Package 'sbl'

July 23, 2025

Type Package

Title Sparse Bayesian Learning for QTL Mapping and Genome-Wide
Association Studies

Version 0.1.0

Author Meiyue Wang and Shizhong Xu

Maintainer Meiyue Wang <mwang024@ucr.edu>

Description Implements sparse Bayesian learning method for QTL mapping and genome-
wide association studies.

License GPL-3

Encoding UTF-8

LazyData true

Depends stats, R (>= 2.10)

RoxygenNote 6.1.0

Date 2018-10-17

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2018-10-26 15:50:03 UTC

Contents

---

gen   *Genotype*

Description

A matrix with dimension of 30*50 containing the numeric genotype indicator for 30 simulated individuals of an F2 family generated from the cross of two inbred lines according to `map` provided in the package.

Usage

gen

Format

A matrix with 30 rows and 50 columns:

ID individual ID

marker name of marker

---

intercept   *Intercept*

Description

Design vector of the intercept for the 30 simulated individuals.

Usage

intercept

Format

A sequence of number containing 30 replicates of number 1.

map 3

---

map   *Genetic map*

Description

The genetic map used in the generation of numeric genotype indicators for 30 simulated individuals of an F2 family generated from the cross of two inbred lines.

Usage

```
map
```

Format

A data frame with 50 rows and 4 variables:

chr chromosome

marker name of marker

cm genetic distance measured in centi-Morgan

effect simulated effect assigned to marker in the simulation experiment

---

phe   *Phenotype*

Description

Observations of the phenotype for 30 simulated individuals.

Usage

```
phe
```

Format

A sequence of number with length equal to 30.

---

sblgwas	*Sparse Bayesian Learning for QTL Mapping and Genome-Wide Association Studies*

Description

The sparse Bayesian learning (SBL) method for quantitative trait locus (QTL) mapping and genome-wide association studies (GWAS) deals with a linear mixed model. This is also a multiple locus model that includes all markers (random effects) in a single model and detect significant markers simultaneously. SBL method adopts coordinate descent algorithm to update parameters by estimating one parameter at a time conditional on the posterior modes of all other parameters. The parameter estimation process requires multiple iterations and the final estimated parameters take the values when the entire program converges.

Usage

```
sblgwas(x, y, z, t = -1, max.iter = 200, min.err = 1e-06)
```

Arguments

x	a design matrix for fixed effects
y	a vector of response variables
z	a design matrix for random effects
t	a number between [-2,0] to control sparseness of the model, default is -1.
max.iter	maximum number of iterations set by user to stop the program, default is 200.
min.err	minimum threshold of mean squared error of random effects estimated from the current and the previous iteration to stop the program, default is 1e-6.

Details

The multiple locus hierarchical linear mixed model of SBL is

$$y = X\beta + Z\gamma + \epsilon$$

where y is an $n * 1$ vector of response variables (observations of the trait); X is an $n * p$ design matrix for fixed effects; β is a $p * 1$ vector of fixed effect; Z is an $n * m$ genotype indicator matrix; γ is an $m * 1$ vector of marker effects and ϵ is an $n * 1$ vector of residual errors with an aassumed $\epsilon \ N(0, \Sigma)$ distribution. Each marker effect, $\gamma[k]$ for marker k, is treated as a random variable following $N(0, \Phi[k])$ distribution, where $\Phi[k]$ is the prior variance. The estimate of $\gamma[k]$ is best linear unbiased prediciton (BLUP). The estimate of $\Phi[k]$ is maximum likelihood estimate (MLE).

Value

iteration	a matrix storing intermediate results of each iteration before the entire program converges, including

`"iter"` iteration indicator

`"error"` mean squared error of random effects estimated from the current and the previous iteration

`"s2"` estimated variance of residual error

"$beta[1] \ldots beta[p]$" estimates of fixed effects

"$gamma[1] \ldots gamma[m]$" estimates of random effects

parm   a vector containing 5 elements: `"iter"`, `"error"`, `"s2"`, `"beta"` and `"df"`

`"iter"` the number of iterations required by program to stop

`"error"` mean square error of random effect estimated from the last iteration before the program stops

`"s2"` estimated variance of residual error

`"beta"` estimate of fixed effect

`"df"` the effective degree of freedom from total random effects

blup   a matrix containing 4 columns: `"gamma"`, `"vg"`, `"wald"` and `"p_wald"`

`"gamma"` estimate of random effect

`"vg"` estimated variance of random effect

`"wald"` Wald statistic calculated as γ^2/Φ

`"p_wald"` the *p*-value of Wald statistic following Chi-squared distribution with 1 degree of freedom

Author(s)

Meiyue Wang and Shizhong Xu
Maintainer: Meiyue Wang <mwang024@ucr.edu>

Examples

```
# Load example data from sbl package
data(gen)
data(phe)
data(intercept)

# Run sblgwas() to perform association study of example data
# setting t = 0 leads to the most sparse model
fit<-sblgwas(x=intercept, y=phe, z=gen, t=0)
```

```
my.blup<-fit$blup

# setting t = -2 leads to the least sparse model
fit<-sblgwas(x=intercept, y=phe, z=gen, t=-2)
my.blup<-fit$blup
```

Index