

Package ‘CKLRT’

September 13, 2018

Type Package

Title Composite Kernel Machine Regression Based on Likelihood Ratio Test

Version 0.2.3

Author Ni Zhao [aut], Haoyu Zhang [aut, cre]

Maintainer Haoyu Zhang <andrew.haoyu@gmail.com>

Description Composite Kernel Machine Regression based on Likelihood Ratio Test (CKLRT): in this package, we develop a kernel machine regression framework to model the overall genetic effect of a SNP-set, considering the possible GE interaction. Specifically, we use a composite kernel to specify the overall genetic effect via a nonparametric function and we model additional covariates parametrically within the regression framework. The composite kernel is constructed as a weighted average of two kernels, one corresponding to the genetic main effect and one corresponding to the GE interaction effect. We propose a likelihood ratio test (LRT) and a restricted likelihood ratio test (RLRT) for statistical significance. We derive a Monte Carlo approach for the finite sample distributions of LRT and RLRT statistics. (N. Zhao, H. Zhang, J. Clark, A. Maity, M. Wu. Composite Kernel Machine Regression based on Likelihood Ratio Test with Application for Combined Genetic and Gene-environment Interaction Effect (Submitted).)

License GPL-3

Depends R (>= 3.1.0),

Imports Rcpp (>= 0.12.17), mgcv, MASS, nlme, compiler, stats

LinkingTo RcppEigen, Rcpp

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation yes

Repository CRAN

Date/Publication 2018-09-13 21:40:02 UTC

R topics documented:

omniLRT_fast	2
omniRLRT_fast	3

Index	5
--------------	----------

omniLRT_fast	<i>Composite kernel machine regression based likelihood ratio test.</i>
--------------	---

Description

Composite kernel machine regression based likelihood ratio test. The approximate method for likelihood ratio test tend to be too conservative for small alpha values. We recommend not using it in GWAS

Usage

```
omniLRT_fast(y, X, K1, K2, N = 10000, length.lambda = 200,
             length.rho = 21)
```

Arguments

y	: y is the vecgtor of the continous outcomes.
X	: X denotes the additional covariates.
K1	: K1 is the first kernel corresponding to the genetic main effect.
K2	: K2 is the second kernel corresponding to the genetic and environment interaction effect.
N	: N is th total number of randomly generated normal variables used to generate the emprical null distribution of LRT. Default value is 10,000.
length.lambda	: the length of lambda. Dafult value is 200. The values of lambda are all more than 0.
length.rho	: the length of rho. Default value is 21. The values of rho are between 0 and 1.

Value

the result is a list containing three elements. 1. p.dir is the p-value of likelihood ratio test based on emprical distrition. 2. p.aud is the p-value by approximating the null distribution as a mixture of a point mass at zero with probability b and weighted chi square distribution with d degrees of freedom with probality of 1-b. 3. LR is the likelihood ratio test statistics.

Examples

```

set.seed(6)
n = 50 # the number of observations
X = rnorm(n) # the other covariates
p = 2 # two snp in a gene will be simulated
G = runif(n*p) < 0.5
G = G + runif(n*p) < 0.5
G = matrix(G, n,p) #genetic matrix
E = (runif(n) < 0.5)^2 #enviroment effect
y = rnorm(n) + G[,1] * 0.3 #observations
omniRLRT_fast(y, X = cbind(X, E),K1 = G %% t(G),K2 = (G*E) %% t(G * E))

```

omniRLRT_fast	<i>Composite kernel machine regression based restricted likelihood ratio test</i>
---------------	---

Description

Composite kernel machine regression based restricted likelihood ratio test

Usage

```

omniRLRT_fast(y, X, K1, K2, N = 10000, length.rho = 200,
length.lambda = 21)

```

Arguments

y	vector of the continous outcomes.
X	the additional covariates.
K1	the first kernel corresponding to the genetic main effect.
K2	the second kernel corresponding to the genetic and environment interaction effect.
N	total number of randomly generated normal variables used to generate the empirical null distribution of LRT. Default value is 10,000.
length.rho	the length of rho. Default value is 21. The values of rho are between 0 and 1.
length.lambda	the length of lambda. Dafult value is 200. The values of lambda are all more than 0.

Value

the result is a list containing three elements. 1. p.dir is the p-value of restricted likelihood ratio test based on empirical distrition. 2. p.aud is the p-value by approximating the null distribution as a mixture of a point mass at zero with probability b and weighted chi square distribution with d degrees of freedom with probality of 1-b. 3. LR is the likelihood ratio test statistics.

Examples

```
set.seed(6)
n = 50 # the number of observations
X = rnorm(n) # the other covariates
p = 2 # two snp in a gene will be simulated
G = runif(n*p) < 0.5
G = G + runif(n*p) < 0.5
G = matrix(G, n,p) #genetic matrix
E = (runif(n) < 0.5)^2 #enviroment effect
y = rnorm(n) + G[,1] * 0.3 #observations
omniRLRT_fast(y, X = cbind(X, E),K1 = G %*% t(G),K2 = (G*E) %*% t(G * E))
```

Index

omniLRT_fast, [2](#)
omniRLRT_fast, [3](#)