Package ‘RXshrink’

February 19, 2015

Title Maximum Likelihood Shrinkage via Generalized Ridge or Least Angle Regression

Version 1.0-8

Date 2014-1-13

Author Bob Obenchain <wizbob@att.net>

Maintainer Bob Obenchain <wizbob@att.net>

Depends R (>= 1.8.0), lars

Description Identify and display TRACEs for a specified shrinkage path and determine the extent of shrinkage most likely, under normal distribution theory, to produce an optimal reduction in MSE Risk in estimates of regression (beta) coefficients.

License GPL (>= 2)

URL http://www.r-project.org, http://members.iquest.net/~softrx

NeedsCompilation no

Repository CRAN

Date/Publication 2014-01-14 07:14:08

R topics documented:

RXshrink-package .................................................. 2
haldport ............................................................... 3
longley2 ............................................................... 4
mpg ................................................................. 4
plot.RXlarlso ....................................................... 5
plot.RXridge ......................................................... 6
plot.RXtrisk ......................................................... 7
plot.RXtsimu ........................................................ 8
plot.RXuclars ........................................................ 9
RXlarlso ............................................................. 10
RXridge ............................................................. 12
RXtrisk ............................................................. 14
RXtsimu ............................................................. 16
RXuclars ............................................................. 18
Description

The functions in this package augment the basic calculations of Generalized Ridge and Least Angle Regression with important visualization tools. Specifically, they display TRACEs of normal-distribution-theory Maximum Likelihood estimates of the key quantities that completely characterize the effects of shrinkage on the MSE Risk of fitted coefficients.

Details

Package: RXshrink
Type: Package
Version: 1.0-7
Date: 2011-12-24
License: GNU GENERAL PUBLIC LICENSE, Version 2, June 1991

RXridge() calculates and displays TRACEs for the Q-shaped shrinkage path, including the M-extent of shrinkage along that path, that are most likely under normal distribution theory to yield optimal reductions in MSE Risk.

When regression parameters have specified, KNOWN numerical values, RXtrisk() calculates and displays the corresponding True MSE Risk profiles and RXtsimu() first simulates Y-outcome data then calculates true Squared Error Losses associated with Q-shape shrinkage.

RXlarlso() calls the Efron/Hastie lars() R-function to perform Least Angle Regression then augments these calculations with Maximum Likelihood TRACE displays like those of RXridge().

RXuclars() applies Least Angle Regression to the uncorrelated components of a possibly ill-conditioned set of X-variables using a closed-form expression for the lars/lasso shrinkage delta factors that exits in this special case.

Author(s)

Bob Obenchain <wizbob@att.net>

References


**Description**

Heat evolved during setting of 13 cement mixtures of four basic ingredients. Each ingredient percentage appears to be rounded down to a full integer. The sum of the four mixture percentages varies from a maximum of 99% to a minimum of 95%. If all four regressor X-variables always summed to 100%, the centered X-matrix would then be of rank only 3. Thus, the regression of heat on four X-percentages is ill-conditioned, with an approximate rank deficiency of MCAL = 1.

**Usage**

data(haldport)

**Format**

A data frame with 13 observations on the following 5 variables.

- `p3ca` Integer percentage of 3CaO.Al2O3 in the mixture.
- `p3cs` Integer percentage of 3CaO.SiO2 in the mixture.
- `p4caf` Integer percentage of 4CaO.Al2O3.Fe2O3 in the mixture.
- `p2cs` Integer percentage of 2CaO.SiO2 in the mixture.
- `heat` Heat (cals/gm) evolved in setting, recorded to nearest tenth.

**Details**

The (RXshrink) haldport data are identical to the (MASS) cement data except for variable names.

**Source**


**References**

Description

Data from the "Employment and Training Report of the President, 1976" compiled by Art Hoerl, University of Delaware. Some slightly different numerical values than those used by Longley(1967) ...plus the added years of 1963-1975. Longley(1967) used only data from the 16 years of 1947 through 1962.

Usage

data(longley2)

Format

A data frame of 7 variables collected for 29 consecutive years; no NAs.

- **GNP.deflator** GNP price deflation index.
- **Unemployed** Unemployment percentage.
- **Armed.Forces** Size of the Armed Forces.
- **Population** Total U.S. Population.
- **Year** 1947 to 1975.
- **Employed** employment.
- **GNP** Gross National Product.

References


---

Description

Performance data on 32 autos from Motor Trends magazine.

Usage

data(mpg)
**plot.RXlarlso**

**Format**

A data frame of 11 variables collected for 32 automobiles; no NAs.

- **cylnds**: number of cylinders.
- **cubins**: cubic inches of engine displacement.
- **hpower**: engine horsepower.
- **weight**: total auto weight in pounds.
- **mpg**: miles per gallon.
- **shape**: engine shape (1=Straight, 0=V).
- **transm**: transmission type.
- **speeds**: number of forward speeds.
- **carbs**: number of carburetors.
- **fratio**: final drive ratio.
- **qmilt**: quarter mile time.

**References**


---

**plot.RXlarlso** 

*Plot method for RXlarlso objects*

**Description**

Plot trace displays for RXlarlso regression coefficients. The default is to display all five traces on one page with no legends.

**Usage**

```r
## S3 method for class 'RXlarlso'
plot(x, trace = "all", trkey = FALSE, ...)  # S3 method for class 'RXlarlso'
```

**Arguments**

- **x**: Output list object of class RXlarlso.
- **trace**: One of seven possible options: "all" to display 5 traces in one graph, "seq" to display 5 full-sized traces in sequence in response to user prompts, "coef" to display only the estimated beta coefficient trace, "rmse" to display only the estimated relative MSE risk trace, "exev" to display only the estimated excess MSE eigenvalue (OLS minus larlso), "infd" to display only the estimated inferior direction cosine trace, or "spat" to display only the delta-factor pattern trace.
- **trkey**: If TRUE, display a crude legend at the bottom of each trace plot.
- **...**: Optional argument(s) passed on to plot().
Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxlobj <- RXlarlso(form, data=longley2)
plot(rxlobj)

---

plot.RXridge  

**Plot method for RXridge objects**

Description

Plot trace displays for RXridge shrinkage regression coefficients. The default is to display all five traces on one page with no legends.

Usage

```r
## S3 method for class 'RXridge'
plot(x, trace = "all", trkey = FALSE, ...)
```

Arguments

- `x`  
  Output list object of class RXridge.
- `trace`  
  One of seven possible options: "all" to display 5 traces in one graph, "seq" to display 5 full-sized traces in sequence in response to user prompts, "coef" to display only the estimated beta coefficient trace, "rmse" to display only the estimated relative MSE risk trace, "exev" to display only the estimated excess MSE eigenvalue (OLS minus larlso), "infd" to display only the estimated inferior direction cosine trace, or "spat" to display only the delta-factor pattern trace.
- `trkey`  
  If TRUE, display a crude legend at the bottom of each trace plot.
- `...`  
  Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>
plot.RXtrisk

Examples

data(longleyR)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxrobj <- RXridge(form, data=longleyR)
plot(rxrobj)

plot.RXtrisk Plots method for RXtrisk objects

Description

Plot trace displays for RXtrisk true regression coefficients. The default is to display all five traces
on one page with no legends.

Usage

## S3 method for class 'RXtrisk'
plot(x, trace = "all", trkey = FALSE, 

Arguments

x Output list object of class RXtrisk.
trace One of seven possible options: "all" to display 5 traces in one graph, "seq" to
display 5 full-sized traces in sequence in response to user prompts, "coef" to
display only the estimated beta coefficient trace, "rmse" to display only the esti-
mated relative MSE risk trace, "exev" to display only the estimated excess MSE
eigenvalue (OLS minus larlso), "infd" to display only the estimated inferior di-
rection cosine trace, or "spat" to display only the shrinkage (delta) factor pattern
trace.
trkey If TRUE, display a crude legend at the bottom of each trace plot.
... Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
# set true regression parameter values not unlike their OLS estimates from RXridge().
trugam <- matrix(c(0.5,-.1,.2,.2,.2,-.2,.6,1)
trusig <- 0.04
rxtobj <- RXtrisk(form, data=longley2, trugam, trusig, Q=-1.5)
plot(rxtobj)
plot.RXtsimu

Plot method for RXtsimu objects

Description

Plot trace displays for RXtsimu simulated regression coefficients. The default is to display all three traces on one page with no legends.

Usage

```
## S3 method for class 'RXtsimu'
plot(x, trace = "all", trkey = FALSE, ...)  
```

Arguments

- `x`: Output list object of class RXtsimu.
- `trace`: One of five possible options: "all" to display 3 traces in one graph, "seq" to display 3 full-sized traces in sequence in response to user prompts, "coef" to display only the estimated shrunken beta coefficient trace, "rsel" to display only the relative squared error loss trace, or "spat" to display only the shrinkage (delta) factor pattern.
- `trkey`: If TRUE, display a crude legend at the bottom of each trace plot.
- `...`: Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

```r
data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Force+Population+Year+Employed
# set true regression parameter values not unlike their OLS estimates from RXridge().
trugam <- matrix(c(0, -1, 1, 1, -1, 0.6, 0.6), 2)
trusig <- 0.04
rxsobj <- RXtsimu(form, data=longley2, trugam, trusig, Q=-1.5)
plot(rxsobj)
```
Description

Plot trace displays for RXuclars regression coefficients. The default is to display all five traces on one page with no legends.

Usage

```r
## S3 method for class 'RXuclars'
plot(x, trace = "all", trkey = FALSE, ...)$
```

Arguments

- `x` Output list object of class RXuclars.
- `trace` One of seven possible options: "all" to display 5 traces in one graph, "seq" to display 5 full-sized traces in sequence in response to user prompts, "coef" to display only the estimated shrunken beta coefficient trace, "rmse" to display only the estimated relative MSE risk trace, "exev" to display only the estimated excess MSE eigenvalue (OLS minus ridge) trace, "infd" to display only the estimated inferior direction cosine trace, or "spat" to display only the shrinkage (delta) factor pattern trace.
- `trkey` If TRUE, display a crude legend at the bottom of each trace plot.
- `...` Optional argument(s) passed on to `plot()`.

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

```r
data(longley2)
form <- GNP-GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxuobj <- RXuclars(form, data=longley2)
plot(rxuobj)
```
Description

Identify whether least angle regression estimates are generalized ridge shrinkage estimates and generate TRACE displays for estimates that do correspond to ridge shrinkage factors between 0.00 and 0.99.

Usage

RXlarlso(form, data, rscale = 1, type = "lar", trace = FALSE, 
eps = .Machine$double.eps, omdmin = 9.9e-13, ...)

Arguments

form A regression formula \([y \sim x_1 + x_2 + \ldots]\) suitable for use with lm().
data Data frame containing observations on all variables in the formula.
rscale One of three possible choices (0, 1 or 2) for rescaling of variables as they are being "centered" to remove non-essential ill-conditioning: 0 implies no rescaling; 1 implies divide each variable by its standard error; 2 implies rescale as in option 1 but re-express answers as in option 0.
type One of "lasso", "lar" or "forward.stagewise" for function lars(). Names can be abbreviated to any unique substring. Default in RXlarlso() is "lar".
trace If TRUE, lars() function prints out its progress.
eps The effective zero for lars().
omdmin Strictly positive minimum allowed value for one-minus-delta (default = 9.9e-013.)
... Optional argument(s) passed on to the lars() function from the lars R-package.

Details

RXlarlso() calls the Efron/Hastie lars() function to perform Least Angle Regression on X-variables that have been centered and possibly rescaled but which may be (highly) correlated. Maximum likelihood TRACE displays paralleling those of RXridge are also computed and (optionally) plotted.

Value

An output list object of class RXlarlso:

form The regression formula specified as the first argument.
data Name of the data.frame object specified as the second argument.
p Number of regression predictor variables.
n Number of complete observations after removal of all missing values.
\textit{RXlarlso} 11

- \textbf{r2} Numerical value of R-square goodness-of-fit statistic.
- \textbf{s2} Numerical value of the residual mean square estimate of error.
- \textbf{prinstat} Listing of principal statistics.
- \textbf{gmat} Orthogonal matrix of direction cosines for regressor principal axes.
- \textbf{lars} An object of class lars.
- \textbf{coef} Matrix of shrinkage-ridge regression coefficient estimates.
- \textbf{risk} Matrix of MSE risk estimates for fitted coefficients.
- \textbf{exev} Matrix of excess MSE eigenvalues (ordinary least squares minus ridge.)
- \textbf{infd} Matrix of direction cosines for the estimated inferior direction, if any.
- \textbf{spat} Matrix of shrinkage pattern multiplicative delta factors.
- \textbf{mlik} Listing of criteria for maximum likelihood selection of M-extent-of-shrinkage.
- \textbf{sext} Listing of summary statistics for all M-extents-of-shrinkage.

\textbf{Author(s)}

Bob Obenchain <wizbob@att.net>

\textbf{References}


\textbf{See Also}

\textit{RXuclars}.

\textbf{Examples}

\begin{verbatim}
data(longley2) form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed rxlobj <- RXlarlso(form, data=longley2) rxlobj names(rxlobj) plot(rxlobj)
\end{verbatim}
**RXridge**

*Maximum Likelihood Shrinkage in Regression*

**Description**

Identify and display TRACES for the Q-shaped shrinkage path, including the M-extent of shrinkage along that path, that are most likely under normal distribution theory to yield optimal reductions in MSE Risk.

**Usage**

RXridge(form, data, rscale = 1, Q = "qmse", steps = 8, nq = 21, qmax = 5, qmin = -5, omdmin = 9.9e-13)

**Arguments**

- **form**: A regression formula [y~x1+x2+...] suitable for use with lm().
- **data**: Data frame containing observations on all variables in the formula.
- **rscale**: One of three possible choices (0, 1 or 2) for rescaling of variables as they are being "centered" to remove non-essential ill-conditioning: 0 implies no rescaling; 1 implies divide each variable by its standard error; 2 implies rescale as in option 1 but re-express answers as in option 0.
- **Q**: Shape parameter that controls the curvature of the shrinkage path through regression-coefficient likelihood space (default = "qmse" implies use the value found most likely to be optimal.) Use Q = 0 to specify Hoerl-Kennard "ordinary" ridge regression.
- **steps**: Number of equally spaced values per unit change along the horizontal M-extent-of-shrinkage axis for estimates to be calculated and displayed in TRACES (default = 8.)
- **nq**: Number of equally spaced values on the lattice of all possible values for shrinkage Q-shape between the "qmin" and "qmax" parameter settings (default = 21.)
- **qmax**: Maximum allowed Q-shape (default = +5.)
- **qmin**: Minimum allowed Q-shape (default = -5.)
- **omdmin**: Strictly positive minimum allowed value for one-minus-delta (default = 9.9e-013.)

**Details**

Illconditioned and/or nearly multicollinear regression models are unlikely to produce Ordinary Least Squares (OLS) regression coefficient estimates that are very close, numerically, to their unknown true values. Specifically, OLS estimates can then tend to have "wrong" numerical signs and/or unreasonable relative magnitudes, while shrunken (generalized ridge) estimates chosen to maximize their likelihood of reducing Mean Squared Error (MSE) Risk (expected loss) can be much more stable and reasonable, numerically. On the other hand, because only OLS estimates are guaranteed to be minimax when risk is matrix valued (truly multivariate), no guarantee of an actual reduction in MSE Risk is necessarily associated with shrinkage.
Value

An output list object of class RXridge:

- **form**: The regression formula specified as the first argument.
- **data**: Name of the data.frame object specified as the second argument.
- **p**: Number of regression predictor variables.
- **n**: Number of complete observations after removal of all missing values.
- **r2**: Numerical value of R-square goodness-of-fit statistic.
- **s2**: Numerical value of the residual mean square estimate of error.
- **prinstat**: Listing of principal statistics.
- **crlqstat**: Listing of criteria for maximum likelihood selection of path Q-shape.
- **qmse**: Numerical value of Q-shape most likely to be optimal.
- **qp**: Numerical value of the Q-shape actually used for shrinkage.
- **coef**: Matrix of shrinkage-ridge regression coefficient estimates.
- **risk**: Matrix of MSE risk estimates for fitted coefficients.
- **exev**: Matrix of excess MSE eigenvalues (ordinary least squares minus ridge.)
- **infd**: Matrix of direction cosines for the estimated inferior direction, if any.
- **spat**: Matrix of shrinkage pattern multiplicative delta factors.
- **mlik**: Listing of criteria for maximum likelihood selection of M-extent-of-shrinkage.
- **sext**: Listing of summary statistics for all M-extents-of-shrinkage.

Author(s)

Bob Obenchain <wizbob@att.net>

References


See Also

RXtrisk and RXtsimu.
Examples

```r
data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxrobj <- RXridge(form, data=longley2)
rxrobj
names(rxrobj)
plot(rxrobj)
```

---

RXtrisk

*True MSE Risk of Shrinkage Resulting from Known Regression Parameters*

Description

By specifying numerical values for regression parameters (uncorrelated components and error sigma) that usually are unknown, these functions can calculate and display True MSE Risk statistics associated with shrinkage along a given Q-shaped path.

Usage

```
RXtrisk(form, data, trug, trus, Q = 0, rscale = 1, steps = 8,
        qmax = 5, qmin = -5)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>form</td>
<td>A regression formula ([y = x_1 + x_2 + ...]) suitable for use with <code>lm()</code>.</td>
</tr>
<tr>
<td>data</td>
<td>Data frame containing observations on all variables in the formula.</td>
</tr>
<tr>
<td>trug</td>
<td>Column vector of numerical values for the true uncorrelated components of the regression coefficient vector.</td>
</tr>
<tr>
<td>trus</td>
<td>Numerical value for the true error standard deviation, Sigma.</td>
</tr>
<tr>
<td>Q</td>
<td>Numerical value for the shape parameter controlling shrinkage path curvature. Default shape is Q = 0 for Hoerl-Kennard &quot;ordinary&quot; ridge regression.</td>
</tr>
<tr>
<td>rscale</td>
<td>One of three possible choices (0, 1 or 2) for rescaling of variables as they are being &quot;centered&quot; to remove non-essential ill-conditioning: 0 implies no rescaling; 1 implies divide each variable by its standard error; 2 implies rescale as in option 1 but re-express answers as in option 0.</td>
</tr>
<tr>
<td>steps</td>
<td>Number of equally spaced values per unit change along the horizontal M-extent-of-shrinkage axis where estimates are calculated and displayed in TRACES (default = 8.)</td>
</tr>
<tr>
<td>qmax</td>
<td>Maximum allowed Q-shape (default = +5.)</td>
</tr>
<tr>
<td>qmin</td>
<td>Minimum allowed Q-shape (default = -5.)</td>
</tr>
</tbody>
</table>
Details

The RXridge() functions calculate maximum likelihood estimates (corrected, if necessary, so as to have correct range) for typical statistical inference situations where regression parameters are unknowns. In sharp contrast with this usual situation, the RXtrisk() functions show exactly how expected regression coefficients and true Mean Squared Error Risk actually do change with shrinkage when regression parameters take on specified, KNOWN numerical values.

Value

An output list object of class RXtrisk:

- `form` The regression formula specified as the first argument.
- `data` Name of the data.frame object specified as the second argument.
- `trug` Vector of numerical values for the true uncorrelated gamma components.
- `trus` Numerical value for the true error standard deviation, Sigma.
- `qp` Numerical value of the Q-shape actually used for shrinkage.
- `p` Number of regression predictor variables.
- `n` Number of complete observations after removal of all missing values.
- `prinstat` Listing of principal statistics.
- `coef` Matrix of expected shrinkage-ridge regression coefficients.
- `rmse` Matrix of true MSE risk values for shrunken coefficients.
- `exev` Matrix of true excess eigenvalues (ordinary least squares minus ridge.)
- `infd` Matrix of direction cosines for the true inferior direction, if any.
- `spat` Matrix of shrinkage pattern delta factors.
- `sext` Listing of summary statistics for all M-extents-of-shrinkage.

Author(s)

Bob Obenchain <wizbob@att.net>

See Also

RXridge and RXtsimu.

Examples

data(mpg)
form <- mpg-cylnds+cubins+hpower+weight
rxrobj <- RXridge(form, data=mpg)
# define true parameter values.
trugam <- matrix(c(-5,-.1,.1,-.6),4,1)
trusig <- 0.4
# create true shrinkage MSE risk scenario.
trumse <- RXtrisk(form, data=mpg, trugam, trusig, Q=-1, steps=4)
plot(trumse)
True Squared Error LOSS of Shrinkage for a Simulated Response Y-vector

Description

By specifying numerical values for regression parameters (uncorrelated components and error sigma) that usually are unknown, these functions allow the user to simulate response data and display the True Squared Error Loss associated with shrinkage along a given Q-shaped path.

Usage

RXtsimu(form, data, trug, trus, Q = 0, rscale = 1, steps = 8, seed = 0123, qmax = 5, qmin = -5)

Arguments

- form: A regression formula [y~x1+x2+...] suitable for use with lm().
- data: Data frame containing observations on all variables in the formula.
- trug: Column vector of numerical values for the true uncorrelated components of the regression coefficient vector.
- trus: Numerical value for the true error standard deviation, Sigma.
- Q: Numerical value for the shape parameter controlling the shrinkage path curvature. Default shape is Q = 0 for Hoerl-Kennard "ordinary" ridge regression.
- rscale: One of three possible choices (0, 1 or 2) for rescaling of variables as they are being "centered" to remove non-essential ill-conditioning: 0 implies no rescaling; 1 implies divide each variable by its standard error; 2 implies rescale as in option 1 but re-express answers as in option 0.
- steps: Number of equally spaced values per unit change along the horizontal M-extent-of-shrinkage axis where estimates are calculated and displayed in TRACEs (default = 8.)
- seed: Seed for random number generator. To get two different simulated response vectors and different sets of coefficients and SE losses, invoke the RXtsimu() function twice with different seed values. (default = 0123.)
- qmax: Maximum allowed Q-shape (default = +5.)
- qmin: Minimum allowed Q-shape (default = -5.)

Details

The RXridge() functions calculate maximum likelihood estimates (corrected, if necessary, so as to have correct range) for typical statistical inference situations where regression parameters are unknowns. In sharp contrast with this usual situation, the RXtsimu() functions show exactly how regression coefficient estimates and their true Squared Error Losses change with shrinkage for a simulated response Y-vector generated using given numerical values for regression parameters. In
fact, it is interesting to compare the output from RXtrisk() and RXtsimu() for given regression parameters to the corresponding output from RXridge() on the data.frame in which any original response Y-vector has been replaced by the ydat object from the RXtsimu() output list.

Value

An output list object of class RXtsimu:

- **form**: The regression formula specified as the first argument.
- **data**: Name of the data.frame object specified as the second argument.
- **trug**: Vector of numerical values for the true uncorrelated gamma components.
- **trus**: Numerical value for the true error standard deviation, Sigma.
- **qp**: Numerical value of the Q-shape actually used for shrinkage.
- **p**: Number of regression predictor variables.
- **n**: Number of complete observations after removal of all missing values.
- **prinstat**: Listing of principal statistics.
- **ydat**: Matrix with simulated Y-responses in its first column and the true expected values of those responses in its second column.
- **coef**: Matrix of shrinkage-ridge regression coefficient estimates.
- **rsel**: Matrix of true relative SE losses in regression coefficient estimates.
- **spat**: Matrix of shrinkage pattern multiplicative delta factors.
- **sext**: Listing of summary statistics for all M-extents-of-shrinkage.

Author(s)

Bob Obenchain <wizbob@att.net>

See Also

RXridge and RXtrisk.

Examples

data(haldport)
form <- heat~p3ca+p3cs+p4caf+p2cs
rxrobj <- RXridge(form, data=haldport)
plot(rxrobj)
# define true parameter values.
trugam <- matrix(c(NXLNPLNSLNU),4,1)
trusig <- 2.2
# create true shrinkage MSE risk scenario.
trumse <- RXtrisk(form, data=haldport, trugam, trusig, Q=5)
# calculate true shrinkage squared error losses.
trusim <- RXtsimu(form, data=haldport, trugam, trusig, Q=5)
haldpsim <- haldport
haldpsim[,5] <- trusim$ydat[,1]
rxsobj <- RXridge(form, data=haldpsim) # analysis as if parameters unknown
plot(rxsobj)
RXuclars

Maximum Likelihood Least Angle Regression on Uncorrelated X-Components

Description

Apply least angle regression estimation to the uncorrelated components of a possibly ill-conditioned linear regression model and generate normal-theory maximum likelihood TRACE displays.

Usage

RXuclars(form, data, rscale = 1, type = "lar", trace = FALSE, eps = .Machine$double.eps, omdmin = 9.9e-13, ...)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>form</td>
<td>A regression formula ([y \sim x_1 + x_2 + \ldots]) suitable for use with lm().</td>
</tr>
<tr>
<td>data</td>
<td>Data frame containing observations on all variables in the formula.</td>
</tr>
<tr>
<td>rscale</td>
<td>One of three possible choices (0, 1 or 2) for rescaling of variables as they are being &quot;centered&quot; to remove non-essential ill-conditioning: 0 implies no rescaling; 1 implies divide each variable by its standard error; 2 implies rescale as in option 1 but re-express answers as in option 0.</td>
</tr>
<tr>
<td>type</td>
<td>One of &quot;lasso&quot;, &quot;lar&quot; or &quot;forward.stagewise&quot; for function lars(). Names can be abbreviated to any unique substring. Default in RXlarlso() is &quot;lar&quot;.</td>
</tr>
<tr>
<td>trace</td>
<td>If TRUE, lars() function prints out its progress.</td>
</tr>
<tr>
<td>eps</td>
<td>The effective zero for lars().</td>
</tr>
<tr>
<td>omdmin</td>
<td>Strictly positive minimum allowed value for one-minus-delta (default = 9.9e-013.)</td>
</tr>
</tbody>
</table>

Details

RXuclars() applies Least Angle Regression to the uncorrelated components of a possibly ill-conditioned set of X-variables. A closed-form expression for the lars/lasso shrinkage delta factors exits in this case: \(\Delta(i) = \max(0,1-k/\text{abs}[\text{PC}(i)])\), where \(\text{PC}(i)\) is the principal correlation between \(Y\) and the \(i\)-th principal coordinates of \(X\). Note that the \(k\)-factor in this formulation is limited to a subset of \([0,1]\). \(\text{MCAL}=0\) occurs at \(k=0\), while \(\text{MCAL} = P\) results when \(k\) is the maximum absolute principal correlation.

Value

An output list object of class RXuclars:

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>form</td>
<td>The regression formula specified as the first argument.</td>
</tr>
<tr>
<td>data</td>
<td>Name of the data.frame object specified as the second argument.</td>
</tr>
</tbody>
</table>
p  Number of regression predictor variables.
n  Number of complete observations after removal of all missing values.
r2  Numerical value of R-square goodness-of-fit statistic.
s2  Numerical value of the residual mean square estimate of error.
prinstat  Listing of principal statistics.
gmat  Orthogonal matrix of direction cosines for regressor principal axes.
lars  An object of class lars.
coef  Matrix of shrinkage-ridge regression coefficient estimates.
risk  Matrix of MSE risk estimates for fitted coefficients.
exev  Matrix of excess MSE eigenvalues (ordinary least squares minus ridge.)
infd  Matrix of direction cosines for the estimated inferior direction, if any.
spat  Matrix of shrinkage pattern multiplicative delta factors.
mlik  Listing of criteria for maximum likelihood selection of M-extent-of-shrinkage.
sext  Listing of summary statistics for all M-extents-of-shrinkage.

Author(s)

Bob Obenchain <wizbob@att.net>

References


See Also

RXlarlso.

Examples

data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxuobj <- RXuclars(form, data=longley2)
rxuobj
plot(rxuobj)
Index

*Topic **datasets**
  - haldport, 3
  - longley2, 4
  - mpg, 4
*Topic **hplot**
  - plot.RXlarlso, 5
  - plot.RXridge, 6
  - plot.RXtrisk, 7
  - plot.RXtsimu, 8
  - plot.RXuclars, 9
  - RXlarlso, 10
  - RXridge, 12
  - RXtrisk, 14
  - RXtsimu, 16
  - RXuclars, 18
*Topic **methods**
  - plot.RXlarlso, 5
  - plot.RXridge, 6
  - plot.RXtrisk, 7
  - plot.RXtsimu, 8
  - plot.RXuclars, 9
*Topic **package**
  - RXshrink-package, 2
*Topic **regression**
  - RXlarlso, 10
  - RXridge, 12
  - RXtrisk, 14
  - RXtsimu, 16
  - RXuclars, 18

haldport, 3
longley2, 4
mpg, 4
plot.RXlarlso, 5
plot.RXridge, 6
plot.RXtrisk, 7
plot.RXtsimu, 8
plot.RXuclars, 9
RXlarlso, 10, 19
RXridge, 12, 15, 17
RXshrink-package, 2
RXtrisk, 13, 14, 17
RXtsimu, 13, 15, 16
RXuclars, 11, 18