Package ‘freebird’

July 22, 2025

Title Estimation and Inference for High Dimensional Mediation and Surrogate Analysis

Version 1.0

Description Estimates and provides inference for quantities that assess high dimensional mediation and potential surrogate markers including the direct effect of treatment, indirect effect of treatment, and the proportion of treatment effect explained by a surrogate/mediator; details are described in Zhou et al (2022) <doi:10.1002/sim.9352> and Zhou et al (2020) <doi:10.1093/biomet/asaa016>. This package relies on the optimization software 'MOSEK', <`https://www.mosek.com`>.

Depends R (>= 3.5.0), scalreg, Rmosek, Matrix

Imports stats, MASS

License GPL (>= 2)

RoxygenNote 7.1.1

Encoding UTF-8

NeedsCompilation yes

Author Ruixuan Zhou [aut, cph],
Dave Zhao [aut, cph],
Layla Parast [cre]

Maintainer Layla Parast <`parast@austin.utexas.edu`>

Repository CRAN

Date/Publication 2022-09-27 10:20:05 UTC

Contents

`hilma` *Estimation and Inference for High Dimensional Mediation Analysis*

Description

This function implements the estimation and inference for the indirect effect in high dimensional linear mediation analysis models. It provides estimates and p-values under both incomplete mediation, where a direct effect may exist, as well as complete mediation, where the direct effect is known to be absent.

Usage

```
hilma(
  Y,
  G,
  S,
  mediation_setting = "incomplete",
  tuning_method = "uniform",
  lam_list = NA,
  min.ratio = 0.1,
  n.lambda = 5,
  center = TRUE
)
```

Arguments

`Y`	The n-dimensional outcome vector.
`G`	The n by p mediator matrix. p can be larger than n.
`S`	The n by q exposure matrix. q can be 1, and q < n is required.
`mediation_setting`	Either ‘incomplete’ or ‘complete’
`tuning_method`	‘uniform’ or ‘aic’, the default is ‘uniform’
`lam_list`	tuning parameter for uniform tuning or list of tuning parameter for aic tuning
`min.ratio`	the ratio of the minimum lambda to the maximum
`n.lambda`	number of tuning parameters to choose from
`center`	center the data or not, the default is TRUE

Value

A list with components:

`beta_hat`	estimated indirect effect
`alpha1_hat`	estimated direct effect
`pvalue_beta_hat`	the p value for testing the significance of the indirect effect
`lambda_used`	lambda used during optimization

Author(s)

Ruixuan Zhou

Examples

```
n = 30
p = 50
q = 2
G = MASS::mvrnorm(n, rep(0,p), diag(p))
S = as.matrix(MASS::mvrnorm(n, rep(0,q), diag(q)))
Y = as.matrix(rnorm(n))
out = hilma(Y,G,S, mediation_setting = 'complete', tuning_method = 'uniform', lam_list = 0.2)
out
```

`ptehd` *Proportion of treatment effect explained by high-dimensional surrogates*

Description

Estimates the proportion of the treatment effect explained by the indirect effect via high-dimensional surrogates.

Usage

```
ptehd(Yt, Yc, St, Sc, lambda_range = c(0, 1))
```

Arguments

`Yt`	The n-dmensional outcome vector in the treatment group.
`Yc`	The n-dmensional outcome vector in the control group.
`St`	The n x p matrix of surrogates in the treatment group.
`Sc`	The n x p matrix of surrogates in the treatment group.
`lambda_range`	Min and max of range of range of tuning parameter to use during the constrained l1 optimization step.

Value

A list with components:

`est_id`	Estimate of indirect effect, defined as $\int E(Y\|S = s, Z = 1)dF(s\|Z = 1) - \int E(Y\|S = s, Z = 0)dF(s\|Z = 0)$
`sd_id`	Standard deviation of indirect effect estimate
`est_total`	Estimate of total effect
`sd_total`	Standard deviation of total effect estimate
`V`	Covariance matrix of (est_id, est_total)

`est_R`	Estimate of proportion of treatment effect explained by surrogates
`sd_R`	Standard deviation of proportion estimate
`lambda_used`	lambda used during optimization

Author(s)

Ruixuan Zhou

Examples

```
n = 10
St = replicate(n, rnorm(20, mean = 1))
Sc = replicate(n, rnorm(20))
Yt = 1 + rowSums(St) / 2 + rnorm(n)
Yc = rowSums(Sc) / 3 + rnorm(n)
# Requires installation of mosek to run
## Not run:
out = ptehd(Yt, Yc, St, Sc)

## End(Not run)
```

Index