

Package: GraphMRcML (via r-universe)

June 6, 2026

Type Package

Title Causal Network Inference via Mendelian Randomization with cML and Network Deconvolution

Version 0.1.0

Description Combines Mendelian randomization (constrained maximum likelihood, cML) and network deconvolution for inference of causal networks from GWAS summary data, as described in Lin, Xue, and Pan (2023) <[doi:10.1371/journal.pgen.1010762](https://doi.org/10.1371/journal.pgen.1010762)>.

License GPL (>= 3)

URL <https://github.com/ZhaotongL/GraphMRcML>,
<https://mrcieu.r-universe.dev/GraphMRcML>

Encoding UTF-8

LazyData true

Imports dplyr, grDevices, igraph, MASS, pbmcapply, stats

Config/roxygen2/version 8.0.0

Config/pak/sysreqs libglpk-dev libxml2-dev

Repository <https://mrcieu.r-universe.dev>

Date/Publication 2026-05-07 18:57:42 UTC

RemoteUrl <https://github.com/ZhaotongL/GraphMRcML>

RemoteRef refs/pull/2/head

RemoteSha a0942082654a1efe87fdc01a47cf4f1966ac0c45

Contents

cML_estimate_O	2
cML_estimate_random_O	3
cML_SdTheta_O	4
Generate_Perturb	4
Graph_Estimate	5
Graph_Perturb	6

Graph_Screen	7
loglik	8
mr_cML_DP_O	9
mr_cML_O	10
plot_graph	11
subset_Graph_d1	12

Index	13
--------------	-----------

cML_estimate_0	<i>Constrained maximum likelihood estimate (single start)</i>
----------------	---

Description

Coordinate-descent estimation of the causal effect under the cML model with a fixed number of invalid IVs.

Usage

```
cML_estimate_0(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K,
  initial_theta = 0,
  initial_mu = rep(0, length(b_exp)),
  maxit = 100,
  rho = 0,
  t = 0
)
```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for 'b_exp'.
se_out	Numeric vector of standard errors for 'b_out'.
K	Integer; number of invalid IVs to allow.
initial_theta	Numeric; starting value for the causal effect.
initial_mu	Numeric vector; starting values for the SNP-exposure effects (length matching 'b_exp').
maxit	Integer; maximum number of coordinate-descent iterations.
rho	Numeric; correlation between 'b_exp' and 'b_out' due to sample overlap.
t	Numeric; instrument-selection threshold on the z-scale.

Value

A list with elements ‘theta’ (estimated causal effect), ‘b_vec’ (estimated true SNP-exposure effects) and ‘r_vec’ (estimated invalid-IV effects).

cML_estimate_random_0 *Constrained MLE with random restarts*

Description

Runs ‘cML_estimate_O’ from one or more random starting points and returns the solution achieving the smallest negative log-likelihood.

Usage

```
cML_estimate_random_0(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K,
  random_start = 0,
  maxit = 100,
  rho = 0,
  t = 0,
  var_est = 1
)
```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for ‘b_exp’.
se_out	Numeric vector of standard errors for ‘b_out’.
K	Integer; number of invalid IVs to allow.
random_start	Integer; number of additional random starts (in addition to the default zero start).
maxit	Integer; maximum number of coordinate-descent iterations.
rho	Correlation between ‘b_exp’ and ‘b_out’ due to sample overlap.
t	Instrument-selection threshold on the z-scale.
var_est	Integer in ‘1:4’ selecting the variance estimator returned by ‘cML_SdTheta_O’: ‘1’ cMLE, ‘2’ cMLE robust, ‘3’ MPLE, ‘4’ MPLE robust.

Value

A list with elements ‘theta’, ‘se’, ‘l’ (negative log-likelihood) and ‘r_est’ (invalid-IV indicator vector).

cML_SdTheta_0	<i>Standard error of the cML causal-effect estimate</i>
---------------	---

Description

Computes model-based and robust (sandwich) standard errors for both the constrained MLE and the profile MLE of the causal effect.

Usage

```
cML_SdTheta_0(b_exp, b_out, se_exp, se_out, theta, b_vec, r_vec, rho, t)
```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for 'b_exp'.
se_out	Numeric vector of standard errors for 'b_out'.
theta	Estimated causal effect from 'cML_estimate_O'.
b_vec	Estimated true SNP-exposure effects from 'cML_estimate_O'.
r_vec	Estimated invalid-IV effects from 'cML_estimate_O'.
rho	Correlation between 'b_exp' and 'b_out' due to sample overlap.
t	Instrument-selection threshold on the z-scale.

Value

A list with elements 'cMLE_se', 'cMLE_robust_se', 'MPLE_se' and 'MPLE_robust_se' (any of which may be 'NaN' if the corresponding variance is non-positive or non-invertible).

Generate_Perturb	<i>Generate one LD-aware perturbed copy of the GWAS estimates</i>
------------------	---

Description

Adds correlated Gaussian noise to 'b_mat' using the per-block square-root matrices from 'Graph_Screen', preserving cross-trait correlation and within-block LD.

Usage

```
Generate_Perturb(b_mat, se_mat, n_vec, rho_mat, DP_mat_list)
```

Arguments

b_mat	Numeric matrix of GWAS effect estimates (rows = SNPs, columns = traits).
se_mat	Numeric matrix of standard errors matching 'b_mat'.
n_vec	Integer vector of GWAS sample sizes (one per trait).
rho_mat	'N x N' correlation matrix between GWAS estimates across traits.
DP_mat_list	List of LD-block perturbation matrices produced by 'Graph_Screen'.

Value

A numeric matrix of the same shape as 'b_mat' containing the perturbed effect estimates.

Graph_Estimate	<i>Estimate the pairwise causal-effect graph</i>
----------------	--

Description

For every pair of traits '(i, j)', runs 'mr_cML_O' in both directions using the screened IV indices, fills in the corresponding off-diagonal entries of the observed-effect graph, and applies network deconvolution to obtain the direct-effect graph.

Usage

```
Graph_Estimate(
  b_mat,
  se_mat,
  n_vec,
  rho_mat,
  IJ_snp_list,
  t,
  random_start = 10
)
```

Arguments

b_mat	Numeric matrix of GWAS effect estimates (rows = SNPs, columns = traits).
se_mat	Numeric matrix of standard errors matching 'b_mat'.
n_vec	Integer vector of GWAS sample sizes (one per trait).
rho_mat	'N x N' correlation matrix between GWAS estimates across traits.
IJ_snp_list	Per-pair list of screened IV indices, as returned by 'Graph_Screen'.
t	Instrument-selection threshold on the z-scale, passed through to 'mr_cML_O'.
random_start	Integer; number of random starts per cML fit.

Value

A list with 'obs_graph', 'obs_graph_se', 'obs_graph_pval' and the network-deconvolved 'dir_graph'.

Description

Top-level routine: screens IVs once via ‘Graph_Screen’, then repeatedly generates LD-aware perturbed datasets and fits the pairwise causal-effect graph on each one. Perturbation replicates are run in parallel via [pbmccapply::pbmccapply()].

Usage

```
Graph_Perturb(
  b_mat,
  se_mat,
  n_vec,
  rho_mat,
  IV_list,
  R_list,
  c_vec = rep(1, length(n_vec)),
  sig.cutoff = 5e-08,
  num_pert = 100,
  random_start = 10,
  seed = 0,
  trait_vec = NULL,
  curse = F
)
```

Arguments

b_mat	Numeric matrix of GWAS effect estimates (rows = SNPs, columns = traits). Row names must be SNP rsids.
se_mat	Numeric matrix of standard errors matching ‘b_mat’.
n_vec	Integer vector of GWAS sample sizes (one per trait).
rho_mat	‘N x N’ correlation matrix between GWAS estimates across traits.
IV_list	List of length ‘N*(N-1)/2’ giving candidate IV rsids for each trait pair.
R_list	List of LD blocks (see ‘Graph_Screen’).
c_vec	Numeric vector of LDSC inflation factors (length ‘N’).
sig.cutoff	Numeric; GWAS-significance threshold for instrument selection.
num_pert	Integer; number of data-perturbation replicates.
random_start	Integer; number of random starts per cML fit.
seed	Integer; RNG seed.
trait_vec	Optional character vector of trait names; defaults to ‘colnames(b_mat)’.
curse	Logical; if ‘TRUE’, applies the "curse-of-winner" correction by setting the cML threshold to ‘-qnorm(sig.cutoff/2)’.

Value

A list of per-perturbation observed graphs, their standard errors and p-values, the deconvolved direct-effect graphs, and 'trait_vec'.

Graph_Screen

Screen instruments and prepare LD-aware perturbation matrices

Description

For each pair of traits, restricts the candidate IV set to SNPs that are GWAS-significant for the exposure (and not more strongly associated with the outcome), and pre-computes the LD-block square-root matrices used to generate joint perturbations of the GWAS estimates. Optionally inflates the standard errors using LDSC-style factors.

Usage

```
Graph_Screen(
  b_mat,
  se_mat,
  n_vec,
  IV_list,
  R_list,
  rho_mat,
  c_vec = rep(1, length(n_vec)),
  sig.cutoff = 5e-08
)
```

Arguments

b_mat	Numeric matrix of GWAS effect estimates with one column per trait. Row names must be SNP rsids.
se_mat	Numeric matrix of standard errors matching 'b_mat'.
n_vec	Integer vector of GWAS sample sizes (one per trait).
IV_list	List of length 'N*(N-1)/2' giving candidate IV rsids for each trait pair, in order '(1,2), (1,3), ..., (N-1,N)'.
R_list	List of LD blocks. Each element has 'R' (an LD correlation matrix with rsid row/column names) and 'snp' (the rsids in the block).
rho_mat	'N x N' correlation matrix between GWAS estimates across traits (e.g. from bivariate LDSC).
c_vec	Numeric vector of inflation factors (length 'N') applied as 'sqrt(max(1, c))' to each trait's standard errors.
sig.cutoff	Numeric; GWAS-significance threshold for instrument selection.

Value

A list with the per-pair screened IV indices ('IJ_snp_list'), the per-block perturbation square-root matrices ('DP_mat_list') and the (possibly inflated) 'b_mat' and 'se_mat'.

loglik	<i>Profile log-likelihood for the cML model</i>
--------	---

Description

Computes the (profile) log-likelihood used by the constrained maximum likelihood (cML) estimator for two-sample Mendelian randomization with overlapping samples.

Usage

```
loglik(b_exp, b_out, se_exp, se_out, b_t, theta_t, r_vec_t, rho, t = 0)
```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for 'b_exp'.
se_out	Numeric vector of standard errors for 'b_out'.
b_t	Numeric vector of true SNP-exposure effects (length matching 'b_exp').
theta_t	Numeric scalar; candidate causal effect.
r_vec_t	Numeric vector of invalid-IV pleiotropic effects.
rho	Numeric scalar; correlation between 'b_exp' and 'b_out' due to sample overlap.
t	Numeric scalar; significance threshold (on the z-scale) used for instrument selection. Defaults to '0'.

Value

A numeric scalar giving the log-likelihood value.

 mr_cML_DP_O

MR-cML with data perturbation

Description

Runs ‘mr_cML_O’ and supplements the result with data-perturbation (DP) standard errors and p-values that account for selection of the invalid-IV set.

Usage

```

mr_cML_DP_O(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K_vec = 0:(length(b_exp) - 2),
  random_start = 0,
  random_start_pert = 0,
  maxit = 100,
  num_pert = 100,
  random_seed = 0,
  n,
  rho = 0,
  c1 = 1,
  c2 = 1,
  t = 0
)

```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for ‘b_exp’.
se_out	Numeric vector of standard errors for ‘b_out’.
K_vec	Integer vector; candidate numbers of invalid IVs.
random_start	Integer; number of random starts on the original data.
random_start_pert	Integer; number of random starts per perturbed dataset.
maxit	Integer; maximum coordinate-descent iterations.
num_pert	Integer; number of data-perturbation replicates.
random_seed	Integer; if non-zero, used to seed the RNG.
n	Integer; GWAS sample size used in the BIC penalty.
rho	Correlation between ‘b_exp’ and ‘b_out’ due to sample overlap.
c1, c2	Numeric inflation factors (≥ 1) applied to ‘se_exp’ and ‘se_out’ respectively.
t	Instrument-selection threshold on the z-scale.

Value

A list combining the original ‘mr_cML_O’ output with the DP point estimates, standard errors and p-values for both cML-BIC and cML-MA-BIC.

mr_cML_O	<i>Mendelian randomization via cML with BIC selection</i>
----------	---

Description

Runs the cML estimator across a grid of candidate values for the number of invalid IVs and combines the results with BIC and BIC model averaging (cML-BIC and cML-MA-BIC).

Usage

```
mr_cML_O(  
  b_exp,  
  b_out,  
  se_exp,  
  se_out,  
  K_vec = 0:(length(b_exp) - 2),  
  random_start = 0,  
  maxit = 100,  
  random_seed = 0,  
  n,  
  rho = 0,  
  t = 0,  
  var_est = 1  
)
```

Arguments

b_exp	Numeric vector of SNP-exposure effect estimates.
b_out	Numeric vector of SNP-outcome effect estimates.
se_exp	Numeric vector of standard errors for ‘b_exp’.
se_out	Numeric vector of standard errors for ‘b_out’.
K_vec	Integer vector; candidate numbers of invalid IVs.
random_start	Integer; number of random starts per ‘K’.
maxit	Integer; maximum coordinate-descent iterations.
random_seed	Integer; if non-zero, used to seed the RNG.
n	Integer; GWAS sample size used in the BIC penalty.
rho	Correlation between ‘b_exp’ and ‘b_out’ due to sample overlap.
t	Instrument-selection threshold on the z-scale.
var_est	Integer in ‘1:4’ selecting the variance estimator (see ‘cML_estimate_random_O’).

Value

A list with cML-MA-BIC and cML-BIC point estimates, standard errors and p-values, the indices of the selected invalid IVs, and the negative log-likelihood and BIC sequences over ‘K_vec’.

plot_graph	<i>Build an igraph object for a GraphMRcML summary graph</i>
------------	--

Description

Turns a GraphMRcML mean-effect matrix and matching p-value matrix into a directed ‘igraph’ graph, colouring edges by sign and significance and suggesting per-edge widths. By default, edges with p-value below ‘0.05 / Me’ are drawn in solid colour; edges with p-value below ‘thres1’ are drawn in a lighter shade.

Usage

```
plot_graph(G_mean, G_pval, Me, thres1 = 0.05, Bonferroni = FALSE)
```

Arguments

G_mean	Numeric matrix of mean effects, e.g. ‘obs_graph_mean’ or ‘dir_graph_mean’ from ‘subset_Graph_d1’.
G_pval	Numeric matrix of p-values matching ‘G_mean’.
Me	Integer; effective number of tests from ‘subset_Graph_d1’. Used to set the primary significance threshold ‘0.05 / Me’ when ‘Bonferroni = FALSE’.
thres1	Numeric; secondary p-value threshold for light-coloured edges. Set to ‘0’ to suppress.
Bonferroni	Logical; if ‘TRUE’ (or ‘Me’ is ‘NULL’), use the Bonferroni threshold ‘0.05 / (N^2 - N)’ instead.

Value

A list with ‘p’ (an ‘igraph’ graph) and ‘edge_width’ (a numeric vector of suggested edge widths).

subset_Graph_d1

*Summarise GraphMRcML perturbation results for a subset of traits***Description**

Subsets the per-perturbation graphs from ‘Graph_Perturb’ to a chosen set of traits, iteratively re-runs network deconvolution to enforce zero diagonals on the direct-effect graph, drops replicates whose deconvolved graph has a spectral radius above one, and returns means, standard deviations and p-values for both the observed-effect and direct-effect graphs together with an effective number of tests ‘Me’.

Usage

```
subset_Graph_d1(
  dp_list,
  keep_trait,
  B = 2000,
  check = TRUE,
  show = TRUE,
  maxit = 10000
)
```

Arguments

<code>dp_list</code>	A list as returned by ‘Graph_Perturb’.
<code>keep_trait</code>	Character vector of trait names to retain.
<code>B</code>	Integer; number of perturbation replicates to sample (without replacement) from ‘dp_list’.
<code>check</code>	Logical; if ‘TRUE’, drop replicates whose direct-effect graph has a spectral radius greater than one.
<code>show</code>	Logical; if ‘FALSE’ and every replicate fails the spectral check, the direct-effect summaries are returned as ‘NULL’.
<code>maxit</code>	Integer; maximum iterations of the diagonal-zeroing loop in network deconvolution.

Value

A list with ‘dp_res’ (summaries), ‘dp_list’ (filtered per-replicate graphs), ‘warning’, ‘warn_len’, ‘conv_len’ and ‘total_B’.

Index

cML_estimate_0, [2](#)
cML_estimate_random_0, [3](#)
cML_SdTheta_0, [4](#)

Generate_Perturb, [4](#)
Graph_Estimate, [5](#)
Graph_Perturb, [6](#)
Graph_Screen, [7](#)

loglik, [8](#)

mr_cML_DP_0, [9](#)
mr_cML_0, [10](#)

plot_graph, [11](#)

subset_Graph_d1, [12](#)