Package 'CateSelection'

February 19, 2015

Type Package

Title Categorical Variable Selection Methods

Version 1.0

Date 2014-10-28

Author Yi Xu and Jixiang Wu

Maintainer Yi Xu <statxy@gmail.com>

Description A multi-factor dimensionality reduction based forward selection method for genetic association mapping.

License GPL (>= 2)

Depends R(>= 2.10)

Repository CRAN

NeedsCompilation no

Date/Publication 2014-10-28 21:58:29

R topics documented:

`MDRforward-package` *A MDR based forward selection method*

Description

A MDR (multi-factor dimensionality reduction) based forward selection method for genetic association mapping.

Details

Package:	MDRforward
Type:	Package
Version:	1.0
Date:	2014-03-04
License: GPL(>= 2)	

Author(s)

Yi Xu and Jixiang Wu
Maintainer: Yi Xu <statxy@gmail.com>

References

Marylyn D. Ritchie, Lance W. Hahn, Nady Roodi, L. Renee Bailey, William D. Dupont, Fritz F. Parl, and Jason H. Moore (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. American Journal of Human Genetics 69: 138-147.

Yi Xu, Jixiang Wu, Detecting epistatic effects among SNP markers associated with three barley traits by a MDR based forward selection method (unpublished).

Yi Xu, Jixiang Wu, Detecting higher-order interactions of SNP markers associated with three barley agronomic traits (unpublished).

`data1` *A simulated dataset*

Description

A data frame with simulated genotype and phenotype

Usage

```
data(data1)
```

Format

A data frame with 100 observations on the following 21 variables.

The first column is the phenotypic data, and other 20 columns are the genotypic data.

Details

The phenotypic data was generated using the first four predictive variables(x1-x4) by interacting with x1 and x2, x3 and x4, resepectively.

Examples

```
data(data1)
```

`data2` *A simulated dataset*

Description

A data frame with simulated genotype and phenotype

Usage

```
data(data2)
```

Format

A data frame with 100 observations on the following 16 variables.

The first column is the phenotypic data, and other 15 columns are the genotypic data.

Details

The phenotypic data was generated using the first six predictive variables(x1-x6) by interacting with x1 x2 and x3, x4 x5 and x6, resepectively.

Examples

```
data(data2)
```

`MDR.high.forward` *MDR based selection methods for higher-order interacations*

Description

MDR based three-stage selection methods for higher-order interacations

Usage

```
MDR.high.forward(x, y, order = NULL, trace = NULL, alpha = NULL, beta = NULL,
pvalue = NULL, r2 = NULL, ...)
```

Arguments

`x`	A matrix of genotypic data/genetic markers (predictor variables), where the rows are the samples and the columns are the predictors.
`y`	A numeric vector of phenotypic data (response variable).
`order`	The order of interaction. Default is 3.
`trace`	Show computations? Default FALSE.
`alpha`	Cutoff value for the difference (D1) of coefficient of determination between single modles with and without MRD interactions in the first stage. Default is 0.1.
`beta`	Cutoff value for the difference (D2) of coefficient of determination between single modle with p interactions and single model with (p-1) interactions in the second stage. Default is 0.05.
`pvalue`	Cutoff value for p-value in the third stage. Default is 0.01.
`r2`	Cutoff value for the difference of coefficient of determination in the third stage. Default is 0.02.
`...`	Other arguments for future methods.

Value

It returns a matrix with the index of selected interactive predictors, and the corresponding adjusted coefficient of determination.

References

Yi Xu, Jixiang Wu, Detecting higher-order interactions of SNP markers associated with three barley agronomic traits (unpublished).

Examples

```
data(data2)
y <- data2[,1]
x <- data2[,-1]
res <- MDR.high.forward(x,y,order=3)
res
```

`MDR.sing.mod` *Single Model with and without MDR interactions*

Description

Computes coefficient of determination and p-value for both single model with and without MDR interactions

Usage

```
MDR.sing.mod(x, y, order = NULL, trace = NULL, ...)
```

Arguments

`x`	A matrix of genotypic data/genetic markers (predictor variables), where the rows are the samples and the columns are the predictors.
`y`	A numeric vector of phenotypic data (response variable).
`order`	The order of interaction. Default is 2.
`trace`	Show computations? Default FALSE.
`...`	Other arguments for future methods.

Value

It returns a matrix with the index of interactive predictors, and the corresponding p-value, and adjusted coefficient of determination both for single models with and without MRD interactions.

References

Yi Xu, Jixiang Wu, Detecting higher-order interactions of SNP markers associated with three barley agronomic traits (unpublished).

Examples

```
data(data2)
y <- data2[,1]
x <- data2[,-1]
res <- MDR.sing.mod(x,y,order=3)
res
```

`MDR.stage.forward` *MDR based forward selection method*

Description

MDR based forward selection method for association mapping

Usage

```
MDR.stage.forward(x, y, order = NULL, s1.rsquared = NULL, s1.pvalue = NULL,
s2.rsquared = NULL, s2.pvalue = NULL, max.step = NULL, trace = NULL, ...)
```

Arguments

`x`	A matrix of genotypic data/genetic markers (predictor variables), where the rows are the samples and the columns are the predictors.
`y`	A numeric vector of phenotypic data (response variable).
`order`	The order of interaction. Default is 2.
`s1.rsquared`	Cutoff value for coefficient of determination in the first stage. Default is 0.02.
`s1.pvalue`	Cutoff value for p-value in the first stage. Default is 0.01.
`s2.rsquared`	Cutoff value for coefficient of determination in the second stage. Default is 0.02.
`s2.pvalue`	Cutoff value for p-value in the second stage. Default is 0.01.
`max.step`	The maximum selection step. Default is 100.
`trace`	Show computations? Default FALSE.
`...`	Other arguments for future methods.

Value

It returns a matrix with the index of selected interactive predictors, and the corresponding adjusted coefficient of determination for each step.

References

Yi Xu, Jixiang Wu, Detecting epistatic effects among SNP markers associated with three barley traits by a MDR based forward selection method (unpublished).

Examples

```
data(data1)
y <- data1[,1]
x <- data1[,-1]
res <- MDR.stage.forward(x,y,order=2)
res
```

`sing.mod` *Single MDR Model*

Description

Computes coefficient of determination and p-value for each single (marginal) MDR model

Usage

```
sing.mod(x, y, order = NULL, alpha = NULL, beta = NULL, delete = NULL, trace = NULL, ...)
```

Arguments

`x`	A matrix of genotypic data/genetic markers (predictor variables), where the rows are the samples and the columns are the predictors.
`y`	A numeric vector of phenotypic data (response variable).
`order`	The order of interaction. Default is 2.
`alpha`	Cutoff value for p-value. Default is 0.01. If "delete" is TURE, the interactive predictors in the model with the p-value greater than "alpha" will be removed in the final results.
`beta`	Cutoff value for coefficient of determination. Default is 0.01. If "delete" is TURE, the interactive predictors in the model with the coefficient of determination smaller than "beta" will be removed in the final results.
`delete`	Logistic value. Default FALSE; TRUE will remove the interactive predictors according two cutoff values "alpha" and "beta".
`trace`	Show computations? Default FALSE.
`...`	Other arguments for future methods.

Value

It returns a matrix with the index of interactive predictors, and the corresponding p-value, coefficient of determination, and adjusted coefficient of determination.

References

Yi Xu, Jixiang Wu, Detecting epistatic effects among SNP markers associated with three barley traits by a MDR based forward selection method (unpublished).

Examples

```
data(data1)
y <- data1[,1]
x <- data1[,-1]
res <- sing.mod(x,y,order=2)
res
```

Index