Package ‘spselect’

October 14, 2022

Type Package
Title Selecting Spatial Scale of Covariates in Regression Models
Version 0.0.1
Date 2016-08-29
Author Lauren Grant, David Wheeler
Maintainer Lauren Grant <pacele@vcu.edu>
Depends tester, magic, pracma
Description Fits spatial scale (SS) forward stepwise regression, SS incremental forward stagewise regression, SS least angle regression (LARS), and SS lasso models. All area-level covariates are considered at all available scales to enter a model, but the SS algorithms are constrained to select each area-level covariate at a single spatial scale.
License GPL (>= 2)
LazyLoad yes
NeedsCompilation no
Repository CRAN
Date/Publication 2016-08-29 18:53:24

R topics documented:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>spselect-package</td>
<td>2</td>
</tr>
<tr>
<td>lars.ss</td>
<td>3</td>
</tr>
<tr>
<td>lasso.ss</td>
<td>4</td>
</tr>
<tr>
<td>stagewise.ss</td>
<td>6</td>
</tr>
<tr>
<td>stepwise.ss</td>
<td>7</td>
</tr>
<tr>
<td>X</td>
<td>8</td>
</tr>
<tr>
<td>X.3D</td>
<td>9</td>
</tr>
<tr>
<td>y</td>
<td>10</td>
</tr>
</tbody>
</table>

Index 11
Description

Fits spatial scale (SS) forward stepwise regression, SS incremental forward stagewise regression, SS least angle regression (LARS), and SS lasso models. All area-level covariates are considered at all available scales to enter a model, but the SS algorithms are constrained to select each area-level covariate at a single spatial scale.

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>spselect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>0.0.1</td>
</tr>
<tr>
<td>Date:</td>
<td>2016-08-29</td>
</tr>
<tr>
<td>License:</td>
<td>GPL (&gt;=2)</td>
</tr>
<tr>
<td>LazyLoad:</td>
<td>yes</td>
</tr>
</tbody>
</table>

Author(s)

Lauren Grant, David Wheeler

Maintainer: Lauren Grant <pacele@vcu.edu>

References


Examples

```r
data(y)
data(X.3D)
y.name <- "y"
ss <- c("ind", "ss1", "ss2")
mod_forward.step.ss_1 <- stepwise.ss(y, X.3D, y.name, ss, 1)
```
### Description

This function fits a spatial scale (SS) LARS model.

### Usage

```r
lars.ss(y, X, ss, a.lst, S.v, C.v, col.plot, verbose=TRUE, plot=TRUE)
```

### Arguments

- `y`: A numeric response vector
- `X`: A data frame of numeric variables
- `ss`: A vector of names to identify the different levels of covariates available as potential candidates for model input
- `a.lst`: A list of identity matrices, where each column indicates a particular level or spatial scale for a specified covariate (e.g., ss1_x2)
- `S.v`: A vector of positive integers, where each number denotes the number of spatial scales associated with a particular covariate
- `C.v`: A vector, where all values are initialized to 0
- `col.plot`: A vector of colors (corresponding to each SS) used in the coefficient path plot
- `verbose`: If TRUE, details are printed as the algorithm progresses
- `plot`: If TRUE, a coefficient path plot is generated

### Details

This function estimates coefficients using the SS LARS modeling approach. The function also provides summary details and plots a coefficient path plot.

### Value

A list with the following items:

- `beta`: Regression coefficient estimates from all set of model solutions
- `beta.aic`: Regression coefficient estimates from final model
- `ind.v`: Vector of indices to denote the corresponding columns of X associated with each active predictor
- `aic.v`: Vector of Akaike information criterion (AIC) values
- `stack.ss`: Vector of indices to indicate the level at which each covariate enters the model

### Author(s)

Lauren Grant, David Wheeler
References


Examples

data(y)
data(X)

names.X <- colnames(X)

ss <- c("ind", "ss1", "ss2")

a.lst <- list(NULL)
a.lst[[1]] <- 1
dim(a.lst[[1]]) <- c(1,1)
dimnames(a.lst[[1]]) <- list(NULL, names.X[1])
a.lst[[2]] <- diag(2)
dimnames(a.lst[[2]]) <- list(NULL, names.X[c(2,3)])
a.lst[[3]] <- diag(2)
dimnames(a.lst[[3]]) <- list(NULL, names.X[c(4,5)])

S.v <- c(1,2,2)
C.v <- rep(0,length(a.lst))

mod_LARS.ss <- lars.ss(y, X, ss, a.lst, S.v, C.v, c("black", "red", "green"))
S.v  A vector of positive integers, where each number denotes the number of spatial scales associated with a particular covariate
C.v  A vector, where all values are initialized to 0
col.plot  A vector of colors (corresponding to each SS) used in the coefficient path plot
verbose  If TRUE, details are printed as the algorithm progresses
plot  If TRUE, a coefficient path plot is generated

Details
This function estimates coefficients using the SS lasso modeling approach. The function also provides summary details and plots a coefficient path plot.

Value
A list with the following items:

beta  Regression coefficient estimates from all set of model solutions
beta.aic  Regression coefficient estimates from final model
ind.v  Vector of indices to denote the corresponding columns of X associated with each active predictor
aic.v  Vector of Akaike information criterion (AIC) values
stack.ss  Vector of indices to indicate the level at which each covariate enters the model

Author(s)
Lauren Grant, David Wheeler

References

Examples
```r
data(y)
data(X)

names.X <- colnames(X)

ss <- c("ind", "ss1", "ss2")

a.lst <- list(NULL)
a.lst[[1]] <- 1
dim(a.lst[[1]]) <- c(1,1)
dimnames(a.lst[[1]]) <- list(NULL, names.X[1])

a.lst[[2]] <- diag(2)
dimnames(a.lst[[2]]) <- list(NULL, names.X[c(2,3)])```
a.lst[[3]] <- diag(2)
dimnames(a.lst[[3]]) <- list(NULL, names.X[c(4,5)])

S.v <- c(1,2,2)
C.v <- rep(0,length(a.lst))

mod_lasso.ss <- lasso.ss(y, X, ss, a.lst, S.v, C.v, c("black", "red", "green"))

stagewise.ss <- lasso.ss(y, X, X.3D, ss, increment, tolerance, col.plot, verbose=TRUE, plot=TRUE)

**Description**

This function fits a spatial scale (SS) incremental forward stagewise regression model.

**Usage**

```r
stagewise.ss(y, X, X.3D, ss, increment, tolerance, col.plot, verbose=TRUE, plot=TRUE)
```

**Arguments**

- `y` A numeric response vector
- `X` A data frame of numeric variables
- `X.3D` A 3-D or stacked array of numeric variables, where each stack represents a particular level of covariates (i.e., individual- and area-level variables at more than one spatial scale). In cases where values are only present for a covariate at certain levels, that covariate is assigned missing values at all other levels.
- `ss` A vector of names to identify the different levels of covariates available as potential candidates for model input
- `increment` A positive step size
- `tolerance` A small, positive value used as a stopping criterion when none of the predictors are correlated with the residuals. The algorithm stops if the overall maximum correlation is less than a specified tolerance.
- `col.plot` A vector of colors (corresponding to each SS) used in the coefficient path plot
- `verbose` If TRUE, details are printed as the algorithm progresses
- `plot` If TRUE, a coefficient path plot is generated

**Details**

This function estimates coefficients using the SS forward stagewise regression approach. The function also provides summary details and plots a coefficient path plot.

**Value**

A list with the following items:

- `beta.final` Regression coefficient estimates from final model
- `stack.ss` Vector of indices to indicate the level at which each covariate enters the model
Author(s)
Lauren Grant, David Wheeler

References

Examples
data(y)
data(X)
data(X.3D)
ss <- c("ind", "ss1", "ss2")
mod_forward.stage.ss_0.1 <- stagewise.ss(y, X, X.3D, ss, 0.1, 0.1, c("black", "red", "green"))

Description
This function fits a spatial scale (SS) forward stepwise regression model.

Usage
stepwise.ss(y, X.3D, y.name, ss, epsilon, verbose=TRUE)

Arguments
y  A numeric response vector
X.3D  A 3-D or stacked array of numeric variables, where each stack represents a particular level of covariates (i.e., individual- and area-level variables at more than one spatial scale). In cases where values are only present for a covariate at certain levels, that covariate is assigned missing values at all other levels.
y.name  A name for y
ss  A vector of names to identify the different levels of covariates available as potential candidates for model input
epsilon  A positive value used as a stopping criterion when there is inadequate improvement in the model’s performance. The algorithm stops if the difference in the Akaike information criterion (AIC) between the current model and the proposed model is less than epsilon.
verbose  If TRUE, details are printed as the algorithm progresses

Details
This function estimates coefficients using the SS forward stepwise regression approach. The function also estimates the model fit and provides summary details.
Value

A list with the following items:

- **beta.final**: Regression coefficient estimates from final model
- **aic.final**: AIC for final model
- **summary.final**: Summary output of final model
- **stack.ss**: Vector of indices to indicate the level at which each covariate enters the model

Author(s)

Lauren Grant, David Wheeler

References


Examples

```r
data(y)
data(X.3D)
y.name <- "y"
ss <- c("ind", "ss1", "ss2")
mod_forward.step.ss_1 <- stepwise.ss(y, X.3D, y.name, ss, 1)
```

Description

Simulated input data

Usage

```r
data(X)
```

Format

A data frame with 20 observations on the following 5 variables:

- **x1**: a numeric vector
- **ss1_x2**: a numeric vector
- **ss2_x2**: a numeric vector
- **ss1_x3**: a numeric vector
- **ss2_x3**: a numeric vector
Details

The data consist of simulated variables, including an individual-level covariate (x1) and two area-level covariates (x2, x3) available at two different spatial scales (ss1, ss2).

Examples

data(X)

X.3D

Input data X.3D

Description

Simulated input data (in stacked array format)

Usage

data(X.3D)

Format

A 20x3x3 stacked array with the following stacks:

ind a numeric array containing an individual-level variable (x1)
ss1 a numeric array containing area-level variables (x2, x3) available at ss1
ss2 a numeric array containing area-level variables (x2, x3) available at ss2

Details

The data consist of simulated variables, including an individual-level covariate (x1) and two area-level covariates (x2, x3) available at two different spatial scales (ss1, ss2). The data are in the form of a 3-D or stacked array, where each stack represents a particular level of covariates, including spatial scale. The first stack contains the individual-level variable; the second and third stacks contain the area-level variables at the ss1 and ss2 levels, respectively. Note that in cases where values are only present for a covariate at certain levels, that covariate is assigned missing values at all other levels.

References


Examples

data(X.3D)
Response data \( y \)

**Description**
Simulated response data

**Usage**
data(\( y \))

**Format**
A numeric response vector with 20 observations

**Examples**
data(\( y \))
Index

* datasets
  X, 8
  X.3D, 9
  y, 10

* models
  lars.ss, 3
  lasso.ss, 4
  stagewise.ss, 6
  stepwise.ss, 7

* package
  spselect-package, 2

lars.ss, 3
lasso.ss, 4

spselect(spselect-package), 2
spselect-package, 2
stagewise.ss, 6
stepwise.ss, 7

X, 8
X.3D, 9

y, 10