# Package ‘TPCselect’

October 12, 2022

**Type**  Package  
**Title**  Variable Selection via Threshold Partial Correlation  
**Version**  0.8.1  
**License**  GPL-3  
**Encoding**  UTF-8  
**RoxygenNote**  7.2.1  
**Imports**  stats, corpcor, psych, MASS, KernSmooth  
**Suggests**  testthat (>= 3.0.0)  
**Config/testthat/edition**  3  
**NeedsCompilation**  no  
**Author**  Cynthia Shao [aut, trl, cre]  
Runze Li [aut]  
**Maintainer**  Cynthia Shao <cynyu.shao@gmail.com>  
**Repository**  CRAN  
**Date/Publication**  2022-09-15 08:00:12 UTC  

## R topics documented:

- `generate_toy_pldata`  .......................................................... 2  
- `TPC`  .................................................................................. 2  
- `TPC_BIC`  ........................................................................... 3  
- `TPC_pl`  ............................................................................... 4  
- `TPC_pl_BIC`  ....................................................................... 5  

**Index**  7
generate_toy_pldata  A function to generate toy partial linear model data

Description
A function to generate toy partial linear model data

Usage
generate_toy_pldata()

Value
a list that contains:
• ysample - the response variable
• x_cov - the linear predictors
• times - the non-linear term

TPC  Variable Selection via Thresholded Partial Correlation

Description
These are the main selection functions with fixed significance level s and constant. The function TPC implements the thresholded partial correlation (TPC) approach to selecting important variables in linear models of Li et al. (2017). The function TPC_p1 implements the thresholded partial correlation approach to selecting important variables in partial linear models of Liu et al. (2018). This function also extends the PC-simple algorithm of Bühlmann et al. (2010) to partial linear models.

Usage
TPC(y, x, s = 0.05, constant = 1, method = "threshold")
TPCselect(y, x, s = 0.05, constant = 1, method = "threshold")

Arguments
y  response vector;
x  covariate matrix;
s  a numeric value that used as significance level(s) for partial correlation test.
constant  a value that used as the tuning constant for partial correlation test. constant is treated as 1 when method is "simple".
method  the method to be used; default set as method = "threshold"; "simple" is also available.
TPC_BIC

Value

TPC.object a TPC object, which extends the lm object. New attributes are:

• beta - the fitted coefficients
• selected_index - the selected coefficients indices

Examples

#generate sample data
p = 200
n = 200
truebeta <- c(c(3,1.5,0,0,2),rep(0,p-5))
rho = 0.3
sigma = matrix(0,p+1,p+1)
for(i in 1:(p+1)){
  for(j in 1:(p+1)){
    sigma[i,j] = rho^abs(i-j)
  }
}
x_error = 0.9*MASS::mvrnorm(n,rep(0,p+1),sigma) + 0.1*MASS::mvrnorm(n,rep(0,p+1),9*sigma)
x = x_error[,1:p]
error = x_error[,p+1]
y = x%*%truebeta + error

#perform variable selection via partial correlation
TPC.fit = TPC(y,x,0.05,1,method="threshold")
TPC.fit$beta

TPC_BIC

Variable Selection via Thresholded Partial Correlation

Description

Use BIC to select the best s and constant over grids.

Usage

TPC_BIC(y, x, s = 0.05, constant = 1, method = "threshold")

Arguments

y response vector;
x covariate matrix;
s a value or a vector that used as significance level(s) for partial correlation test. BIC will be used to select the best s.
constant a value or a vector that used as the tuning constant for partial correlation test. BIC will be used to select the best constant. constant is treated as 1 when method is "simple".

method the method to be used; default set as method = "threshold"; "simple" is also available.

Value

TPC.object a TPC object, which extends the lm object. New attributes are:

- beta - the fitted coefficients
- selected_index - the selected coefficients indices

Examples

#generate sample data
p = 200
n = 200
truebeta <- c(c(3,1.5,0,0,2),rep(0,p-5))
rho = 0.3
sigma = matrix(0,p+1,p+1)
for(i in 1:(p+1)){
  for(j in 1:(p+1)){
    sigma[i,j] = rho^abs(i-j)
  }
}
x_error = 0.9*MASS::mvrnorm(n,rep(0,p+1),sigma) + 0.1*MASS::mvrnorm(n,rep(0,p+1),9*sigma)
x = x_error[,1:p]
error = x_error[,p+1]
y = x%*%truebeta + error

#perform variable selection via partial correlation
TPC.fit = TPC_BIC(y,x,0.05,c(1,1.5),method="threshold")
TPC.fit$beta
**Usage**

TPC_pl(y, x, u = NULL, s = 0.05, constant = 1, method = "threshold", ...)

**Arguments**

- **y**: response vector;
- **x**: covariate matrix;
- **u**: non-parametric variable, should be a vector;
- **s**: a numeric value or vector that used as the significance level(s) for the partial correlation tests;
- **constant**: a value that used as the tuning constant for partial correlation test. constant is treated as 1 when method is "simple";
- **method**: the method to be used; default set as method = "threshold"; "simple" is also available.
- **...**: smoothing parameters and functions: kernel, degree, and bandwidth h.

**Value**

TPC.object a TPC object, which extends the lm object. New attributes are:

- **beta**: the fitted coefficients
- **selected_index**: the selected coefficients indices

**Examples**

#generate partial linear data
samples <- generate_toy_pldata()
y <- samples[[1]]
x <- samples[[2]]
times <- samples[[3]]

#perform variable selection via partial correlation
TPC.fit = TPC_pl(y,x,times,0.05,1,method="threshold")
TPC.fit$beta

---

**TPC_pl_BIC**

Variable Selection via Thresholded Partial Correlation

**Description**

Use BIC to select the best s and constant over grids.

**Usage**

TPC_pl_BIC(y, x, u = NULL, s = 0.05, constant = 1, method = "threshold", ...)

---
Arguments

\( y \)  
response vector;

\( x \)  
covariate matrix;

\( u \)  
non-parametric variable, should be a vector;

\( s \)  
a value or a vector that used as significance level(s) for partial correlation test. BIC will be used to select the best \( s \).

\( \text{constant} \)  
a value or a vector that used as the tuning constant for partial correlation test. BIC will be used to select the best constant. \( \text{constant} \) is treated as 1 when method is "simple".

\( \text{method} \)  
the method to be used; default set as \( \text{method} = \text{"threshold"} \); "simple" is also available.

\( \ldots \)  
smoothing parameters and functions: kernel, degree, and bandwidth \( h \).

Value

TPC.object a TPC object, which extends the \( \text{lm} \) object. New attributes are:

- beta - the fitted coefficients
- selected_index - the selected coefficients indices

Examples

```r
#generate partial linear data
samples <- generate_toy_pldata()
y <- samples[[1]]
x <- samples[[2]]
times <- samples[[3]]

#perform variable selection via partial correlation
TPC.fit = TPC_pl_BIC(y,x,times,0.05,c(1,1.5),method="threshold")
TPC.fit$beta
```
Index

generate_toy_pldata, 2

TPC, 2
TPC_BIC, 3
TPC_pl, 4
TPC_pl_BIC, 5
TPCselect (TPC), 2