Package ‘ridge’

February 27, 2017

Title  Ridge Regression with Automatic Selection of the Penalty Parameter

Description  Linear and logistic ridge regression functions. Additionally includes special functions for genome-wide single-nucleotide polymorphism (SNP) data.

Version  2.2

Date  2017-02-19

Author  Steffen Moritz, Erika Cule

Maintainer  Steffen Moritz <steffen.moritz10@gmail.com>

Type  Package

BugReports  https://github.com/steffenmoritz/ridge/issues

URL  http://github.com/steffenmoritz/ridge

Repository  CRAN

Depends  R (>= 3.0.1)

Imports  stats, graphics, grDevices, utils

License  GPL-2

SystemRequirements  Gnu Scientific Library version >= 1.14

NeedsCompilation  yes

RoxygenNote  6.0.1

Date/Publication  2017-02-27 18:25:53

R topics documented:

ridge-package .......................................................... 2
GenBin ................................................................. 2
GenCont ............................................................... 3
Gorman ................................................................. 4
Hald ..................................................................... 4
linearRidge ........................................................... 5
linearRidgeGenotypes .................................................. 8
linearRidgeGenotypesPredict ........................................ 10
Description

R package for fitting linear and logistic ridge regression models.

Details

This package contains functions for fitting linear and logistic ridge regression models, including functions for fitting linear and logistic ridge regression models for genome-wide SNP data supplied as file names when the data are too big to read into R.

For a complete list of functions, use help(package="ridge").

Author(s)

Steffen Moritz, Erika Cule

GenBin

Simulated genetic data with a binary phenotypes

Description

Simulated genetic data at 15 SNPs, together with simulated binary phenotypes

Usage

data(GenBin)

Format

GenBin is a saved R matrix with 500 rows and 15 columns. The first column is the phenotype and columns 2-15 contain the genotypes. Each row represents an individual. The same data are stored in flat text files in GenBin_genotypes and GenBin_phenotypes (in the directory extdata (in the installed package) or inst/extdata (in the source)).

Source

Simulated using FREGENE
References


Examples

data(GenBin)

---

GenCont

Simulated genetic data with continuous outcomes

Description

Simulated genetic data with continuous outcomes.

Usage

data(GenCont)

Format

GenCont is a saved R matrix with 500 rows and 13 columns. The first column is the phenotypes and columns 2-13 contain the genotypes. Each row represents an individual. The same data are stored in flat text files in GenCont_genotypes and GenCont_phenotypes (in the directory extdata (in the installed package) or inst/extdata (in the source)).

Details

Genotypes were simulated using FREGENE.

References


Examples

data(GenCont)
The Ten-Factor data first described by Gorman and Toman (1966).

Description
A Ten-Factor data set first described by Gorman and Toman (1966) and used by Hoerl and Kennard (1970) (and others) to investigate regression problems.

Usage
data(Gorman)

Format
Numeric matrix.

Details
The first column is the response on the log scale, the remaining columns are the predictors.

Source

References

Examples
data(Gorman)

Hald data

Description
The Hald data as used by Hoerl, Kennard and Baldwin (1975). These data are also in package wle.

Usage
data(Hald)
**Format**

Numeric matrix.

**Details**

The first column is the response and the remaining four columns are the predictors.

**References**


**Examples**

```r
data(Hald)
```

---

**Description**

Fits a linear ridge regression model. Optionally, the ridge regression parameter is chosen automatically using the method proposed by Cule et al (2012).

**Usage**

```r
linearRidge(formula, data, lambda = "automatic", nPCs = NULL, scaling = c("corrForm", "scale", "none"), ...)
```

```r
## S3 method for class 'ridgeLinear'
coef(object, all.coef = FALSE, ...)
```

```r
## S3 method for class 'ridgeLinear'
plot(x, y = NULL, ...)
```

```r
## S3 method for class 'ridgeLinear'
predict(object, newdata, na.action = na.pass, all.coef = FALSE, ...)
```

```r
## S3 method for class 'ridgeLinear'
print(x, all.coef = FALSE, ...)
```

```r
## S3 method for class 'ridgeLinear'
summary(object, all.coef = FALSE, ...)
```

```r
## S3 method for class 'summary.ridgeLinear'
print(x, digits = max(3, getOption("digits") - 3),
      signif.stars = getOption("show.signif.stars"), ...)
```
Arguments

- **formula**: a formula expression as for regression models, of the form `response ~ predictors`. See the documentation of `formula` for other details.
- **data**: an optional data frame in which to interpret the variables occurring in `formula`.
- **lambda**: A ridge regression parameter. May be a vector. If `lambda` is "automatic" (the default), then the ridge parameter is chosen automatically using the method of Cule et al (2012).
- **npcs**: The number of principal components to use to choose the ridge regression parameter, following the method of Cule et al (2012). It is not possible to specify both `lambda` and `npcs`.
- **scaling**: The method to be used to scale the predictors. One of "corrform" (the default) scales the predictors to correlation form, such that the correlation matrix has unit diagonal. "scale" Standardizes the predictors to have mean zero and unit variance. "none" No scaling.
- **object**: A ridgeLinear object, typically generated by a call to `linearRidge`.
- **newdata**: An optional data frame in which to look for variables with which to predict. If omitted, the fitted values are used.
- **na.action**: function determining what should be done with missing values in `newdata`. The default is to predict NA.
- **all.coef**: Logical. Should results be returned for all ridge regression penalty parameters (`all.coef = TRUE`), or only for the ridge parameter chosen automatically using the method of Cule et al?
- **x**: An object of class `ridgeLinear` (for the `print.ridgeLinear` and `plot.ridgeLinear` functions) or an object of class `summary.ridgeLinear` (for the `print.summary.ridgeLinear` function)
- **y**: Dummy argument for compatibility with the default plot method. Ignored.
- **digits**: minimum number of significant digits to be used for most numbers
- **signif.stars**: logical; if `TRUE`, P-values are additionally encoded visually as significance stars in order to help scanning of long coefficient tables. It defaults to the `show.signif.stars` slot of options.
- **...**: Additional arguments to be passed to or from other methods.

Details

If an intercept is present in the model, its coefficient is not penalised. If you want to penalise an intercept, put in your own constant term and remove the intercept.

Value

An object of class "ridgeLinear", with components:

- **automatic**: Logical. Was `lambda` chosen automatically?
- **call**: The matched call.
- **coef**: A named vector of fitted coefficients.
df

A vector of degrees of freedom of the model fit, degrees of freedom for variance, and residual degrees of freedom of the fitted model.

Inter

Was an intercept included?

isScaled

Were the predictors scaled before the model was fitted?

lambda

The ridge regression parameter(s).

scales

The scales used to standardize the predictors.

terms

The terms object used.

x

The scaled predictor matrix.

xm

A vector of means of the predictors.

y

The response.

ym

The mean of the response.

And optionally the components

max.nPCs

The maximum number of principal components for which a ridge regression parameter was computed.

chosen.nPCs

The number of principal components used to compute the ridge parameter.

Author(s)

Erika Cule

References


See Also

logisticRidge

Examples

data(GenCont)
mod <- linearRidge(Phenotypes ~ ., data = as.data.frame(GenCont))
summary(mod)
linearRidgeGenotypes  Fits linear ridge regression models for genome-wide SNP data.

Description

Fits linear ridge regression models for genome-wide SNP data. The SNP genotypes are not read into R but file names are passed the code directly, enabling the analysis of genome-wide scale SNP data sets.

Usage

linearRidgeGenotypes(genotypesfilename, phenotypesfilename, lambda = -1, thinfilename = NULL, betafile = NULL, approxfilename = NULL, permfilename = NULL, intercept = TRUE, verbose = FALSE)

Arguments

genotypesfilename  character string: path to file containing SNP genotypes coded 0, 1, 2. See Input file formats.

phenotypesfilename  character string: path to file containing phenotypes. See Input file formats.

lambda  (optional) shrinkage parameter. If not provided, the default denotes automatic choice of the shrinkage parameter using the method of Cule & De Iorio (2012).

thinfilename  (optional) character string: path to file containing three columns: SNP name, chromosome and SNP position. See Input file formats. (See details.)

betafile  (optional) character string: path to file where the output will be written. See Output file formats.

approxfilename  (optional) character string: path to fine where the approximate test p-values will be written. Approximate p-values are not computed unless this argument is given. Approximate p-values are computed using the method of Cule et al (2011). See Output file formats.

permfilename  (optional) character string: path to file where the permutation test p-values will be written. Permutation test p-values are not computed unless this argument is given. (See warning). See Output file formats.

intercept  Logical: Should the ridge regression model be fitted with an intercept? (Defaults to TRUE)

verbose  Logical: If TRUE, additional information is printed to the R output as the code runs. Defaults to FALSE.

Details

If a file thin is supplied, and the shrinkage parameter lambda is being computed automatically based on the data, then this file is used to thin the SNP data by SNP position. If this file is not supplied, SNPs are thinned automatically based on number of SNPs.
Value

The vector of fitted ridge regression coefficients. If betafilename is given, the fitted coefficients are written to this file as well as being returned. If approxfilename and/or permfilename are given, results of approximate test p-values and/or permutation test p-values are written to the files given in their arguments.

Input file formats

**genotypesfilename**: A header row, plus one row for each individual, one SNP per column. The header row contains SNP names. SNPs are coded as 0, 1, 2 for minor allele count. Missing values are not accommodated. Invariant SNPs in the data cause an error, please remove these from the file before calling the function.

**phenofilename**: A single column of phenotypes with the individuals in the same order as those in the file genotypesfilename.

**thin**: (optional) Three columns and the same number of rows as there are SNPs in the file genotypesfilename, one row per SNP. First column: SNP names (must match names in genotypesfilename); second column: chromosome; third column: SNP position in BP.

Output file formats

All output files are optional. Whether or not betafilename is provided, fitted coefficients are returned to the R workspace. If betafilename is provided, fitted coefficients are written to the file specified (in addition).

**betafilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is fitted coefficients. If intercept = TRUE (the default) then the first row is the fitted intercept (with the name Intercept in the first column).

**approxfilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is approximate p-values.

**permfilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is permutation p-values.

Warning

When data are large, the permutation test p-values may take a very long time to compute. It is recommended not to request permutation test p-values (using the argument permfilename) when data are large.

Author(s)

Erika Cule

References

See Also

linearRidge for fitting linear ridge regression models when the data are small enough to be read into R. logisticRidge and logisticRidgeGenotypes for fitting logistic ridge regression models.

Examples

```r
## Not run:
genotypesfile <- system.file("extdata", "GenCont_genotypes.txt", package = "ridge")
phenotypesfile <- system.file("extdata", "GenCont_phenotypes.txt", package = "ridge")
beta_linearRidgeGenotypes <- linearRidgeGenotypes(genotypesfilename = genotypesfile,
                                               phenotypesfilename = phenotypesfile)
# compare to output of linearRidge
data(GenCont)  ## Same data as in GenCont_genotypes.txt and GenCont_phenotypes.txt
beta_linearRidge <- linearRidge(Phenotypes ~ ., data = as.data.frame(GenCont))
cbind(round(coef(beta_linearRidge), 6), beta_linearRidgeGenotypes)

## End(Not run)
```

linearRidgeGenotypesPredict

Predict phenotypes from genome-wide SNP data based on a file of coefficients

Description

Predict phenotypes from genome-wide SNP data based on a file of coefficients. Genotypes and fitted coefficients are provided as filenames, allowing the computation of fitted probabilities when SNP data are too large to be read into R.

Usage

```r
linearRidgeGenotypesPredict(genotypesfilename, betafilename, phenotypesfilename = NULL,
                            verbose = FALSE)
```

Arguments

- `genotypesfilename`: character string: path to file containing SNP genotypes coded 0, 1, 2. See Input file formats.
- `betafilename`: character string: path to file containing fitted coefficients. See Input file formats.
- `phenotypesfilename`: (optional) character string: path to file in which to write out the predicted phenotypes. See Output file formats. Whether or not this argument is supplied, the fitted coefficients are also returned by the function.
- `verbose`: Logical: If TRUE, additional information is printed to the R output as the code runs. Defaults to FALSE.
Value

A vector of fitted values, the same length as the number of individuals whose data are in genotypesfilename. If phenotypesfilename is supplied, the fitted values are also written there.

Input file formats

**genotypesfilename**: A header row, plus one row for each individual, one SNP per column. The header row contains SNP names. SNPs are coded as 0, 1, 2 for minor allele count. Missing values are not accommodated.

**betafilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is fitted coefficients. If the coefficients include an intercept then the first row of betafilename should contain it with the name Intercept in the first column. An Intercept thus labelled will be used appropriately in predicting the phenotypes. SNP names must match those in genotypesfilename. The format of betafilename is that of the output of linearRidgeGenotypes, meaning linearRidgeGenotypesPredict can be used to predict using coefficients fitted using linearRidgeGenotypes (see the example).

Output file format

Whether or not phenotypesfilename is provided, predicted phenotypes are returned to the R workspace. If phenotypesfilename is provided, predicted phenotypes are written to the file specified (in addition).

**phenotypesfilename**: One column, containing predicted phenotypes, one individual per row.

Author(s)

Erika Cule

References


See Also

linearRidgeGenotypes for model fitting, logisticRidgeGenotypes and logisticRidgeGenotypesPredict for corresponding functions to fit and predict on SNP data with binary outcomes.

Examples

```r
# Not run:
genotypesfile <- system.file("extdata","GenCont_genotypes.txt",package = "ridge")
phenotypesfile <- system.file("extdata","GenCont_phenotypes.txt",package = "ridge")
betafile <- tempfile(pattern = "beta", fileext = ".dat")
beta_linearRidgeGenotypes <- linearRidgeGenotypes(genotypesfilename = genotypesfile,
                                           phenotypesfilename = phenotypesfile,
                                           betafilename = betafile)
pred_phen_geno <- linearRidgeGenotypesPredict(genotypesfilename = genotypesfile,
                                              phenotypesfilename = phenotypesfile,
                                              betafilename = betafile)
```
logisticRidge

Logistic ridge regression.

Description

Fits a logistic ridge regression model. Optionally, the ridge regression parameter is chosen automatically using the method proposed by Cule et al (2012).

Usage

logisticRidge(formula, data, lambda = "automatic", npcs = NULL, scaling = c("corrForm", "scale", "none"), ...)

## S3 method for class 'ridgeLogistic'
coef(object, all.coef = FALSE, ...)

## S3 method for class 'ridgeLogistic'
plot(x, y = NULL, ...)

## S3 method for class 'ridgeLogistic'
predict(object, newdata = NULL, type = c("link", "response"),
  na.action = na.pass, all.coef = FALSE, ...)

## S3 method for class 'ridgeLogistic'
print(x, all.coef = FALSE, ...)

## S3 method for class 'ridgeLogistic'
summary(object, all.coef = FALSE, ...)

## S3 method for class 'summary.ridgeLogistic'
print(x, digits = max(3,getOption("digits") - 3),
  signif.stars = getOption("show.signif.stars"), ...)

Arguments

formula a formula expression as for regression models, of the form response ~ predictors. See the documentation of formula for other details.
data  an optional data frame in which to interpret the variables occurring in formula.

lambda  A ridge regression parameter. If lambda is "automatic" (the default), then the
         ridge parameter is chosen automatically using the method of Cule et al (2012).

nPCs  The number of principal components to use to choose the ridge regression pa-
         rameter, following the method of Cule et al (2012). It is not possible to specify
         both lambda and nPCs.

scaling  The method to be used to scale the predictors. One of "corrform" (the default)
         scales the predictors to correlation form, such that the correlation matrix has
         unit diagonal. "scale" Standardizes the predictors to have mean zero and unit
         variance. "none" No scaling.

object  A ridgeLogistic object, typically generated by a call to linearRidge.

newdata  An optional data frame in which to look for variables with which to predict. If
         omitted, the fitted values are used.

type  the type of prediction required. The default predictions are of log-odds (probabil-
         ities on logit scale) and type = "response" gives the predicted probabilities.

na.action  function determining what should be done with missing values in newdata. The
         default is to predict NA.

all.coef  Logical. Should results be returned for all ridge regression penalty parameters
         (all.coef = TRUE), or only for the ridge parameter chosen automatically using
         the method of Cule et al?

x  An object of class ridgeLogistic (for the print.ridgeLogistic and plot.ridgeLogistic
function) or an object of class summary.ridgeLogistic (for the print.summary.ridgeLogistic
function)

y  Dummy argument for compatibility with the default plot method. Ignored.

digits  minimum number of significant digits to be used for most numbers

signif.stars  logical; if TRUE, P-values are additionally encoded visually as significance stars
         in order to help scanning of long coefficient tables. It defaults to the show.signif.stars
         slot of options.

...  Additional arguments to be passed to or from other methods.

Details

If an intercept is present in the model, its coefficient is not penalised. If you want to penalise an
intercept, put in your own constant term and remove the intercept.

Value

An object of class "ridgeLogistic", with components:

automatic  Was lambda chosen automatically?

call  The matched call.

coef  A named vector of fitted coefficients.

df  A vector of degrees of freedom of the model fit and degrees of freedom for
     variance.
logisticRidgeGenotypes

Inter    Was in antercept included?
isScaled Were the predictors scaled before the model was fitted?
lambda   The ridge regression parameter.
scales   The scales used to standardize the predictors.
terms    The terms object used.
x        The scaled predictor matrix.
xm       A vector of means of the predictors.
y        The response.

And optionally the component

nPCs    The number of principal components used to compute the ridge regression parameter.

Author(s)

Erika Cule

References


See Also

linearRidge

Examples

data(GenBin)
mod <- logisticRidge(Phenotypes ~ ., data = as.data.frame(GenBin))
summary(mod)

logisticRidgeGenotypes

Fits logistic ridge regression models for genome-wide SNP data.

Description

Fits logistic ridge regression models for genome-wide SNP data. The SNP genotypes are not read into R but file names are passed to the code directly, enabling the analysis of genome-wide SNP data sets which are too big to be read into R.
Usage

logisticRidgeGenotypes(genotypesfilename, phenotypesfilename, lambda = -1, thinfilename = NULL, betafilename = NULL, approxfilename = NULL, permfilename = NULL, intercept = TRUE, verbose = FALSE)

Arguments

genotypesfilename
character string: path to file containing SNP genotypes coded 0, 1, 2. See Input file formats.

phenotypesfilename
character string: path to file containing phenotypes. See Input file formats.

lambda
(optional) shrinkage parameter. If not provided, the default denotes automatic choice of the shrinkage parameter using the method of Cule & De Iorio (2012).

thinfilename
(optional) character string: path to file containing three columns: SNP name, chromosome and SNP position. See Input file formats. (See details.)

betafilename
(optional) character string: path to file where the output will be written. See Output file formats.

approxfilename
(optional) character string: path to fine where the approximate test p-values will be written. Approximate p-values are not computed unless this argument is given. Approximate p-values are computed using the method of Cule et al (2011). See Output file formats.

permfilename
(optional) character string: path to file where the permutation test p-values will be written. Permutation test p-values are not computed unless this argument is given. (See warning). See Output file formats.

intercept
Logical: Should the ridge regression model be fitted with an intercept? Defaults to TRUE.

verbose
Logical: If TRUE, additional information is printed to the R output as the code runs. Defaults to FALSE.

Details

If a file thin is supplied, and the shrinkage parameter lambda is being computed automatically based on the data, then this file is used to thin the SNP data by SNP position. If this file is not supplied, SNPs are thinned automatically based on number of SNPs.

Value

The vector of fitted ridge regression coefficients. If betafilename is given, the fitted coefficients are written to this file as well as being returned. If approxfilename and/or permfilename are given, results of approximate test p-values and/or permutation test p-values are written to the files given in their arguments.
Input file formats

**genotypesfilename**: A header row, plus one row for each individual, one SNP per column. The header row contains SNP names. SNPs are coded as 0, 1, 2 for minor allele count. Missing values are not accommodated. Invariant SNPs in the data cause an error, please remove these from the file before calling the function.

**phenofilename**: A single column of phenotypes with the individuals in the same order as those in the file genotypesfilename. Phenotypes must be coded as 0 or 1.

**thin**: (optional) Three columns and the same number of rows as there are SNPs in the file genotypesfilename, one row per SNP. First column: SNP names (must match names in genotypesfilename); second column: chromosome; third column: SNP position in BP.

Output file formats

All output files are optional. Whether or not betafilename is provided, fitted coefficients are returned to the R workspace. If betafilename is provided, fitted coefficients are written to the file specified (in addition).

**betafilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is fitted coefficients. If intercept = TRUE (the default) then the first row is the fitted intercept (with the name Intercept in the first column).

**approxfilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is approximate p-values.

**permfilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is permutation p-values.

Warning

When data are large, the permutation test p-values may take a very long time to compute. It is recommended not to request permutation test p-values (using the argument permfilename) when data are large.

Author(s)

Erika Cule

References


See Also

logisticRidge for fitting logistic ridge regression models when the data are small enough to be read into R. linearRidge and linearRidgeGenotypes for fitting linear ridge regression models.
Examples

```r
### Not run:
genotypesfile <- system.file("extdata","GenBin_genotypes.txt",package = "ridge")
phenotypesfile <- system.file("extdata","GenBin_phenotypes.txt",package = "ridge")
beta_logisticRidgeGenotypes <-
logisticRidgeGenotypesPredict(genotypesfilename = genotypesfile, phenotypesfilename = phenotypesfile)
### compare to output of logisticRidge
data(GenBin) ### Same data as in GenBin_genotypes.txt and GenBin_phenotypes.txt
beta_logisticRidge <- logisticRidge(Phenotypes ~ ., data = as.data.frame(GenBin))
cbind(round(coef(beta_logisticRidge), 6), beta_logisticRidgeGenotypes)
### End(Not run)
```

logisticRidgeGenotypesPredict

_Predict fitted probabilities from genome-wide SNP data based on a file of coefficients_

Description

Predict fitted probabilities from genome-wide SNP data based on a file of coefficients. Genotypes and fitted coefficients are provided as filenames, allowing the computation of fitted probabilities when SNP data are too large to be read into R.

Usage

```r
logisticRidgeGenotypesPredict(genotypesfilename, betafilename, phenotypesfilename = NULL, verbose = FALSE)
```

Arguments

- **genotypesfilename**
  - character string: path to file containing SNP genotypes coded 0, 1, 2. See Input file formats.

- **betafilename**
  - character string: path to file containing fitted coefficients. See Input file formats.

- **phenotypesfilename**
  - (optional) character string: path to file in which to write out the fitted probabilities. See Output file formats. Whether or not this argument is supplied, the fitted coefficients are also returned by the function.

- **verbose**
  - Logical: If TRUE, additional information is printed to the R output as the code runs. Defaults to FALSE.

Value

A vector of fitted probabilities, the same length as the number of individuals whose data are in genotypesfilename. If phenotypesfilename is supplied, the fitted probabilities are also written there.
Input file formats

**genotypesfilename**: A header row, plus one row for each individual, one SNP per column. The header row contains SNP names. SNPs are coded as 0, 1, 2 for minor allele count. Missing values are not accommodated.

**betafilename**: Two columns: First column is SNP names in same order as in genotypesfilename, second column is fitted coefficients. If the coefficients include an intercept then the first row of betafilename should contain it with the name Intercept in the first column. An Intercept thus labelled will be used appropriately in predicting the phenotypes. SNP names must match those in genotypesfilename. The format of betafilename is that of the output of linearRidgeGenotypes, meaning linearRidgeGenotypesPredict can be used to predict using coefficients fitted using linearRidgeGenotypes (see the example).

Output file format

Whether or not phenotypesfilename is provided, fitted probabilities are returned to the R workspace. If phenotypesfilename is provided, fitted probabilities are written to the file specified (in addition).

**phenotypesfilename**: One column, containing fitted probabilities, one individual per row.

Author(s)

Erika Cule

References


See Also

logisticRidgeGenotypes for model fitting. linearRidgeGenotypes and linearRidgeGenotypesPredict for corresponding functions to fit and predict on SNP data with continuous outcomes.

Examples

```r
## Not run:
genotypesfile <- system.file("extdata","GenBin_genotypes.txt",package = "ridge")
phenotypesfile <- system.file("extdata","GenBin_phenotypes.txt",package = "ridge")
betafile <- tempfile(pattern = "beta", fileext = ".dat")
beta_logisticRidgeGenotypes <- logisticRidgeGenotypes(genotypesfilename = genotypesfile,
    phenotypesfilename = phenotypesfile,
    betafilename = betafile)
pred_phen_geno <- logisticRidgeGenotypesPredict(genotypesfilename = genotypesfile,
    phenotypesfilename = phenotypesfile,
    betafilename = betafile)
## compare to output of logisticRidge
data(GenBin) ## Same data as in GenBin_genotypes.txt and GenBin_phenotypes.txt
beta_logisticRidge <- logisticRidge(Phenotypes ~ ., data = as.data.frame(GenBin))
pred_phen <- predict(beta_logisticRidge, type="response")
print(cbind(pred_phen_geno, pred_phen))
```
## Description

Functions for computing, printing and plotting p-values for ridgeLinear and ridgeLogistic models. The p-values are computed using the significance test of Cule et al (2011).

## Usage

```r
pvals(x, 
## S3 method for class 'ridgeLinear'
pvals(x, 
## S3 method for class 'ridgeLogistic'
pvals(x, 
## S3 method for class 'pvalsRidgeLinear'
print(x, digits = max(3, getOption("digits") - 3),
signif.stars = getOption("show.signif.stars"), all.coef = FALSE, ...)

## S3 method for class 'pvalsRidgeLogistic'
print(x, digits = max(3, getOption("digits") - 3),
signif.stars = getOption("show.signif.stars"), all.coef = FALSE, ...)

## S3 method for class 'pvalsRidgeLinear'
plot(x, y = NULL, ...)

## S3 method for class 'pvalsRidgeLogistic'
plot(x, y = NULL, ...)
```

## Arguments

- `x` For the `pvals` methods, an object of class "ridgeLinear" or "ridgeLogistic", typically from a call to "linearRidge" or "logisticRidge". For the print and plot methods, an object of class "pvalsRidgeLinear" or "pvalsRidgeLogistic", typically from a call to "pvals".
- `digits` minimum number of significant digits to be used for most numbers
- `signif.stars` logical; if TRUE, P-values are additionally encoded visually as significance stars in order to help scanning of long coefficient tables. It defaults to the show.signif.stars slot of options.
all.coef Logical. Should p-values for all the ridge regression parameters be printed, or only the one from the ridge parameter chosen using the method of Cule et al (2012)
y Dummy argument for compatibility with the default plot method. Ignored.

Details

Standard errors, test statistics and p-values are computed using coefficients and data on the scale that was used to fit them. If the coefficients were standardized before the model was fitted, then the p-values relate to the scaled data.

Value

For the pvals methods, an object of class "pvalsRidgeLinear" or "pvalsRidgeLogistic" which is a list with elements

coef The (scaled) regression coefficients
se The standard errors of the regression coefficients
tstat The test statistic of the regression coefficients
pval The p-values of the regression coefficients
isScaled Were the data scaled before the regression coefficients were fitted?

For the print methods, the argument x is returned invisibly.

Author(s)

Erika Cule

References


See Also

linearRidge, logisticRidge

Examples

data(GenBin)
mod <- logisticRidge(Phenotypes ~ ., data = as.data.frame(GenBin))
pvalsMod <- pvals(mod)
print(pvalsMod)
print(pvalsMod, all.coef = TRUE)
plot(pvalsMod)
Index

*Topic datasets
   - GenBin, 2
   - GenCont, 3
   - Gorman, 4
   - Hald, 4

   coef.ridgeLinear (linearRidge), 5
   coef.ridgeLogistic (logisticRidge), 12

   GenBin, 2
   GenBin_genotypes (GenBin), 2
   GenBin_phenotypes (GenBin), 2
   GenCont, 3
   GenCont_genotypes (GenCont), 3
   GenCont_phenotypes (GenCont), 3
   Gorman, 4

   Hald, 4

   linearRidge, 5, 10, 14, 16
   linearRidgeGenotypes, 8, 11, 16, 18
   linearRidgeGenotypesPredict, 10, 18
   logisticRidge, 7, 10, 12, 16
   logisticRidgeGenotypes, 10, 11, 14, 18
   logisticRidgeGenotypesPredict, 11, 17

   plot.pvalsRidgeLinear (pvals), 19
   plot.pvalsRidgeLogistic (pvals), 19
   plot.ridgeLinear (linearRidge), 5
   plot.ridgeLogistic (logisticRidge), 12
   predict.ridgeLinear (linearRidge), 5
   predict.ridgeLogistic (logisticRidge), 12

   print.pvalsRidgeLinear (pvals), 19
   print.pvalsRidgeLogistic (pvals), 19
   print.ridgeLinear (linearRidge), 5
   print.ridgeLogistic (logisticRidge), 12
   print.summary.ridgeLinear (linearRidge), 5
   print.summary.ridgeLogistic (logisticRidge), 12

   pvals, 19

   ridge-package, 2

   summary.ridgeLinear (linearRidge), 5
   summary.ridgeLogistic (logisticRidge), 12

   terms, 7, 14