

# Package ‘comets’

February 26, 2024

**Type** Package

**Title** Covariance Measure Tests for Conditional Independence

**Version** 0.0-1

**Description** Covariance measure tests for conditional independence testing against conditional covariance and nonlinear conditional mean alternatives. Contains versions of the generalised covariance measure test (Shah and Peters, 2020, <doi:10.1214/19-aos1857>) and projected covariance measure test (Lundborg et al., 2023, <doi:10.48550/arXiv.2211.02039>). Applications can be found in Kook and Lundborg (2024, <doi:10.48550/arXiv.2402.14416>).

**Imports** ranger, glmnet, Formula

**License** GPL-3

**Encoding** UTF-8

**RoxygenNote** 7.2.3

**Suggests** testthat (>= 3.0.0), ggplot2, tidyr, dplyr

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Lucas Kook [aut, cre],  
Anton Rask Lundborg [ctb]

**Maintainer** Lucas Kook <lucasheinrich.kook@gmail.com>

**Repository** CRAN

**Date/Publication** 2024-02-26 16:30:15 UTC

## R topics documented:

comet	2
gcm	3
pcm	4

<b>Index</b>	<b>7</b>
--------------	----------

---

`comet`*Covariance measure tests with formula interface*

---

**Description**

Covariance measure tests with formula interface

**Usage**

```
comet(formula, data, test = c("gcm", "pcm"), ...)
```

**Arguments**

<code>formula</code>	Formula of the form $Y \sim X \mid Z$ for testing Y independent of X given Z.
<code>data</code>	Data.frame containing the variables in formula.
<code>test</code>	Character string; "gcm" or "pcm".
<code>...</code>	Additional arguments passed to test.

**Details**

Formula-based interface for the generalised and projected covariance measure tests.

**Value**

Object of class "gcm" or "pcm" and "htest". See [gcm](#) and [pcm](#) for details.

**References**

Kook, L. & Lundborg A. R. (2024). Algorithm-agnostic significance testing in supervised learning with multimodal data. arXiv preprint. doi: [10.48550/arXiv.2402.14416](https://doi.org/10.48550/arXiv.2402.14416)

**Examples**

```
tn <- 3e2
df <- data.frame(y = rnorm(tn), x1 = rnorm(tn), x2 = rnorm(tn), z = rnorm(tn))
comet(y ~ x1 + x2 | z, data = df, test = "gcm")
```

**Description**

Generalised covariance measure test using random forests

**Usage**

```
gcm(
  Y,
  X,
  Z,
  alternative = c("two.sided", "less", "greater"),
  reg_YonZ = "rf",
  reg_XonZ = "rf",
  args_XonZ = NULL,
  ...
)
```

**Arguments**

Y	Vector of response values. Can be supplied as a numeric vector or a single column matrix.
X	Matrix or data.frame of covariates.
Z	Matrix or data.frame of covariates.
alternative	A character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less"
reg_YonZ	Character string or function specifying the regression for Y on Z.
reg_XonZ	Character string or function specifying the regression for X on Z.
args_XonZ	Additional arguments passed to reg_XonZ.
...	Additional arguments passed to reg_YonZ

**Details**

The generalised covariance measure test tests whether the conditional covariance of Y and X given Z is zero.

**Value**

Object of class 'gcm' and 'htest' with the following components:

statistic	The value of the test statistic.
p.value	The p-value for the hypothesis

parameter	In case $X$ is multidimensional, this is the degrees of freedom used for the chi-squared test.
hypothesis	String specifying the null hypothesis .
null.value	String specifying the null hypothesis.
method	The string "Generalised covariance measure test".
data.name	A character string giving the name(s) of the data.
rY	Residuals for the Y on Z regression.
rX	Residuals for the X on Z regression.

## References

Rajen D. Shah, Jonas Peters "The hardness of conditional independence testing and the generalised covariance measure," The Annals of Statistics, 48(3), 1514-1538. doi: [10.1214/19aos1857](https://doi.org/10.1214/19aos1857)

## Examples

```
X <- matrix(rnorm(3e2), ncol = 2)
colnames(X) <- c("X1", "X2")
Z <- matrix(rnorm(3e2), ncol = 2)
colnames(Z) <- c("Z1", "Z2")
Y <- rnorm(150) # X[, 2] + Z[, 2] + rnorm(150)
(gcm1 <- gcm(Y, X, Z))
```

---

pcm

*Projected covariance measure test for conditional mean independence*

---

## Description

Projected covariance measure test for conditional mean independence

## Usage

```
pcm(
  Y,
  X,
  Z,
  rep = 1,
  est_vhat = TRUE,
  reg_YonXZ = "rf",
  reg_YonZ = "rf",
  args_YonXZ = NULL,
  args_YonZ = NULL,
  mtry = identity,
  ...
)
```

**Arguments**

Y	Vector of response values. Can be supplied as a numeric vector or a single column matrix.
X	Matrix or data.frame of covariates.
Z	Matrix or data.frame of covariates.
rep	Number of repetitions with which to repeat the PCM test
est_vhat	Logical; whether to estimate the variance functional
reg_YonXZ	Character string or function specifying the regression for Y on X and Z, default is "rf" for random forest.
reg_YonZ	Character string or function specifying the regression for Y on Z, default is "rf" for random forest.
args_YonXZ	Arguments passed to reg
args_YonZ	Arguments passed to reg
mtry	Argument passed to ranger
...	Additional arguments passed to ranger

**Details**

The projected covariance measure test tests whether the conditional mean of Y given X and Z depends on X.

**Value**

Object of class 'pcm' and 'htest' with the following components:

statistic	The value of the test statistic.
p.value	The p-value for the hypothesis
parameter	In case X is multidimensional, this is the degrees of freedom used for the chi-squared test.
hypothesis	Null hypothesis of conditional mean independence.
null.value	Null hypothesis of conditional mean independence.
method	The string "Projected covariance measure test".
data.name	A character string giving the name(s) of the data.
check.data	A data.frame containing the residuals for plotting.

**References**

Lundborg, A. R., Kim, I., Shah, R. D., & Samworth, R. J. (2022). The Projected Covariance Measure for assumption-lean variable significance testing. arXiv preprint. doi: [10.48550/arXiv.2211.02039](https://doi.org/10.48550/arXiv.2211.02039)

**Examples**

```
n <- 150
X <- matrix(rnorm(2 * n), ncol = 2)
colnames(X) <- c("X1", "X2")
Z <- matrix(rnorm(2 * n), ncol = 2)
colnames(Z) <- c("Z1", "Z2")
Y <- rnorm(n) # X[, 2] + Z[, 2] + rnorm(1e3)
(pcm1 <- pcm(Y, X, Z))
```

# Index

comet, 2

gcm, 2, 3

pcm, 2, 4