# Package 'iQTL'

August 3, 2010

**Type** Package

**Title** Genome-wide QTL Mapping via H-likelihood

**Version** 2010.1

**Date** 2010-07-26

**Author** Xia Shen

**Maintainer** Xia Shen <xia.shen@lcb.uu.se>

**Description** This package provides an R implementation of the double hierarchical generalized linear model (DHGLM) particularly for analyzing genome-wide marker data.

**License** Unlimited

**LazyLoad** yes

**Depends** VGAM

## R topics documented:



---

`iQTL-package` *Genome-wide QTL Mapping via H-likelihood*

---

### Description

This package provides an R implementation of the double hierarchical generalized linear model (DHGLM) particularly for analyzing genome-wide marker data.

### Details

| | |
|---|---|
| Package: | iQTL |
| Type: | Package |
| Version: | 2010.1 |
| Date: | 2010-07-26 |
| License: | Unlimited |
| LazyLoad: | yes |
| Depends: | VGAM |

## Author(s)

Xia Shen

Maintainer: Xia Shen <xia.shen@lcb.uu.se>

## References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

## See Also

h.GWAS, h.testQTL

## Examples

```
## Not run:

## data

data(snpid)
data(pt)
data(gt)
data(pedi)
data(info)
marker.id <- info[snpid,1]

## set parameters

rho <- .9

## create spatial correlation matrix
## note that there are 5 chromosomes

border.idx <- c(0, 1971, 4032, 6080, 8068, 10029)
nmarker.chr <- numeric(5)
for (i in 1:5) nmarker.chr[i] <- sum((marker.id > border.idx[i])*
(marker.id <= border.idx[i + 1]))
ac.mat <- matrix(0, dim(gt)[2], dim(gt)[2])
blockborder <- c(0, cumsum(nmarker.chr))
for (i in 1:5) ac.mat[(blockborder[i] + 1):blockborder[i + 1],
(blockborder[i] + 1):blockborder[i + 1]] <-
rho**(toeplitz(1:nmarker.chr[i]) - 1)
```

```
## fixed effect design matrix

X <- model.matrix(~factor(pedi[,4]))
dimnames(X) <- dimnames(pt) <- dimnames(gt) <- c(NULL, NULL)

## analyze QT

GWAS <- h.GWAS(y = pt[,2], X = X, Z = gt, rho = rho, ac.mat = ac.mat,
phi.start = 2, plotting = FALSE)

par(mfrow = c(2, 1))
plot(marker.id, GWAS$v, xlab = 'Marker index', ylab = 'Marker effect',
type = 'p', cex = .6, col = 4, bty = 'n')
abline(v = border.idx, lty = 2, col = 'green4')
plot(marker.id, GWAS$lambda, xlab = 'Marker index',
ylab = 'Marker dispersion', type = 'l', col = 2, bty = 'n')
abline(v = border.idx, lty = 2, col = 'green4')


## End(Not run)
```

---

`gt` *Genotypes of The Example Data*

---

## Description

The genotypes of the example data with in the coded format.

## Usage

```
data(gt)
```

## Format

The matrix is 2326 (individuals) by 1628 (SNPs).

## Details

The SNP markers are selected from the original dataset simulated for the 14th QTLMAS workshop, Poznan, Poland, 2010. See 'Source' and 'References'.

## Source

http://jay.up.poznan.pl/qtlmas2010/dataset.html

## References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

h.GWAS *Genome-wide Association via H-likelihood*

## Description

This is an R implementation of the double hierarchical generalized linear model (DHGLM) particularly for analyzing genome-wide marker data.

## Usage

```
h.GWAS(y, X, Z, family = 'normal', link = 'logit', X.disp = NULL,
       ac.mat = NULL, rho = 0, phi.start = 1, lambda.start = 0.1,
       alpha.start = 1, conv.crit = 1e-05, max.iter = 200, plotting = TRUE)
```

## Arguments

| | |
|---|---|
| y | a numeric vector of the phenotypes. |
| X | a design matrix of the fixed effects. |
| Z | an incidence matrix of the markers across the genome. Each element in the matrix gives the genotype of the corresponding individual at the corresponding marker. See 'Details' for more about the coding. |
| family | a string indicating the distribution family of the phenotypes. 'normal' and 'binary' is available. See 'Details' for more information. |
| link | a string giving the link function for binary phenotypes. 'logit','probit', and 'cloglog' are implemented and only avalable when family = 'binary'. |
| X.disp | a design matrix of the fixed effects of the dispersion part of the model, i.e. for modeling the marker-specific variance. |
| ac.mat | an autocorrelation matrix for the markers, which is used for smoothing the marker specific variances. If NULL, an spatial correlation matrix is created with the parameter rho. See 'Details' for more information. |
| rho | a numeric value in [0, 1) defining the spatial correlation for two adjacent markers. Only activated when ac.mat = NULL. See Details for more information. |
| phi.start | a numeric value giving the IWLS starting value of the dispersion parameter of the phenotypes. When family = 'binary', this parameter is always fixed to be 1. |
| lambda.start | a numeric value giving the IWLS starting value of the dispersion parameter of the marker effects. |
| alpha.start | a numeric value giving the IWLS starting value of the dispersion parameter of the marker-specific variances. See 'References'. |
| conv.crit | a numeric value giving the convergence tolerance of the IWLS algorithm. |
| max.iter | an integer restricting the maximum number of IWLS interations. |
| plotting | a logic value specifying whether a figure will be produced after each IWLS iteration. The figures record the convergence behaviors of marker effects and marker-specific variances and will be saved in a single PDF file. |

## Details

Typically, for an intercross, at a particular marker, there are three kinds of genotypes: AA, AB, and BB. Coding of `Z` can be simply set to be the number of 'A' copies, i.e. 2 for 'AA', 1 for 'AB', and 0 for 'BB'. For a backcross, 1 and 0 can be used for coding heterozygotes and homozygotes, respectively.

When `family = 'binary'`, a link function is applied for linearizing the binary phenotypes to a continuous scale. The algorithm is not easy to converge depending on the starting values. `plotting = TRUE` is recommended to observe results for a certain number of iterations.

It is recommended that `ac.mat` is created according to the linkage disequilibium for each pair of markers across the genome. If `ac.mat = NULL`, the algorithm will create a spatial correlation matrix using argument `rho`. By default, `rho = 0` indicating no autocorrelation for the marker-specific variances. See 'References'.

## Value

| | |
|---|---|
| `phi` | the dispersion parameter of the phenotypes. When `family = 'binary'`, this parameter is always fixed to be 1. |
| `alpha` | the dispersion parameter of the marker-specific variances. See 'References'. |
| `lambda` | the marker-specific variances. |
| `beta` | the estimated fixed effects of the phenotypes. |
| `gamma` | the estimated fixed effects of the marker-specific variances. |
| `v` | the marker effects. |
| `b` | the estimated random effects of marker-specific variances. |
| `niter` | the number of iterations taken in the IWLS algorithm. |

## Author(s)

Xia Shen

## References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

## See Also

`iQTL-package`

## Examples

```
## Not run:

## data

data(snpid)
data(pt)
data(gt)
data(pedi)
data(info)
marker.id <- info[snpid,1]
```

```
## set parameters

rho <- .9

## create spatial correlation matrix
## note that there are 5 chromosomes

border.idx <- c(0, 1971, 4032, 6080, 8068, 10029)
nmarker.chr <- numeric(5)
for (i in 1:5) nmarker.chr[i] <- sum((marker.id > border.idx[i])*
(marker.id <= border.idx[i + 1]))
ac.mat <- matrix(0, dim(gt)[2], dim(gt)[2])
blockborder <- c(0, cumsum(nmarker.chr))
for (i in 1:5) ac.mat[(blockborder[i] + 1):blockborder[i + 1],
(blockborder[i] + 1):blockborder[i + 1]] <-
rho**(toeplitz(1:nmarker.chr[i]) - 1)

## fixed effect design matrix

X <- model.matrix(~factor(pedi[,4]))
dimnames(X) <- dimnames(pt) <- dimnames(gt) <- c(NULL, NULL)

## analyze QT

GWAS <- h.GWAS(y = pt[,2], X = X, Z = gt, rho = rho, ac.mat = ac.mat,
phi.start = 2, plotting = FALSE)

par(mfrow = c(2, 1))
plot(marker.id, GWAS$v, xlab = 'Marker index', ylab = 'Marker effect',
type = 'p', cex = .6, col = 4, bty = 'n')
abline(v = border.idx, lty = 2, col = 'green4')
plot(marker.id, GWAS$lambda, xlab = 'Marker index',
ylab = 'Marker dispersion', type = 'l', col = 2, bty = 'n')
abline(v = border.idx, lty = 2, col = 'green4')


## End(Not run)
```

---

`h.testQTL` *Score test for QTL existence*

---

## Description

This function performs QTL existence test via a score statistic.

## Usage

```
h.testQTL(y, Z, sig.level = .05)
```

## Arguments

| | |
|---|---|
| `y` | a numeric vector of the phenotypes. |

| | |
|---|---|
| Z | an incidence matrix of the markers across the genome. Each element in the matrix gives the genotype of the corresponding individual at the corresponding marker. See 'Details' for more about the coding. |
| sig.level | a numeric value between 0 and 1 specifying the significance level of QTL existence test. |

## Details

Typically, for an intercross, at a particular marker, there are three kinds of genotypes: AA, AB, and BB. Coding of Z can be simply set to be the number of 'A' copies, i.e. 2 for 'AA', 1 for 'AB', and 0 for 'BB'. For a backcross, 1 and 0 can be used for coding heterozygotes and homozygotes, respectively.

## Value

Screen print-out giving the score test statistic value and the p-value.

## Author(s)

Xia Shen

## See Also

iQTL-package

## Examples

```
## Not run:

## data

data(pt)
data(gt)

## test QTL existence using the last 10 F3 individuals

h.testQTL(pt[2317:2326,2], gt[2317:2326,]) # p-value = 0 < 0.05

## change the phenotypes and re-test

set.seed(123)
h.testQTL(rnorm(10), gt[2317:2326,]) # p-value = 0.7383826 > 0.05


## End(Not run)
```

---

# info *Marker Information of The Example Data*

---

## Description

The marker information of the example data.

### Usage

```
data(info)
```

### Format

Column 1: marker ID. Column 2: chromosome. Column 3: chromosomal position in base pairs.

### Details

The markers are from the original dataset simulated for the 14th QTLMAS workshop, Poznan, Poland, 2010. See 'Source' and 'References'.

### Source

http://jay.up.poznan.pl/qtlmas2010/dataset.html

### References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

---

## pedi *The Pedigree Structure of The Example Data*

---

### Description

The pedigree information of the example data.

### Usage

```
data(pedi)
```

### Format

Column 1: individual ID. Column 2: sire ID. Column 3: dam ID. Column 4: sex.

### Details

The pedigree is from the original dataset simulated for the 14th QTLMAS workshop, Poznan, Poland, 2010. See 'Source' and 'References'.

### Source

http://jay.up.poznan.pl/qtlmas2010/dataset.html

### References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

## pt *Phenotypes of The Example Data*

### Description

The phenotypes of the example data.

### Usage

```
data(pt)
```

### Format

Column 1: individual ID. Column 2: the simulated quantitative trait. Column 3: the simulated binary trait.

### Details

The phenotypes are from the original dataset simulated for the 14th QTLMAS workshop, Poznan, Poland, 2010. See 'Source' and 'References'.

### Source

http://jay.up.poznan.pl/qtlmas2010/dataset.html

### References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

## snpid *Indices of The Selected SNPs of The Example Data*

### Description

The indices of the selected SNPs of the example data.

### Usage

```
data(snpid)
```

### Format

a numeric vector containing SNP indices.

### Details

Every 6 SNPs were selected from the original dataset simulated for the 14th QTLMAS workshop, Poznan, Poland, 2010. See 'Source' and 'References'.

## Source

http://jay.up.poznan.pl/qtlmas2010/dataset.html

## References

Xia Shen, Lars Ronnegard and Orjan Carlborg (2010): Hierarchical likelihood opens a new way of estimating genetic values using dense marker maps. (14th QTLMAS, Poznan, Poland) SUBMITTED.

# Index