Package ‘hJAM’

February 20, 2020

Encoding UTF-8
Type Package
Title Hierarchical Joint Analysis of Marginal Summary Statistics
Version 1.0.0
Author Lai Jiang <jian848@usc.edu>
Maintainer Lai Jiang <jian848@usc.edu>
License MIT + file LICENSE
LazyData true
RoxygenNote 6.1.1
Suggests knitr, rmarkdown
VignetteBuilder knitr
URL https://github.com/lailylajiang/hJAM
BugReports https://github.com/lailylajiang/hJAM/issues
Imports ggplot2, ggpubr, dplyr, reshape2
NeedsCompilation no
Repository CRAN
Date/Publication 2020-02-20 14:50:05 UTC

R topics documented:

betas.Gy ................................................................. 2
conditional_A ........................................................... 2
get_cond_A ............................................................... 3
get_cond_alpha .......................................................... 4
Gl ........................................................................... 5
hJAM_egger .............................................................. 5
hJAM_Lnreg .............................................................. 7
marginal_A ............................................................... 8
output.format ............................................................ 8
print.hJAM_egger ..................................................... 9
print.hJAM_Lnreg ..................................................... 9
SNPs_heatmap ......................................................... 10
SNPs_info ............................................................... 10
SNPs_scatter_plot .................................................... 11

Index

betas.Gy

Example beta list of hJAM

Description

Example beta list of hJAM

Usage

betas.Gy

Format

The betas.Gy is the beta vector in the hJAM model: the association estimates between 210 SNPs and myocardial infarction. The summary data was collected from UK Biobank (n=459,324).

References


conditional_A

Example conditional A matrix of hJAM

Description

Example conditional A matrix of hJAM

Usage

conditional_A
**get_cond_A**

**Format**

The conditional $A$ is the conditional estimates alpha matrix in the hJAM model: the association estimates between 210 SNPs and body mass index (BMI) and type 2 diabetes (T2D). The summary data was collected from GIANT consortium ($n=339,224$) and DIAGRAM+GERA+UKB ($n=659316$) for BMI and T2D, respectively. We converted it from marginal $A$, using get_cond_A function in hJAM package.

**References**


---

**get_cond_A**

Compute conditional Z matrix

**Description**

The get_cond_A function is to get the conditional A matrix by using marginal A matrix

**Usage**

```r
get_cond_A(marginal_A, Gl, N.Gx, ridgeTerm = FALSE)
```

**Arguments**

- `marginal_A`: the marginal effects of SNPs on the exposures (Gx).
- `Gl`: the reference panel (Gl), such as 1000 Genome
- `N.Gx`: the sample size of each Gx. It can be a scalar or a vector. If there are multiple X’s from different Gx, it should be a vector including the sample size of each Gx. If all alphas are from the same Gx, it could be a scalar.
- `ridgeTerm`: ridgeTerm = TRUE when the matrix L is singular. Matrix L is obtained from the cholesky decomposition of G0’G0. Default as FALSE.

**Value**

A matrix with conditional estimates which are converted from marginal estimates using the JAM model.

**Author(s)**

Lai Jiang
**get_cond_alpha**

**Examples**

```r
data(Gl)
data(betas.Gy)
data(marginal_A)
get_cond_A(marginal_A = marginal_A, Gl = Gl, N.Gx = c(339224, 659316), ridgeTerm = TRUE)
```

**Description**

The `get_cond_alpha` function is to compute the conditional alpha vector for each X. If only one X in the model, please use `get_cond_alpha` instead of `get_cond_A`. A sub-step in the `get_cond_A` function.

**Usage**

```r
get_cond_alpha(alphas, Gl, N.Gx, ridgeTerm = FALSE)
```

**Arguments**

- `alphas`: the marginal effects of SNPs on one exposure (Gx).
- `Gl`: the reference panel (Gl), such as 1000 Genome.
- `N.Gx`: the sample size of the Gx. It can be a scalar.
- `ridgeTerm`: `ridgeTerm = TRUE` when the matrix L is singular. Matrix L is obtained from the cholesky decomposition of G0'G0. Default as FALSE.

**Value**

A vector with conditional estimates which are converted from marginal estimates using the JAM model.

**Author(s)**

Lai Jiang

**References**


**Examples**

```r
data(Gl)
data(betas.Gy)
data(marginal_A)
get_cond_alpha(alphas = marginal_A[, 1], Gl = Gl, N.Gx = 339224, ridgeTerm = TRUE)
```
**Gl**  
Example reference data of hJAM

**Description**  
The real data example from hJAM paper

**Usage**  
Gl

**Format**  
The Gl object is a data matrix with 2467 individual of 210 SNPs from 1000 Genome project.

**References**  

**hJAM_egger**  
Fit hJAM with Egger regression

**Description**  
The hJAM_egger function is to get the results from the hJAM model with Egger regression. It is for detecting potential pleiotropy

**Usage**  
hJAM_egger(betas.Gy, N.Gy, Gl, A, ridgeTerm = FALSE)

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>betas.Gy</td>
<td>The betas in the paper: the marginal effects of SNPs on the phenotype (Gy)</td>
</tr>
<tr>
<td>N.Gy</td>
<td>The sample size of Gy</td>
</tr>
<tr>
<td>Gl</td>
<td>The reference panel (Gl), such as 1000 Genome</td>
</tr>
<tr>
<td>A</td>
<td>The A matrix in the paper: the marginal/conditional effects of SNPs on the exposures (Gx)</td>
</tr>
<tr>
<td>ridgeTerm</td>
<td>ridgeTerm = TRUE when the matrix L is singular. Matrix L is obtained from the cholesky decomposition of G0’G0. Default as FALSE.</td>
</tr>
</tbody>
</table>
Value

An object of the hJAM with egger regression results.

**Exposure**  The intermediates, such as the modifiable risk factors in Mendelian Randomization and gene expression in transcriptome analysis.

**numSNP**  The number of SNPs that the user use in the instrument set.

**Estimate**  The conditional estimates of the associations between intermediates and the outcome.

**StdErr**  The standard error of the conditional estimates of the associations between intermediates and the outcome.

**Lower.CI**  The lower bound of the 95% confidence interval of the estimates.

**Upper.CI**  The upper bound of the 95% confidence interval of the estimates.

**Pvalue**  The p value of the estimates with a type-I error equals 0.05.

**Est.Int**  The intercept of the regression of intermediates on the outcome.

**StdErr.Int**  The standard error of the intercept of the regression of intermediates on the outcome.

**Lower.CI.Int**  The lower bound of the 95% confidence interval of the intercept.

**Upper.CI.Int**  The upper bound of the 95% confidence interval of the intercept.

**Pvalue.Int**  The p value of the intercept with a type-I error equals 0.05.

An object of hJAM with egger regression results.

**Author(s)**

Lai Jiang

**References**


**Examples**

data(Gl)
data(betas.Gy)
data(conditional_A)
hJAM_egger(betas.Gy = betas.Gy, Gl = Gl, N.Gy = 459324, A = conditional_A, ridgeTerm = TRUE)
**hJAM_lnreg**

Fit hJAM with linear regression

**Description**

The hJAM function is to get the results from the hJAM model using input data.

**Usage**

```
hJAM_lnreg(betas.Gy, N.Gy, Gl, A, ridgeTerm = FALSE)
```

**Arguments**

- `betas.Gy`: The betas in the paper: the marginal effects of SNPs on the phenotype (Gy)
- `N.Gy`: The sample size of Gy
- `Gl`: The reference panel (Gl), such as 1000 Genome
- `A`: The A matrix in the paper: the marginal/conditional effects of SNPs on the exposures (Gx)
- `ridgeTerm`: ridgeTerm = TRUE when the matrix L is singular. Matrix L is obtained from the cholesky decomposition of G0'G0. Default as FALSE.

**Value**

An object of the hJAM with linear regression results.

- **Exposure**: The intermediates, such as the modifiable risk factors in Mendelian Randomization and gene expression in transcriptome analysis.
- **numSNP**: The number of SNPs that the user use in the instrument set.
- **Estimate**: The conditional estimates of the associations between intermediates and the outcome.
- **StdErr**: The standard error of the conditional estimates of the associations between intermediates and the outcome.
- **Lower.CI**: The lower bound of the 95% confidence interval of the estimates.
- **Upper.CI**: The upper bound of the 95% confidence interval of the estimates.
- **Pvalue**: The p value of the estimates with a type-I error equals 0.05.

**Author(s)**

Lai Jiang

**References**

Examples

data(Gl)
data(betas.Gy)
data(conditional_A)
hJAM_lnreg(betas.Gy = betas.Gy, Gl = Gl, N.Gy = 459324, A = conditional_A, ridgeTerm = TRUE)

marginal_A

Example marginal A matrix of hJAM

Description
Example marginal A matrix of hJAM

Usage
marginal_A

Format

The marginal_A is the marginal estimates alpha matrix in the hJAM model: the association estimates between 210 SNPs and body mass index (BMI) and type 2 diabetes (T2D). The summary data was collected from GIANT consortium (n=339,224) and DIAGRAM+GERA+UKB (n=659316) for BMI and T2D, respectively.

References


output.format

Keep the output as three digits

Description
Keep the output as three digits

Usage
output.format(x, ...)

Arguments

x   input
...
other options you want to put in
print.hJAM_egger

Author(s)
Lai Jiang

---

Description
Print out for hJAM_egger

Usage
```r
## S3 method for class 'hJAM_egger'
print(x, ...)
```

Arguments
- `x` - input
- `...` - other options you want to put in

Author(s)
Lai Jiang

---

print.hJAM_lnreg  

Description
Print out for hJAM_lnreg

Usage
```r
## S3 method for class 'hJAM_lnreg'
print(x, ...)
```

Arguments
- `x` - input
- `...` - other options you want to put in

Author(s)
Lai Jiang
SNPs_heatmap

Description
To generate the heatmap of all the SNPs that the user use in the analysis

Usage
SNPs_heatmap(Gl)

Arguments
Gl The reference panel (Gl) of the SNPs that the user use in the analysis, such as 1000 Genome

Author(s)
Lai Jiang

Examples
data(Gl)
t = SNPs_heatmap(Gl = Gl)
t

SNPs_info

Example SNPs' information of hJAM

Description
Example SNPs' information of hJAM

Usage
SNPs_info

Format
The SNPs_info is the information of the 210 SNPs that we used in this data example. It includes three columns: the rsID, major allele, and minor allele frequency of each SNP. The minor allele frequencies were calculated in the 503 European-ancestry subjects in 1000 Genome project.

References
SNPs_scatter_plot

---

**SNPs_scatter_plot**  Scatter plot for all the SNPs used in the analysis

**Description**

To generate the scatter plot of all the SNPs that the user use in the analysis

**Usage**

```r
SNPs_scatter_plot(A, betas.Gy, num_X)
```

**Arguments**

- `A`: The effects of SNPs on the exposures (Gx).
- `betas.Gy`: The betas in the paper: the marginal effects of SNPs on the phenotype (Gy)
- `num_X`: The number of intermediates in the research question.

**Value**

A set of scatter plots with x-axis being the conditional $\alpha$ estimates for each intermediate and y-axis being the $\beta$ estimates.

**Author(s)**

Lai Jiang

**Examples**

```r
data(conditional_A)
data(betas.Gy)
t = SNPs_scatter_plot(A = conditional_A, betas.Gy = betas.Gy, num_X = 2)
t```

---
Index

*Topic datasets
  betas.Gy, 2
  conditional_A, 2
  Gl, 5
  marginal_A, 8
  SNPs_info, 10

betas.Gy, 2
conditional_A, 2
get_cond_A, 3
get_cond_alpha, 4
Gl, 5

hJAM_egger, 5
hJAM_lnreg, 7

marginal_A, 8
output.format, 8

print.hJAM_egger, 9
print.hJAM_lnreg, 9

SNPs_heatmap, 10
SNPs_info, 10
SNPs_scatter_plot, 11