Package 'CMDMeasure'

February 10, 2018

Title Conditional Mean Dependence Measures via Energy Statistics

Version 1.0.0

Date 2018-01-30

Description Implementation of conditional mean dependence measures and conditional mean independence tests in Shao, X., and Zhang, J. (2014) <doi:10.1080/01621459.2014.887012> and Park, T., et al. (2015) <doi:10.1214/15-EJS1047>.

Depends R (>= 3.4.0)

Suggests testthat (>= 2.0.0), energy (>= 1.7-0)

License GPL (>= 2)

LazyData true

RoxygenNote 6.0.1

Collate 'CMDMeasure-package.R' 'functions.R' 'cmdm_test.R' 'mdc.R' 'mdd.R' 'pmdc.R' 'pmdd.R'

R topics documented:

CMDMeasure-package 2
mdc 2
mdd 3
pmdc 4
pmdd 5
xxx 6

Index **8**

CMDMeasure-package *Conditional Mean Dependence Measures via Energy Statistics*

Description

CMDMeasure: A package for mutual dependence measures via energy statistics

Details

The CMDMeasure package provides measures of conditional mean dependence and tests of conditional mean independence.

Measuring conditional mean dependence

The conditional mean dependence measures include:

- asymmetric measure $\mathcal{R}_n$ based on distance covariance $\mathcal{V}_n$
- symmetric measure $\mathcal{S}_n$ based on distance covariance $\mathcal{V}_n$
- complete measure $\mathcal{Q}_n$ based on complete V-statistics
- simplified complete measure $\mathcal{Q}_n^\star$ based on incomplete V-statistics
- asymmetric measure $\mathcal{J}_n$ based on complete measure $\mathcal{Q}_n$
- simplified asymmetric measure $\mathcal{J}_n^\star$ based on simplified complete measure $\mathcal{Q}_n^\star$
- symmetric measure $\mathcal{I}_n$ based on complete measure $\mathcal{Q}_n$
- simplified symmetric measure $\mathcal{I}_n^\star$ based on simplified complete measure $\mathcal{Q}_n^\star$

Testing conditional mean independence

The conditional mean independence tests based on the conditional mean dependence measures are implemented as permutation tests.

Author(s)

Ze Jin <zj58@cornell.edu>, Shun Yao <shunyao2@illinois.edu>,
Xiaofeng Shao <xshao@illinois.edu>, David S. Matteson <matteson@cornell.edu>

mdc *Martingale Difference Correlation*

Description

`mdc` measures conditional mean dependence of `Y` given `X`, where each contains one variable (univariate) or more variables (multivariate).

Usage

```
mdc(X, Y, center = "U")
```

Arguments

X	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Y	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
center	The approach for centering, including • U: U-centering which leads to an unbiased estimator; • D: double-centering which leads to a biased estimator.

Value

mdc returns the value of squared martingale difference correlation.

References

Shao, X., and Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. Journal of the American Statistical Association, 109(507), 1302-1318. http://dx.doi.org/10.1080/01621459.2014.887012.

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. Electronic Journal of Statistics, 9(1), 1492-1517. http://dx.doi.org/10.1214/15-EJS1047.

Examples

```
# X, Y are 10 x 2 matrices with 10 samples and 2 variables
X <- matrix(rnorm(10 * 2), 10, 2)
Y <- matrix(rnorm(10 * 2), 10, 2)

mdc(X, Y, center = "U")
mdc(X, Y, center = "D")
```

---

mdd *Martingale Difference Divergence*

Description

mdd measures conditional mean dependence of Y given X, where each contains one variable (univariate) or more variables (multivariate).

Usage

```
mdd(X, Y, compute = "C", center = "U")
```

Arguments

X	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Y	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
compute	The method for computation, including

- C: computation implemented in C code;
- R: computation implemented in R code.

center The approach for centering, including

- U: U-centering which leads to an unbiased estimator;
- D: double-centering which leads to a biased estimator.

Value

mdd returns the value of squared martingale difference divergence.

References

Shao, X., and Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. Journal of the American Statistical Association, 109(507), 1302-1318. http://dx.doi.org/10.1080/01621459.2014.887012.

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. Electronic Journal of Statistics, 9(1), 1492-1517. http://dx.doi.org/10.1214/15-EJS1047.

Examples

```
# X, Y are vectors with 10 samples and 1 variable
X <- rnorm(10)
Y <- rnorm(10)

mdd(X, Y, compute = "C")
mdd(X, Y, compute = "R")

# X, Y are 10 x 2 matrices with 10 samples and 2 variables
X <- matrix(rnorm(10 * 2), 10, 2)
Y <- matrix(rnorm(10 * 2), 10, 2)

mdd(X, Y, center = "U")
mdd(X, Y, center = "D")
```

---

pmdc *Partial Martingale Difference Correlation*

Description

pmdc measures conditional mean dependence of Y given X conditioning on Z, where each contains one variable (univariate) or more variables (multivariate).

Usage

```
pmdc(X, Y, Z)
```

Arguments

X	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Y	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Z	A vector, matrix or data frame, where rows represent samples, and columns represent variables.

Value

pmdc returns the value of squared partial martingale difference correlation.

References

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. Electronic Journal of Statistics, 9(1), 1492-1517. http://dx.doi.org/10.1214/15-EJS1047.

Examples

```
# X, Y, Z are 10 x 2 matrices with 10 samples and 2 variables
X <- matrix(rnorm(10 * 2), 10, 2)
Y <- matrix(rnorm(10 * 2), 10, 2)
Z <- matrix(rnorm(10 * 2), 10, 2)

pmdc(X, Y, Z)
```

---

pmdd *Partial Martingale Difference Divergence*

Description

pmdd measures conditional mean dependence of Y given X conditioning on Z, where each contains one variable (univariate) or more variables (multivariate).

Usage

```
pmdd(X, Y, Z)
```

Arguments

X	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Y	A vector, matrix or data frame, where rows represent samples, and columns represent variables.
Z	A vector, matrix or data frame, where rows represent samples, and columns represent variables.

Value

pmdd returns the value of squared partial martingale difference divergence.

References

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. Electronic Journal of Statistics, 9(1), 1492-1517. http://dx.doi.org/10.1214/15-EJS1047.

Examples

```
# X, Y, Z are vectors with 10 samples and 1 variable
X <- rnorm(10)
Y <- rnorm(10)
Z <- rnorm(10)

pmdd(X, Y, Z)

# X, Y, Z are 10 x 2 matrices with 10 samples and 2 variables
X <- matrix(rnorm(10 * 2), 10, 2)
Y <- matrix(rnorm(10 * 2), 10, 2)
Z <- matrix(rnorm(10 * 2), 10, 2)

pmdd(X, Y, Z)
```

---

xxx *Mutual Dependence Measures*

Description

cmdm measures mutual dependence of all components in X, where each component contains one variable (univariate) or more variables (multivariate).

Usage

```
xxx(x, y)
```

Arguments

x	A matrix or data frame, where rows represent samples, and columns represent variables.
y	A matrix or data frame, where rows represent samples, and columns represent variables.

Value

cmdm returns a list including the following components:

stat	The value of the mutual dependence measure.
dist	The distances between all components from all samples.

References

Shao, X., and Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. Journal of the American Statistical Association, 109(507), 1302-1318. https://arxiv.org/abs/1709.02532.

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. Electronic Journal of Statistics, 9(1), 1492-1517. https://arxiv.org/abs/1709.02532.

Examples

```
# X, Y is a 10 x 3 matrix with 10 samples and 3 variables
X <- matrix(rnorm(10 * 3), 10, 3)
Y <- matrix(rnorm(10 * 3), 10, 3)
```

Index

CMDMeasure *(*CMDMeasure-package*)*, 2
CMDMeasure-package, 2

mdc, 2
mdd, 3

pmdc, 4
pmdd, 5

xxx, 6