Package 'RepeatABEL'

August 19, 2016

Type Package

Title GWAS for Multiple Observations on Related Individuals

Version 1.1

Date 2016-08-19

Author Lars Ronnegard

Maintainer Lars Ronnegard <lrn@du.se>

Description Performs genome-
wide association studies on individuals that are both related and have repeated measurements.

License GPL

Imports methods, stats

Depends R (>= 2.10), hglm, GenABEL

NeedsCompilation no

Repository CRAN

Date/Publication 2016-08-19 14:09:51

R topics documented:

RepeatABEL-package *GWAS for repeated observations on related individuals*

Description

The RepeatABEL package computes score statistic based p-values for a linear mixed model including random polygenic effects and a random effect for repeated measurements. A p-value is computed for each marker and the null hypothesis tested is that the additive marker effect is zero.

Details

Package:	RepeatABEL
Type:	Package
Version:	1.1
Date:	2016-08-19
License:	GPL
Depends:	hglm, GenABEL

The core function is rGLS that requires an GenABEL object as input and produces an GenABEL object.

Author(s)

Lars Ronnegard

Maintainer: Lars Ronnegard <lrn@du.se>

References

Ronnegard et al. (2016). **Increasing the power of genome wide association studies in natural populations using repeated measures: evaluation and implementation.**. *Methods in Ecology and Evolution*, **7**, 792-799.

Husby et al. (2015) **Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life history trait**. *Proceedings of the Royal Society B: Biological Sciences*, 282(1806), 20150156.

Ronnegard, Shen & Alam (2010). **hglm: A Package for Fitting Hierarchical Generalized Linear Models**. *The R Journal*, **2**(2), 20-28.

See Also

rGLS

compute.GRM *Computes a Genetic Relationship Matrix from a GenABEL object*

Description

Two alternative methods for GRM computations implemented.

Usage

```
compute.GRM(gen.data, method = "GenABEL")
```

Arguments

gen.data	The GenABEL object.
method	Method to be used.

Author(s)

Lars Ronnegard

constructV *Constructs the (co)variance matrix for y*

Description

Constructs the (co)variance matrix for y.

Usage

```
constructV(Z, RandC, ratio)
```

Arguments

Z	The incidence matrix for the random effects column binded with the Cholesky of the GRM
RandC	The number of columns in the two matrices combined in Z.
ratio	The ratios between random effect variances and the residual variance.

Author(s)

Lars Ronnegard

Create_gwaa_scan *Creates a scan.gwaa object*

Description

Creates a scan.gwaa object from a GenABEL object and p-values.

Usage

```
Create_gwaa_scan(data, P1df, SNP.eff)
```

Arguments

data	A gwaa.data object
P1df	P-values computed from external analysis
SNP.eff	Estimated additive SNP effects

Author(s)

Lars Ronnegard

flycatchers *10,000 SNPs from 849 collared flycatchers*

Description

A GenABEL object of class gwaa.data. Only ID and sex included in phdata(flycatchers).

Usage

```
data(flycatchers)
```

Examples

```
data(flycatchers)
summary(flycatchers)
```

gen.data *Example GenABEL data*

Description

Example GenABEL data

Usage

```
data(gen.data)
```

Examples

```
data(gen.data)
summary(gen.data)
```

Phen.Data *Example phenotype data*

Description

Data with multiple observations per individual

Usage

```
data(Phen.Data)
```

Examples

```
data(Phen.Data)
summary(Phen.Data)
```

preFitModel *Fits a linear mixed model (without fixed SNP effects) and computes the fitted variance-covariance matrix for later use in the rGLS function.*

Description

Uses a GenABEL object and phenotype data as input. The model is fitted using the `hglm` function in the hglm package.

Usage

```
preFitModel(fixed = y ~ 1, random = ~1 | id, id.name = "id", genabel.data,
  phenotype.data, corStruc = NULL, GRM = NULL, Neighbor.Matrix = NULL)
```

Arguments

fixed	A formula including the response and fixed effects
random	A formula for the random effects
id.name	The column name of the IDs in phen.data
genabel.data	An GenABEL object including marker information. This object has one observation per individual.
phenotype.data	A data frame including the repeated observations and IDs.
corStruc	A list specifying the correlation structure for each random effect. The options are: "Ind" for iid random effects, "GRM" for a correlation structure given by a genetic relationship matrix, or "CAR" for a spatial correlation structure given by a Conditional Autoregressive model specified by a neighborhood matrix.
GRM	A genetic relationship matrix. If not specified whilst the "GRM" option is given for corStruc then the GRM is computed internally within the function.
Neighbor.Matrix	A neighborhood matrix having non-zero value for an element (i,j) where the observations i and j come from neighboring locations. The diagonal elements should be zero.

Value

Returns a list including the fitted hglm object fitted.hglm, the variance-covariance matrix V and the ratios between estimated variance components for the random effects divided by the residual variance, ratio.

Author(s)

Lars Ronnegard

Examples

```
####### FIRST EXAMPLE USING GRM #############
 data(Phen.Data) #Phenotype data with repeated observations
 data(gen.data) #GenABEL object including IDs and marker genotypes
 GWAS1 <- rGLS(y ~ age + sex, genabel.data = gen.data, phenotype.data = Phen.Data)
 plot(GWAS1, main="")
 summary(GWAS1)
 #Summary for variance component estimation without SNP effects
 summary(GWAS1@call$hglm)
 #The same results can be computed using the preFitModel as follows
 fixed = y ~ age + sex
 Mod1 <- preFitModel(fixed, random=~1|id, genabel.data = gen.data,
                     phenotype.data = Phen.Data, corStruc=list( id=list("GRM","Ind") ))
 GWAS1b <- rGLS(fixed, genabel.data = gen.data,
                phenotype.data = Phen.Data, V = Mod1$V)
 plot(GWAS1b, main="Results using the preFitModel function")
 ####### SECOND EXAMPLE USING CAR #############
 # Add a fake nest variable to the data just to run the example
```

```
#In this example there are 6 nests and 60 observations per nest
Phen.Data$nest <- rep(1:6, each=60)
#A model including polygenic effects, permanent environmental effects,
#and nest effect as random
Mod2 <- preFitModel(fixed, random=~1|id + 1|nest, genabel.data = gen.data,
       phenotype.data = Phen.Data, corStruc=list( id=list("GRM","Ind"), nest=list("Ind")) )
GWAS2 <- rGLS(fixed, genabel.data = gen.data, phenotype.data = Phen.Data, V = Mod2$V)
plot(GWAS2)
#Similar to previous plot because the nest effect variance component is almost 0.
###################
#Construct a fake nighbourhood matrix
D = matrix(0,6,6)
D[1,2] = D[2,1] = 1
D[5,6] = D[6,5] = 1
D[2,4] = D[4,2] = 1
D[3,5] = D[5,3] = 1
D[1,6] = D[6,1] = 1
D[3,4] = D[4,3] = 1
#The matrix shows which pair of nests that can be considered as neighbours
image(Matrix(D), main="Neighbourhood matrix")
Mod3 <- preFitModel(fixed, random=~1|id + 1|nest, genabel.data = gen.data,
         phenotype.data = Phen.Data, corStruc=list( id=list("GRM","Ind"),
                                    nest=list("CAR")), Neighbor.Matrix=D )
GWAS2b <- rGLS(fixed, genabel.data = gen.data,
               phenotype.data = Phen.Data, V = Mod3$V)
plot(GWAS2b)
```

---

rGLS *GWAS for Studies having Repeated Measurements on Related Individuals*

Description

It is used to perform genome-wide association studies on individuals that are both related and have repeated measurements. The function computes score statistic based p-values for a linear mixed model including random polygenic effects and a random effect for repeated measurements. A p-value is computed for each marker and the null hypothesis tested is a zero additive marker effect.

Usage

```
rGLS(formula.FixedEffects = y ~ 1, genabel.data, phenotype.data,
  id.name = "id", GRM = NULL, V = NULL, memory = 1e+08)
```

Arguments

`formula.FixedEffects`	Formula including the response variable and cofactors as fixed effects.
`genabel.data`	An GenABEL object including marker information. This object has one observation per individuals.

phenotype.data	A data frame including the repeated observations and IDs.
id.name	The column name of the IDs in phen.data
GRM	An optional genetic relationship matrix (GRM) can be included as input. Otherwise the GRM is computed within the function.
V	An optional (co)variance matrix can be included as input. Otherwise it is computed using the hglm function.
memory	Used to optimize computations. The maximum number of elements in a matrix that can be stored efficiently.

Details

A generalized squares (GLS) is fitted for each marker given a (co)variance matrix V. The computations are made fast by transforming the GLS to an ordinary least-squares (OLS) problem using an eigen-decomposition of V. The OLS are computed using QR-factorization. If V is not specified then a model including random polygenic effects and permanent environmental effects is fitted (using the hglm package) to compute V. A GenABEL object (scan.gwaa class) is returned (including also the hglm results). Let e.g. GWAS1 be an object returned by the rGLS function. Then a Manhattan plot can be produced by calling plot(GWAS1) and the top SNPs using summary(GWAS1). Both of these functions are generic GenABEL functions.
The results from the fitted linear mixed model without any SNP effect included are produced by calling summary(GWAS1@call$hglm).

Author(s)

Lars Ronnegard

Examples

```
data(Phen.Data) #Phenotype data with repeated observations
 data(gen.data) #GenABEL object including IDs and marker genotypes
 GWAS1 <- rGLS(y ~ age + sex, genabel.data = gen.data, phenotype.data = Phen.Data)
 plot(GWAS1, main="")
 summary(GWAS1)
 #Summary for variance component estimation without SNP effects
 summary(GWAS1@call$hglm)
```

---

simulate_PhenData *Simulation function for the RepeatABEL package.*

Description

The function takes a GenABEL object as input and generates simulated phenotypic values for related individuals having repeated obserevations.

Usage

```
simulate_PhenData(formula.FixedEffects = y ~ 1, genabel.data, n.obs,
  SNP.eff = NULL, SNP.nr = NULL, beta = NULL, VC = c(1, 1, 1),
  GRM = NULL, sim.gamma = FALSE)
```

Arguments

formula.FixedEffects	A formula including the name of the simulated variable as response, and cofactors as fixed effects.
genabel.data	A GenABEL object of class gwaa.data.
n.obs	A vector including the number of observations per individual. The length of n.obs must be equal to the number if individuals in genabel.data.
SNP.eff	The size of a simulated SNP.effect.
SNP.nr	The SNP genotype that the SNP effect is simulated on. SNP.nr=i is the i:th SNP.
beta	The simulated fixed effects. Must be equal to the number of cofactors simulated (including the intercept term).
VC	A vector of length 3 including the simulated variances of the polygenic effect, permanent environmental effect and residuals, respectively.
GRM	An optional input where the Genetic Relationship Matrix can be given. Otherwise it is computed using the GenABEL package.
sim.gamma	A logical parameter specifying whether the residuals shuld be simulated from a gamma distribution or not. If specified as TRUE then residuals are drawn from a gamma distribution with variance equal to the residual variance specified in VC[3]

Value

Returns a data frame including the simulated phenotypic values, cofactors and IDs.

Author(s)

Lars Ronnegard

Examples

```
data(gen.data)
 #Simulate 4 observations per individual
 set.seed(1234)
 Phen.Sim <- simulate_PhenData(y ~ age, genabel.data=gen.data,
                n.obs=rep(4, nids(gen.data)), SNP.eff=1, SNP.nr=1000, VC=c(1,1,1))
 GWAS1 <- rGLS(y ~ age, genabel.data = gen.data, phenotype.data = Phen.Sim)
 plot(GWAS1, main="Simulated Data Results")
```

SmoothSNPmatrix *Imputes column means to missing genotypes*

Description

Imputes column means to missing genotypes.

Usage

```
SmoothSNPmatrix(SNP)
```

Arguments

SNP	A matrix including SNP coding.

Author(s)

Lars Ronnegard

thinned.flycatchers *10,000 SNPs from 200 collared flycatchers*

Description

A GenABEL object of class gwaa.data. Only ID and sex included in phdata(thinned.flycatchers).

Usage

```
data(thinned.flycatchers)
```

Examples

```
data(thinned.flycatchers)
summary(thinned.flycatchers)
```

Index