

Package ‘GLM2020’

March 30, 2020

Type Package

Title GWAS using GLM

Version 0.1.0

Author Who wrote it

Maintainer The package maintainer <yourself@somewhere.net>

Description More about what it does (maybe more than one line)

Use four spaces when indenting paragraphs within the Description.

License What license is it under?

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Suggests knitr, rmarkdown

Depends psych

VignetteBuilder knitr

R topics documented:

GLM2020-package	1
cofactor.pca.cor	2
GWASbyGLM	3

Index	4
--------------	----------

GLM2020-package	<i>A tool for implementing genome-wide association analysis using a general linear model.</i>
-----------------	---

Description

This package offers several features for efficiently carrying out genome-wide association studies (GWAS) using general linear models (GLM). Users may input phenotypic and genotypic data, as well as user-specified cofactors (e.g., location collected) and principal components calculated from genotype data. In addition to implementing GWAS using a GLM, this package offers functionality for automatically removing genotypic principal components that exhibit correlation with user-specified cofactors.

Usage

GLM2020-package

Format

An object of class character of length 1.

cofactor.pca.cor

Correlation between cofactors and principal components.

Description

Test for correlations between user-specified cofactors and principal components calculated from genotype data. Automatically remove principal components linearly dependent (correlated) with user-specified cofactors.

Usage

```
cofactor.pca.cor(U, G)
```

Arguments

- | | |
|---|---|
| U | A numeric matrix containing user-specified cofactors. Dimensions are n rows (individuals) by t columns (cofactors). |
| G | A numeric matrix containing genotype data. Dimensions are n rows (individuals) by m columns (genetic markers). |

Details

When U is unspecified, cofactor.pca.cor will return a list of 1 object. With U unspecified, function will carry out principal components analysis identically to the native R function prcomp(), and cofactor.pca.cor will return principal components scores in \$cov. \$cov is a numeric matrix containing all principal components and individual scores. Dimensions are n rows (individuals) by t columns (principal components).

When U is specified, cofactor.pca.cor will return a list of 3 objects. \$orig_pc is a numeric matrix containing all original principal components and individual scores. \$cov is a numeric matrix containing user-specified cofactors and all principal components not correlated with the user-specified cofactors. Dimensions are n rows (individuals) by t columns (cofactors). \$removed is a character matrix indicating which principal components were removed.

The \$cov matrix is intended for use as the "C" argument in the GWASbyGLM function included in this package.

Value

A list of 1 or 3 objects.

U unspecified: 1 object. \$cov, a numeric matrix containing all principal components and individual scores.

U specified: 3 objects. \$orig_pc, a numeric matrix containing all original principal components \$cov, a numeric matrix containing user-specified cofactors and retained principal components. \$removed, a matrix indicating which principal components were removed.

GWASbyGLM*Genome-wide association analysis using a general linear model.*

Description

Genome-wide association analysis using a general linear model.

Usage

```
GWASbyGLM(y, G, C, NC)
```

Arguments

- | | |
|----|--|
| y | A numeric matrix containing phenotype data. Dimensions are n rows (individuals) by 1 column. |
| G | A numeric matrix containing genotype data. Dimensions are n rows (individuals) by m columns (genetic markers). |
| C | A numeric matrix containing covariate data. Dimensions are n rows (individuals) by t columns (covariates). The expected input for this parameter is the \$cov numeric matrix returned from the cofactor.pca.cor function included in this package. |
| NC | An integer specifying the number of covariates to retain for analysis. |

Value

A numeric matrix containing a p-value for each genetic marker. Dimensions are 1 row by m columns (genetic markers).

Index

*Topic **datasets**

GLM2020-package, [1](#)

cofactor.pca.cor, [2](#)

GLM2020-package, [1](#)

GWASbyGLM, [3](#)