Package 'QSCAN'

May 15, 2020

Type Package

Title Simultaneous Detection of Signal Regions Using Quadratic Scan Statistics

Version 1.0.0

Date 2020-05-14

Author Zilin Li [aut, cre]

Maintainer Zilin Li <li@hsph.harvard.edu>

Description R package for performing Q-SCAN procedure in whole genome squencing studies.

License GPL-3

Copyright See COPYRIGHTS for details.

Imports Rcpp, RcppArmadillo, Matrix

Encoding UTF-8

LazyData true

Depends R (>= 3.0.0)

LinkingTo Rcpp, RcppArmadillo

RoxygenNote 7.1.0

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

`M_SCAN` *Mean SCAN (M-SCAN) statistic based procedure*

Description

The `M_SCAN` function takes in genotype, phenotype and covariates and detect the association between a quantitative/dichotomous phenotype and a variant-set in a sequence by using quadratic scan statistic based procedure.

Usage

```
M_SCAN(
  genotype,
  phenotype,
  X,
  family,
  Lmax,
  Lmin,
  steplength = 1,
  times = 2000,
  alpha = 0.05,
  f = 0
)
```

Arguments

genotype	an n*p genotype matrix (dosage matrix) of the target sequence, where n is the sample size and p is the number of variants.
phenotype	an n*1 phenotype vector, where n is the sample size.
X	an n*q covariates matrix, where n is the sample size and q is the number of covariates.
family	a description of the error distribution and link function to be used in the model. This can be a character string naming a family function, a family function or the result of a call to a family function. (See family for details of family functions). Can be either "gaussian" for continuous phenotype or "binomial" for binary phenotype.
Lmax	maximum number of variants in searching windows.
Lmin	minimum number of variants in searching windows.
steplength	difference of number of variants in searching windows, that is, the number of variants in searching windows are Lmin, Lmin+steplength, Lmin+steplength,...,Lmax. (Default is 1).
times	a number of pesudo-residuals (default = 2000).
alpha	familty-wise/genome-wide significance level. (Default is 0.05).
f	an overlap fraction, which controls for the overlapping proportion of of detected regions. For example, when f=0, the detected regions are non-overlapped with each other, and when f=1, we keep every susceptive region as detected regions. (Default is 0.)

Value

The function returns a list with the following members:

SCAN_res: A matrix that summarized the significant region detected by Q-SCAN. The first column is the quadratic scan statistic of the detected region. The next two columns are the location of the detected region (in sense of variants order). The last column is the family-wise/genome-wide error rate of the detected region. The result (0,0,0,1) means there is no significant region.

SCAN_top1: A vector of length 4 which summarized the top 1 region detected by M-SCAN. The first element is the quadratic scan statistic of the detected region The next two elements are the location of the detected region (in sense of variants order). The last element is the family-wise/genome-wide p-value.

SCAN_thres: Empirical threshold of Q-SCAN for controlling the family-wise type I error at alpha level.

SCAN_thres_boot: A vector of Monte Carlo simulation sample for generating the empirical threshold. The 1-alpha quantile of this vector is the empirical threshold.

Q_SCAN *Quadratic SCAN (Q-SCAN) statistic based procedure*

Description

The Q_SCAN function takes in genotype, phenotype and covariates and detect the association between a quantitative/dichotomous phenotype and a variant-set in a sequence by using quadratic scan statistic based procedure.

Usage

```
Q_SCAN(
  genotype,
  phenotype,
  X,
  family,
  Lmax,
  Lmin,
  steplength = 1,
  times = 2000,
  alpha = 0.05,
  f = 0
)
```

Arguments

genotype	an n*p genotype matrix (dosage matrix) of the target sequence, where n is the sample size and p is the number of variants.
phenotype	an n*1 phenotype vector, where n is the sample size.
X	an n*q covariates matrix, where n is the sample size and q is the number of covariates.
family	a description of the error distribution and link function to be used in the model. This can be a character string naming a family function, a family function or the result of a call to a family function. (See family for details of family functions). Can be either "gaussian" for continuous phenotype or "binomial" for binary phenotype.
Lmax	maximum number of variants in searching windows.
Lmin	minimum number of variants in searching windows.
steplength	difference of number of variants in searching windows, that is, the number of variants in searching windows are Lmin, Lmin+steplength, Lmin+steplength,...,Lmax. (Default is 1).
times	a number of pesudo-residuals (default = 2000).
alpha	familty-wise/genome-wide significance level. (Default is 0.05).

`f` an overlap fraction, which controls for the overlapping proportion of of detected regions. For example, when f=0, the detected regions are non-overlapped with each other, and when f=1, we keep every susceptive region as detected regions. (Default is 0.)

Value

The function returns a list with the following members:

`SCAN_res`: A matrix that summarized the significant region detected by Q-SCAN. The first column is the quadratic scan statistic of the detected region. The next two columns are the location of the detected region (in sense of variants order). The last column is the family-wise/genome-wide error rate of the detected region. The result (0,0,0,1) means there is no significant region.

`SCAN_top1`: A vector of length 4 which summarized the top 1 region detected by Q-SCAN. The first element is the quadratic scan statistic of the detected region The next two elements are the location of the detected region (in sense of variants order). The last element is the family-wise/genome-wide p-value.

`SCAN_thres`: Empirical threshold of Q-SCAN for controlling the family-wise type I error at alpha level.

`SCAN_thres_boot`: A vector of Monte Carlo simulation sample for generating the empirical threshold. The 1-alpha quantile of this vector is the empirical threshold.

Index