = \gamma_1 X_{1i} + \gamma_3 X_{3i} + \cdots + \gamma_p X_{pi} + u_i.$

In general, there are p auxiliary regressions and the dependent variable is omitted in each auxiliary regression. Let R_g^2 be the coefficient of determination of the model and R_j^2 the jth coefficient of determination of the jth auxiliary regression.

Value

Returns an object of class `MTest`. An object of class `MTest` is a list containing at most the following components:

- `pval_vif` p values for vif test;
- `pval_klein` p values for klein test;
- `Bvals` A $nboot \times (p+1)$ matrix where rows are the number of bootstap samples and the columns are $R^2_{g_{boot}}$ and $R^2_{j_{boot}}$ which are estimates of estimates of R^2_g and R^2_j, see Section **Details**
- `vif.tot` Observed VIF values;
- `R.tot` Observed R^2_g and R^2_j values;
- `nsam` sample size used in bootstrap procedure.

Author(s)

Víctor Morales Oñate, <victor.morales@uv.cl>, https://sites.google.com/site/moralesonatevictor/,https://www.linkedin.com/in/vmoralesonate/ Bolívar Morales Oñate, <bmoralesonate@gmail.com>, https://sites.google.com/site/moralesonatevictor/

References

Morales-Oñate, V., and Morales-Oñate, B. (2023). *MTest: a Bootstrap Test for Multicollinearity*. Revista Politécnica, 51(2), 53–62. doi:10.33333/rp.vol51n2.05

Examples

```
library(MTest)
data(simDataMTest)
m1 <- lm(y~.,data = simDataMTest)

boot.sol <- MTest(m1,trace=FALSE,seed = 1,nboot = 50)
boot.sol$pval_vif
boot.sol$pval_klein
head(boot.sol$Bvals)
print(boot.sol)
```

`pairwiseKStest` *pairwiseKStest*

Description

Returns the p-value of the columns of `X` (pairwisely).

Usage

```
pairwiseKStest(X,alternative="greater")
```

Arguments

X	Numeric; a matrix (Bvals output from MTest function) whose columns are to be compared.
alternative	String; letter of the value, but the argument name must be given in full. See 'ks.test' for the meanings of the possible values.

Details

Using a pairwise Kolmogorov-Smirnov (KS) test of a given matrix X. In particular, if X is the Bvals output from MTest function, pairwiseKStest establishes a guide for an educated removal of variables that are causing multicolli-nearity.

Note that the matrix $B_{n_{boot} \times (p+1)}$ (which is Bvals output from MTest function) allow us to inspect results in detail and make further tests such as boxplots, pariwise Kolmogorov-Smirnov (KS) of the predictors and so on.

Value

Returns an object of class pairwiseKStest. An object of class pairwiseKStest is a list containing at most the following components:

KSpwMatrix	p-values matrix of pairwise KS testing;
alternative	Character; indicates the alternative hypothesis.
Suggestion	Character; indicates row sums (or col sums) of KSpwMatrix suggesting the removal order in case that is the strategy for dealing with multicollinearity.

Author(s)

Víctor Morales Oñate, <victor.morales@uv.cl>, https://sites.google.com/site/moralesonatevictor/,https://www.linkedin.com/in/vmoralesonate/ Bolívar Morales Oñate, <bmoralesonate@gmail.com>, https://sites.google.com/site/moralesonatevictor/

References

Morales-Oñate, V., and Morales-Oñate, B. (2023). *MTest: a Bootstrap Test for Multicollinearity*. Revista Politécnica, 51(2), 53–62. doi:10.33333/rp.vol51n2.05

Examples

```
library(MTest)
data(simDataMTest)
pairwiseKStest(X=simDataMTest)
```

plot.MTest *Plot density or empirical cumulative distribution from MTest*

Description

Plot density or empirical cumulative distribution from Bvals in MTest output.

Usage

```
## S3 method for class 'MTest'
plot(x, type=1,plotly = FALSE,...)
```

Arguments

x	an object of the class "MTest"
type	Numeric; 1 if density, 2 if ecdf plot is returned
plotly	Logical; if FALSE, a ggplotly plot is returned
...	other arguments to be passed to the function ggplot

Details

This function plots density or empirical cumulative distribution function from MTest bootstrap replications.

Value

Produces a plot. No values are returned.

See Also

MTest for procedure and examples.

simDataMTest *Simulated data for MTest*

Description

This data set helps testing functions in MTest package, the generating process is documented in the reference.

Usage

```
simDataMTest
```

Format

A dataframe containing 10000 observations and four columns.

References

Morales-Oñate, V., and Morales-Oñate, B. (2023). *MTest: a Bootstrap Test for Multicollinearity*. Revista Politécnica, 51(2), 53–62. doi:10.33333/rp.vol51n2.05

Index