

Package ‘MRTL’

January 3, 2025

Type Package

Title Tran-Ancestry Mendelian Randomization using Transfer Learning

Version 0.1.1

Author Yihe Yang

Maintainer Yihe Yang <xy1234@case.edu>

Description This package performs tran-ancestry multivariate Mendelian randomization. The central idea is using transfer learning to gain power from European population. It is characterized by the removal of measurement error bias caused by the estimation error of GWAS effect size estimates using an unbiased estimating function in measurement error analysis. It also utilizes the minimum concave penalty (MCP) to dynamically detect and remove potential pleiotropy, making the causal effect robust to pleiotropy.

License MIT LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 4.4.0),

Imports MASS,
data.table,
FDREstimation,
susieR,
MRBEEEX,
CppMatrix

RoxygenNote 7.3.2

Remotes harryiheyang/MRBEEEX,
harryiheyang/CppMatrix,
stephenslab/susieR

ORCID 0000-0001-6563-3579

Contents

errorCov	2
filter_align	2
Joint.test	3
MRBEE_IMRP_UV	4
MRBEE_IPOD_SuSiE	5
MRBEE_TL	6

MRcML_IPOD	8
MRcML_IPOD_SuSiE	9
MRcML_TL	10
TR_MRBEE_IMRP_UV	12

Index	14
--------------	-----------

errorCov	<i>Estimate Error Covariance Matrix Using GWAS Insignificant Effects</i>
----------	--

Description

This function estimates the error covariance matrix by subsampling a proportion of insignificant GWAS effects and calculating their correlation coefficients.

Usage

```
errorCov(  
  ZMatrix,  
  Zscore.cutoff = 2,  
  subsampling.ratio = 0.1,  
  subsampling.time = 1000  
)
```

Arguments

- ZMatrix A matrix of Z-scores for exposure and outcome, with the outcome GWAS in the last column.
- Zscore.cutoff The cutoff for significance. Defaults to 2.
- subsampling.ratio The proportion of effects to subsample for each iteration. Defaults to 0.1.
- subsampling.time The number of subsampling iterations. Defaults to 1000.

Value

A matrix representing the estimated error covariance.

filter_align	<i>Filter and Align GWAS Data to a Reference Panel</i>
--------------	--

Description

The filter_align function processes a list of GWAS summary statistics data frames, harmonizes alleles according to a reference panel, removes duplicates, and aligns data to common SNPs. It's used to prepare data for further analysis such as LDSC.

Usage

```
filter_align(gwas_data_list, ref_panel, allele_match = TRUE)
```

Arguments

- `gwas_data_list` A list of data.frames where each data.frame contains GWAS summary statistics for a trait. Each data.frame should include columns for SNP identifiers, Z-scores of effect size estimates, sample sizes (N), effect allele (A1), and reference allele (A2).
- `ref_panel` A data.frame containing the reference panel data. It must include columns for SNP, A1, and A2.
- `allele_match` Logical. Whether to match alleles. Default TRUE.

Details

The function performs several key steps: adjusting alleles according to a reference panel, removing duplicate SNPs, and aligning all GWAS data frames to a set of common SNPs. This is often a necessary preprocessing step before performing genetic correlation and heritability analyses.

Value

A list of data.frames, each corresponding to an input GWAS summary statistics data frame, but filtered, harmonized, and aligned to the common SNPs found across all data frames.

Examples

```
## Not run:
# Assuming GWAS_List and ref_panel are already defined:
GWAS_List <- filter_align(GWAS_List, ref_panel)

## End(Not run)
```

Joint.test

Perform Joint Test on GWAS Summary Data

Description

This function conducts a joint test on GWAS summary data, calculating Chi-squared values and corresponding P-values based on allele-aligned Z-scores and the correlation matrix of estimation errors for these Z-scores.

Usage

```
Joint.test(bZ, RZ)
```

Arguments

- `bZ` A matrix of allele-aligned Z-scores from GWAS summary data, where each row represents a different genetic variant and each column represents a different trait or study. The matrix dimensions are $n \times p$, with n being the number of variants and p being the number of traits or studies.
- `RZ` The correlation matrix of estimation errors for the Z-scores in 'bZ'. This matrix should be $p \times p$, where p is the number of traits or studies.

Value

A data frame with two columns: 'Chi2', containing the Chi-squared values for each genetic variant, and 'P', containing the corresponding P-values.

Examples

```
## Not run:
# Example usage:
# Assuming bZ is your matrix of allele-aligned Z-scores and RZ is the correlation matrix
result <- Joint.test(bZ, RZ)

## End(Not run)
```

MRBEE_IMRP_UV

*Univariable MR Estimation using MRBEE Method***Description**

This function performs univariable Mendelian Randomization (MR) analysis using the MRBEE method.

Usage

```
MRBEE_IMRP_UV(
  by,
  bx,
  byse,
  bxse,
  Rxy,
  max.iter = 30,
  max.eps = 1e-04,
  pv.thres = 0.05,
  var.est = "variance",
  FDR = T,
  adjust.method = "Sidak"
)
```

Arguments

by	Vector of GWAS effect sizes for the outcome (n x 1).
bx	Vector of GWAS effect sizes for the exposure (n x 1).
byse	Vector of standard errors (SE) for the GWAS effect sizes of the outcome (n x 1).
bxse	Vector of SE for the GWAS effect sizes of the exposure (n x 1).
Rxy	Correlation matrix (p+1 x p+1) of the exposures and outcome, with the outcome being the last.
max.iter	Maximum number of iterations for the estimation process. Defaults to 30.
max.eps	Tolerance level for convergence. Defaults to 1e-4.
pv.thres	P-value threshold for pleiotropy detection. Defaults to 0.05.

<code>var.est</code>	Method for estimating the standard error in the pleiotropy test. Can be "robust", "variance", or "ordinal".
<code>FDR</code>	Logical indicating whether to apply False Discovery Rate (FDR) correction. Defaults to TRUE.
<code>adjust.method</code>	Method for estimating q-values, defaults to "Sidak".

Value

A list containing the estimated causal effect, its covariance, and pleiotropy

MRBEE_IPOD_SuSiE	<i>Estimate Causal Effect with MRBEE and SuSiE</i>
------------------	--

Description

This function estimates the causal effect using a bias-correction estimating equation, considering potential pleiotropy and measurement errors, and using SuSiE to select the causal exposures.

Usage

```
MRBEE_IPOD_SuSiE(
  by,
  bX,
  byse,
  bXse,
  Rxy,
  tauvec = seq(3, 30, 3),
  admm.rho = 3,
  Lvec = c(1:6),
  ebic.theta = 1,
  ebic.gamma = 2,
  susie.iter = 200,
  pip.thres = 0.3,
  max.iter = 50,
  max.eps = 1e-04,
  reliability.thres = 0.8,
  ridge.diff = 100,
  sampling.time = 100,
  sampling.iter = 10
)
```

Arguments

<code>by</code>	A vector (n x 1) of the GWAS effect size of outcome.
<code>bX</code>	A vector (n x 1) of the liner predictor of the transferred exposure's effect.
<code>byse</code>	A vector (n x 1) of the GWAS effect size SE of outcome.
<code>bXse</code>	A vector (n x 1) of the standard error of the liner predictor of the transferred exposure's effect.

Rxy	A matrix (p+1 x p+1) of the correlation matrix of the p exposures and outcome. The first one should be the transferred linear predictor and last one should be the outcome.
tauvec	The candidate vector of tuning parameters for the MCP penalty function. Default is seq(3, 30, by=3).
admm.rho	The tuning parameter in the nested ADMM algorithm. Default is 3.
Lvec	A vector of the number of single effects used in SuSiE. Default is c(1:6).
ebic.gamma	A scale of tuning parameter of horizontal pleiotropy in extended BIC. Default is 2.
susie.iter	A scale of the maximum number of iterations used in SuSiE. Default is 200.
pip.thres	A scale of PIP threshold for calibrating causality used in SuSiE. Default is 0.3.
max.iter	Maximum number of iterations for causal effect estimation. Default is 50.
max.eps	Tolerance for stopping criteria. Default is 1e-4.
reliability.thres	A scale of threshold for the minimum value of the reliability ratio. If the original reliability ratio is less than this threshold, only part of the estimation error is removed so that the working reliability ratio equals this threshold. Default is 0.8.
ridge.diff	A scale of parameter on the differences of causal effect estimate in one credible set. Defaults to 10.
sampling.time	A scale of number of subsampling in estimating the standard error. Default is 100.
sampling.iter	A scale of iteration in subsampling in estimating the standard error. Default is 10.
ebic.delta	A scale of tuning parameter of causal effect estimate in extended BIC. Default is 1.

Value

A list containing the estimated causal effect, its covariance, and pleiotropy.

MRBEE_TL

Estimate Non-Transferable Causal Effect with MRBEE and SuSiE

Description

This function estimates the non-transferable causal effect using a bias-correction estimating equation, considering potential pleiotropy and measurement errors, and using SuSiE to select the non-transferable causal effect.

Usage

```
MRBEE_TL(
  by,
  bX,
  byse,
  bXse,
```

```

Rxy,
theta.source,
theta.source.cov,
tauvec = seq(3, 30, 3),
admm.rho = 3,
Lvec = c(1:6),
ebic.delta = 1,
ebic.gamma = 2,
transfer.coef = 1,
susie.iter = 200,
pip.thres = 0.3,
max.iter = 50,
max.eps = 1e-04,
reliability.thres = 0.8,
ridge.diff = 100,
sampling.time = 100,
sampling.iter = 10
)

```

Arguments

by	A vector (n x 1) of the GWAS effect size of outcome.
bX	A matrix (n x p) of the GWAS effect sizes of p exposures.
byse	A vector (n x 1) of the GWAS effect size SE of outcome.
bXse	A matrix (n x p) of the GWAS effect size SEs of p exposures.
Rxy	A matrix (p+1 x p+1) of the correlation matrix of the p exposures and outcome. The first one should be the transferred linear predictor and last one should be the outcome.
theta.source	A vector (p x 1) of the causal effect estimate learning from the source data.
theta.source.cov	A matrix (p x p) of the covariance matrix of the causal effect estimate learning from the source data.
tauvec	The candidate vector of tuning parameters for the MCP penalty function. Default is seq(3, 30, by=3).
admm.rho	The tuning parameter in the nested ADMM algorithm. Default is 3.
Lvec	A vector of the number of single effects used in SuSiE. Default is c(1:6).
ebic.delta	A scale of tuning parameter of causal effect estimate in extended BIC. Default is 1.
ebic.gamma	A scale of tuning parameter of horizontal pleiotropy in extended BIC. Default is 2.
transfer.coef	A scale of transfer.coef of theta.source to theta.target. Default is 1.
susie.iter	A scale of the maximum number of iterations used in SuSiE. Default is 200.
pip.thres	A scale of PIP threshold for calibrating causality used in SuSiE. Default is 0.3.
max.iter	Maximum number of iterations for causal effect estimation. Default is 50.
max.eps	Tolerance for stopping criteria. Default is 1e-4.
reliability.thres	A scale of threshold for the minimum value of the reliability ratio. If the original reliability ratio is less than this threshold, only part of the estimation error is

	removed so that the working reliability ratio equals this threshold. Default is 0.8.
ridge.diff	A scale of parameter on the differences of causal effect estimate in one credible set. Defaults to 10.
sampling.time	A scale of number of subsampling in estimating the standard error. Default is 100.
sampling.iter	A scale of iteration in subsampling in estimating the standard error. Default is 10.

Value

A list containing the estimated causal effect, its covariance, and pleiotropy.

MRcML_IPOD	<i>Estimate Causal Effect with MRcML and SuSiE</i>
------------	--

Description

This function estimates the causal effect using a bias-correction estimating equation, considering potential pleiotropy and measurement errors, and using SuSiE to select the causal exposures.

Usage

```
MRcML_IPOD(
  by,
  bX,
  byse,
  bXse,
  Rxy,
  tau = 9,
  admm.rho = 3,
  max.iter = 10,
  max.eps = 1e-04,
  sampling.time = 100,
  sampling.iter = 10,
  theta.ini
)
```

Arguments

by	A vector (n x 1) of the GWAS effect size of outcome.
bX	A vector (n x 1) of the liner predictor of the transferred exposure's effect.
byse	A vector (n x 1) of the GWAS effect size SE of outcome.
bXse	A vector (n x 1) of the standard error of the liner predictor of the transferred exposure's effect.
Rxy	A matrix (p+1 x p+1) of the correlation matrix of the p exposures and outcome. The first one should be the transferred linear predictor and last one should be the outcome.
tau	A scale of tuning parameter in MCP. Defaults to 9.

admm.rho	A scale of tuning parameter in ADMM. Defaults to 3.
max.iter	Maximum number of iterations for causal effect estimation. Defaults to 10.
max.eps	Tolerance for stopping criteria. Defaults to 1e-4.
sampling.time	A scale of number of subsampling in estimating the standard error.
sampling.iter	A scale of iteration in subsampling in estimating the standard error.
theta.ini	An initial estimate of causal effect.
adjust.method	Method for estimating q-value. Defaults to "Sidak".

Value

A list containing the estimated causal effect, its covariance, and pleiotropy.

MRcML_IPOD_SuSiE	<i>Estimate Causal Effect with MRcML and SuSiE</i>
------------------	--

Description

This function estimates the causal effect using a bias-correction estimating equation, considering potential pleiotropy and measurement errors, and using SuSiE to select the causal exposures.

Usage

```
MRcML_IPOD_SuSiE(
  by,
  bX,
  byse,
  bXse,
  Rxy,
  L = min(10, ncol(bX)),
  tau = 9,
  admm.rho = 3,
  susie.iter = 200,
  pip.thres = 0.2,
  max.iter = 10,
  max.eps = 1e-04,
  ridge.diff = 100,
  sampling.time = 100,
  sampling.iter = 10,
  theta.ini = F
)
```

Arguments

by	A vector (n x 1) of the GWAS effect size of outcome.
bX	A vector (n x 1) of the liner predictor of the transferred exposure's effect.
byse	A vector (n x 1) of the GWAS effect size SE of outcome.
bXse	A vector (n x 1) of the standard error of the liner predictor of the transferred exposure's effect.

Rxy	A matrix (p+1 x p+1) of the correlation matrix of the p exposures and outcome. The first one should be the transferred linear predictor and last one should be the outcome.
L	A scale of the number of single effects used in SuSiE. Defaults to 5.
tau	A scale of tuning parameter in MCP. Defaults to 9.
admm.rho	A scale of tuning parameter in ADMM. Defaults to 3.
susie.iter	A scale of the maximum number of iterations used in SuSiE. Defaults to 200.
pip.thres	A scale of PIP threshold for calibrating causality used in SuSiE.
max.iter	Maximum number of iterations for causal effect estimation. Defaults to 10.
max.eps	Tolerance for stopping criteria. Defaults to 1e-4.
ridge.diff	A scale of parameter on the differences of causal effect estimate in one credible set. Defaults to 10.
sampling.time	A scale of number of subsampling in estimating the standard error. Defaults to 100.
sampling.iter	A scale of iteration in subsampling in estimating the standard error. Defaults to 10.
theta.ini	An initial estimate of causal effect.

Value

A list containing the estimated causal effect, its covariance, and pleiotropy.

MRcML_TL

Estimate Non-Transferable Causal Effect with MRcML and SuSiE

Description

This function estimates the non-transferable causal effect using a conditional score method, considering potential pleiotropy and measurement errors, and using SuSiE to select the non-transferable causal effect.

Usage

```
MRcML_TL(
  by,
  bX,
  byse,
  bXse,
  Rxy,
  L = min(10, ncol(bX)),
  tau = 9,
  admm.rho = 3,
  susie.iter = 200,
  pip.thres = 0.2,
  max.iter = 10,
  max.eps = 1e-04,
  ridge.diff = 100,
  sampling.time = 100,
```

```

    sampling.iter = 10,
    theta.ini = F,
    theta.source,
    theta.source.cov,
    transfer.coef = 1
)

```

Arguments

by	A vector (n x 1) of the GWAS effect size of outcome.
bX	A vector (n x 1) of the liner predictor of the transferred exposure's effect.
byse	A vector (n x 1) of the GWAS effect size SE of outcome.
bXse	A vector (n x 1) of the standard error of the liner predictor of the transferred exposure's effect.
Rxy	A matrix (p+1 x p+1) of the correlation matrix of the p exposures and outcome. The first one should be the transferred linear predictor and last one should be the outcome.
L	A scale of the number of single effects used in SuSiE. Defaults to 5.
tau	A scale of tuning parameter in MCP. Defaults to 9.
admm.rho	A scale of tuning parameter in ADMM. Defaults to 3.
susie.iter	A scale of the maximum number of iterations used in SuSiE. Defaults to 200.
pip.thres	A scale of PIP threshold for calibrating causality used in SuSiE.
max.iter	Maximum number of iterations for causal effect estimation. Defaults to 10.
max.eps	Tolerance for stopping criteria. Defaults to 1e-4.
ridge.diff	A scale of parameter on the differences of causal effect estimate in one credible set. Defaults to 10.
sampling.time	A scale of number of subsampling in estimating the standard error. Defaults to 100.
sampling.iter	A scale of iteration in subsampling in estimating the standard error. Defaults to 10.
theta.ini	An initial estimate of causal effect.
theta.source	A vector (p x 1) of the causal effect estimate learning from the source data.
theta.source.cov	A matrix (p x p) of the covariance matrix of the causal effect estimate learning from the source data.
transfer.coef	A scale of transfer.coef of theta.source to theta.target. Default is 1.

Value

A list containing the estimated causal effect, its covariance, and pleiotropy.

TR_MRBEE_IMRP_UV	<i>Estimate Non-Transferable Causal Effect with Univariable MRBEE and SuSiE</i>
------------------	---

Description

This function performs univariable Mendelian Randomization (MR) analysis using the MRBEE method.

Usage

```
TR_MRBEE_IMRP_UV(
  by,
  bx,
  byse,
  bxse,
  Rxy,
  theta.source,
  susie.iter = 100,
  pip.thres = 0.3,
  max.iter = 30,
  max.eps = 1e-04,
  pv.thres = 0.05,
  var.est = "variance",
  FDR = T,
  adjust.method = "Sidak"
)
```

Arguments

by	Vector of GWAS effect sizes for the outcome (n x 1).
bx	Vector of GWAS effect sizes for the exposure (n x 1).
byse	Vector of standard errors (SE) for the GWAS effect sizes of the outcome (n x 1).
bxse	Vector of SE for the GWAS effect sizes of the exposure (n x 1).
Rxy	Correlation matrix (p+1 x p+1) of the exposures and outcome, with the outcome being the last.
theta.source	A vector (p x 1) of the causal effect estimate learning from the source data.
susie.iter	A scale of the maximum number of iterations used in SuSiE.
pip.thres	A scale of PIP threshold for calibrating causality used in SuSiE.
max.iter	Maximum number of iterations for the estimation process. Defaults to 30.
max.eps	Tolerance level for convergence. Defaults to 1e-4.
pv.thres	P-value threshold for pleiotropy detection. Defaults to 0.05.
var.est	Method for estimating the standard error in the pleiotropy test. Can be "robust", "variance", or "ordinal".
FDR	Logical indicating whether to apply False Discovery Rate (FDR) correction. Defaults to TRUE.
adjust.method	Method for estimating q-values, defaults to "Sidak".

Value

A list containing the estimated causal effect, its covariance, and pleiotropy

Index

`errorCov`, [2](#)

`filter_align`, [2](#)

`Joint.test`, [3](#)

`MRBEE_IMRP_UV`, [4](#)

`MRBEE_IPOD_SuSiE`, [5](#)

`MRBEE_TL`, [6](#)

`MRcML_IPOD`, [8](#)

`MRcML_IPOD_SuSiE`, [9](#)

`MRcML_TL`, [10](#)

`TR_MRBEE_IMRP_UV`, [12](#)