Package 'EstHer'

October 12, 2022

Type Package

Title Estimation of Heritability in High Dimensional Sparse Linear Mixed Models using Variable Selection

Version 1.0

Date 2015-07-16

Author Anna Bonnet and Celine Levy-Leduc

Maintainer Anna Bonnet `<anna.bonnet@agroparistech.fr>`

Description Our method is a variable selection method to select active components in sparse linear mixed models in order to estimate the heritability. The selection allows us to reduce the size of the data sets which improves the accuracy of the estimations. Our package also provides a confidence interval for the estimated heritability.

License GPL-2

Depends R (>= 2.10), glmnet, parallel, MASS

Imports Rcpp (>= 0.11.5)

LinkingTo Rcpp, RcppArmadillo

NeedsCompilation yes

Repository CRAN

Date/Publication 2015-07-23 16:32:59

R topics documented:

`EstHer-package` *Estimation of Heritability in High Dimensional Sparse Linear Mixed Models using Variable Selection*

Description

Our method is a variable selection method whicg can be used to select active components in sparse linear mixed models in order to estimate the heritability. The selection allows us to reduce the size of the data sets which improves the accuracy of the estimations. Our package also provides a confidence interval for the estimated heritability.

Details

Package:	EstHer
Type:	Package
Version:	1.0
Date:	2015-07-16
License:	GPL-2

~~ An overview of how to use the package, including the most important functions ~~

Author(s)

Anna Bonnet and Celine Levy-Leduc

Maintainer: Anna Bonnet <anna.bonnet@agroparistech.fr> ~~ The author and/or maintainer of the package ~~

References

A.Bonnet, C.Levy-Leduc, E.Gassiat, R.Toro, T.Bourgeron. "Improving heritability estimation by a variable selection approach in sparse high dimensional linear mixed models", arXiv:1507.06245, 2015.

`bootstrap_corr` *Computation of a confidence interval for the estimated heritability.*

Description

This function is based on a non parametric bootstrap technique to compute a confidence interval for the heritability. The strength of this method is that it can deal with correlated observations.

Usage

```
bootstrap_corr(Y,Z,K,eta_hat,sigma2_hat,level,nb_cores)
```

Arguments

`Y`	Vector of observations.
`Z`	Matrix of genetic informations.
`K`	number of subsamples of Y used to apply our bootstrap technique.
`eta_hat`	Estimator of the heritability.
`sigma2_hat`	Estimator of the variance involving the variances of the two random parts of the model.
`level`	Percentage of values which will be removed from the estimated heritabilities to build a confidence interval.
`nb_cores`	Number of cores of the computer. It is used for parallelizing the computations.

Value

`CI_up`	Upper bound of the confidence interval for the estimated heritability
`CI_low`	Lower bound of the confidence interval for the estimated heritability

Author(s)

Anna Bonnet and Celine Levy-Leduc

`estim_herit` *Estimation of heritability in linear mixed models.*

Description

The function computes heritability when the random effects follow either a Gaussian distribution or a mixture of a Dirac and a Gaussian distribution.

Usage

```
estim_herit(Y, Z)
```

Arguments

`Y`	vector of observations of size n
`Z`	matrix with n rows and N columns

Value

`heritability`	Heritability

Author(s)

Anna Bonnet

References

The method is developped in the paper "Heritability estimation in high dimensional linear mixed models" (A.Bonnet,vE.Gassiat, C.Levy-Leduc,2014)

`prod_cpp` *Additional function*

Description

This function is used for efficiently computing ZZ'.

Usage

```
prod_cpp(Z)
```

Arguments

`Z`	A large matrix having n rows and N columns.

Value

The product ZZ'.

Author(s)

Anna Bonnet et Celine Levy-Leduc

`Selvar` *Estimation of heritability in high dimensional sparse linear mixed models using variable selection.*

Description

This function selects active components in sparse linear mixed models in order to estimate heritability. The selection allows us to reduce the size of the data sets which improves the accuracy of the estimations. Our package also provides a confidence interval for the estimated heritability.

Usage

```
Selvar(Y,Z,X,thresh_vect,nb_boot=80,nb_repli=50,CI_level=0.95,nb_cores=1)
```

Arguments

`Y`	Vector of observations of size n.
`Z`	Matrix with genetic information of size n x N.
`X`	Matrix of fixed effects of size n x d.
`thresh_vect`	Vector of thresholds in the stability selection step: the higher the threshold, the smallest the set of selected components.
`nb_boot`	Number of subsamples of Y to apply our bootstrap technique. The value by default is 80.
`nb_repli`	Number of replications in the stability selection. The value by default is 50.
`CI_level`	Level of the confidence interval for the estimation of the heritability. The value by default is 0.95.
`nb_cores`	Number of cores of the computer. It is used for parallelizing the computations. The value by default is 1.

Value

`heritability`	Estimation of the heritability
`CI_up`	Upper bound of the confidence interval for the estimated heritability
`CI_low`	Lower bound of the confidence interval for the estimated heritability
`selec_ind`	Indexes of the columns of the selected components

Author(s)

Anna Bonnet and Celine Levy-Leduc

Examples

```
library(EstHer)
data(Y)
data(W)
data(X)
Z=scale(W,center=TRUE,scale=TRUE)
res=Selvar(Y,Z,X,thresh_vect=c(0.7,0.8,0.9),nb_boot=80,nb_repli=50,CI_level=0.95,nb_cores=1)
res$heritability
res$CI_low
res$CI_up
```

`W` *Genetic matrix*

Description

Matrix of size n x N with n=200 and N=5000

Usage

```
data("W")
```

Format

The format is: num [1:200, 1:5000] 0 0 1 0 0 0 0 0 1 0 ...

Examples

```
data(W)
## maybe str(W) ; plot(W) ...
```

`X` *Matrix of fixed effects*

Description

Matrix of fixed effects of size n x d where n=200 and d=2

Usage

```
data("X")
```

Format

The format is: num [1:200, 1:2] 1 1 1 1 1 1 1 1 1 1 ...

Examples

```
data(X)
## maybe str(X) ; plot(X) ...
```

`Y` *Vector of observations*

Description

Vector of observations of size n=200.

Usage

```
data("Y")
```

Format

The format is: num [1:200, 1] 6.196076 7.225162 9.662264 3.134030 5.404111 ...

Examples

```
data(Y)
## maybe str(Y) ; plot(Y) ...
```

Index