Package ‘geneSLOPE’

June 20, 2025

Type Package

Title Genome-Wide Association Study with SLOPE

Version 0.38.3

Date 2025-06-19

Description Genome-wide association study (GWAS) performed with SLOPE, short for Sorted L-One Penalized Estimation, a method for estimating the vector of coefficients in linear model. In the first step of GWAS, SNPs are clumped according to their correlations and distances. Then, SLOPE is performed on data where each clump has one representative.

License GPL-3

URL `https://github.com/psobczyk/geneSLOPE`

BugReports `https://github.com/psobczyk/geneSLOPE/issues`

Depends R (>= 3.1.3), SLOPE

Imports ggplot2, bigmemory, grid, utils, stats

Suggests shiny, knitr, rmarkdown, testthat

VignetteBuilder knitr

Repository CRAN

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Author Damian Brzyski [aut],
Christine Peterson [aut],
Emmanuel J. Candes [aut],
Malgorzata Bogdan [aut],
Chiara Sabatti [aut],
Piotr Sobczyk [cre, aut]

Maintainer Piotr Sobczyk <`pj.sobczyk@gmail.com`>

Date/Publication 2025-06-20 04:40:02 UTC

Contents

clumpingResult . . . 2
clump_snps . . . 3
create_lambda . . . 4
geneSLOPE . . . 4
gui_geneSLOPE . . . 5
identify_clump . . . 6
identify_clump.clumpingResult . . . 6
identify_clump.selectionResult . . . 7
phenotypeData . . . 7
plot.selectionResult . . . 8
print.clumpingResult . . . 8
print.phenotypeData . . . 9
print.screeningResult . . . 9
print.selectionResult . . . 10
read_phenotype . . . 10
screeningResult . . . 11
screen_snps . . . 11
selectionResult . . . 12
select_snps . . . 13
summary.clumpingResult . . . 14
summary.phenotypeData . . . 14
summary.screeningResult . . . 15
summary.selectionResult . . . 15

Index **16**

`clumpingResult` *clumpingResult class*

Description

A result of procedure for snp clumping produced by `clump_snps`

Details

Always a named list of eleven elements

1. `X` numeric matrix, consists of one snp representative for each clump
2. `y` numeric vector, phenotype
3. `SNPnumber` numeric vector, which columns in SNP matrix `X_all` are related to clumps representatives
4. `SNPclumps` list of numeric vectors, which columns in SNP matrix `X_all` are related to clump members
5. `X_info` data.frame, mapping information about SNPs from .map file. Copied from the result of screening procedure.

6. selectedSnpsNumbers numeric vector, which rows of X_info matrix are related to selected clump representatives
7. X_all numeric matrix, all the snps that passed screening procedure
8. numberOfSnps numeric, total number of SNPs before screening procedure
9. selectedSnpsNumbersScreening numeric vector, which rows of X_info data.frame are related to snps that passed screening
10. pVals numeric vector, p-values from marginal tests for each snp
11. pValMax numeric, p-value used in screening procedure

See Also

screeningResult clump_snps

clump_snps *Clumping procedure for SLOPE*

Description

Clumping procedure performed on SNPs, columns of matrix X, from object of class screeningResult, which is an output of function screen_snps. SNPs are clustered based on their correlations. For details see package vignette.

Usage

```
clump_snps(screenResult, rho = 0.5, pValues = NULL, verbose = TRUE)
```

Arguments

screenResult object of class screeningResult

rho numeric, minimal correlation between two SNPs to be assigned to one clump

pValues numeric vector, p-values for SNPs computed outside geneSLOPE, eg. with EMMAX

verbose logical, if TRUE (default) progress bar is shown

Value

object of class clumpingResult

create_lambda *Lambda sequences for SLOPE*

Description

Computes λ sequences for SLOPE according to several pre-defined methods.

Usage

```
create_lambda(n, p, fdr = 0.2, method = c("bhq", "gaussian"))
```

Arguments

`n`	number of observations
`p`	number of variables
`fdr`	target False Discovery Rate (FDR)
`method`	method to use for computing λ (see Details)

Details

The following methods for computing λ are supported:

- `bhq`: Computes sequence inspired by Benjamini-Hochberg (BHq) procedure
- `gaussian`: Computes modified BHq sequence inspired by Gaussian designs

geneSLOPE *Genome-Wide Association Study with SLOPE*

Description

Package geneSLOPE performes genome-wide association study (GWAS) with **SLOPE**, short for Sorted L-One Penalized Estimation. SLOPE is a method for estimating the vector of coefficients in linear model. For details about it see references.

Details

GWAS is splitted into three steps.

- In the first step data is read using **bigmemory** package and immediately screened using marginal tests for each SNP
- SNPs are clumped based on their correlations
- SLOPE is performed on data where each clump has one representative (therefore we ensure that variables in linear model are not strogly correlated)

Version: 0.38.3

Author(s)

Malgorzata Bogdan, Damian Brzyski, Emmanuel J. Candes, Christine Peterson, Chiara Sabatti, Piotr Sobczyk

Maintainer: Piotr Sobczyk <pj.sobczyk@gmail.com>

References

SLOPE – Adaptive Variable Selection via Convex Optimization, Malgorzata Bogdan, Ewout van den Berg, Chiara Sabatti, Weijie Su and Emmanuel Candes

See Also

Useful links:

- https://github.com/psobczyk/geneSLOPE
- Report bugs at https://github.com/psobczyk/geneSLOPE/issues

Examples

```
famFile <- system.file("extdata", "plinkPhenotypeExample.fam", package = "geneSLOPE")
mapFile <- system.file("extdata", "plinkMapExample.map", package = "geneSLOPE")
snpsFile <- system.file("extdata", "plinkDataExample.raw", package = "geneSLOPE")
phe <- read_phenotype(filename = famFile)
screening.result <- screen_snps(snpsFile, mapFile, phe, pValMax = 0.05, chunkSize = 1e2)
clumping.result <- clump_snps(screening.result, rho = 0.3, verbose = TRUE)
slope.result <- select_snps(clumping.result, fdr=0.1)

## Not run:
gui_geneSLOPE()

## End(Not run)
```

---

gui_geneSLOPE *GUI for GWAS with SLOPE*

Description

A graphical user interface for performing Genome-wide Association Study with SLOPE

Usage

```
gui_geneSLOPE()
```

Details

requires installing **shiny** package

Value

null

identify_clump *identify_clump*

Description

identify_clump

Usage

```
identify_clump(x, ...)
```

Arguments

x	appropiate class object
...	other arguments

Details

Enable interactive selection of snps in plot. Return clump number.

identify_clump.clumpingResult *Identify clump number in clumpingResult class plot*

Description

Identify clump number in clumpingResult class plot

Usage

```
## S3 method for class 'clumpingResult'
identify_clump(x, ...)
```

Arguments

x	clumpingResult class object
...	Further arguments to be passed to or from other methods. They are ignored in this function.

identify_clump.selectionResult

Identify clump number in selectionResult class plot

Description

Identify clump number in selectionResult class plot

Usage

```
## S3 method for class 'selectionResult'
identify_clump(x, ...)
```

Arguments

`x`	selectionResult class object
`...`	Further arguments to be passed to or from other methods. They are ignored in this function.

phenotypeData

phenotypeData class

Description

Phenotype data

Details

Always a named list of two elements

1. `y` numeric vector, phenotype
2. `yInfo` data.frame, additional information about observations provied in .fam file

See Also

`read_phenotype`

plot.selectionResult *Plot selectionResult class object*

Description

Plot selectionResult class object

Usage

```
## S3 method for class 'selectionResult'
plot(x, chromosomeNumber = NULL, clumpNumber = NULL, ...)
```

Arguments

`x`	selectionResult class object
`chromosomeNumber`	optional parameter, only selected chromosome will be plotted
`clumpNumber`	optional parameter, only SNPs from selected clump will be plotted
`...`	Further arguments to be passed to or from other methods. They are ignored in this function.

print.clumpingResult *Print clumpingResult class object*

Description

Print clumpingResult class object

Usage

```
## S3 method for class 'clumpingResult'
print(x, ...)
```

Arguments

`x`	clumpingResult class object
`...`	Further arguments to be passed to or from other methods. They are ignored in this function.

`print.phenotypeData` *Print phenotypeData class object*

Description

Print phenotypeData class object

Usage

```
## S3 method for class 'phenotypeData'
print(x, ...)
```

Arguments

`x`	phenotypeData class object
`...`	Further arguments to be passed to or from other methods. They are ignored in this function.

`print.screeningResult` *Print function for class screeningResult class*

Description

Print function for class screeningResult class

Usage

```
## S3 method for class 'screeningResult'
print(x, ...)
```

Arguments

`x`	screeningResult class object
`...`	Further arguments to be passed to or from other methods. They are ignored in this function.

print.selectionResult *Print selectionResult class object*

Description

Print selectionResult class object

Usage

```
## S3 method for class 'selectionResult'
print(x, ...)
```

Arguments

x	selectionResult class object
...	Further arguments to be passed to or from other methods. They are ignored in this function.

Value

Nothing.

read_phenotype *Read phenotype from .fam file*

Description

Reading phenotype data from file. It is assumed, that data is given in .fam file. In this format, first column is family id (FID), second is individual id (IID), third is Paternal individual ID (PAT), fourth is Maternal individual ID (MAT), fifth is SEX and sixth and last is PHENOTYPE. If file has only four columns, then it is assumed that PAT and MAT columns are missing. If there is only one column, then it is assumed that only phenotype is provided.

Usage

```
read_phenotype(filename, sep = " ", header = FALSE, stringAsFactors = FALSE)
```

Arguments

filename	character, name of file with phenotype
sep	character, field seperator in file
header	logical, does first row of file contain variables names
stringAsFactors	logical, should character vectors be converted to factors?

Value

object of class phenotypeData

screeningResult *screeningResult class*

Description

A result of procedure for snp clumping produced by screen_snps

Details

Always a named list of eight elements

1. X numeric matrix, consists of snps that passed screening
2. y numeric vector, phenotype
3. X_info data.frame, SNP info from .map file
4. pVals numeric vector, p-values from marginal tests for each snp
5. numberOfSnps numeric, total number of SNPs in .raw file
6. selectedSnpsNumbers numeric vector, which rows of X_info data.frame are related to snps that passed screening
7. pValMax numeric, p-value used in screening procedure
8. phenotypeInfo data.frame, additional information about observations provied in phenotypeData object

See Also

phenotypeData screen_snps

screen_snps *Reading and screening SNPs from .raw file and*

Description

Reading .raw file that was previously exported from PLINK - see details. Additional information about SNP mapping is read from .map file.

Usage

```
screen_snps(
  rawFile,
  mapFile = "",
  phenotype,
  pValMax = 0.05,
  chunkSize = 100,
  verbose = TRUE
)
```

Arguments

`rawFile`	character, name of .raw file
`mapFile`	character, name of .map file
`phenotype`	numeric vector or an object of class `phenotypeData`
`pValMax`	numeric, p-value threshold value used for screening
`chunkSize`	integer, number of snps that will be processed together. The bigger chunkSize is, the faster function works but computer might run out of RAM
`verbose`	if TRUE (default) information about progress is printed

Details

Exporting data from PLINK To import data to R, it needs to be exported from PLINK using the option "–recodeAD" The PLINK command should therefore look like `plink --file input --recodeAD --out output`. For more information, please refer to: `https://zzz.bwh.harvard.edu/plink/dataman.shtml`

Value

object of class `screeningResult`

`selectionResult` *selectionResult class*

Description

A result of applying SLOPE to matrix of SNPs obtained by clumping produced. Result of function `select_snps`

Details

Always a named list of eighteen elements

1. `X` numeric matrix, consists of one snp representative for each clump selected by SLOPE
2. `effects` numeric vector, coefficients in linear model build on snps selected by SLOPE
3. `R2` numeric, value of R-squared in linear model build on snps selected by SLOPE
4. `selectedSNPs` which columns in matrix `X_all` are related to snps selected by SLOPE
5. `y` selectedClumps list of numeric vectors, which columns in SNP matrix `X_all` are related to clump members selected by SLOPE
6. `lambda` numeric vector, lambda values used by SLOPE procedure
7. `y` numeric vector, phenotype
8. `clumpRepresentatives` numeric vector, which columns in SNP matrix `X_all` are related to clumps representatives
9. `clumps` list of numeric vectors, which columns in SNP matrix `X_all` are related to clump members

10. `X_info` data.frame, mapping information about SNPs from .map file. Copied from the result of clumping procedure
11. `X_clumps` numeric matrix, consists of one snp representative for each clump
12. `X_all` numeric matrix, all the snps that passed screening procedure
13. `selectedSnpsNumbers` numeric vector, which rows of `X_info` data.frame are related to snps that were selected by SLOPE
14. `clumpingRepresentativesNumbers` numeric vector, which rows of `X_info` data.frame are related to snps that are clump represenatives
15. `screenedSNPsNumbers` numeric vector, which rows of `X_info` data.frame are related to snps that passed screening
16. `numberOfSnps` numeric, total number of SNPs before screening procedure
17. `pValMax` numeric, p-value used in screening procedure
18. `fdr` numeric, false discovery rate used by `SLOPE`

See Also

`screeningResult` `clumpingResult` `select_snps` `SLOPE`

`select_snps` *GWAS with SLOPE*

Description

Performs GWAS with SLOPE on given snp matrix and phenotype. At first clumping procedure is performed. Highly correlated (that is stronger than parameter *rho*) snps are clustered. Then SLOPE is used on snp matrix which contains one representative for each clump.

Usage

```
select_snps(
  clumpingResult,
  fdr = 0.1,
  type = c("slope", "smt"),
  lambda = "gaussian",
  sigma = NULL,
  verbose = TRUE
)
```

Arguments

`clumpingResult` clumpProcedure output

`fdr` numeric, False Discovery Rate for SLOPE

`type` method for snp selection. slope (default value) is SLOPE on clump representatives, smt is Benjamini-Hochberg procedure on single marker test p-values for clump representatives `lambda` lambda for SLOPE. See `create_lambda`

`sigma` numeric, sigma for SLOPE

`verbose` logical, if TRUE progress bar is printed

Value

object of class `selectionResult`

Examples

```
## Not run:
slope.result <- select_snps(clumping.result, fdr=0.1)

## End(Not run)
```

`summary.clumpingResult` *Summary clumpingResult class object*

Description

Summary clumpingResult class object

Usage

```
## S3 method for class 'clumpingResult'
summary(object, ...)
```

Arguments

`object` clumpingResult class object

`...` Further arguments to be passed to or from other methods. They are ignored in this function.

`summary.phenotypeData` *Summary phenotypeData class object*

Description

Summary phenotypeData class object

Usage

```
## S3 method for class 'phenotypeData'
summary(object, ...)
```

Arguments

`object` phenotypeData class object

`...` Further arguments to be passed to or from other methods. They are ignored in this function.

`summary.screeningResult` *Summary function for class screeningResult*

Description

Summary function for class screeningResult

Usage

```
## S3 method for class 'screeningResult'
summary(object, ...)
```

Arguments

`object` screeningResult class object

`...` Further arguments to be passed to or from other methods. They are ignored in this function.

`summary.selectionResult` *Summary selectionResult class object*

Description

Summary selectionResult class object

Usage

```
## S3 method for class 'selectionResult'
summary(object, clumpNumber = NULL, ...)
```

Arguments

`object` selectionResult class object

`clumpNumber` number of clump to be summarized

`...` Further arguments to be passed to or from other methods. They are ignored in this function.

Index

bigmemory, *4*

clump_snps, *2, 3*, 3
clumpingResult, 2, *3*, *13*
create_lambda, 4, *14*

geneSLOPE, 4
geneSLOPE-package *(*geneSLOPE*)*, 4
gui_geneSLOPE, 5

identify_clump, 6
identify_clump.clumpingResult, 6
identify_clump.selectionResult, 7

phenotypeData, 7, *11, 12*
plot.selectionResult, 8
print.clumpingResult, 8
print.phenotypeData, 9
print.screeningResult, 9
print.selectionResult, 10

read_phenotype, *7*, 10

screen_snps, *3*, *11*, 11
screeningResult, *3*, 11, *12, 13*
select_snps, *12, 13*, 13
selectionResult, 12, *14*
shiny, *5*
SLOPE, *4*, *13*
summary.clumpingResult, 14
summary.phenotypeData, 14
summary.screeningResult, 15
summary.selectionResult, 15