Package 'LAVAKnock'

May 21, 2024

Type Package

Title Local genetic correlation via knockoffs reduces confounding due to cross-trait assortative mating

Version 0.1

Author Shiyang Ma, Iuliana Ionita-Laza

Maintainer Shiyang Ma <mashiyang1991@sjtu.edu.cn>

Description Functions for LAVAKnock to conduct local genetic correlation analysis via knockoffs.

License GPL-3

Depends R(>= 3.6.0)

Imports
SKAT, Matrix, MASS, SPAtest, CompQuadForm, irlba, matrixsampling, corpcor, GhostKnockoff

NeedsCompilation no

Repository CRAN

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

R topics documented:

Example.Ghostknockoff-generation

Data example of LAVA-Knock Z-scores with Ghostknockoff generation.

Description

This example dataset contains the Z-scores for two phenotypes and the genotype matrix of a unit: chr1:58387484-70992905. Due to computational and memory size constaints, we use 100 snps as an example to generation knockoff Z-scores using Ghostknockoff.

Usage

```
data("Ghostknockoff_generation.example")
```

Format

An object of class `list` of length 5.

Examples

```
data(Ghostknockoff_generation.example)
chr=Ghostknockoff_generation.example$chr
unit_start=Ghostknockoff_generation.example$unit_start
unit_end=Ghostknockoff_generation.example$unit_end
print(paste0(chr,chr,:,unit_start,-,unit_end))
#"chr1:58387484-70992905"
Zscore_0_unit=Ghostknockoff_generation.example$Zscore_0_unit
dim(Zscore_0_unit)
G_unit=Ghostknockoff_generation.example$G_unit
dim(G_unit)
```

Example.LAVAKnock   *Data example of LAVA-Knock knockoff filter for multiple testing.*

Description

This example dataset contains the LAVAKnock bivariateresults on 1,468 windows. Each window has an original p-values and M=5 knockoff p-values

Usage

```
data("Results_LAVAKnock.example")
```

Format

An object of class `data.frame` with 1468 rows and 7 columns.

Examples

```
data("Results_LAVAKnock.example")
dim(Results_LAVAKnock.example) #1468     7
```

```
Example.LAVAKnock_bivariate
```
Data example for LAVA-Knock bivariate local genetic correlation analysis on one window.

Description

This example dataset contains a 100-Kb window on chr1:64,344,227-64,444,227, its genotype matrix G_window with n=20000 samples and p=10 variants, with the original and M=5 knockoff Z-scores of 10 variants and 2 phenotypes

Usage

```
data("LAVAKnock_bivariate.example")
```

Format

An object of class `list` of length 4.

Examples

```
data("LAVAKnock_bivariate.example")
window=LAVAKnock_bivariate.example$window
window
G_window=LAVAKnock_bivariate.example$G_window
dim(G_window) #20000    10
Zscore_pheno1_window=LAVAKnock_bivariate.example$Zscore_pheno1_window
dim(Zscore_pheno1_window) #10  6
Zscore_pheno2_window=LAVAKnock_bivariate.example$Zscore_pheno2_window
```

```
Example.LAVAKnock_univariate
```
Data example for LAVA-Knock univariate heritability test on one locus.

Description

This example dataset contains a locus on chr1:64,344,227-65,894,184, its genotype matrix G_locus with n=20000 samples and p=197 variants, with the original Z-score of 197 variants and 2 phenotypes

Usage

```
data("LAVAKnock_univariate.example")
```

Format

An object of class `list` of length 3.

Examples

```
data("LAVAKnock_univariate.example")

Zscore_0=LAVAKnock_univariate.example$Zscore_0
G_locus=LAVAKnock_univariate.example$G_locus
locus=LAVAKnock_univariate.example$locus
print(locus)
dim(G_locus) #20000  197
dim(Zscore_0) #197   2
```

Ghostknockoff_generation

Conduct LAVA-Knock Ghostknockoff generation for a unit.

Description

This function generate the knockoff Z-scores for a unit using Ghostknockoff with (shrinkage) empirical LD matrix.

Usage

```
Ghostknockoff_generation(G_unit, Zscore_0_unit, LD.threshold = 0.75, n = 20000)
```

Arguments

G_unit	The genotype matrix or the reference genotype for a unit, a n*p matrix
Zscore_0_unit	Zscores of a unit, a p*2 matrix for two phenotypes
LD.threshold	LD.threshold for single-linkage hierarchical clustering to filter out highly correlated variants. The recommended level is 0.75.
n	sample size.

Value

Zscore_pheno1	5 knockoff Zscores for phenotype 1.
Zscore_pheno2	5 knockoff Zscores for phenotype 2.

Examples

```
data(Ghostknockoff_generation.example)
Zscore_0_unit=Ghostknockoff_generation.example$Zscore_0_unit
dim(Zscore_0_unit)
G_unit=Ghostknockoff_generation.example$G_unit
dim(G_unit)

Results_Ghostknockoff_generation=Ghostknockoff_generation(G_unit,
Zscore_0_unit,LD.threshold=0.75,n=20000)
dim(Results_Ghostknockoff_generation$Zscore_pheno1)
dim(Results_Ghostknockoff_generation$Zscore_pheno2)
```

LAVAKnock   *Conduct LAVA-Knock knockoff filter for multiple windows.*

Description

This function conducts the knockoff filter of LAVA-Knock on multiple windows and select significant windows.

Usage

```
LAVAKnock(
  M = 5,
  p0 = Results_LAVAKnock.example$pval.orginal,
  p_ko = cbind(Results_LAVAKnock.example$pval.knockoff1,
    Results_LAVAKnock.example$pval.knockoff2, Results_LAVAKnock.example$pval.knockoff3,
    Results_LAVAKnock.example$pval.knockoff4, Results_LAVAKnock.example$pval.knockoff5),
  fdr = 0.1,
  window_id = Results_LAVAKnock.example$window,
  Rej.Bound = 20000
)
```

Arguments

M	number of knockoffs. The recommended number is 5.
p0	p-values of the original Z-scores of multiple windows, which is a w*1 vector for w windows.
p_ko	matrix of the p-values for M knockoffs, which is a w*M vector for w windows.
fdr	target FDR level. The recommended level is 0.1.
window_id	id of the windows considered for multiple tesing.
Rej.Bound	calculate ratios for top Rej.Bound tau values. The recommended level is 20000.

Value

W	knockoff statistics for each window.
Qvaluel	Qvalue for each window.
W.threshold	threshold of the W statistics with target fdr level.
window_sign	Significant windows with q-values less then the fdr threshold.

Examples

```
Results_LAVAKnock=LAVAKnock(M=5,p0=Results_LAVAKnock.example$pval.orginal,
p_ko=cbind(Results_LAVAKnock.example$pval.knockoff1,
          Results_LAVAKnock.example$pval.knockoff2,
          Results_LAVAKnock.example$pval.knockoff3,
          Results_LAVAKnock.example$pval.knockoff4,
          Results_LAVAKnock.example$pval.knockoff5),
fdr = 0.1,window_id=Results_LAVAKnock.example$window,Rej.Bound=20000)
Results_LAVAKnock$W.threshold#1.071494
sum(Results_LAVAKnock$W>=Results_LAVAKnock$W.threshold) #278
sum(Results_LAVAKnock$Qvalue<=0.1) #278
```

```
print(paste0(length(Results_LAVAKnock$window_sign),  detected windows under M=5 knockoffs))
#278 detected windows under M=5 knockoffs
```

LAVAKnock_bivariate   *Conduct window-level LAVA-Knock bivariate local genetic correlation analysis.*

Description

This function conducts the window-level bivariate local genetic correlation test on one window.

Usage

```
LAVAKnock_bivariate(
  Zscore_pheno1_window = Zscore_pheno1_window,
  Zscore_pheno2_window = Zscore_pheno2_window,
  G_window = G_window,
  chr = chr,
  window_start,
  window_end,
  n = 20000,
  prune.thresh = 99
)
```

Arguments

Zscore_pheno1_window

> The original and M=5 knockoffs Z-scores of a window under phenotype 1, which is a p*6 matrix for p variants.

Zscore_pheno2_window

> The original and M=5 knockoffs Z-scores of a window under phenotype 2, which is a p*6 matrix for p variants.

G_window   The genotype matrix for a window, which is a n*p matrix.

chr   chromosome.

window_start   window start.

window_end   window end.

n   sample size.

prune.thresh   Pruning threshold of sungular value decomposition. The recommended level is to explain 99 percent of total variance.

Value

rg.orginal   local genetic correlation estimate of original Z-score.

pval.orginal   bivariate p-value of local genetic correlation test for original Z-score.

rg.knockoff   local genetic correlation estimate of knockoff Z-score.

pval.knockoff   bivariate p-value of local genetic correlation test for knockoff Z-score.

Examples

```
data("LAVAKnock_bivariate.example")
window=LAVAKnock_bivariate.example$window
G_window=LAVAKnock_bivariate.example$G_window
Zscore_pheno1_window=LAVAKnock_bivariate.example$Zscore_pheno1_window
Zscore_pheno2_window=LAVAKnock_bivariate.example$Zscore_pheno2_window
chr=window$chr
window_start=window$start
window_end=window$end
Results_LAVAKnock_bivariate_window=LAVAKnock_bivariate(Zscore_pheno1_window=Zscore_pheno1_window,
Zscore_pheno2_window=Zscore_pheno2_window,
G_window=G_window,chr=chr,window_start,window_end,n=20000,prune.thresh=99)
Results_LAVAKnock_bivariate_window
```

LAVAKnock_univariate   *Conduct loci-level univariate heritability test.*

Description

This function conducts the loci-level univariate heritability test on one locus.

Usage

```
LAVAKnock_univariate(
  Zscore = Zscore_0,
  G_locus = G_locus,
  chr,
  locus_start,
  locus_end,
  n = 20000,
  prune.thresh = 99
)
```

Arguments

Zscore	The original Z-score of a locus, which is a p*2 matrix for p variants and 2 phenotypes.
G_locus	The genotype matrix or the reference genotype for a locus, which is a n*p matrix.
chr	chromosome.
locus_start	locus start.
locus_end	locus end.
n	sample size.
prune.thresh	Pruning threshold of sungular value decomposition. The recommended level is to explain 99 percent of total variance.

Value

    `h2_phe1.orginal`

                  heritability estimate of phenotype 1.

    `h2_phe2.orginal`

                  heritability estimate of phenotype 2.

    `p_uni_phe1.orginal`

                  p-value of univariate heritability test of phenotype 1.

    `p_uni_phe2.orginal`

                  p-value of univariate heritability test of phenotype 2.

Examples

```
data("LAVAKnock_univariate.example")
Zscore_0=LAVAKnock_univariate.example$Zscore_0
G_locus=LAVAKnock_univariate.example$G_locus
locus=LAVAKnock_univariate.example$locus
chr=locus$chr
locus_start=locus$start
locus_end=locus$stop
Results_univariate_locus=LAVAKnock_univariate(Zscore=Zscore_0,
G_locus=G_locus,chr,locus_start,locus_end,n=20000,prune.thresh=99)
Results_univariate_locus
```

Index