Package ‘MWLasso’

August 9, 2016

Type Package

Title Penalized Moving-Window Lasso Method for Genome-Wide Association Studies

Version 1.3.1

Date 2016-07-27

Author Minli Bao

Maintainer Minli Bao <minli-bao@uiowa.edu>

Description The Moving-Window Lasso (MWLasso) method for genome-wide association studies. A window scans the design matrix. For predictors in the same window, their coefficients estimates are smoothed.

Suggests penalized

License GPL

NeedsCompilation no

Repository CRAN

Date/Publication 2016-08-09 01:23:16

R topics documented:

`MWLasso` *Penalized Moving-Window Lasso Method for Genome-Wide Association Studies*

Description

Find coefficients for a regression model with moving-window Lasso penalty (MWLasso). The model is marginalized so that missing values can be accommendated.

Usage

```
MWLasso(X, Y, lambda, eta, d, method='linear', epson = 1e-10, M = 100)
MW_parameters(X, Y, d, number, gamma2, method="linear", epson = 1e-10, M = 100)
```

Arguments

X The design matrix which can include missing values.

Y The response variable.

lambda The L1 penalty tuning parameter.

eta The moving-window penalty tuning parameter.

d The size of the moving-window penalty.

method The regression method. Method can be linear or logistic.

number The pre-determined number of non-zero coefficients.

gamma2 The proportion of the tuning parameter for moving-window part and L1 penalty part. gamma2 is the ratio of lambda/(lambda+eta).

epson Convergence criterion. The iteration will stop if the relative change is smaller than epson.

M The maximum number of iterations.

Details

The function minimizes 1/(2n)*MLS + lambda*L1 + eta/(2(d-1))*MW. Here MLS is the marginalized least squares, L1 is the L1 penalty in Lasso, and MW is the moving-window penalty.

Value

MWLasso returns:

beta The coefficients estimates.

MW_parameters returns:

lambda The tuning parameter for L1 penalty.

eta The tuning parameter for moving-window penalty.

Thus under such lambda and eta, MWLasso will return beta which satisfies the pre-determined number of non-zero coefficients.

Author(s)

Minli Bao <minli-bao@uiowa.edu>

References

Genome-wide association studies using a penalized moving-window regression. Minli Bao and Kai Wang. Submitted for review.

Examples

```
# In example 1, feature matrix X and response variable Y are randomly generated.
# Example 1:
n = 100
p = 200
X = matrix(rnorm(n*p, mean = 0, sd = 1), nrow = n, ncol = p)
X[2,1] = NA; X[80:90,1]= NA;
Y =  rnorm(n, mean = 0, sd = 1)

number = 20
gamma2 = 0.1
d = 4
params = MW_parameters(X, Y, d, number, gamma2, method="linear")
lambda = params[1]
eta = params[2]
beta.hat = MWLasso(X, Y, lambda, eta, d, method="linear")

# In example 2, the nki70 dataset is from the "penalized" package
# (https://cran.r-project.org/web/packages/penalized/penalized.pdf).
# The response variable Y is "event", the disease indicator.
# The feature matrix X is "TSPYL5...C20orf46", the gene expression
# measurements of 70 prognostic genes.
# Example 2:
require("penalized")
data(nki70)
data = nki70
Y = data$event
data$time <- NULL
data$event <- NULL
data$Diam <- NULL
data$N <- NULL
data$ER <- NULL
data$Grade <- NULL
data$Age <- NULL
X = data.matrix(data)

number = 20
gamma2 = 0.1
d = 4
params = MW_parameters(X, Y, d, number, gamma2, method="linear")
lambda = params[1]
eta = params[2]
beta.hat = MWLasso(X, Y, lambda, eta, d, method="linear")
```

Index