

GGEE Vignette

Virginie Stanislas

2021-05-27

1) Introduction

The GGEE package implements the “Gene-Gene Eigen Epistasis” approach to detect epistasis at the gene level in genome wide association studies (GWAS). This approach compute interaction variables for each gene pair then uses a penalized regression method based on group lasso to select the significant main or interaction effects.

The two main functions of this package are `BuiltEpiVar` and `GLmodel`. `BuiltEpiVar` allows to calculate interaction variables under four different interaction variable modeling approaches. The Eigen-epistasis approach find for each gene pair a component defined as the linear combination of gene markers (SNPs) having the highest correlation with the phenotype. The three other modeling approaches are inspired by previous literature proposals, they compute interaction variable using: Principal Component Analysis (PCA), Partial Least squares (PLS) or Canonical-Correlation Analysis (CCA). `GLmodel` fits a group lasso model on the genetic data set enhanced by interaction variables then uses a screen and clean procedure in order to compute p-values for each group. A group is either made with the SNPs from a given gene or of interaction terms relative to a given gene-pair interaction.

Additionally, the package allows to generate genotype and phenotype data under two phenotypic models.

2) Generating genotype and phenotype data

The GGEE package allows to generate gene structured data and associated continuous phenotype according to the model :

$$\mathbf{y} = \mathbf{X}^T \boldsymbol{\beta} + \mathbf{Z}^T \boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

Where $\mathbf{y} \in \mathbb{R}^n$ denotes the vector of trait values for n individuals, $\mathbf{X} \in \{1, 2, 3\}^{n \times p}$ represents the SNP matrix, \mathbf{Z} the matrix gathering interaction variables and $\boldsymbol{\epsilon} \in \mathbb{R}^n$ a gaussian error term. The columns of \mathbf{X} are structured on G non overlapping genes. Each gene is described by a given number of SNPs p_g where $\sum_g p_g = p$. The matrix of interaction \mathbf{Z} is structured into $G(G-1)/2$ submatrices each submatrice being the group of interaction variables for a specific pair of genes.

The two functions `simGeno` and `simPheno` allows to respectively simulate genotype and phenotype data.

```
library(GGEE)
sizeGenesMain <- rep(6,2) # 2 genes with 6 SNPs
sizeGenesPair <- rep(6,2) # 2 genes with 6 SNPs
sizeGenesRemain <- rep(6,4) # 4 genes with 6 SNPs
SameMainPair <- FALSE # Specify that genes with interaction effects will not have main effects
N<- 600
causalSNPnb <- 2
corr <- 0.8
MAFcausalSNP=0.2
```

```
Geno <- simGeno(N=N, corr=corr, sizeGenesMain=sizeGenesMain, sizeGenesPair=sizeGenesPair,
  sizeGenesRemain=sizeGenesRemain, SameMainPair=SameMainPair, MAFcausalSNP=MAFcausalSNP,
  causalSNPnb=causalSNPnb)
```

With these parameters the function `simGeno` simulate a data set of 6 genes, each one composed of 6 SNPs, for 600 individuals. The 2 first genes are considered to have main effects and the gene 3 and gene 4 to have an interaction effect. For the four causal genes their 2 first SNPs are considered as causal variants. Rather than a defined number of causal SNPs by causal gene, it is possible to use a portion of causal SNPs with the option `causalSNPportion`. In this case the option `causalSNPnb` has to be `NULL`. In both cases, the SNPs considered as causal are the first listed in the gene. The MAF of each SNP is randomly set between the values `minMAF` and `maxMaf` (by default `minMAF=0.05` and `maxMaf=0.5`). For the causal SNPs the MAF value correspond to `MAFcausalSNP`. The correlation between SNPs belonging to the same gene is set to `corr=0.8`.

The output of `Geno` contain the following elements :

- The 600×48 genotype matrix X
- The list `listGenesSNP` that indicate the names of the SNPs composing each gene
- The vectors `MainEff` and `GenePair` which give the names of genes having respectively main or interaction effects. The size of the vector `GenePair` is an even number, the pairs being defined with genes successively taken two by two along the vector.
- The vector `MAF` which give the minor allele frequency observed for each simulated SNP

```
Geno$X[1:5,1:8]
```

```
##      Gene.1.SNP.1 Gene.1.SNP.2 Gene.1.SNP.3 Gene.1.SNP.4 Gene.1.SNP.5
## [1,]           1           1           1           1           1
## [2,]           1           1           1           1           1
## [3,]           2           3           2           3           3
## [4,]           2           1           2           2           3
## [5,]           1           1           2           1           2
##      Gene.1.SNP.6 Gene.2.SNP.1 Gene.2.SNP.2
## [1,]           1           1           1
## [2,]           1           1           1
## [3,]           3           1           1
## [4,]           2           1           1
## [5,]           1           2           3
```

```
Geno$listGenesSNP
```

```
## $Genes1
## [1] "Gene.1.SNP.1" "Gene.1.SNP.2" "Gene.1.SNP.3" "Gene.1.SNP.4" "Gene.1.SNP.5"
## [6] "Gene.1.SNP.6"
##
## $Genes2
## [1] "Gene.2.SNP.1" "Gene.2.SNP.2" "Gene.2.SNP.3" "Gene.2.SNP.4" "Gene.2.SNP.5"
## [6] "Gene.2.SNP.6"
##
## $Genes3
## [1] "Gene.3.SNP.1" "Gene.3.SNP.2" "Gene.3.SNP.3" "Gene.3.SNP.4" "Gene.3.SNP.5"
## [6] "Gene.3.SNP.6"
##
## $Genes4
```

```
## [1] "Gene.4.SNP.1" "Gene.4.SNP.2" "Gene.4.SNP.3" "Gene.4.SNP.4" "Gene.4.SNP.5"
## [6] "Gene.4.SNP.6"
##
## $Genes5
## [1] "Gene.5.SNP.1" "Gene.5.SNP.2" "Gene.5.SNP.3" "Gene.5.SNP.4" "Gene.5.SNP.5"
## [6] "Gene.5.SNP.6"
##
## $Genes6
## [1] "Gene.6.SNP.1" "Gene.6.SNP.2" "Gene.6.SNP.3" "Gene.6.SNP.4" "Gene.6.SNP.5"
## [6] "Gene.6.SNP.6"
##
## $Genes7
## [1] "Gene.7.SNP.1" "Gene.7.SNP.2" "Gene.7.SNP.3" "Gene.7.SNP.4" "Gene.7.SNP.5"
## [6] "Gene.7.SNP.6"
##
## $Genes8
## [1] "Gene.8.SNP.1" "Gene.8.SNP.2" "Gene.8.SNP.3" "Gene.8.SNP.4" "Gene.8.SNP.5"
## [6] "Gene.8.SNP.6"
```

```
Geno$MainEff
```

```
## [1] "Genes1" "Genes2"
```

```
Geno$GenePair
```

```
## [1] "Genes3" "Genes4"
```

```
Geno$MAF
```

```
## Gene.1.SNP.1 Gene.1.SNP.2 Gene.1.SNP.3 Gene.1.SNP.4 Gene.1.SNP.5 Gene.1.SNP.6
## 0.20833333 0.20333333 0.23416667 0.37000000 0.49583333 0.42250000
## Gene.2.SNP.1 Gene.2.SNP.2 Gene.2.SNP.3 Gene.2.SNP.4 Gene.2.SNP.5 Gene.2.SNP.6
## 0.18750000 0.18416667 0.40083333 0.49083333 0.47333333 0.37000000
## Gene.3.SNP.1 Gene.3.SNP.2 Gene.3.SNP.3 Gene.3.SNP.4 Gene.3.SNP.5 Gene.3.SNP.6
## 0.18250000 0.18333333 0.35500000 0.44000000 0.37416667 0.17083333
## Gene.4.SNP.1 Gene.4.SNP.2 Gene.4.SNP.3 Gene.4.SNP.4 Gene.4.SNP.5 Gene.4.SNP.6
## 0.20916667 0.20583333 0.42666667 0.40666667 0.22250000 0.30500000
## Gene.5.SNP.1 Gene.5.SNP.2 Gene.5.SNP.3 Gene.5.SNP.4 Gene.5.SNP.5 Gene.5.SNP.6
## 0.19833333 0.20416667 0.14833333 0.14250000 0.07166667 0.18000000
## Gene.6.SNP.1 Gene.6.SNP.2 Gene.6.SNP.3 Gene.6.SNP.4 Gene.6.SNP.5 Gene.6.SNP.6
## 0.21166667 0.19916667 0.39500000 0.44666667 0.23916667 0.12166667
## Gene.7.SNP.1 Gene.7.SNP.2 Gene.7.SNP.3 Gene.7.SNP.4 Gene.7.SNP.5 Gene.7.SNP.6
## 0.17750000 0.18333333 0.38333333 0.21583333 0.18583333 0.30166667
## Gene.8.SNP.1 Gene.8.SNP.2 Gene.8.SNP.3 Gene.8.SNP.4 Gene.8.SNP.5 Gene.8.SNP.6
## 0.20833333 0.24166667 0.13250000 0.09583333 0.14750000 0.24083333
```

Once the genotype matrix obtained, phenotype values can be simulated through the function `simPheno`. The function takes as parameters:

- the outputs of the `simGeno` function,
- two vectors of possible values for coefficients β and γ ,

- the number or portion of causal SNPs to consider by gene (It has to be the same value than the one chosen for `simGeno`),
- a r^2 value that calibrate the difficulty of the problem,
- a value for the intercept β_0 (default $\beta_0 = 0$),
- the model to consider to simulate interaction effects:

- "SNPproduct" : $Y_i = \beta_0 + \sum_g \beta_g \left(\sum_{k \in \mathcal{C}} X_{ik}^g \right) + \sum_{rs} \gamma_{rs} \left(\sum_{(j,k) \in \mathcal{C}^2} X_{ij}^r X_{ik}^s \right) + \epsilon_i$
- "PCproduct": $Y_i = \beta_0 + \sum_g \beta_g \left(\sum_{k \in \mathcal{C}} X_{ik}^g \right) + \sum_{rs} \gamma_{rs} C_{i1}^r C_{i1}^s + \epsilon_i$

where \mathcal{C} and \mathcal{C}^2 are respectively the set of causal SNPs and causal interactions, and ϵ_i a random Gaussian variable. For each causal gene g a coefficient β_g is assigned to the standardized sum of the causal SNPs. for the interactions, in the first model "SNPproduct", all the causal SNPs from a causal pair (r, s) are pairwise multiplied and the interaction of the causal pair is represented by the standardized sum of the products. In the second model "PCproduct", the interaction is represented by the standardized product of the first PCA component $C_{.1}^r$ of gene r and the first PCA component $C_{.1}^s$ of gene s . The computation of PCA components is realized on the whole gene and not only on the causal SNPs.

```
# possible values for coef Beta or Gamma
pvBeta <- c(2,2)
pvGamma <- c(2,2)
r2 <- 0.4

Pheno <- simPheno(X=Geno$X, listGenes=Geno$listGenesSNP, MainEff=Geno$MainEff, GenePair=Geno$GenePair,
  model="SNPproduct", pvBeta=pvBeta, pvGamma=pvGamma, r2=r2, causalSNPnb=causalSNPnb)
```

The outputs of the function `simPheno` includes

- the vector of phenotype continuous values y ,
- the matrix G of the simulated main effects, each column represent one causal gene and correspond to the standardized sum of its causal SNPs,
- the matrix GG of the simulated interaction effects, each column represent one causal interaction defined depending of the selected model,
- values for the coefficient of determination R^2 when considering the model containing only simulated interaction effects R2I or only simulated main effects R2S or both simulated main and interaction effects R2T,
- a list `caract` with the characteristic of the simulation. The information about the part of the coefficient of determination R^2 hat can be attributed to either interaction effects $p_{R_I^2} = \frac{R_I^2}{R_T^2}$ or main effects

$p_{R_M^2} = \frac{R_M^2}{R_T^2}$ is given.

```
head(Pheno$y)
```

```
##           [,1]
## V1 -1.6656498
## V2 -2.6395653
## V3 10.2166795
## V4 -0.4442796
## V5  2.6853271
## V6 -0.2173680
```

```
head(Pheno$G)
```

```
##           Genes1      Genes2
## V1 -0.8177057 -0.7590355
## V2 -0.8177057 -0.7590355
## V3  2.1617887 -0.7590355
## V4  0.1754591 -0.7590355
## V5 -0.8177057  2.3043364
## V6 -0.8177057  0.2620885
```

```
head(Pheno$GG)
```

```
##           X.Genes3.Genes4
## [1,]      -0.90933006
## [2,]       0.06713175
## [3,]       2.02005537
## [4,]       1.04359356
## [5,]      -0.90933006
## [6,]       1.04359356
```

```
Pheno[c("R2T", "R2I", "R2S")]
```

```
## $R2T
## [1] 0.4208872
##
## $R2I
## [1] 0.1587528
##
## $R2S
## [1] 0.2579582
```

```
Pheno$caract
```

```
## $MainEff
## [1] "Genes1" "Genes2"
##
## $nbSNPbyMainEff
## Genes1 Genes2
##      6      6
##
## $Coef_MainEff
## Genes1 Genes2
##      2      2
##
## $causalSNPMainEff
##           Genes1      Genes2
## [1,] "Gene.1.SNP.1" "Gene.2.SNP.1"
## [2,] "Gene.1.SNP.2" "Gene.2.SNP.2"
##
## $GenePair
## [1] "Genes3" "Genes4"
```

```

##
## $nbSNPbyInterGene
## Genes3 Genes4
##      6      6
##
## $Coef_GenePair
## X.Genes3.Genes4
##      2
##
## $causalSNPInter
##      Genes3      Genes4
## [1,] "Gene.3.SNP.1" "Gene.4.SNP.1"
## [2,] "Gene.3.SNP.2" "Gene.4.SNP.2"
##
## $GeneRemain
## [1] "Genes5" "Genes6" "Genes7" "Genes8"
##
## $nbSNPparGeneRemain
## Genes5 Genes6 Genes7 Genes8
##      6      6      6      6
##
## $beta0
## [1] 0
##
## $r2
## [1] 0.4
##
## $causalSNPportion
## NULL
##
## $causalSNPnb
## [1] 2
##
## $R2T
## [1] 0.4208872
##
## $PartR2I
## [1] 37.71862
##
## $PartR2S
## [1] 61.28915

```

3) The G-GEE method

Once genotype and phenotype data are obtained we can apply the G-GEE approach to seek for interaction effects. The first step is to create interaction variables from each gene couple, the second is to test for potential main or interaction effects.

Interaction variable modeling can be done with the function `BuiltEpiVar`. The function takes as parameters the matrix of genotype X , the vector of phenotypic traits y , a list `listGenesSNP` that indicate the names of the SNPs composing each gene and `nbcomp` the number of components to consider to compute interaction variables.

Subjects need to be randomly divided into two equal-sized sets for training and testing. The training group

is used to construct the interaction variables and to estimate the Group LASSO coefficients. The test group is used for the cleaning step to compute the permuted p-values. `idSubs` is a list containing the indices of the samples used in the test or training set.

Four different methods can be use to create interaction variables :

- "GGEE" which find for each gene pair its Eigen-epistasis Component that maximize the correlation between all possible SNP-SNP interactions and the phenotype.
- "PCA" which first compute PCA on each gene of the pair and represent the interaction with component products.
- "PLS" interaction variables are defined by components that maximize the covariance between the two genes and the phenotype.
- "CCA" interaction variables are here represent by the product of pairwise components obtained by a canonical correlation analysis on the gene pair.

Here we show an example using "GGEE"option. As this method can compute only one interaction by gene couple, the parameter `nbcomp` doesn't need to be used.

```
# Distribution of the samples in tre training or test set:
n <- dim(Geno$X)[1]
portionTrain <- 1/2
idTrain <- sample(1:n, size=n*portionTrain)
idTest <- (1:n)[-idTrain]
idSubs <- list(idTrain=idTrain[order(idTrain)], idTest=idTest)

Int <- BuiltEpiVar(Geno$X, Pheno$y, method="GGEE", listGenesSNP=Geno$listGenesSNP, idSubs=idSubs)

## [1] "Genes1.Genes2"
## [1] "Genes1.Genes3"
## [1] "Genes1.Genes4"
## [1] "Genes1.Genes5"
## [1] "Genes1.Genes6"
## [1] "Genes1.Genes7"
## [1] "Genes1.Genes8"
## [1] "Genes2.Genes3"
## [1] "Genes2.Genes4"
## [1] "Genes2.Genes5"
## [1] "Genes2.Genes6"
## [1] "Genes2.Genes7"
## [1] "Genes2.Genes8"
## [1] "Genes3.Genes4"
## [1] "Genes3.Genes5"
## [1] "Genes3.Genes6"
## [1] "Genes3.Genes7"
## [1] "Genes3.Genes8"
## [1] "Genes4.Genes5"
## [1] "Genes4.Genes6"
## [1] "Genes4.Genes7"
## [1] "Genes4.Genes8"
## [1] "Genes5.Genes6"
## [1] "Genes5.Genes7"
## [1] "Genes5.Genes8"
## [1] "Genes6.Genes7"
```

```
## [1] "Genes6.Genes8"
## [1] "Genes7.Genes8"
```

Int is a list composed of the interaction variable matrices (XIntTrain and XIntTest) and a vector interLength indicating the number of interaction variables for each couple.

```
head(Int$XIntTrain)
```

```
##      X.Genes1.Genes2 X.Genes1.Genes3 X.Genes1.Genes4 X.Genes1.Genes5
## 1          2.842293          16.009877          11.962926          22.1486831
## 2          3.649989          -4.824618          -5.068192          -3.1323427
## 3         -2.145704           1.020637          -0.409683           0.2010701
## 4         -6.515337          -4.886101          -6.358666          -5.3675740
## 5          7.235486          -3.489032          -3.687091           0.1634352
## 6          1.960990           4.699701          -3.870751          -0.3492141
##      X.Genes1.Genes6 X.Genes1.Genes7 X.Genes1.Genes8 X.Genes2.Genes3
## 1          0.8654665           3.1407445           2.4291335           2.4192310
## 2         -3.1483067           1.8447779          -2.3620184           0.8851905
## 3         -1.2591928           0.3231216          -0.1688638          -5.1913841
## 4         -5.1197632          -1.5956425          -5.5191339          -5.1913841
## 5         -3.2577827           2.5525862           0.2432952           0.9557768
## 6          1.3691033           0.7041937          -1.7489043           8.0810049
##      X.Genes2.Genes4 X.Genes2.Genes5 X.Genes2.Genes6 X.Genes2.Genes7
## 1          0.03056739           5.477477           -5.043486           -4.023663
## 2          0.79986076           5.504423           4.655974           15.540585
## 3         -5.95229634          -5.905356          -6.228144          -5.563685
## 4         -6.47497487          -5.905356          -5.601562          -2.008244
## 5          0.79250966           7.303516           1.063624           11.277896
## 6         -2.42071700           2.012531           4.534757           3.689319
##      X.Genes2.Genes8 X.Genes3.Genes4 X.Genes3.Genes5 X.Genes3.Genes6
## 1         -4.4845405           10.3954101           20.614199           0.2825474
## 2          7.1574825          -5.9531312          -4.418589          -4.3387627
## 3         -5.9442970          -4.0687253          -3.726548          -4.5345531
## 4         -5.9442970          -4.7478626          -3.726548          -3.7350364
## 5          7.3464409          -5.9531312          -3.760987          -5.6319004
## 6          0.4533293          -0.5136557           4.512969           7.2704230
##      X.Genes3.Genes7 X.Genes3.Genes8 X.Genes4.Genes5 X.Genes4.Genes6
## 1          2.3953317           1.761148           14.718975           -1.182310
## 2         -0.4999212          -3.738319          -4.451534          -3.542525
## 3         -3.5760429          -4.040693          -4.626487          -4.800166
## 4          1.0159135          -4.040693          -5.157345          -4.160784
## 5         -2.0834250          -3.633204          -3.850348          -4.915501
## 6          6.3094786           2.851426          -3.850348          -1.680015
##      X.Genes4.Genes7 X.Genes4.Genes8 X.Genes5.Genes6 X.Genes5.Genes7
## 1          0.2600623          -0.4782762          -0.7270401           5.62225511
## 2         -0.4580027          -3.6271521           0.3101075           2.92007544
## 3         -4.4734857          -4.7305312           0.3425674          -3.96345680
## 4         -1.5178362          -5.1993127           0.9104001           0.23895542
## 5         -2.0761992          -3.5229152          -0.7529312           2.01593939
## 6         -2.7667721          -4.3918918           0.3625704          -0.03549921
##      X.Genes5.Genes8 X.Genes6.Genes7 X.Genes6.Genes8 X.Genes7.Genes8
## 1          3.0578733           3.5276436           4.7186843          -0.5803333
## 2         -1.2888318          -1.0053611           2.2040843           1.6820634
```



```
## 3      -3.9839280      3.2122174      4.7186843      -1.3172557
## 4      -3.9839280     -0.4220853      4.2078917      -0.8321881
## 5      -0.7252913      0.9732950      3.0669132      0.2281915
## 6      -1.2949293     -2.4057272      0.8212111     -0.5021014
```

```
head(Int$XIntTest)
```

```
##      X.Genes1.Genes2 X.Genes1.Genes3 X.Genes1.Genes4 X.Genes1.Genes5
## 1      -5.967755     -4.8861007     -6.3586661     -3.0236035
## 2      -5.962644     -4.2726247     -2.0325187     -2.2002319
## 3      -2.950755      0.4747500      6.6571172      0.3993538
## 4      -4.504261     -2.6974830     -0.2193723      0.4249796
## 5      4.510817      -1.3363338      4.8741445      0.3621475
## 6      -2.145704     -0.2325953     -0.4096830      1.6700855
##      X.Genes1.Genes6 X.Genes1.Genes7 X.Genes1.Genes8 X.Genes2.Genes3
## 1      -5.741741     -5.1184257     -4.753821      -4.434644
## 2      -5.741741     -5.7344716     -1.475204      -3.757405
## 3      -2.292610      7.5016579      -1.431697      -4.904203
## 4      1.728724      -5.3165109     -0.568457      -1.601254
## 5      9.028464      0.5441405      6.812913      -3.157553
## 6      -1.259193      8.1436229      4.989518      -5.766669
##      X.Genes2.Genes4 X.Genes2.Genes5 X.Genes2.Genes6 X.Genes2.Genes7
## 1      -5.904534     -2.458145      -5.633046      -4.891196
## 2      -1.288516     -1.477724      -5.731449      -5.611457
## 3      -1.889357     -5.211278      -6.228144      -1.024858
## 4      1.404841      2.443721      4.455977      -4.677552
## 5      1.604198      -1.790546      5.471716      -1.700332
## 6      -5.952296     -5.183294      -6.228144      -1.839552
##      X.Genes2.Genes8 X.Genes3.Genes4 X.Genes3.Genes5 X.Genes3.Genes6
## 1      -4.3755176     -4.747863      -0.6119675     -4.5345531
## 2      -0.7762809      1.926996      1.6563074      -3.9251579
## 3      -5.9442970      2.097092      -2.1637133     -3.9606444
## 4      1.2894559      4.616694      5.6765483      7.6377268
## 5      3.7837141      -3.097914      -5.1337924     -0.7904484
## 6      -3.4597977     -4.748777      -3.7362665     -5.0909634
##      X.Genes3.Genes7 X.Genes3.Genes8 X.Genes4.Genes5 X.Genes4.Genes6
## 1      -3.5760429     -3.084745      -2.821439      -4.915501
## 2      -3.7591752      2.642804      6.105233      -1.182310
## 3      3.4111224      -3.365747      1.816067      -1.066975
## 4      -2.9798408      4.103091      10.963911      12.665098
## 5      -4.9583802     -2.095290      -1.624091      5.139696
## 6      0.0504172      -1.810711      -3.903418      -4.800166
##      X.Genes4.Genes7 X.Genes4.Genes8 X.Genes5.Genes6 X.Genes5.Genes7
## 1      -4.9200052     -4.434573055    0.07737444     -1.4884510
## 2      -1.1760677      7.225843441    -1.67684776     -1.8085284
## 3      10.0262270     -0.009494709    -0.03778843      3.0358276
## 4      -0.3905274      9.306756150    -4.48621713     -0.1570287
## 5      -1.6530716      3.242242678    0.10598561     -3.9634568
## 6      -0.3762300     -2.243723824    -0.37257538      1.4281253
##      X.Genes5.Genes8 X.Genes6.Genes7 X.Genes6.Genes8 X.Genes7.Genes8
## 1      -0.8426261      3.2122174      4.0023104     -1.30089979
## 2      5.2407419      3.8075636      0.9161479      0.03298621
## 3      -3.4748149      0.2505608      4.7186843      0.53742432
## 4      9.9665618      -2.3862890     -12.7917958     0.41567985
```

```
## 5      -1.3624014      -1.0307395      -8.7925148      -2.40249051
## 6      -0.5729821       0.3790278       2.1433518       1.00887721
```

```
Int$interLength
```

```
## X.Genes1.Genes2 X.Genes1.Genes3 X.Genes1.Genes4 X.Genes1.Genes5 X.Genes1.Genes6
##              1              1              1              1              1
## X.Genes1.Genes7 X.Genes1.Genes8 X.Genes2.Genes3 X.Genes2.Genes4 X.Genes2.Genes5
##              1              1              1              1              1
## X.Genes2.Genes6 X.Genes2.Genes7 X.Genes2.Genes8 X.Genes3.Genes4 X.Genes3.Genes5
##              1              1              1              1              1
## X.Genes3.Genes6 X.Genes3.Genes7 X.Genes3.Genes8 X.Genes4.Genes5 X.Genes4.Genes6
##              1              1              1              1              1
## X.Genes4.Genes7 X.Genes4.Genes8 X.Genes5.Genes6 X.Genes5.Genes7 X.Genes5.Genes8
##              1              1              1              1              1
## X.Genes6.Genes7 X.Genes6.Genes8 X.Genes7.Genes8
##              1              1              1
```

Test for potential main or interaction effects is done with the function `GLmodel`. Parameters include `nlambda`, the length of the grid of possible lambda values, `limitLambda` the number of the largest lambda values among the grid to consider for the cross validation and `lambda.cri` the criteria for lambda selection (minimum or oneSE value).

```
res <- GLmodel(Geno$X, Pheno$y, XIntTrain=Int$XIntTrain, XIntTest=Int$XIntTest, idSubs =idSubs,
               interLength=Int$interLength, listGenesSNP=Geno$listGenesSNP, nlambda=100,
               limitLambda=25, lambda.cri="min")
```

```
##
## LinReg()
## Lambda = 1700.427 1347.559 1067.917 846.3063 670.6833 531.505 421.2086 333.8006 264.5313 209.6365 160.0000
```

The outputs of `GLmodel` contain:

- the matrix `res_GL.min` giving for each SNP an interaction variable the group lasso coefficient values at the optimal lambda level,
- the matrix `pval.adj` that give adjusted pvalues of each variable with nonzero group lasso coefficient.

```
res
```

```
## $res_GL.min
##              Coefs
## Genes1      0.46931689
## Genes1      0.60012778
## Genes1      0.10274726
## Genes1      0.11641116
## Genes1     -0.09214164
## Genes1      0.02705877
## Genes2      0.12778085
## Genes2      0.16455911
## Genes2      0.07588303
## Genes2      0.03102738
```

## Genes2	0.03922373
## Genes2	0.02100725
## Genes3	0.00000000
## Genes3	0.00000000
## Genes3	0.00000000
## Genes3	0.00000000
## Genes3	0.00000000
## Genes3	0.00000000
## Genes3	0.00000000
## Genes4	0.46346578
## Genes4	0.31336447
## Genes4	-0.09349572
## Genes4	-0.43056254
## Genes4	-0.17383212
## Genes4	0.26891543
## Genes5	0.00000000
## Genes5	0.00000000
## Genes5	0.00000000
## Genes5	0.00000000
## Genes5	0.00000000
## Genes5	0.00000000
## Genes5	0.00000000
## Genes6	0.00000000
## Genes6	0.00000000
## Genes6	0.00000000
## Genes6	0.00000000
## Genes6	0.00000000
## Genes6	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes7	0.00000000
## Genes8	0.00000000
## Genes8	0.00000000
## Genes8	0.00000000
## Genes8	0.00000000
## Genes8	0.00000000
## Genes8	0.00000000
## X.Genes1.Genes2	0.83384999
## X.Genes1.Genes3	0.26799883
## X.Genes1.Genes4	0.00000000
## X.Genes1.Genes5	0.00000000
## X.Genes1.Genes6	0.00000000
## X.Genes1.Genes7	0.00000000
## X.Genes1.Genes8	0.00000000
## X.Genes2.Genes3	0.65729551
## X.Genes2.Genes4	0.13813063
## X.Genes2.Genes5	0.61851544
## X.Genes2.Genes6	0.00000000
## X.Genes2.Genes7	0.00000000
## X.Genes2.Genes8	0.00000000
## X.Genes3.Genes4	1.37070486
## X.Genes3.Genes5	0.00000000
## X.Genes3.Genes6	0.00000000

```

## X.Genes3.Genes7 0.00000000
## X.Genes3.Genes8 0.00000000
## X.Genes4.Genes5 0.00000000
## X.Genes4.Genes6 0.00000000
## X.Genes4.Genes7 0.06318480
## X.Genes4.Genes8 0.00000000
## X.Genes5.Genes6 0.00000000
## X.Genes5.Genes7 0.00000000
## X.Genes5.Genes8 0.00000000
## X.Genes6.Genes7 0.00000000
## X.Genes6.Genes8 0.24888980
## X.Genes7.Genes8 0.34509050
##
## $pval.adj
##          pval.adj
## Genes1      0.0840
## Genes2      0.0480
## Genes4      0.0480
## X.Genes1.Genes2 0.6696
## X.Genes1.Genes3 0.6696
## X.Genes2.Genes3 1.0000
## X.Genes2.Genes4 0.6696
## X.Genes2.Genes5 0.6696
## X.Genes3.Genes4 0.0720
## X.Genes4.Genes7 0.6696
## X.Genes6.Genes8 0.6696
## X.Genes7.Genes8 0.7440
##
## $vc
## $vc$cv.error
##          [,1]      [,2]      [,3]
## [1,] 30.75648 33.03971 35.32294
## [2,] 29.14004 31.29819 33.45634
## [3,] 26.50541 28.46265 30.41990
## [4,] 23.89746 25.60569 27.31391
## [5,] 21.67059 23.23029 24.78999
## [6,] 20.82074 22.30426 23.78778
## [7,] 20.14538 21.60362 23.06185
## [8,] 19.65181 21.11676 22.58170
## [9,] 19.42867 20.84001 22.25134
## [10,] 19.17885 20.58852 21.99820
## [11,] 18.87906 20.30049 21.72191
## [12,] 18.92215 20.36571 21.80927
## [13,] 18.10613 19.52432 20.94251
## [14,] 18.44898 19.92265 21.39633
## [15,] 18.66172 20.13256 21.60341
## [16,] 18.80640 20.31178 21.81715
## [17,] 18.64620 20.10557 21.56495
## [18,] 19.47557 21.05050 22.62542
## [19,] 19.09278 20.60736 22.12194
## [20,] 21.63301 23.43910 25.24519
## [21,] 22.01281 23.79843 25.58405
## [22,] 21.52700 23.23733 24.94766
## [23,] 21.07266 22.79796 24.52325

```

```
## [24,] 22.09436 23.90944 25.72453
## [25,] 21.59251 23.40295 25.21338
##
## $vc$lambda
## [1] 1700.426576 1347.558982 1067.917447 846.306314 670.683281 531.505030
## [7] 421.208647 333.800649 264.531306 209.636536 166.133370 131.657854
## [13] 104.336598 82.684970 65.526424 51.928571 41.152504 32.612656
## [19] 25.844972 20.481698 16.231396 12.863103 10.193790 8.078404
## [25] 6.401997
##
## $vc$lambda.min
## [1] 104.3366
##
## $vc$lambda.oneSE
## [1] 264.5313
##
## $vc$id.lambda.min
## [1] 13
##
##
## $id_varSign
## [1] 1 2 4 9 10 16 17 18 22 29 35 36
##
## attr("class")
## [1] "GLmodel"
```

The `GGEE` package contains a plot function `plot.GLmodel`. The function takes as parameter a `GLmodel` object and provides a representation of cross validation results. It depicts the value of the cross validation error for each lambda considered and thus allows to identify the optimal lambda values depending of the criteria of interest (minimal or oneSE). This plot allows to verify that enough lambda values was considered for the cross validation. If the curve doesn't show a clear minimal value the parameter `limitLambda` of the `GLmodel` has to be enlarged.

```
plotGLmodel(res)
```

