

Package: MrDAG (via r-universe)

September 23, 2026

Type Package

Title MrDAG: Bayesian causal graphical model for joint Mendelian randomization analysis of multiple exposures and outcomes

Version 0.1.1

Date 2025-04-01

Maintainer Leonardo Bottolo <lb664@cam.ac.uk>

Description This package performs Mendelian randomization for multiple exposures and outcomes with Bayesian structure learning and causal effects estimation. The directionality of the causal effects between the exposures and the outcomes is assumed known, i.e., the exposures can only be potential causes of the outcomes and no reverse causation is allowed.

License GPL-2 | file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

RdMacros Rdpack

Imports gRbase, mvtnorm, Rdpack, utils

Depends R (>= 3.5.0)

LazyData true

Config/pak/sysreqs libglpk-dev libxml2-dev

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

Date/Publication 2025-04-04 05:06:20 UTC

RemoteUrl <https://github.com/lb664/MrDAG>

RemoteRef HEAD

RemoteSha 9b37d111ac6e9b33c54ee6f54de3b5801b46e8a0

Contents

get_causaleffect	2
get_causaleffects	4
get_edgeprob	6
LBT2MD_data	7
MD2LBT_data	8
MrDAG	9

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

get_causaleffect	<i>Estimation of the (average) marginal causal effects under intervention on an exposure</i>
------------------	--

Description

Estimation of the (average) marginal causal effect under intervention on a trait (target) and measured on another one (response) based on the Directed Acyclic Graphs (DAGs) explored by MrDAG algorithm

Usage

```
get_causaleffect(output, response, target, BMA = TRUE, CI = 0.95)
```

Arguments

output	Output produced by MrDAG algorithm
response	Trait (response) where the effect of the intervention on another trait (target) is measured. The index refers to position of the trait (response) in the output produced by MrDAG algorithm
target	Trait (target) under intervention. The index refers to position of the trait (response) in the output produced by MrDAG algorithm
BMA	If TRUE, Bayesian Model Averaging of the estimated (average) causal effect across all explored DAGs in the visited Completed Partially DAG (CPDAG) (Chickering (2002)) or Essential Graph (EG) (Andersson et al. (1997)) is performed
CI	Level (0.95 default) of the credible interval (CI) of the (average) causal effect. It is calculated based on suitable quantiles of the estimated (average) causal effect across all explored DAGs

Value

The value returned is a list object `list(causalEffect, causalEffect_LL, causalEffect_UL, group, BMA, CI)`

- **causalEffect** Estimate of the (average) causal effect under intervention on an trait (target) on another one (response) based on the DAGs explored by MrDAG algorithm. Feedback loop causal effect of the same trait is not allowed (NA). Note that the estimated causal effect corresponds to the marginal causal effect of a target on a response *without* considering the effects of the others traits
- **causalEffect_LL** Lower limit of the CI of the (average) causal effect. It is calculated as the $(1 - \text{CIboot})/2$ % quantile of the (average) causal effect across all explored DAGs
- **causalEffect_UL** Upper limit of the CI of the (average) causal effect. It is calculated as the $1 - [(1 - \text{CIboot})/2]$ % quantile of the (average) causal effect across all explored DAGs
- **group** If MrDAGcheck in MrDAG algorithm contains the indices of the traits in data that define the outcomes and the exposures, group coincides with this list
- **BMA** Logical option
- **CI Level** of the credible interval option

References

Andersson SA, Madigan D, Perlman MD (1997). “A characterization of Markov equivalence classes for acyclic digraphs.” *Ann. Stat.*, **25**(2), 505–541. doi:10.1214/aos/1031833662.

Chickering DM (2002). “Learning equivalence classes of bayesian-network structures.” *J. Mach. Learn. Res.*, **2**, 445–498. <http://www.ai.mit.edu/projects/jmlr/papers/volume2/chickering02a/chickering02a.pdf>.

Examples

```
# Example: Estimation of the (average) causal effects under intervention on lifestyle and
# behavioural exposures, and measured on mental health phenotypes. 708 independent Instrumental
# Variables (IVs) were selected to be associated at genome-wide significance with six lifestyle
# and behavioural traits after pruning or clumping which are considered exposures of the risk
# of seven mental health outcome phenotypes

# After loading the data set

data(LBT2MD_data)

# the indices of the traits that define the outcomes and exposures are provided in a list object

MrDAGcheck <- NULL
MrDAGcheck$Y_idx <- 1 : 7      # Mental health phenotypes
MrDAGcheck$X_idx <- 8 : 13     # Lifestyle and behavioural traits

# MrDAG algorithm is run to generate 1,000 posterior samples of all unknowns

output <- MrDAG(data = LBT2MD_data,
  niter = 7500, burnin = 2500, thin = 5, tempMax = 20, pp = 0.01,
  MrDAGcheck = MrDAGcheck, fileName = NULL)

# Finally, the posterior (average) causal effects and credible intervals (CI) of the intervention
# on ALC and measured on SCZ are estimated
```

```
causalEffect <- get_causaleffect(output, 7, 8) # Intervention on ALC(8) and measured on SCZ(7)

causalEffect <- get_causaleffect(output, 7, c(4, 8)) # Intervention on ALC(8) and measured on
# SCZ(7) after conditioning on BD(4)
```

get_causaleffects	<i>Estimation of the (average) causal effects under intervention on the exposures</i>
-------------------	---

Description

Estimation of the (average) causal effects under intervention on the exposures (targets) and measured on the outcomes (responses) based on the Directed Acyclic Graphs (DAGs) explored by MrDAG algorithm

Usage

```
get_causaleffects(
  output,
  ord = NULL,
  BMA = TRUE,
  CI = 0.95,
  progress = FALSE,
  plot = FALSE
)
```

Arguments

output	Output produced by MrDAG algorithm
ord	Indices of the traits in data that define the outcomes and the exposures. They specify the order of appearance of the traits when printing the (average) causal effect. If ord = NULL, the order is the same as MrDAGcheck in MrDAG output
BMA	If TRUE, Bayesian Model Averaging of the estimated (average) causal effect across all explored DAGs in the visited Completed Partially DAG (CPDAG) (Chickering (2002)) or Essential Graph (EG) (Andersson et al. (1997)) is performed
CI	Level (0.95 default) of the credible interval (CI) of the (average) causal effect. It is calculated based on suitable quantiles of the estimated (average) causal effect across all explored DAGs
progress	Logical value set as FALSE by default to print on the screen the progress of the causal effects estimation
plot	Logical value (default FALSE). If TRUE, get_causaleffects generates the output to be used in a plot

Value

The value returned is a list object `list(causalEffects, causalEffects_LL, causalEffects_UL, group, ord, BMA, CI)`

- `causalEffects` Estimate of the (average) causal effects under intervention on the exposures based on the DAGs explored by MrDAG algorithm. Feedback loop causal effect of the same outcome is not allowed (NA)
- `causalEffects_LL` Lower limit of the CI of the (average) causal effects. It is calculated as the $(1 - \text{CIboot})/2$ % quantile of the (average) causal effect across all explored DAGs
- `causalEffects_UL` Upper limit of the CI of the (average) causal effects. It is calculated as the $1 - [(1 - \text{CIboot})/2]$ % quantile of the (average) causal effect across all explored DAGs
- `group` If MrDAGcheck in MrDAG algorithm contains the indices of the traits in data that define the outcomes and the exposures, group coincides with this list
- `ord` Indices of the traits in data that specify the outcomes and the exposures. It might differ from group if ord has been specified and it is different from MrDAGcheck
- `BMA` Logical option
- `CI` Level of the credible interval option

References

Andersson SA, Madigan D, Perlman MD (1997). “A characterization of Markov equivalence classes for acyclic digraphs.” *Ann. Stat.*, **25**(2), 505–541. doi:10.1214/aos/1031833662.

Chickering DM (2002). “Learning equivalence classes of bayesian-network structures.” *J. Mach. Learn. Res.*, **2**, 445–498. <http://www.ai.mit.edu/projects/jmlr/papers/volume2/chickering02a/chickering02a.pdf>.

Examples

```
# Example: Estimation of the (average) causal effects under intervention on lifestyle and
# behavioural exposures, and measured on mental health phenotypes. 708 independent Instrumental
# Variables (IVs) were selected to be associated at genome-wide significance with six lifestyle
# and behavioural traits after pruning or clumping which are considered exposures of the risk
# of seven mental health outcome phenotypes
```

```
# After loading the data set
```

```
data(LBT2MD_data)
```

```
# the indices of the traits that define the outcomes and exposures are provided in a list object
```

```
MrDAGcheck <- NULL
```

```
MrDAGcheck$Y_idx <- 1 : 7    # Mental health phenotypes
```

```
MrDAGcheck$X_idx <- 8 : 13   # Lifestyle and behavioural traits
```

```
# MrDAG algorithm is run to generate 1,000 posterior samples of all unknowns
```

```
output <- MrDAG(data = LBT2MD_data,
                niter = 7500, burnin = 2500, thin = 5, tempMax = 20, pp = 0.01,
```

```

MrDAGcheck = MrDAGcheck, fileName = NULL)

# Finally, the posterior (average) causal effects and 90% credible intervals are estimated

ord <- c(8 : 13, 1 : 7)
causalEffects <- get_causaleffects(output, ord = ord, CI = 0.90, progress = TRUE)

```

get_edgeprob

Estimation of the Posterior Probability of Edge Inclusion

Description

Estimation of the Posterior Probability of Edge Inclusion (PPEI) based on the Directed Acyclic Graphs (DAGs) explored by MrDAG algorithm

Usage

```
get_edgeprob(output, ord = NULL)
```

Arguments

output	Output produced by MrDAG algorithm
ord	Indices of the traits in data that define the outcomes and the exposures. They specify the order of appearance of the traits when printing the PPEI. If ord = NULL (default), the order is the same as MrDAGcheck in MrDAG output. Note that despite ord ordering, the outcomes always precede the exposures

Value

The value returned is a list object `list(edgeProb, group, ord)`

- `edgeProb` Estimate of the PPEI between each trait. Feedback loop of the same trait is not allowed (NA). If a partition of the traits between outcomes and exposures is specified, the PPEIs between outcomes and exposures are not considered (NA)
- `group` If MrDAGcheck in MrDAG algorithm contains the indices of the traits in data that define the outcomes and the exposures, group coincides with this list
- `ord` Indices of the traits in data that specify the outcomes and the exposures. It might differ from group if ord has been specified and it is different from MrDAGcheck

Examples

```

# Example: Estimation of the Posterior Probabilities of Edge Inclusion (PPEIs) of lifestyle and
# behavioural traits, and mental health phenotypes. 708 independent Instrumental Variables
# (IVs) were selected to be associated at genome-wide significance with six lifestyle and
# behavioural traits after pruning or clumping which are considered exposures of the risk of
# seven mental health outcome phenotypes

# After loading the data set

```

```

data(LBT2MD_data)

# the indices of the traits that define the outcomes and exposures are provided in a list object

MrDAGcheck <- NULL
MrDAGcheck$Y_idx <- 1 : 7    # Mental health phenotypes
MrDAGcheck$X_idx <- 8 : 13   # Lifestyle and behavioural traits

# MrDAG algorithm is run to generate 1,000 posterior samples of all unknowns

output <- MrDAG(data = LBT2MD_data,
                 niter = 7500, burnin = 2500, thin = 5, tempMax = 20, pp = 0.01,
                 MrDAGcheck = MrDAGcheck, fileName = NULL)

# Finally, PPEIs are calculated and presented with lifestyle and behavioural traits first,
# followed by mental health phenotypes

ord <- c(8 : 13, 1 : 7)
PPEI <- get_edgeprob(output, ord = ord)

```

LBT2MD_data

MrDAG data set: Lifestyle and behavioural exposures that might impact mental health phenotypes

Description

The data set contains lifestyle and behavioural traits that are considered exposures of the risk of mental health phenotypes. As outcomes, seven mental health phenotypes are considered, including (in alphabetic order) attention deficit hyperactivity disorder (ADHD), anorexia nervosa (AN), autism spectrum disorder (ASD), bipolar disorder (BD), cognition (COG), major depressive disorder (MDD) and schizophrenia (SCZ). As exposures, six lifestyle and behavioural traits that have previously been investigated for their protective/risk effects on mental health are considered, including (in alphabetic order) alcohol consumption (ALC), education (in years) (EDU), leisure screen time (LST), physical activity (PA), lifetime smoking index (SM) and sleep duration (SP)

Usage

```
LBT2MD_data
```

Format

A data frame consisting of 708 independent Instrumental Variables (IVs) selected to be associated at genome-wide significance with the exposures after pruning or clumping. For details, see Zuber et al. (2025)

References

Zuber V, Cronjé HT, Cai N, Gill D, Bottolo L (2025). “Bayesian causal graphical model for joint Mendelian randomization analysis of multiple exposures and outcomes.” *Am. J. Hum. Genet.*.. doi:10.1016/j.ajhg.2025.03.005. Published online.

Examples

```
# Example:
```

```
data(LBT2MD_data)
head(LBT2MD_data)
```

MD2LBT_data	<i>MrDAG data set: Mental health phenotypes that might impact lifestyle and behavioural traits</i>
-------------	--

Description

The data set contains mental health phenotypes that are considered exposures of the risk of lifestyle and behavioural traits. As outcomes, six lifestyle and behavioural traits are considered, including (in alphabetic order) alcohol consumption (ALC), education (in years) (EDU), leisure screen time (LST), physical activity (PA), lifetime smoking index (SM) and sleep duration (SP). As exposures, seven mental health phenotypes are considered, including (in alphabetic order) attention deficit hyperactivity disorder (ADHD), anorexia nervosa (AN), autism spectrum disorder (ASD), bipolar disorder (BD), cognition (COG), major depressive disorder (MDD) and schizophrenia (SCZ)

Usage

```
MD2LBT_data
```

Format

A data frame consisting of 470 independent Instrumental Variables (IVs) selected to be associated at genome-wide significance with the exposures after pruning or clumping. For details, see Zuber et al. (2025)

References

Zuber V, Cronjé HT, Cai N, Gill D, Bottolo L (2025). “Bayesian causal graphical model for joint Mendelian randomization analysis of multiple exposures and outcomes.” *Am. J. Hum. Genet.*.. doi:10.1016/j.ajhg.2025.03.005. Published online.

Examples

```
# Example:
```

```
data(MD2LBT_data)
head(MD2LBT_data)
```

MrDAG

*MrDAG: Bayesian causal graphical model for joint Mendelian randomization analysis of multiple exposures and outcomes***Description**

Markov chain Monte Carlo (MCMC) implementation of Bayesian multivariable, multi-response Mendelian randomization (MR) model for summary-level data with structure learning and causal effects estimation. The directionality of the causal effects between the exposures and the outcomes is assumed known, i.e., the exposures can only be potential causes of the outcomes and no reverse causation is allowed

Usage

```
MrDAG(
  data,
  niter,
  burnin,
  thin,
  pp = 0.05,
  a = NULL,
  U = NULL,
  tempMax = 10,
  MrDAGcheck = NULL,
  fileName = "MrDAG_object",
  filePath = NULL,
  saveMemory = FALSE,
  seed = 31122021
)
```

Arguments

data	Number of observations (IVs in a summary-level MR design) times the number of traits (both outcomes and exposures). There is a restriction in the order of appearance of the traits in the data matrix: The group of outcomes have to appear first, followed by the group of exposures. See also MrDAGcheck argument
niter	Number of MCMC iterations (including burn-in)
burnin	Number of MCMC iterations to be discarded as burn-in
thin	Parameter that defines how often the MCMC output should be stored, i.e., at every thin-th iteration
pp	Prior probability of edge inclusion (0.05 default) between each pair of nodes (vertices) in the graph
a	Degrees of freedom of the Wishart prior distribution on the precision matrix, i.e., the inverse of the covariance matrix between the traits (m default, where m is the total number of outcomes and exposures)

U	(Proportional to the) expected value of the Wishart prior distribution on the precision matrix, i.e., the inverse of the covariance matrix between the traits. Specifically, proportional to an m-dimensional diagonal matrix as default, where m is the total number of outcomes and exposures
tempMax	Annealing parameter T used to facilitate the convergence of the MCMC algorithm to the target distribution (10 default). Temperature 1/T increases linearly during the burn-in until T=1 at the end of the burn-in
MrDAGcheck	List object that contains the indices of the traits in data that are defined as outcomes and exposures. If NULL (default), no partition of the traits between outcomes and exposures is specified and the MrDAG algorithm performs structure learning between all traits without constraints. Note that if so, this procedure does not correspond to an MR analysis since valid instruments for the exposures are invalid for the outcomes and vice versa
fileName	Name of the file for MrDAG output object ("MrDAG_object" default). If NULL, the output is not saved
filePath	Directory where MrDAG object is saved. If the directory does not exist, the MrDAG algorithm will create it by using the current working directory as the root directory. If the directory is not specified (NULL default), MrDAG object is saved in the current working directory
saveMemory	If logical FALSE (default), the visited graph and the posterior draws from the modified Cholesky decomposition (L,D) (Zuber et al. (2025) and Castelletti and Consonni (2021)) are stored as an array, otherwise they are stored as a list object
seed	Seed to be used in the initialisation of the MCMC algorithm (3112021 default)

Details

For details regarding the model and the algorithm, see Zuber et al. (2025)

Value

The value returned is a list object `list(graphs, L,D, logMargLik, validPropMrDAG, acceptPropDAG, timeMrDAG, hyperPar, samplerPar, opt)`

- `graphs` Explored DAGs that belong to the learned Completed Partially DAGs (CPDAGs) (Chickering (2002)) or Essential Graphs (EGs) (Andersson et al. (1997)). The class of graphs depends on the `saveMemory` option. The number of explored DAGs corresponds to the number of thinned (thin) MCMC iterations (excluding burn-in)
- `L` Posterior samples of the lower triangular matrix of the modified Cholesky decomposition (L,D) (Zuber et al. (2025) and Castelletti and Consonni (2021))
- `D` Posterior samples of the diagonal matrix of the modified Cholesky decomposition (L,D) (Zuber et al. (2025) and Castelletti and Consonni (2021))
- `logMargLik` Log-marginal likelihood of explored DAGs which belong to the Markov Equivalent Classes whose unique representative chain graphs are the EGs learned during MCMC iterations (including burn-in) without thinning
- `validPropMrDAG` If `MrDAGcheck` is different from NULL, the proportion of proposed DAGs that comply with the partial ordering (Perković et al. (2017) implied by the partition of the

traits between exposures and outcomes with directed causal effects only from the former to the latter

- `acceptPropDAG` Proportion of proposed DAGs that are accepted by the Metropolis-Hastings ratio after burn-in. If `MrDAGcheck` is different from `NULL`, `acceptPropDAG` corresponds to the proposed DAGs that comply with the partial ordering
- `timeMrDAG` Time in minutes employed by the MrDAG algorithm to analyse the data
- `hyperPar` List of the hyper-parameters `list(w, a, U, tempMax)` and, if specified, `MrDAGcheck` list
- `samplerPar` List of parameters used in the MCMC algorithm `list(niter, burnin, thin)`
- `opt` List of options used in MrDAG algorithm `list(saveMemory, seed)`

References

Andersson SA, Madigan D, Perlman MD (1997). “A characterization of Markov equivalence classes for acyclic digraphs.” *Ann. Stat.*, **25**(2), 505–541. doi:10.1214/aos/1031833662.

Castelletti F, Consonni G (2021). “Bayesian inference of causal effects from observational data in Gaussian graphical models.” *Biometrics*, **77**, 136–149. doi:10.1111/biom.13281.

Chickering DM (2002). “Learning equivalence classes of bayesian-network structures.” *J. Mach. Learn. Res.*, **2**, 445–498. <http://www.ai.mit.edu/projects/jmlr/papers/volume2/chickering02a/chickering02a.pdf>.

Perković E, Kalisch M, Maathuis MH (2017). “Interpreting and using CPDAGs with background knowledge.” In Elidan G, Kersting K, Ihler AT (eds.), *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI2017)*. <http://auai.org/uai2017/proceedings/papers/120.pdf>.

Zuber V, Cronjé HT, Cai N, Gill D, Bottolo L (2025). “Bayesian causal graphical model for joint Mendelian randomization analysis of multiple exposures and outcomes.” *Am. J. Hum. Genet.*. doi:10.1016/j.ajhg.2025.03.005. Published online.

Examples

```
# Example: Analysis of lifestyle and behavioural exposures that might impact mental health
# phenotypes. 708 independent Instrumental Variables (IVs) were selected to be associated at
# genome-wide significance with six lifestyle and behavioural traits after pruning or clumping,
# which are considered exposures of the risk of seven mental health outcome phenotypes

# After loading the data set

data(LBT2MD_data)

# the indices of the traits that define the outcomes and exposures are provided in a list object

MrDAGcheck <- NULL
MrDAGcheck$Y_idx <- 1 : 7      # Mental health phenotypes
MrDAGcheck$X_idx <- 8 : 13     # Lifestyle and behavioural traits
```

```
# MrDAG algorithm is run to generate 1,000 posterior samples of all unknowns with memory savings  
output <- MrDAG(data = LBT2MD_data,  
               niter = 7500, burnin = 2500, thin = 5, tempMax = 20, pp = 0.01,  
               MrDAGcheck = MrDAGcheck, fileName = NULL)
```

Index

* **datasets**

LBT2MD_data, [7](#)

MD2LBT_data, [8](#)

get_causaleffect, [2](#)

get_causaleffects, [4](#)

get_edgeprob, [6](#)

LBT2MD_data, [7](#)

MD2LBT_data, [8](#)

MrDAG, [9](#)