Package ‘GeneralisedCovarianceMeasure’

March 24, 2022

Type Package

Title Test for Conditional Independence Based on the Generalized Covariance Measure (GCM)

Version 0.2.0

Author Jonas Peters and Rajen D. Shah

Maintainer Jonas Peters <jonas.peters@math.ku.dk>

Description A statistical hypothesis test for conditional independence. It performs nonlinear regressions on the conditioning variable and then tests for a vanishing covariance between the resulting residuals. It can be applied to both univariate random variables and multivariate random vectors. Details of the method can be found in Rajen D. Shah and Jonas Peters: The Hardness of Conditional Independence Testing and the Generalised Covariance Measure, Annals of Statistics 48(3), 1514–1538, 2020.

License GPL-2

Encoding UTF-8

Imports CVST, graphics, kernlab, mgcv, stats, xgboost

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2022-03-24 08:10:05 UTC

R topics documented:

  comp.resids ................................................................. 2
  gcm.test ................................................................. 3
  GeneralisedCovarianceMeasure ........................................ 4

Index 5
### Description

This function is used for the GCM test. Other methods can be added.

### Usage

```r
comp.resids(V, Z, regr.pars, regr.method)
```

### Arguments

- **V**: A (nxp)-dimensional matrix (or data frame) with n observations of p variables.
- **Z**: A (nxp)-dimensional matrix (or data frame) with n observations of p variables.
- **regr.pars**: Some regression methods require a list of additional options.
- **regr.method**: A string indicating the regression method that is used. Currently implemented are "gam", "xgboost", "kernel.ridge", "nystrom". The regression is performed only if not both resid.XonZ and resid.YonZ are set to NULL.

### Value

Vector of residuals.

### References


### Examples

```r
set.seed(1)
n <- 250
Z <- 4*rnorm(n)
X <- 2*sin(Z) + rnorm(n)
res <- comp.resids(X, Z, regr.pars = list(), regr.method = "gam")
```
gcm.test

Test for Conditional Independence Based on the Generalized Covariance Measure (GCM)

Description
Test for Conditional Independence Based on the Generalized Covariance Measure (GCM)

Usage
gcm.test(X, Y, Z = NULL, alpha = 0.05, regr.method = "xgboost",
    regr.pars = list(), plot.residuals = FALSE, nsim = 499L,
    resid.XonZ = NULL, resid.YonZ = NULL)

Arguments
X A (nxp)-dimensional matrix (or data frame) with n observations of p variables.
Y A (nxp)-dimensional matrix (or data frame) with n observations of p variables.
Z A (nxp)-dimensional matrix (or data frame) with n observations of p variables.
alpha Significance level of the test.
regr.method A string indicating the regression method that is used. Currently implemented are "gam", "xgboost", "kernel.ridge". The regression is performed only if not both resid.XonZ and resid.YonZ are set to NULL.
regr.pars Some regression methods require a list of additional options.
plot.residuals A Boolean indicating whether some plots should be shown.
nsim An integer indicating the number of bootstrap samples used to approximate the null distribution of the test statistic.
resid.XonZ It is possible to directly provide the residuals instead of performing a regression. If set to NULL, the regression method specified in regr.method is used.
resid.YonZ It is possible to directly provide the residuals instead of performing a regression. If set to NULL, the regression method specified in regr.method is used.

Value
The function tests whether X is conditionally independent of Y given Z. The output is a list containing
• p.value: P-value of the test.
• test.statistic: Test statistic of the test.
• reject: Boolean that is true iff p.value < alpha.

References
Examples

```r
set.seed(1)
n <- 250
Z <- 4*rnorm(n)
X <- 2*sin(Z) + rnorm(n)
Y <- 2*sin(Z) + rnorm(n)
Y2 <- 2*sin(Z) + X + rnorm(n)
gcm.test(X, Y, Z, regr.method = "gam")
gcm.test(X, Y2, Z, regr.method = "gam")
```

GeneralisedCovarianceMeasure

*Package for testing conditional independence based on the Generalized Covariance Measure (GCM)*

Description

Contains the function `gcm.test` that can be used for performing a conditional independence test based on the GCM.

Author(s)

Jonas Peters `<jonas.peters@math.ku.dk>`, Rajen D. Shah

References

Index

comp.resids, 2

gcm.test, 3
GeneralisedCovarianceMeasure, 4
GeneralisedCovarianceMeasure-package
   (GeneralisedCovarianceMeasure), 4