Package 'HIMA2.0'

May 07, 2021

Type Package

Title High-dimensional Mediation Analysis and its application in genome-wide DNA methylation data.

Authors Chamila Perera, Haixiang Zhang, Yinan Zheng, Lifang Hou, Annie Qu, Lei Liu

Description We propose a new method to estimate and test high-dimensional mediation effects based on sure independent screening and de-biased Lasso inference procedure. A multiple-testing procedure is used to accurately control the false discovery rate (FDR) when testing high-dimensional mediation hypotheses.

References

Perera C, Zhang H, Zheng Y, Hou L, Qu A, Liu L. High-dimensional Mediation Analysis and its application in genome-wide DNA methylation data. PLoS Comput Biol. 2020. Submitted.

Depends R (>= 3.4) hdi, HDMT, MASS

HIMA2.0   High-dimensional Mediation Analysis Version 2

Description
HIMA2.0 is used to estimate and test high-dimensional mediation effects.

Usage
HIMA2.0 (X, Y, M, Z)

Arguments

X	a vector of exposure.
Y	a vector of outcome. Can be either continuous or binary (0-1).
M	a data.frame or matrix of high-dimensional mediators. Rows represent samples, columns represent variables.
Z	a data.frame or matrix of covariates dataset for testing the association M ~ X and Y ~ M.

Value
A data.frame containing mediation testing results of selected mediators.
- M: Selected mediators

- alpha: coefficient estimates of exposure (X) –> mediators (M).
- alpha_SE: Standard error estimates of exposure (X) –> mediators (M).
- beta: coefficient estimates of mediators (M) –> outcome (Y) (adjusted for exposure).
- beta_SE: Standard error estimates of mediators (M) –> outcome (Y) (adjusted for exposure).
- alpha*beta: mediation effect.
- p_val: statistical significance of the mediator.

Example

```
n <- 400 # sample size
p <- 1000 # the dimension of mediators
q<-2 #the number of adjusted covariates

# the regression coefficients alpha (exposure --> mediators)
alpha <- matrix(0,1,p)

# the regression coefficients beta (mediators --> outcome)
beta <- matrix(0,1,p)

# the first five markers are true mediators.
alpha[1:5] <- c(0.20,0.25,0.15,0.30,0.35)
beta[1:5] <- c(0.20,0.25,0.15,0.30,0.35)

alpha[6] <- 0.1
beta[7] <- 0.1

##Regression coefficients eta (covariates --> outcome)
eta <- matrix(0.3,p,q)
## the regression coefficients gamma (exposure --> outcome)
gamma <- matrix(0.5,1,1)
## the regression coefficients delta (covariates --> mediator)
delta <- matrix(0.5,1,q)

##Correlation matrix
sigma_e <- matrix(0,p,p)
rho <- 0.25  # the correlation of X
for (i in 1:p) {
 for (j in 1:p) {
   sigma_e[i,j]=(rho^(abs(i-j)));
 }
}
```

```
# Generate simulation data
simdat = simHIMA2(n, p,q,alpha, beta, seed=1234)

# HIMA2 output
hima2.0.fit <- HIMA2.0(X=simdat$X, Y=simdat$Y, M=simdat$M, Z=simdat$Z)
hima2.0.fit
```

---

simHIMA2.0   Simulation Data Generator for High-dimensional Mediation Analysis2

Description

simHIMA2.0 is used to generate simulation data for high-dimensional mediation analysis.

Usage

simHIMA2.0(n, p, alpha, beta, binaryOutcome = FALSE, seed)

Arguments

n	an integer specifying sample size.
p	an integer specifying the dimension of mediators.
alpha	a numeric vector specifying the regression coefficients alpha (exposure → mediators).
beta	a numeric vector specifying the regression coefficients beta (mediators → outcome).
binaryOutcome	logical. Should the simulated outcome variable be binary?
seed	an integer specifying a seed for random number generation.

Example

```
n <- 400 # sample size
p <- 1000 # the dimension of covariates
q<-2 #the number of adjusted covariates

# the regression coefficients alpha (exposure --> mediators)
alpha <- matrix(0,1,p)

# the regression coefficients beta (mediators --> outcome)
beta <- matrix(0,1,p)
```

```r
# the first five markers are true mediators.
alpha[1:5] <- c(0.20,0.25,0.15,0.30,0.35)
beta[1:5] <- c(0.20,0.25,0.15,0.30,0.35)

alpha[6] <- 0.5
beta[7] <- 0.5

##Regression coefficients eta (covariates --> outcome)
eta <- matrix(0.3,p,q)
## the regression coefficients gamma (exposure --> outcome)
gamma <- matrix(0.5,1,1)
## the regression coefficients delta (exposure --> mediator)
delta <- matrix(0.5,1,q)

sigma_e <- matrix(0,p,p)
rho <- 0.25  # the correlation of X
for (i in 1:p) {
  for (j in 1:p) {
    sigma_e[i,j]=(rho^(abs(i-j)));
  }
}

# Generate simulation data
simdat = simHIMA2.0(n, p, q, alpha, beta, seed=1234)
```