Package 'GWASmeta'

February 24, 2023

Type Package

Title Rapid calculation of ABF for GWAS meta-analysis

Version 0.1.0

Author Sun Jianle

Maintainer The package maintainer <sjl-2017@sjtu.edu.cn>

Description Providing three methods to get optimal ABF in Bayesian GWAS meta-analysis including Subset-Exhaustive, MCMC and shotgun stochastic search.

License GPL

Encoding UTF-8

LazyData true

Imports MASS,
dplyr

RoxygenNote 7.2.1

R topics documented:

---

exh_abf   *ABF calculation for GWAS meta-analysis through Subset-Exhaustive*

Description

Subset-Exhaustive to get the ABFs in all subsets for a single SNP and a phenotype in GWAS meta-analysis.

Usage

```
exh_abf(
  betas,
  ses,
  prior.sigma = 0.3,
  prior.cor = "indep",
  prior.rho = NA,
  cryptic.cor = NA,
  log = FALSE,
  log10 = FALSE,
  study.names = NULL,
  na.rm = FALSE,
  tolerance = 0
)
```

Arguments

betas	a numeric vector of observed effect sizes of a single SNP in a set of studies. Each element of the vector is assumed to correspond to a study.
ses	a numeric vector of standard errors corresponding to those in betas. It should have the same length as betas.
prior.sigma	the prior on true effect sizes for each SNP in each study. It can be a flat value, set for each study (i.e. a vector whose length is equal to the number of studies in the meta-analysis) or set for each study and SNP (i.e. a matrix of same dimension as betas).
prior.cor	a square matrix whose row and column numbers are the same as the number of studies. Its elements are the pairwise correlations between true effect sizes of the studies. It can take values "indep" (independent effects), "fixed" (fixed effects), "correlated" (correlated effects, which requires the prior.rho parameter to be set), as well as individual matrices. If betas and ses are matrices, the same prior.cor will be applied to every row (representing every SNP).
prior.rho	either a single value or the upper triangle of a correlation matrix for the prior.cor matrix when it is set to "correlated". If this value is set, but prior.cor is not set to "correlated", this parameter will be ignored.
cryptic.cor	a square matrix whose row and coumn numbers are the same as the number of studies. If the studies in the meta-analysis are not independent of each other, it may be necessary to set this parameter so that the covariance in null effects is accounted for.
log	sets whether the answer should be given in log space.
log10	sets whether the answer should be given in log10 space.
study.names	if set, the output will label the columns with the study names.
na.rm	if there are NAs in the data, these are removed and the calculation is performed with the remaining data. This happens regardless of how this parameter is set. By default, the output will include a column of NAs for the study with the missing data. Changing this parameter to TRUE removes this column.
tolerance	for the ABF calculation, this can be lowered (or raised, if necessary) if the answers are not what was expected. Should probably never be altered, but is there in case it is needed. Default is 1e-1000.

Value

a data frame containing ABF calculated in all subsets.

---

mcmc_abf	*ABF calculation for GWAS meta-analysis through MCMC*

Description

using MCMC algorithm to get the optimal ABF for a single SNP and a phenotype in GWAS meta-analysis.

Usage

```
mcmc_abf(
  betas,
  ses,
  prior.sigma = 0.3,
  prior.cor = "indep",
  prior.rho = NA,
  cryptic.cor = NA,
  log = FALSE,
  log10 = FALSE,
  na.rm = FALSE,
  tolerance = 0,
  n.iter = 500
)
```

Arguments

betas
: a numeric vector of observed effect sizes of a single SNP in a set of studies. Each element of the vector is assumed to correspond to a study.

ses
: a numeric vector of standard errors corresponding to those in betas. It should have the same length as betas.

prior.sigma
: the prior on true effect sizes for each SNP in each study. It can be a flat value, set for each study (i.e. a vector whose length is equal to the number of studies in the meta-analysis) or set for each study and SNP (i.e. a matrix of same dimension as betas).

prior.cor
: a square matrix whose row and column numbers are the same as the number of studies. Its elements are the pairwise correlations between true effect sizes of the studies. It can take values "indep" (independent effects), "fixed" (fixed effects), "correlated" (correlated effects, which requires the prior.rho parameter to be set), as well as individual matrices. If betas and ses are matrices, the same prior.cor will be applied to every row (representing every SNP).

prior.rho
: either a single value or the upper triangle of a correlation matrix for the prior.cor matrix when it is set to "correlated". If this value is set, but prior.cor is not set to "correlated", this parameter will be ignored.

cryptic.cor
: a square matrix whose row and coumn numbers are the same as the number of studies. If the studies in the meta-analysis are not independent of each other, it may be necessary to set this parameter so that the covariance in null effects is accounted for.

log   sets whether the answer should be given in log space.

log10   sets whether the answer should be given in log10 space.

na.rm   if there are NAs in the data, these are removed and the calculation is performed
                  with the remaining data. This happens regardless of how this parameter is set.
                  By default, the output will include a column of NAs for the study with the miss-
                  ing data. Changing this parameter to TRUE removes this column.

tolerance   for the ABF calculation, this can be lowered (or raised, if necessary) if the an-
                  swers are not what was expected. Should probably never be altered, but is there
                  in case it is needed. Default is 1e-1000.

n.iter   the number of iteration for MCMC. Default is 500.

Value

the optimal ABF calculated

Examples

```
library(GWASmeta)
data(single)
betas <- single$betas
ses <- single$ses
re <- mcmc_abf(betas,ses)
```

---

multi_shotgun_abf   *Title ABF calculation across multiple variants for GWAS meta-*
                           analysis through shotgun stochastic search

---

Description

using shotgun stochastic search algorithm to quickly get the optimal ABF for multiple SNPs and a
phenotype in GWAS meta-analysis.

Usage

```
multi_shotgun_abf(
  df,
  vname = 1,
  vbetas = seq(2, ncol(df), 2),
  vses = seq(3, ncol(df), 2),
  prior.sigma = 0.5,
  prior.cor = "indep",
  prior.rho = NA,
  cryptic.cor = NA,
  log = FALSE,
  log10 = FALSE,
  na.rm = FALSE,
  tolerance = 0,
  n.iter = 50,
  B = 5
)
```

Arguments

df	a data frame containing the names of variants, the effect sizes and standard errors for each variant in each study.
vname	the index of column of variant name in df. Default is 1.
vbetas	a vectors to represent the index of columns of those containing the effect sizes in each study, default is the column 2,4,6,etc.
vses	a vectors to represent the index of columns of those containing the corresponding standard errors in each study, default is the column 3,5,7,etc. The length of vses should be the same as the length of vbetas.
prior.sigma	the prior on true effect sizes for each SNP in each study. It can be a flat value, set for each study (i.e. a vector whose length is equal to the number of studies in the meta-analysis) or set for each study and SNP (i.e. a matrix of same dimension as betas).
prior.cor	a square matrix whose row and column numbers are the same as the number of studies. Its elements are the pairwise correlations between true effect sizes of the studies. It can take values "indep" (independent effects), "fixed" (fixed effects), "correlated" (correlated effects, which requires the prior.rho parameter to be set), as well as individual matrices. If betas and ses are matrices, the same prior.cor will be applied to every row (representing every SNP).
prior.rho	either a single value or the upper triangle of a correlation matrix for the prior.cor matrix when it is set to "correlated". If this value is set, but prior.cor is not set to "correlated", this parameter will be ignored.
cryptic.cor	a square matrix whose row and coumn numbers are the same as the number of studies. If the studies in the meta-analysis are not independent of each other, it may be necessary to set this parameter so that the covariance in null effects is accounted for.
log	sets whether the answer should be given in log space.
log10	sets whether the answer should be given in log10 space.
na.rm	if there are NAs in the data, these are removed and the calculation is performed with the remaining data. This happens regardless of how this parameter is set. By default, the output will include a column of NAs for the study with the missing data. Changing this parameter to TRUE removes this column.
tolerance	for the ABF calculation, this can be lowered (or raised, if necessary) if the answers are not what was expected. Should probably never be altered, but is there in case it is needed. Default is 1e-1000.
n.iter	the number of iteration for MCMC. Default is 50.
B	the largest number of subsets involved in the optimal set. Default is 5.

Value

a data frame containing the following columns:

SNP	Variant Name
ABF	the final ABF value calculated for each variant
model	a string of 0-1 to represents the subset model selected to calculate.
n_studies	the number of studies involved for each variant. If needClean is FALSE, the value will be the same for all variants.
studies_involved	
	a string of 0-1 to represents the studies involved for each variant. 1 represents the corresponding study is involved.

Examples

```
library(GWASmeta)
data(multi)
re <- multi_shotgun_abf(multi)
```

---

shotgun_abf	*ABF calculation for GWAS meta-analysis through shotgun stochastic search*

Description

using shotgun stochastic search algorithm to quickly get the optimal ABF for a single SNP and a phenotype in GWAS meta-analysis.

Usage

```
shotgun_abf(
  betas,
  ses,
  prior.sigma = 0.3,
  prior.cor = "indep",
  prior.rho = NA,
  cryptic.cor = NA,
  log = FALSE,
  log10 = FALSE,
  na.rm = FALSE,
  tolerance = 0,
  n.iter = 50,
  B = 5
)
```

Arguments

betas	a numeric vector of observed effect sizes of a single SNP in a set of studies. Each element of the vector is assumed to correspond to a study.
ses	a numeric vector of standard errors corresponding to those in betas. It should have the same length as betas.
prior.sigma	the prior on true effect sizes for each SNP in each study. It can be a flat value, set for each study (i.e. a vector whose length is equal to the number of studies in the meta-analysis) or set for each study and SNP (i.e. a matrix of same dimension as betas).
prior.cor	a square matrix whose row and column numbers are the same as the number of studies. Its elements are the pairwise correlations between true effect sizes of the studies. It can take values "indep" (independent effects), "fixed" (fixed effects), "correlated" (correlated effects, which requires the prior.rho parameter to be set), as well as individual matrices. If betas and ses are matrices, the same prior.cor will be applied to every row (representing every SNP).
prior.rho	either a single value or the upper triangle of a correlation matrix for the prior.cor matrix when it is set to "correlated". If this value is set, but prior.cor is not set to "correlated", this parameter will be ignored.

cryptic.cor   a square matrix whose row and coumn numbers are the same as the number of studies. If the studies in the meta-analysis are not independent of each other, it may be necessary to set this parameter so that the covariance in null effects is accounted for.

log   sets whether the answer should be given in log space.

log10   sets whether the answer should be given in log10 space.

na.rm   if there are NAs in the data, these are removed and the calculation is performed with the remaining data. This happens regardless of how this parameter is set. By default, the output will include a column of NAs for the study with the missing data. Changing this parameter to TRUE removes this column.

tolerance   for the ABF calculation, this can be lowered (or raised, if necessary) if the answers are not what was expected. Should probably never be altered, but is there in case it is needed. Default is 1e-1000.

n.iter   the number of iteration for MCMC. Default is 50.

B   the largest number of subsets involved in the optimal set. Default is 5.

Value

the optimal ABF calculated

Examples

```
library(GWASmeta)
data(single)
betas <- single$betas
ses <- single$ses
re <- shotgun_abf(betas,ses)
```

---

shotgun_abf_model	*ABF calculation for GWAS meta-analysis through shotgun stochastic search*

Description

using shotgun stochastic search algorithm to quickly get the optimal ABF for a single SNP and a phenotype in GWAS meta-analysis.

Usage

```
shotgun_abf_model(
  betas,
  ses,
  prior.sigma = 0.3,
  prior.cor = "indep",
  prior.rho = NA,
  cryptic.cor = NA,
  log = FALSE,
  log10 = FALSE,
  na.rm = FALSE,
```

```
  tolerance = 0,
  n.iter = 50,
  B = 5
)
```

Arguments

betas   a numeric vector of observed effect sizes of a single SNP in a set of studies.
                    Each element of the vector is assumed to correspond to a study.

ses   a numeric vector of standard errors corresponding to those in betas. It should
                    have the same length as betas.

prior.sigma   the prior on true effect sizes for each SNP in each study. It can be a flat value, set
                    for each study (i.e. a vector whose length is equal to the number of studies in the
                    meta-analysis) or set for each study and SNP (i.e. a matrix of same dimension
                    as betas).

prior.cor   a square matrix whose row and column numbers are the same as the number
                    of studies. Its elements are the pairwise correlations between true effect sizes
                    of the studies. It can take values "indep" (independent effects), "fixed" (fixed
                    effects), "correlated" (correlated effects, which requires the prior.rho parameter
                    to be set), as well as individual matrices. If betas and ses are matrices, the same
                    prior.cor will be applied to every row (representing every SNP).

prior.rho   either a single value or the upper triangle of a correlation matrix for the prior.cor
                    matrix when it is set to "correlated". If this value is set, but prior.cor is not set to
                    "correlated", this parameter will be ignored.

cryptic.cor   a square matrix whose row and coumn numbers are the same as the number of
                    studies. If the studies in the meta-analysis are not independent of each other, it
                    may be necessary to set this parameter so that the covariance in null effects is
                    accounted for.

log   sets whether the answer should be given in log space.

log10   sets whether the answer should be given in log10 space.

na.rm   if there are NAs in the data, these are removed and the calculation is performed
                    with the remaining data. This happens regardless of how this parameter is set.
                    By default, the output will include a column of NAs for the study with the miss-
                    ing data. Changing this parameter to TRUE removes this column.

tolerance   for the ABF calculation, this can be lowered (or raised, if necessary) if the an-
                    swers are not what was expected. Should probably never be altered, but is there
                    in case it is needed. Default is 1e-1000.

n.iter   the number of iteration for MCMC. Default is 50.

B   the largest number of subsets involved in the optimal set. Default is 5.

Value

a list containing:

ABF   the optimal ABF calculated

model   the optimal subset model

Examples

```
library(GWASmeta)
data(single)
betas <- single$betas
ses <- single$ses
re <- shotgun_abf_model(betas,ses)
```

Index