Package ‘ezglm’

February 19, 2015

Title selects significant non-additive interaction between two variables using fast GLM implementation.

Version 1.0

Date 2012-6-20

Author Yi Yang <yiyang@umn.edu>

Maintainer Yi Yang <yiyang@umn.edu>

Description This package implements a simplified version of least squares, and logistic regression for efficiently selecting the significant non-additive interactions between two variables.

License GPL-2

URL http://code.google.com/p/ezglm/

Repository CRAN

Date/Publication 2012-06-21 05:45:54

NeedsCompilation yes

R topics documented:

   ezglm     . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

Index

   ezglm      selects significant non-additive interaction between two variables using fast GLM implementation.

Description

   selects significant non-additive interaction between two variables using fast GLM implementation.

Usage

   ezglm(y, x1, x2, thr = 1, family=c("gaussian","binomial"))
Arguments

y  response variable, of length n. This argument should be quantitative for least squares, and a two-level factor for logistic regression.

x1  the first predictor, of length n.

x2  the second predictor, of length n.

thr  p-value tolerance. Truncate any p-value to 1 if it is larger than thr. Defaults value is 1.

family  a character string specifying the model to use, valid options are:
        • "gaussian" least squares regression (regression),
        • "binomial" logistic regression (classification).
        Default is "gaussian".

Details

Motivated by pairwise gene interaction selection in genome-wide association study (GWAS), this package implements fast and simplified least squares, and logistic regression for efficiently selecting the significant non-additive interactions between two variables. Once a user specifies a response variable y and predictors x1 and x2, for "gaussian" a least squares model \( y \sim x1 + x2 + x1*x2 \) is fitted; for "binomial" a logistic regression \( \text{logit} \sim x1 + x2 + x1*x2 \) is fitted. Users can then select significant x1*x2 term using returned p-value of Wald test.

Value

A matrix of coefficients.

Author(s)

Yi Yang
Maintainer: Yi Yang <yiyang@umn.edu>

Examples

```r
n = 10000
x1 = rnorm(n)
x2 = rnorm(n)
y1 = sample(c(0,1),n,rep=TRUE)
y2 = rnorm(n)

system.time(m1 <- ezglm(y1, x1, x2, 1, family = "binomial"))
m1

system.time(m2 <- glm(y1~x1+x2+x1*x2, family = binomial))
summary(m2)$coef

system.time(m3 <- ezglm(y2, x1, x2, 1, family = "gaussian"))
m3

system.time(m4 <- glm(y2~x1+x2+x1*x2, family = gaussian))
```
ezglm

summary(m4)$coef
Index

* Topic models
  ezglm, 1

* Topic regression
  ezglm, 1

ezglm, 1