Package: GEEaSPU (via r-universe)

October 19, 2026

Type Package

Title Adaptive Association Tests for Multiple Phenotypes using Generalized Estimating Equations (GEE)

Version 1.0.2

Date 2016-08-03

Author Junghi Kim and Wei Pan

Maintainer Junghi Kim <junghikim0@gmail.com>

Description Provides adaptive association tests for SNP level, gene level and pathway level analyses.

License GNU General Public License (>= 3)

Depends R (>= 2.10), gee, stats

Imports Rcpp (>= 0.12.6)

LinkingTo Rcpp, RcppArmadillo

RoxygenNote 5.0.1

NeedsCompilation yes

Repository https://cran.r-universe.dev

Date/Publication 2016-08-04 11:47:48 UTC

RemoteUrl https://github.com/cran/GEEaSPU

RemoteRef HEAD

RemoteSha 56aa48c405a07d22920058d9eafb6123549cb725

Contents

GEEaSPU-package 2
GEEaSPU 2
GEEaSPUpath 4
GEEaSPUset 7

Index **10**

GEEaSPU-package — *Adaptive Association Tests for Multiple Phenotypes using Generalized Estimating Equations (GEE)*

Description

Provides adaptive association tests for SNP level, gene level and pathway level analyses.

Details

Functions for adaptive association tests including GEEaSPU, GEEaSPUset and GEEaSPUpath. These methods use a weighting scheme for testing associations with multiple phenotypes. GEEaSPU function computes a SNP level p value; GEEaSPUset function can be used for the SNP-set (gene) level association test, while GEEaSPUpath function is for the gene-set (pathway) level analysis.

Author(s)

Junghi Kim and Wei Pan

References

Kim J, Zhang Y, and Pan W (2016). "Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data." *Genetics* 203(2), 715–731.

Zhang Y, Xu Z, Shen, X and Pan W (2014). "Testing for association with multiple traits in generalized estimation equations, with application to neuroimaging data." *NeuroImage* 96, 309–325.

GEEaSPU — *An adaptive sum of powered score (SPU) test for single SNP and mutiple phenotypes*

Description

Tests whether a SNP is associated with multivariate outcomes; provides a series of p-values of GEE-SPU(pow) and GEE-aSPU tests.

Usage

```
GEEaSPU(pheno, geno, Z = NULL, model = "gaussian", corstr = "independence",
       pow = c(1:8, Inf), n.perm = 1000, null.type = "perm", score.test = FALSE)
```

Arguments

pheno	a numeric phenotype matrix with each row as a different individual and each column as a separate phenotype.
geno	a numeric vector with each element for an individual genotype.
Z	a numeric covariate matrix with each row as a different individual and each column as a covariated to be adjested.
model	a character string specifying the model of the phenotypes. Models supported are "gaussian" for a quantitative trait and "binomial" for a binary trait (default = "gaussian").
corstr	a character string specifying the correlation structure of phenotypes. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured" (default = "independence").
pow	a vector of the power weight to be used at a trait level (default = c(1:8, Inf)).
n.perm	a numeric value of number of null statistics (default = 1000).
null.type	a character string specifying how to generate null statistics; "perm" is used when null statistics are generated using permutations and "sim" is used when null statistics are generated using simulations (default = "perm").
score.test	a logical value indicating whether to include GEEaSPU-Score test along with GEE-Score test (default = FALSE). If TRUE, it computes p-values of GEEaSPU-Score and GEE-Score as well as GEEaSPU test.

Details

Adaptive association tests for single SNP and multiple phenotypes using GEE.

Value

a vector of p-values from GEE-SPU(pow) tests and GEE-aSPU test.

Note

When large SNP-set (namely large gene size) or large number of phenotypes are included, the permuation based test (null.type = "perm") is recommended.

An option "binomial" model only supports the option, null.type = "sim".

Author(s)

Junghi Kim and Wei Pan

References

Kim J, Zhang Y, and Pan W (2016). "Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data." *Genetics* 203(2), 715–731.

Zhang Y, Xu Z, Shen, X and Pan W (2014). "Testing for association with multiple traits in generalized estimation equations, with application to neuroimaging data." *NeuroImage*, 96, 309–325.

Examples

```
# -- simulating phenotypes
# -- n.subjects: number of subjects
# -- n.traits: number of phenotypes
# -- Sigma: covariance matrix of phenotypes (e.g. AR(1))

set.seed(136)
n.subjects <- 100
n.traits <- 3
sigma <- 2; rho <- 0.5
Sigma0 <- diag(n.traits);
Sigma <- sigma * rho^abs(row(Sigma0) - col(Sigma0))
eS <- eigen(Sigma, symmetric = TRUE)
ev <- eS$values
X <- matrix(rnorm(n.traits * n.subjects), n.subjects)
pheno <- X %*% diag(sqrt(pmax(ev, 0)), ncol(Sigma)) %*% eS$vectors


# -- simulating genotype
geno <- rbinom(n = n.subjects, size = 2, prob = 0.2)


# -- Computing the p-value of GEEaSPU test with the permutation based method
Pvl <- GEEaSPU(pheno = pheno, geno = geno, Z = NULL, pow = c(1,2,4,Inf),
        n.perm = 1000, null.type = "perm", score.test = FALSE)


# -- Each element of Pvl is a p value of GEE-SPU(pow) in order
# -- The last element of Pvl is a p value of GEE-aSPU test
Pvl
Pvl[length(Pvl)]


# > Pvl
#     SPU.1     SPU.2     SPU.4   SPU.Inf      aSPU
# 0.1890000 0.4070000 0.3520000 0.3040000 0.2917083


# > Pvl[length(Pvl)]
#      aSPU
# 0.2917083
```

GEEaSPUpath *An adaptive sum of powered score (SPU) test for gene-set (pathway) and mutiple phenotypes*

Description

Tests whether gene-set (pathway) is associated with multivariate outcomes; provides a series of p-values of GEE-SPU(pow, pow2, pow3) and GEEaSPUpath tests.

Usage

```
GEEaSPUpath(pheno, geno, nSNPs, Z = NULL, corstr = "independence",
          pow = c(1,2,4,8), pow2 = c(1,2,4,8), pow3 = c(1,2,4,8), n.perm = 1000)
```

Arguments

pheno	a numeric phenotype matrix with each row as a different individual and each column as a separate phenotype.
geno	a numeric genotype matrix with each row as a different individual and each column as a snp; the SNPs (with the number stored in nSNPs) from one gene are stored consecutively from the first gene.
nSNPs	A numeric vector, whose length matches to the total number of genes; each element of vector indicate the number of SNPs in each gene.
Z	a numeric covariate matrix with each row as a different individual and each column as a covariated to be adjested.
corstr	a character string specifying the correlation structure of phenotypes. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured" (default = "independence").
pow	a vector of the power weight to be used at a SNP level (default = c(1,2,4,8)).
pow2	a vector of the power weight to be used at a trait level (default = c(1,2,4,8)).
pow3	a vector of the power weight to be used at a gene level (default = c(1,2,4,8)).
n.perm	a numeric value of number of null statistics (default = 1000).

Details

Adaptive association tests for gene-set (pathway) and multiple phenotypes using GEE.

Value

a vector of p-values from GEE-SPU(pow, pow2, pow3) tests and GEE-aSPUpath test.

Note

GEEaSPUpath function only supports a case for a quantitative trait (model = "gaussian") and a permutation based test (null.type = "perm").

Author(s)

Junghi Kim and Wei Pan

References

Kim J, Zhang Y, and Pan W (2016). "Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data." *Genetics*, 203(2), 715–731.

Examples

```
# -- simulating phenotypes
# -- n.subjects: number of subjects
# -- n.traits: number of phenotypes
# -- Sigma: covariance matrix of phenotypes (e.g. AR(1))

set.seed(136)
n.subjects <- 100
n.traits <- 3
sigma <- 2; rho <- 0.5
Sigma0 <- diag(n.traits)
Sigma <- sigma * rho^abs(row(Sigma0) - col(Sigma0))
eS <- eigen(Sigma, symmetric = TRUE)
ev <- eS$values
X <- matrix(rnorm(n.subjects * n.traits), n.subjects)
pheno <- X %*% diag(sqrt(pmax(ev, 0)), ncol(Sigma)) %*% eS$vectors


# -- simulating genotype
# -- Assume we have two genes each of which has 3 and 5 SNPs respectively.
# -- n.geno1: number of SNPs included in the gene1
# -- n.geno2: number of SNPs included in the gene2
# -- nSNPs <- c(3,5)

n.geno1 <- 3
n.geno2 <- 5
maf1 <- 0.2
maf2 <- 0.4
gene1 <- matrix(rbinom(n = n.subjects*n.geno1, size = 2, prob = maf1), ncol = n.geno1)
gene2 <- matrix(rbinom(n = n.subjects*n.geno2, size = 2, prob = maf2), ncol = n.geno2)
geno <- cbind(gene1, gene2)


# -- Computing the p-value of GEEaSPUpath test
Pvl <- GEEaSPUpath(pheno = pheno, geno = geno, nSNPs = c(3,5), Z = NULL,
        corstr = "independence", pow = c(1,4,8), pow2 = c(1,4,8),
        pow3 = c(1,4,8), n.perm = 1000)


# -- Each element of Pvl is a p value of GEE-SPU(pow,pow2,pow3) in order
# -- The last element of Pvl is a p value of GEE-aSPUpath test
Pvl
Pvl[length(Pvl)]


# > Pvl
# SPU.1.1.1  SPU.1.1.4  SPU.1.1.8  SPU.1.4.1  SPU.1.4.4  SPU.1.4.8  SPU.1.8.1  SPU.1.8.4
# 0.00900000 0.05600000 0.07000000 0.06200000 0.08300000 0.11200000 0.06100000 0.08200000
#  SPU.1.8.8  SPU.4.1.1  SPU.4.1.4  SPU.4.1.8  SPU.4.4.1  SPU.4.4.4  SPU.4.4.8  SPU.4.8.1
# 0.10600000 0.58100000 0.54300000 0.49200000 0.62400000 0.64000000 0.62700000 0.64900000
#  SPU.4.8.4  SPU.4.8.8  SPU.8.1.1  SPU.8.1.4  SPU.8.1.8  SPU.8.4.1  SPU.8.4.4  SPU.8.4.8
# 0.67100000 0.67500000 0.58300000 0.53700000 0.48100000 0.63400000 0.64600000 0.63800000
```

```
#  SPU.8.8.1  SPU.8.8.4  SPU.8.8.8   aSPUpath
# 0.66000000 0.68100000 0.67900000 0.04395604


# > Pvl[length(Pvl)]
#   aSPUpath
# 0.04395604
```

GEEaSPUset *An adaptive sum of powered score (SPU) test for SNP-set (gene) and mutiple phenotypes*

Description

Tests whether SNP-set (gene) is associated with multivariate outcomes; provides a series of p-values of GEE-SPU(pow, pow2) and GEEaSPUset tests.

Usage

```
GEEaSPUset(pheno, geno, Z = NULL, model = "gaussian", corstr = "independence",
          pow = c(1,2,4,8), pow2 = c(1,2,4,8), n.perm = 1000,
          null.type = "perm", score.test = FALSE)
```

Arguments

pheno	a numeric phenotype matrix with each row as a different individual and each column as a separate phenotype.
geno	a numeric genotype matrix with each row as a different individual and each column as a snp.
Z	a numeric covariate matrix with each row as a different individual and each column as a covariated to be adjested.
model	a character string specifying the model of the phenotypes. Models supported are "gaussian" for a quantitative trait and "binomial" for a binary trait (default = "gaussian").
corstr	a character string specifying the correlation structure of phenotypes. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured" (default = "independence").
pow	a vector of the power weight to be used at a SNP level (default = c(1,2,4,8)).
pow2	a vector of the power weight to be used at a trait level (default = c(1,2,4,8)).
n.perm	a numeric value of number of null statistics (default = 1000).
null.type	a character string specifying how to generate null statistics; "perm" is used when null statistics are generated using permutations and "sim" is used when null statistics are generated using simulations (default = "perm").
score.test	a logical value indicating whether to include GEEaSPU-Score test along with GEE-Score test (default = FALSE). If TRUE, it computes p-values of GEEaSPU-Score and GEE-Score as well as GEEaSPU test.

Details

Adaptive association tests for SNP-set (gene) and multiple phenotypes using GEE.

Value

a vector of p-values from GEE-SPU(pow, pow2) tests and GEE-aSPUset test.

Note

When large SNP-set (namely large gene size) or large number of phenotypes are included, the permuation based test (null.type = "perm") is recommended.

An option "binomial" model only supports the option, null.type="sim".

Author(s)

Junghi Kim and Wei Pan

References

Kim J, Zhang Y, and Pan W (2016). "Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data." *Genetics*, 203(2), 715–731.

Examples

```
# -- simulating phenotypes
# -- n.subjects: number of subjects
# -- n.traits: number of phenotypes
# -- Sigma: covariance matrix of phenotypes (e.g. AR(1))

set.seed(136)
n.subjects <- 100
n.traits <- 3
sigma <- 2; rho <- 0.5
Sigma0 <- diag(n.traits);
Sigma <- sigma * rho^abs(row(Sigma0) - col(Sigma0))
eS <- eigen(Sigma, symmetric = TRUE)
ev <- eS$values
X <- matrix(rnorm(n.subjects * n.traits), n.subjects)
pheno <- X %*% diag(sqrt(pmax(ev, 0)), ncol(Sigma)) %*% eS$vectors


# -- simulating genotype
# -- n.geno: number of SNPs included in the SNP set/gene

n.geno <- 3
maf <- 0.2
geno <- matrix(rbinom(n = n.subjects * n.geno, size = 2, prob = maf), ncol = n.geno)


# -- Computing the p-value of GEEaSPUpath test with the permutation based method
Pvl <- GEEaSPUset(pheno = pheno, geno = geno, Z = NULL, model = "gaussian",
```

```
        corstr = "independence", pow = c(1,4,Inf), pow2 = c(1,4,Inf), n.perm = 1000,
        null.type = "perm", score.test = FALSE)


# -- Each element of Pvl is a p value of GEE-SPU(pow,pow2) in order
# -- The last element of Pvl is a p value of GEE-aSPUset test
Pvl
Pvl[length(Pvl)]


# > Pvl
#     SPU.1.1     SPU.1.4   SPU.1.Inf     SPU.4.1     SPU.4.4   SPU.4.Inf   SPU.Inf.1
#  0.01400000  0.08800000  0.07200000  0.53000000  0.41000000  0.32100000  0.55100000
#   SPU.Inf.4 SPU.Inf.Inf     aSPUset
#  0.48700000  0.41000000  0.04095904


# > Pvl[length(Pvl)]
#    aSPUset
# 0.04095904
```

Index

∗ **SNP based**
GEEaSPU, 2
∗ **SNP-set/gene based**
GEEaSPUset, 7
∗ **gene-set/pathway based**
GEEaSPUpath, 4
∗ **multiple phenotypes**
GEEaSPU, 2
GEEaSPUpath, 4
GEEaSPUset, 7

GEEaSPU, 2
GEEaSPU-package, 2
GEEaSPUpath, 4
GEEaSPUset, 7