Package 'DKAT'

March 3, 2017

Type Package

Title Dual Kernel-based Association Test

Version 1.0

Depends R(>= 3.1.0)

Imports glasso,PearsonDS

Author Xiang Zhan and Michael C. Wu

Maintainer Xiang Zhan <xiangzhan9@gmail.com>

Description This package tests the overall association between high-dimensional structured traits and multiple genetic variants (both common and rare).

License GPL (>=2)

R topics documented:

DKAT *Dual Kernel-based Association Test*

Description

Test the association between genotypes and phenotypes using two kernel matrices.

Usage

```
DKAT(K, L)
```

Arguments

K	A n by n kernel matrix calculated from genotypes
L	A n by n kernel matrix calculated from phenotypes

Details

The DKAT test statistic is the normalized frobenius inner product between two centered kernel matrices, and permutations are used to evaluate the significance based on this DKAT statistic. Instead of explicitly drawing permutations and calculating the permuted DKAT statistics, we approxiamte the permutation null distribution to a Pearson type III distribution. Details of this approximation can be found in Zhan et al., (2017).

Value

A p-value

References

Zhan, X. et al. (2017). Powerful genetic association analysis for common or rare varaints with high–dimensional structured tratis. Genetics, submitted.

Examples

```
X = matrix(rbinom(200,2,0.3), nrow=200,ncol=1)  ## 200*1 SNP vector
Y = matrix(rnorm(1000,0,1),nrow=200,ncol=5) ## 200*5 normal outcomes
W.beta=c(1,1) ## Beta-weights: flat for SNPs and beta(1,25) for RV
KX=wlin.kernel(X,W.beta)
KY=pheno.kernel(Y)
pv=DKAT(KX,KY)
```

---

pheno.kernel *Phenotype Kernel*

Description

Calculates the kernel matrix for multivariate (potentially high-dimensional and structured) phenotypes

Usage

```
pheno.kernel(Y, rho = 0.1)
```

Arguments

Y	Phenotype matrix, each row is a sample and each column is a phenotype
rho	Graphical lasso regularization parameter used in estimating the precision matrix of phenotypes

Details

Let Θ be the graphical lasso estimator of the precision matrix of phenotypes. Then the phenotype kernel matrix is calculated as $K = Y\Theta Y^T$.

Value

A n by n kernel matrix, where n is the number of subjects.

References

Friedman, J. et al. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9, 432–441.
Zhan, X. et al. (2017). Powerful genetic association analysis for common or rare varaints with high–dimensional structured tratis. Genetics, submitted.

`wlin.kernel` *Weighted Linear Kernel*

Description

Calculates the weighted linear kernel matrix for genotypes

Usage

```
wlin.kernel(X, W.beta)
```

Arguments

`X`	Genotype matrix, each row is a sample and each column is a genetic variant
`W.beta`	two-dimensional weights as in the beta density function

Details

Let W=diag($w_1, \ldots, w_p$) be the diagonal matrix containing the weights of the p genetic variants, where $\sqrt{w_j} = beta(MAF_j, a_1, a_2)$, MAF_j is the minor allele frequency of variant j, and (a_1, a_2) are the weights. Then the weighted linear kernel matrix is calculated as $K = XWWX^T$.

Value

A n by n kernel matrix, where n is the number of subjects.

References

Wu, M. C. et al. (2011). Rare–variant association testing for sequencing data with the sequence kernel associaiton test. The American Journal of Human Genetics, 89, 82–93.

Index