Package: RXshrink (via r-universe)

September 2, 2024

Title Maximum Likelihood Shrinkage using Generalized Ridge or Least Angle Regression

Version 2.3

Date 2023-08-15

Author Bob Obenchain

Maintainer Bob Obenchain <wizbob@att.net>

Imports lars, ellipse

Suggests mgcv

Description Functions are provided to calculate and display ridge TRACE Diagnostics for a variety of alternative Shrinkage Paths. While all methods focus on Maximum Likelihood estimation of unknown true effects under normal distribution-theory, some estimates are modified to be Unbiased or to have ``Correct Range'' when estimating either [1] the noncentrality of the F-ratio for testing that true Beta coefficients are Zeros or [2] the ``relative'' MSE Risk (i.e. MSE divided by true sigma-square, where the ``relative'' variance of OLS is known.) The eff.ridge() function implements the ``Efficient Shrinkage Path'' introduced in Obenchain (2022) <Open Statistics>. This ``p-Parameter'' Shrinkage-Path always passes through the vector of regression coefficient estimates Most-Likely to achieve the overall Optimal Variance-Bias Trade-Off and is the shortest Path with this property. Functions eff.aug() and eff.biv() augment the calculations made by eff.ridge() to provide plots of the bivariate confidence ellipses corresponding to any of the p*(p-1) possible ordered pairs of shrunken regression coefficients. Functions for plotting TRACE Diagnostics now have more options.

License GPL-2

URL https://www.R-project.org , http://localcontrolstatistics.org

NeedsCompilation no

Depends R (>= 3.5.0)

Date/Publication 2023-08-07 19:30:05 UTC

Repository https://rlobenchain.r-universe.dev

RemoteUrl https://github.com/cran/RXshrink

RemoteRef HEAD

RemoteSha 8d1a9bd1096da0d6abe0c60b75ed77225a8ad9ee

Contents

RXshrink-package	*Maximum Likelihood (ML) Shrinkage using Generalized Ridge or Least Angle Regression*

Description

The functions in this package augment basic calculations of Generalized Ridge and Least Angle Regression plus Visual Insights from five types of ridge TRACE display: [1] regression coefficients, [2] relative MSE risk, [3] excess eigenvalues, [4] inferior direction cosines, and [5] shrinkage delta-factors. These TRACEs reveal the primary Effects of Shrinkage along ridge paths with 1, 2 or more parameters: shrinkage m-Extent, path q-Shape, and the p shrinkage delta-factors applied to the uncorrelated components of the Ordinary Least Squares estimator. All paths start at the OLS estimate [m = 0] and end at the shrinkage Terminous, (0, 0, ..., 0), where m = p = rank of the centered and rescaled X-matrix. Three different measures of overall Likelihood of minimal MSE risk (Classical Normal-Theory, Empirical Bayes, and Random Coefficients) are monitored to suggest an optimal m-Extent of shrinkage for the given matrix of non-constant x-Variables and the observed y-Outcome vector.

Details

Package:	RXshrink
Type:	Package
Version:	2.3
Date:	2023-08-15
License:	GNU GENERAL PUBLIC LICENSE, Version 2, June 1991

The eff.ridge() function calculates generalized ridge TRACE statistics for the Efficient Shrinkage PATH with p > 1 parameters. This PATH always passes through the Beta coefficient point-estimate that is most likely to achieve optimal MSE risk reductions under normal distribution-theory. This PATH is as Short as Possible; it consists of a Two-Piece Linear function with its single "interior" Knot at the MSE Risk Optimal m-Extent of Shrinkage.

MLboot(), MLcalc() and MLhist() support use of Bootstrap resampling to study both the Bias and the MSE Risk characeristics of non-linear (unrestricted) Generalized Ridge Regression (GRR) estimators.

When true regression parameters have user-specified (KNOWN) numerical values, MLtrue() uses this information and generates a new data.frame that contains a y-Outcome vector of the expected form with "disturbance" terms that are I.I.D. Normal errors-in-measurement. Arguments to MLtrue() must include the "formula" for a desired linear model and a data.frame containing the specified X-variables.

qm.ridge() calculates and displays TRACEs for traditional PATHs defined by just 2-parameters: q-Shape and **m**-Extent of Shrinkage. By default, the search for the Path with most likely **q**-Shape uses a lattice of only 21 values within [-5,+5]. However, lattice searches for both q-Shape and m-Extent are easy to modify using the qmax, qmin, nq and steps arguments to qm.ridge(). The "ordinary" ridge Path of Hoerl and Kennard always uses **q**-Shape = 0, while "uniform" shrinkage corresponds

to **q**-Shape = +1. NONE of these qm-Paths generally achieve Overall Minimum MSE Risk when p > 2 because they restrict attention to a "monotome" (increasing or decreasing) family of "delta" shrinkage-factors.

aug.lars() augments the Efron-Hastie lars() R-function to perform Least Angle Regression with MSE risk calculations and Maximum Likelihood TRACE displays ...like those of eff.ridge() and qm.ridge().

uc.lars() applies Least Angle Regression methods to the **Uncorrelated Components** of a possibly ill-conditioned set of x-Variables. Calculations use a closed-form expression for lars/lasso shrinkage delta-factors that apply because NO Ill-Conditioning is present in these "uc" cases.

correct.signs() displays the Normal-theory maximum likelihood estimate of the regression coefficient vector that minimizes MSE Risk in the UNKNOWN direction of p-space PARALLEL to the true Beta vector. This estimate corrects "wrong-sign" problems in the sense that its coefficients have the same relative magnitudes and numerical signs as those of the "Correlation Form" of the X'y vector.

YonX() displays Shrinkage statistics and graphics for "simple" linear regression (p = 1) models.

RXpredict() makes predictions (i.e. computes "fitted.values") for 6 types of RXshrink estimation ...either at a user-specified m-Extent of Shrinkage or at the Normal-theory "minMSE" m-Extent.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Efron B, Hastie T, Johnstone I, Tibshirani R. (2003) Least angle regression. *Annals of Statistics* **32**, 407-499.

Goldstein M, Smith AFM. (1974) Ridge-type estimators for regression analysis. *J. Roy. Stat. Soc. B* **36**, 284-291. (The 2-parameter shrinkage family.)

Obenchain RL. (1975) Ridge Analysis Following a Preliminary Test of the Shrunken Hypothesis. *Technometrics* **17**, 431-441. doi:10.1080/00401706.1975.10489369

Obenchain RL. (1977) Classical F-tests and Confidence Regions for Ridge Regression. *Technometrics* **19**, 429-439. doi:10.1080/00401706.1977.10489582

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2023) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.3. http://localcontrolstatistics.org

Examples

```
demo(longley2)
```

---

aug.lars	*Maximum Likelihood Estimation of Effects in Least Angle Regression*

Description

These functions perform calculations that determine whether least angle and lasso regression estimates correspond to generalized ridge regression (GRR) estimates (i.e. whether they use shrinkage delta-factors that are both non-negative and strictly less than 1.0). They also estimate the Normal-theory likelihood that MSE risk is minimized and compute diagnostics for display in ridge TRACE plots.

Usage

```
aug.lars(form, data, rscale = 1, type = "lar", trace = FALSE,
    eps = .Machine$double.eps, 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 (after being "centered") to remove all "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 aug.lars() 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.)

Details

aug.lars() 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 eff.ridge() and qm.ridge() are also computed and (optionally) plotted.

Value

An output list object of class aug.lars:

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.
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.
mClk	Most Likely Extent of Shrinkage Observed: best multiple of (1/steps) <= p.
minC	Minimum Observed Value of Normal-theory -2*log(Likelihood-Ratio).

Author(s)

Bob Obenchain <wizbob@att.net>

References

Breiman L. (1995) Better subset regression using the non-negative garrote. *Technometrics* **37**, 373-384.

Efron B, Hastie T, Johnstone I, Tibshirani R. (2003) Least angle regression. *Annals of Statistics* **32**, 407-499.

Hastie T, Efron, B. (2013) lars: Least Angle Regression, Lasso and Forward Stagewise. ver 1.2, https://CRAN.R-project.org/package=lars

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

Tibshirani R. (1996) Regression shrinkage and selection via the lasso. *J. Roy. Stat. Soc. B* **58**, 267-288.

See Also

uc.lars.

Examples

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

---

correct.signs	*Normal-Theory Maximum Likelihood Estimation of Beta Coefficients with "Correct" Signs*

Description

Obenchain(1978) discussed the risk of linear generalized ridge estimators in individual directions within p-dimensional X-space. While shrinkage to ZERO is clearly optimal for all directions strictly ORTHOGONAL to the true BETA, he showed that optimal shrinkage in the UNKNOWN direction PARALLEL to the true BETA is possible. This optimal BETA estimate is of the form k * X'y, where k is the positive scalar given in equation (4.2), page 1118. The correct.signs() function computes this estimate, B(=), that uses GRR delta-shrinkage factors proportional to X-matrix eigenvalues.

Usage

```
correct.signs(form, data)
```

Arguments

form	A regression formula [y~x1+x2+...] suitable for use with lm().
data	Data frame containing observations on all variables in the formula.

Details

Ill-conditioned (nearly multi-collinear) regression models can produce Ordinary Least Squares estimates with numerical signs that differ from those of the X'y vector. This is disturbing because X'y contains the sample **correlations** between the X-predictor variables and y-response variable. After all, these variables have been "centered" by subtracting off their mean values and rescaled to vectors of length one. Besides displaying OLS estimates, the correct.signs() function also displays the "correlation form" of X'y, the estimated delta-shrinkage factors, and the k-rescaled beta-coefficients. Finally, the "Bfit" vector of estimates proportional to B(=) is displayed that minimizes the restricted Residual Sum-of-Squares. This restricted RSS of Bfit cannot, of course, be less than the RSS of OLS, but it can be MUCH less that the RSS of B(=) whenever B(=) shrinkage appears excessive.

Value

An output list object of class "correct.signs":

data	Name of the data.frame object specified as the second argument.
form	The regression formula specified as the first 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 (p by 5) from qm.ridge().

kpb	Maximum likelihood estimate of k-factor in equation (4.2) of Obenchain(1978).
bmf	Rescaling factor for B(=) to minimize the Residual Sum-of-Squares.
signs	Listing of five Beta coefficient statistics (p by 5): OLS, X'y, Delta, B(=) and Bfit.
loff	Lack-of-Fit statistics: Residual Sum-of-Squares for OLS, X'y, B(=) and Bfit.
sqcor	Squared Correlation between the y-vector and its predicted values. The two values displayed are for OLS predictions or for predictions using Bfit, X'y or B(=). These two values are the familiar R^2 coefficients of determination for OLS and Bfit.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.2. http://localcontrolstatistics.org

See Also

eff.ridge, qm.ridge and MLtrue.

Examples

```
data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxcsobj <- correct.signs(form, data=longley2)
rxcsobj
str(rxcsobj)
```

---

eff.aug	*Augment calculations performed by eff.ridge() to prepare for display of eliptical confidence regions for pairs of biased coefficient estimates using plot.eff.biv()*

Description

This function makes classical (rather than Bayesian) Normal distribution-theory calculations of the form proposed in Obenchain(1977). Instead of providing "new" confidence regions for estimable linear functions, Generalized Ridge Regression (GRR) can focus interest on estimates that are within traditional confidence intervals and regions but which deviate reasonably from the centroid of that interval or region.

Usage

```
eff.aug(efobj)
```

Arguments

efobj   An output object of class "eff.ridge".

Value

An output list object of class "eff.aug"...

p	Number of regression predictor variables.
LMobj	The lm() output object for the model fitted using eff.ridge().
bstar	The p by 3 matrix of GRR coefficients. Column 1 contains OLS estimates, the middle column gives optimally biased coefficient estimates corresponding to the "Interior Knot" on all p of the Two-Piece Linear Splines, and column 3 contains all zeros for the Shrinkage Terminus.
mcal	Three increasing measures of shrinkage "Extent". The first is 0 for the OLS (BLUE) estimate, the second is the Maximum Likelihood m-Extent of Shrinkage [PURPLE point], and the third is m = p for Shrinkage to beta = 0. This "shrinkage terminus" [BLACK point] is frequently outside of the eff.biv() plot frame ...allowing the ellipse to be as LARGE as possible.
vnams	Names of all variables actually used in the GRR model.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1977) Classical F-tests and Confidence Regions for Ridge Regression. *Technometrics* **19**, 429-439. doi:10.1080/00401706.1977.10489582

Obenchain RL. (2021) *The Efficient Shrinkage Path: Maximum Likelihood of Minimum MSE Risk.* https://arxiv.org/abs/2103.05161

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

See Also

eff.ridge and meff

eff.biv	*Specify pairs of GRR Coefficient Estimates for display in Bivariate Confidence Regions*

Description

This function specifies which Pair of GRR estimates to display and the single (or dual) Confidence Level(s) of the Ellipse(s) displayed. Reguested confidence levels must both be equal to or greater than 0.05 and less than or equal to 0.95.

Usage

```
eff.biv(efaug, x1 = 1, x2 = 2, conf1 = 0.95, conf2 = 0.50)
```

Arguments

efaug	An output list object of class "eff.aug" for a GRR model with p >= 2 x-Variables.
x1	Integer index value >= 1 and <= p for the x-Coefficient to be displayed on the horizontal axis.
x2	Integer index value >= 1 and <= p for the x-Coefficient to be displayed on the vertical axis. Index x2 must differ from x1 to display a plot.
conf1	This first Confidence level must be >= 0.05 and <= 0.95 to display an Ellipse.
conf2	When the second Confidence level is >= 0.05 and <= 0.95, its Ellipse is displayed. No plot is displayed when both conf1 and conf2 are outside of the [0.05, 0.95] range.

Value

An output list object of class "eff.biv"...

p	Number of regression predictor variables.
LMobj	The lm() output object for the model fitted using eff.ridge().
bstar	The p by 3 matrix of shrunken GRR coefficients. The 3 columns correspond to OLS estimates, optimally shrunken estimates, and estimates shrunken to Zeros.
mcal	Three increasing measures of shrinkage m-Extent: 0 for OLS [BLUE], the Optimal m-Extent at the "Interior Knot" [PURPLE], and m = p [BLACK] at the Shrinkage Terminus.
ellip1	matrix[100, 2] of points on confidence ellipse 1.
conf1	confidence level of ellipse 1 within [0.05, 0.95].
ecor1	Pearson correlation between x1 and x2 coordinates.
ellip2	matrix[100, 2] of points on confidence ellipse 2.
conf2	confidence level of ellipse 2 within [0.05, 0.95].
ecor2	Pearson correlation between x1 and x2 coordinates.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1977) Classical F-tests and Confidence Regions for Ridge Regression. *Technometrics* **19**, 429-439. doi:10.1080/00401706.1977.10489582

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

Murdoch DJ. and Chow ED. (1996). A graphical display of large correlation matrices. *The American Statistician* 50, 178-180.

Murdoch DJ. ellipse: Functions for Drawing Ellipses and Ellipse-Like Confidence Regions. https://CRAN.R-project.org/package=ellipse

See Also

`eff.aug`, `ellipse.lm`, `eff.ridge`

---

eff.ridge	*Efficient Maximum Likelihood (ML) Shrinkage via the Shortest Piecewise Linear-Spline PATH*

Description

Compute and display TRACEs for the p-paramater Shrinkage PATH passing through the (classical) Normal-theory Maximum Likelihood (ML) point-estimate of the Beta coefficient vector. The m-Extent of overall Optimal Shrinkage corresponding to this solution occurs at the only "interior" Knot on the Shrinkage Path and is also marked by the vertical dashed-line drawn on all 5-types of eff.ridge TRACE displays.

Usage

```
eff.ridge(form, data, rscale = 1, steps = 20, ...)
```

Arguments

form	A regression formula [y~x1+x2+...+xp] 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 (after being "centered") to remove all "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 for estimates to be calculated and displayed in TRACES (default = 20.)
...	Optional argument(s)

Details

Ill-conditioned and/or nearly multi-collinear 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 have unreasonable relative magnitudes or "wrong" numerical signs when the number of x-variables is 2 or more. Shrunken (Generalized Ridge Regression) estimates chosen to maximize their likelihood of reducing Mean Squared Error (MSE) Risk (expected Squared Error Loss) can be 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 eff.ridge:

data	Name of the data.frame object specified as the second argument.
form	The regression formula is the first argument.
p	Number of regression x-predictor variables.
n	Number of complete observations after removal of all missing values.
r2	Numerical value of R-squared: proportion of variance explained.
s2	Numerical value of the residual mean square estimate of error.
prinstat	Listing of 5 summary statistics for each of p-Principal Axes.
rscale	Variable re-scaling code of 0, 1 or 2 used in calculations.
data	The data.frame containing all variables listed in the formula.
gmat	Orthogonal Matrix of Direction Cosines for Principal Axes.
coef	Matrix of shrinkage-ridge regression coefficient estimates.
rmse	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 an m-Extent for Shrinkage.
sext	Listing of summary statistics for all M-extents-of-shrinkage.
mStar	Optimal m-Extent of Shrinkage with delta[j] = dMSE[j] on TRACE displays.
mMSE	Minimum MSE Risk estimate.
mClk	Most Likely Observed Extent of Shrinkage: best multiple of (1/steps) <= p.
minC	Minimum Observed Value of Normal-theory -2*log(Likelihood-Ratio).
dMSE	Most Likely to be Optimal-values for Shrinkage: dMSE[j] for j in [1:p].

Author(s)

Bob Obenchain <wizbob@att.net>

References

Thompson JR. (1968) Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association* **63**, 113-122. (The "cubic" estimator.)

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2021) *The Efficient Shrinkage Path: Maximum Likelihood of Minimum MSE Risk.* https://arxiv.org/abs/2103.05161

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.3. http://localcontrolstatistics.org

See Also

MLcalc, meff, correct.signs, MLtrue and RXpredict.

Examples

```
data(longley2)
form <- Employed~GNP+GNP.deflator+Unemployed+Armed.Forces+Population+Year
rxefobj <- eff.ridge(form, data=longley2)
rxefobj          # print shrinkage summary statistics...
plot(rxefobj)    # 5 TRACEs on 1 plot...
str(rxefobj)
```

---

haldport | *Portland Cement data of Hald(1952)*

Description

Heat evolved during setting of 13 cement mixtures of four (or five) ingredients. The first four ingredient percentages appear to be "rounded down" to a full integer. The fifth integer percentage of "other" material assures that the five percentages sum to exactly 100%. However, the "centered" X-matrix resulting from inclusion of all five ingredients would then be Singular (rank=4). In other words, regressing any y-Outcome on only the first four X-variables yields an "ill-conditioned" model that, while having numerical "full rank"=4, actually suffers a effective "rank deficiency" of at least mcal = 1.

Usage

```
data(haldport)
```

Format

A data frame with 13 observations on the following 6 variables.

p3ca Positive Integer percentage of 3CaO.Al2O3 [tricalcium aluminate] in the mixture.

p3cs Positive Integer percentage of 3CaO.SiO2 [tricalcium silicate] in the mixture.

p4caf Positive Integer percentage of 4CaO.Al2O3.Fe2O3 [tetracalcium aluminoferrite] in the mixture.

p2cs Positive Integer percentage of 2CaO.SiO2 [dicalcium silicate] in the mixture.

other Positive Integer percentage of other ingredients 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 and inclusion of the "other" X-variable.

Source

Woods H, Steinour HH, Starke HR. "Effect of composition of Portland cement on heat evolved during hardening. Industrial Engineering and Chemistry 1932; 24: 1207-1214.

References

Hald A. Statistical Theory with Engineering Applications. 1952 (page 647.) New York; Wiley.

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

---

longley2   *Art Hoerl's update of the infamous Longley(1967) benchmark dataset*

Description

Data from the "Employment and Training Report of the President, 1976" compiled by Art Hoerl, University of Delaware. This data.frame contains both [1] some different ("corrected") numerical values from those used by Longley(1967) and [2] 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 covering 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

Longley JW. An appraisal of least-squares programs from the point of view of the user. J. Amer. Statist. Assoc. 1967; 62: 819-841.

Hoerl AE, Kennard RW. Ridge Regression - Present and Future. 12th Symposium on the Interface. 1979. Waterloo, Ontario, Canada.

---

meff *m-Extents of Shrinkage used in eff.ridge() Calculations.*

Description

The meff() function computes the numerical Shrinkage delta-factors corresponding to any desired m-Extent of Shrinkage, meobj, along the "efficient" (shortest) Path used by eff.ridge(). This Two-Piece Linear Path has its single "interior" Knot at the Normal-theory Maximum Likelihood estimate with Minimum MSE Risk: d[j] = dMSE[j] for j = 1, 2, ..., p.

Usage

```
meff(meobj, p, dMSE)
```

Arguments

meobj	A desired m-Extent of Shrinkage along the "efficient" Path.
p	The integer number of non-constant x-variables used in defining the linear model being fitted to ill-conditioned (intercorrelated, confounded) data. Note that p must also be rank of the given X-matrix.
dMSE	Maximum Likelihood [ML] estimates of Shrinkage Delta-Factors leading to minimum MSE risk.

Value

The appropriate scalar value for m and corresponding p by p diagonal matrix d:

meobj	Any desired m-Extent of Shrinkage (a scalar) within [0, p].
d	The p by p diagonal matrix of shrinkage-factors: d[j,j] in [0, 1].

Author(s)

Bob Obenchain <wizbob@att.net>

See Also

[eff.ridge](#) and [eff.aug](#).

MLboot	*Calculate Bootstrap distribution of Unrestricted Maximum Likelihood (ML) point-estimates for a Linear Model.*

Description

Resample With-Replacement from a given data.frame and recompute MSE risk-optimal estimates of Beta-Coefficients and their Relative MSE risks using MLcalcs() to compute ML point-estimates.

Usage

```
MLboot(form, data, reps=100, seed, rscale=1)
```

Arguments

form	Regression formula [y~x1+x2+...] suitable for use with lm().
data	data.frame containing observations on all variables in the formula.
reps	Number of Bootstrap replications: Minimum reps = 10, Default is reps = 100. While reps = 10000 is reasonable for bivariate (p=2) linear models, even that many reps could be excessive for models with p » 2.
seed	Either an Integer between 1 and 999 or else missing to generate a random seed.
rscale	One of three possible choices (0, 1 or 2) for "rescaling" of variables (after being "centered") to remove all "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.

Details

Ill-conditioned and/or nearly multi-collinear linear regression models are unlikely to yield reasonable ML unbiased (OLS) point-estimates. But more reasonable ML "optimally-biased" point-estimates from generalized ridge regression (GRR) typically have questionable MSE risk characteristics because they are complicated non-linear functions of the observed y-outcome vector. Thus the distribution of bootstrap resamples is of considerable interest in both theory and practice.

Value

An output list object of class MLboot:

data	Name of the data.frame object specified as the second argument.
form	The regression formula specified as the first argument.
reps	Number of Bootstrap replications performed.
seed	Seed used to start random number generator.
n	Number of complete observations after removal of all missing values.
p	Number of beta, rmse or dmse estimates resampled.
ols.beta	OLS matrix (reps x p) of unbiased beta-coefficient estimates.
ols.rmse	OLS matrix (reps x p) of beta-coefficient relative variances.
opt.dmse	OPT matrix (reps x p) of delta shrinkage-factors with minimum MSE risk.
opt.beta	OPT matrix (reps x p) of biased beta-coefficient estimates.
opt.rmse	OPT matrix (reps x p) of beta-coefficient relative MSE risks.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Thompson JR. (1968) Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association* **63**, 113-122. (The "cubic" estimator.)

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

See Also

eff.ridge, correct.signs

MLcalc	*Calculate Efficient Maximum Likelihood (ML) point-estimates for a Linear Model that are either Unbiased (OLS) or Most Likely to be Optimally Biased under Normal-distribution theory.*

Description

Compute MSE risk-optimal point-estimates of Beta-Coefficients and their Relative MSE risks. Much of the code for this function is identical to that of eff.ridge(), which computes multiple points along the "Efficient" Shrinkage Path. MLcalc() restricts attention to only two points: [1] the Unbiased OLS (BLUE) vector and [2] the Most Likely to be Optimally Biased [Minimum MSE Risk] vector of estimates.

Usage

```
MLcalc(form, data, rscale = 1)
```

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 (after being "centered") to remove all "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.

Details

Ill-conditioned and/or nearly multi-collinear 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 squared-error loss) can be more stable 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 expected reduction in MSE Risk is necessarily associated with "Optimal" Generalized Ridge Regression shrinkage.

Value

An output list object of class MLcalc:

data   Name of the data.frame object specified as the second argument.

form   The regression formula specified as the first 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.

gmat   Orthogonal Matrix of Direction Cosines for Principal Axes [1:p, 1:p].

beta   Numerical shrinkage-ridge regression coefficient estimates [1:2, 1:p].

rmse   Numerical MSE risk estimates for fitted coefficients [1:2, 1:p].

dMSE   Numerical delta-factors for shrinking OLS components [1:p].

ys   Numerical rescaling factor for y-outcome variable [1, 1].

xs   Numerical rescaling factors for given x-variables [1:p].

Author(s)

Bob Obenchain <wizbob@att.net>

References

Thompson JR. (1968) Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association* **63**, 113-122. (The "cubic" estimator.)

Obenchain RL. (2021) *The Efficient Shrinkage Path: Maximum Likelihood of Minimum MSE Risk.* https://arxiv.org/abs/2103.05161

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

See Also

eff.ridge, MLboot, eff.aug

MLhist   *Plot method for MLboot objects*

Description

Frequency Histogram displays that use both a specified "Middle Percentage" of a MLboot() distribution that contains outliers as well as a proposed total number of bins.

Usage

```
MLhist(x, comp="opt.beta", xvar=1, npct = 95, bins = 50 )
```

Arguments

x	An Output-list object of class "MLboot".
comp	One of five possible choices of MLboot() resampled estimates: "ols.beta", "ols.rmse", "opt.dmse", "opt.beta" or "opt.rmse".
xvar	Column number of estimates to be displayed: 1 <= xvar <= x\$p.
npct	An integer percentage of the simulated scalar estimates to be displayed. This percentage should be at least 66 and at most 100.
bins	This proposed number of histogram "breaks" is only a suggestion; actual break-points will be set to "pretty" values.

Value

An output list object of class MLhist:

x	Character string showing user choices for "comp" and "xvar" arguments.
form	The regression formula specifying the linear model in MLboot().
reps	Number of Bootstrap replications performed.

npct	An integer percentage of the simulated scalar estimates displayed.
rbins	Number of histogram Bins requested.
dbins	Number of histogram Bins actually displayed.
ntot	Total number of rows of scalar estimates.
p	Total number of columns of scalar estimates.
nlo	First (smallest) order statistic displayed in Histogram.
nup	Last (largest) order statistic displayed in Histogram.
noin	Total number of scalar estimates displayed in Histogram.
xmn	Observed Mean Estimate: location of vertical "blue" dashed line on Histogram.

Author(s)

Bob Obenchain <wizbob@att.net>

| MLtrue | *Simulate data for Linear Models with known Parameter values and Normal Errors* |

Description

Using specified numerical values for Parameters, such as the true error-term variance, that usually are unknown, MLtrue() creates a new data.frame that contains variables "Yhat" (the vector of expected values) and "Yvec" (Yhat + IID Normal error-terms) as well as the p given X-variables. Thus, MLtrue() produces "correct" linear models that can be ideal for analysis and bootstraping via methods based on Normal-theory.

Usage

```
MLtrue(form, data, seed, go=TRUE, truv, trub, truc)
```

Arguments

form	A regression formula [y~x1+x2+...] suitable for use with lm().
data	Data frame containing observations on all variables in the formula.
seed	Seed for random number generation between 1 and 1000. When this seed number is missing in a call to MLtrue(), a new seed is generated using runif(). Error terms are then generated using calls to rnorm(), and the seed used is reported in the MLtrue output.list to enable reproducible research.
go	Logical value: go=TRUE starts the simulation using OLS estimates or the numerical values from the (optional) truv, trub and truc arguments to define the "Yhat" and "Yvec" variables; go=FALSE causes MLtrue() to compute OLS estimates that ignore the truv, trueb and truc arguments. Thus, go=FALSE provides a convenient way for users to Print and Examine the MLtrue() output.list before deciding which parameter-settings they wish to actually USE or MODIFY before making a final run with go=TRUE.

truv	Optional: numerical value for the true error variance, sigma^2.
trub	Optional: column vector of numerical values for the true regression coefficients.
truc	Optional: column vector of numerical values for the true uncorrelated components of the regression coefficient vector.

Details

RXshrink functions like eff.ridge() and qm.ridge() calculate maximum likelihood estimates ...either unbiased (BLUE) or "optimally" biased (minimum MSE risk) along a specified "path"... for typical statistical inference situations where true regression parameters are unknown. Furthermore the specified linear-model may be "incorrect" or the error-terms may not be IID Normal variates. In sharp contrast with this usual situation, the MLtrue() function generates a "Yvec" vector of outcomes from a CORRECT model of the given form that does contain IID Normal error-terms and all true parameter values are KNOWN. This makes it interesting to compare plots and output.lists from RXshrink GRR-estimation functions on a "real" data.frame with the corresponding outputs from a MLtrue() data.frame with known parameter-setting that are either identical to or "close" to those estimated from "real" data. WARNING: All output X-variables are "centered" and "rescaled" to have mean ~0 and variance 1. Yhat expected values are "centered" but usually have variance differing from 1. Yvec values are not "centered" and have variance determined by either the original data.frame or the "truv" setting.

Value

An output list object of class MLtrue:

new	A data.frame containing the Yvec and Yhat vectors as well as centered and rescaled versions of the input X-matrix and original Y-vector.
Yvec	An additional copy of Yvec pseudo-data containing Normal error-terms.
Yhat	An additional copy of Yhat linear fitted-values.
seed	Seed actually used in random-number generation; needed for replication.
tvar	True numerical value of sigma^2 ...[1,1] matrix.
tbeta	True numerical values of beta-coefficients ...[p,1] matrix.
tcomp	True numerical values of uncorrelated components ...[p,1] matrix.
useb	Logical: TRUE => tbeta vector was used in simulation. FALSE => tcomp vector was used rather than tbeta.
data	The name of the data.frame object specified in the second argument.
form	The regression formula specified in the first argument. NOTE: The final 11 output items [p, ..., xs] are identical to those in output lists from RXshrink functions like eff.ridge() and qm.ridge().

Author(s)

Bob Obenchain <wizbob@att.net>

See Also

[eff.ridge](), [qm.ridge]() and [MLboot]().

Examples

```
# require(RXshrink)
data(mpg)
form <- mpg~cylnds+cubins+hpower+weight
MLtest <- MLtrue(form, mpg, go=FALSE)   # all other potential arguments "missing"...
MLtest   # print current parameter estimates...
cvec <- c(-0.5, 0.15, -0.16, -0.6)   # define alternative "true" components...
MLout <- MLtrue(form, mpg, truc = cvec ) # Use "truc" input with default go=TRUE...
str(MLout)
formY <- Yhat~cylnds+cubins+hpower+weight   # Formula for true expected Y-outcomes...
lmobj <- lm(formY, MLout$new)
max(abs(lmobj$residuals))   # essentially 0 because linear model is "correct"...
# effobj <- eff.ridge(formY, MLout$new)   ...generates "Error" because RSQUARE=1.
```

---

mpg   *Hocking(1976) Miles Per Gallon data: a Multiple Regression Bench-*
                                       mark

Description

Performance data for the 1973-74 models of 32 autos from the Road Tests Section of *Motor Trends* magazine. NOTE: data(mtcars) loads essentially the same data, but most variable names are then different and have fewer characters.

Usage

```
data(mpg)
```

Format

A data frame of 11 variables on 32 autos; no NAs.

cylnds number of cylinders.

cubins cubic inches of engine displacement.

hpower engine horsepower.

weight total weight in pounds.

mpg miles per gallon.

shape engine shape (0 = V, 1 = Straight).

transm transmission type (0 = Automatic, 1 = Manual).

speeds number of forward speeds.

carbs number of carburetors.

fratio final drive rtaio.

qmilt quarter mile time.

References

Hocking RA. The Analysis and Selection of Variables in Regression. Biometrics 1976; 32: 1-51.

Henderson HV, Velleman PF. Building multiple regression models interactively. Biometrics 1981; 37: 391-411.

---

plot.aug.lars	*Plot method for aug.lars objects*

Description

Plot TRACE displays for aug.lars regression coefficients. The default is to display all Five TRACEs on one page with no legend.

Usage

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

Arguments

x	Output list object of class aug.lars.
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 <- aug.lars(form, data=longley2)
plot(rxlobj)
```

plot.eff.biv	*Plot method for eff.biv objects*

Description

Plot One or Two Bivariate Confidence Ellipses or a TRACE display for efficiently shrunken regression coefficients. The default is to display two ellipses at confidence levels 0.95 and 0.50 on a single plot for the first 2 of p >= 2 coefficients. The projection of the eff.ridge() Path from the OLS solution (at m = 0) to the Knot at the most likely Minimum MSE Risk solution (at m = mStar) ...plus the continuation of this Path towards the Shrinkage Terminus at (0,0) is shown in "red". To show the (outer) Ellipse as large as possible, the (0,0) point may be "off the plot".

Usage

```
   ## S3 method for class 'eff.biv'
plot(x, type = "ellip", ...)
```

Arguments

x	Output list object of class eff.ridge.
type	One of 2 options: "ellip" or "trace". The default option of "ellip" displays the Confidence Ellipse(s) specified by arguments to eff.biv().
...	Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

```
# Cost-Effectiveness inferences using Linear Models and eff.ridge...
## Not run:
  library(ICEinfer)
  data(sepsis)
  ndr <- ICEpref(sepsis$icu, sepsis$qalypres, sepsis$totcost, lambda=50000, beta=0.1)
  sndr <- data.frame(cbind(ndr$pref, sepsis)) # ndr: non-linear diminishing returns...
  form4 <- ndr.pref ~ icu + age + orgfails + apache
  usndra <- eff.aug(eff.ridge(form4, sndr)) # compare ndr of 2 Intensive Care Units...
  plot(efobj <- eff.biv(usndra, 2, 4))
  efobj           # implicit print...
  # plot(efobj, type = "tr")

## End(Not run)
```

---

plot.eff.ridge	*Plot method for eff.ridge objects*

Description

Plot TRACE displays for the Efficient Path. This path is [1] as short as possible and [2] consists of a Two-Piece Linear-Function with a emphSingle Interior Knot at the MSE Optimal m-Extent of Shrinkage. The default is to display all five Traces on one page with no legend. This function requires the number, p, of non-constant x-variables be at most 30.

Usage

```
   ## S3 method for class 'eff.ridge'
plot(x, trace = "all", LP = 0, HH = 0, ... )
```

Arguments

x	RXshrink output list object of class eff.ridge.
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.
LP	The "Legend Position" must be an integer between 0 and 9, inclusive: LP = 0 is the default and yields NO legend display, LP = 1 displays a crude legend at the Lower-Right of each trace plot, ... , LP = 7 displays the legend in the Upper-Right position, . , LP = 9 displays the legend in the Middle-Middle position.
HH	The "Half Height" plot option is an integer between 0 and 2, inclusive: HH = 0 is the default and yields trace plots with either 1 or 3 rows, HH = 1 calls par(mfrow=c(2,1)) before displaying the current trace plot, HH = 2 displays a 2nd (or sebsequent) trace in "Half Height" mode. NOTE that these "Half Height" options allow users to display both a qm.ridge() and an eff.ridge() trace on a single 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
rxefobj <- eff.ridge(form, data=longley2)
plot(rxefobj, "coef", LP=7)
```

---

plot.qm.ridge   *Plot method for qm.ridge objects*

Description

Plot TRACE displays for 2-parameter (q, m) generalized ridge estimators. The default is to display all five Traces on one page with no legend. When the number, p, of non-constant x-variables is at most 30, the lines [straight or curved] displayed on Traces avoid (or postpone) use of lty = 3 [dotted].

Usage

```
   ## S3 method for class 'qm.ridge'
plot(x, trace = "all", LP = 0, HH = 0, ... )
```

Arguments

x	Output list object of class qm.ridge.
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.
LP	The "Legend Position" must be an integer between 0 and 9, inclusive: LP = 0 is the default and yields NO legend display, LP = 1 displays a crude legend at the Lower-Right of each trace plot, ... , LP = 7 displays the legend in the Upper-Right position, . , LP = 9 displays the legend in the Middle-Middle position.
HH	The "Half Height" plot option is an integer between 0 and 2, inclusive: HH = 0 is the default and yields trace plots with either 1 or 3 rows, HH = 1 calls par(mfrow=c(2,1)) before displaying the current trace plot, HH = 2 displays a 2nd (or sebsequent) trace in "Half Height" mode. NOTE that these "Half Height" options allow users to display both a qm.ridge() and an eff.ridge() trace on a single 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
rxrobj <- qm.ridge(form, data=longley2)
plot(rxrobj, "rmse", LP = 5)
```

---

plot.RXpredict   *Plot method for RXpredict objects*

Description

Plot Predicted and/or Fitted.Values for all 5 RXshrink regression estimation methods. The default is to plot Predictions for the y-Outcome variable.

Usage

```
   ## S3 method for class 'RXpredict'
plot(x, fit = "yvecprd", ... )
```

Arguments

x	Output list object of class RXpredict.
fit	One of three possible options:, "yvecprd" to display Predictions of the Observed y-Outcomes in a single plot, "cryprd" to display Fitted.Values for the Centered and Rescaled y-Outcomes, "both" to display "yvecprd" and "cryprd" plots in two rows on one page.
...	Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

Examples

```
data(longley2)
form <- Employed~GNP+GNP.deflator+Unemployed+Armed.Forces+Population+Year
rxefobj <- eff.ridge(form, longley2)
rxefprd <- RXpredict(rxefobj, longley2)
plot(rxefprd)
 # Clearly Biased predictions can still represent an "Optimal" Variance-Bias Trade-Off...
```

---

plot.syxi	*Plot method for syxi objects*

Description

Display plots for Visual Validation of the syxi() approach that attempts to straighten nonlinear relationships between two variables.

Usage

```
   ## S3 method for class 'syxi'
plot(x, type = "xy", ...)
```

Arguments

x	syxi output list object of class syxi.
type	One of two possible options: "xy" to display a plot of the x-predictor versus the y-outcome with the lm() fitted line in RED and the corresponding gam() Spline fit in BLUE, "sy" (or any string different from "xy") displays a plot of s(x) versus y with only the lm() fit in BLUE.
...	Optional argument(s) passed on to plot().

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

---

plot.uc.lars	*Plot method for uc.lars objects*

Description

Plot TRACE displays of lars regression coefficients for the uncorrelated components of the X-matrix. The default option of trace = "all" displays all Five TRACEs on one page.

Usage

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

Arguments

x	Output list object of class uc.lars.
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-right 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
rxucobj <- uc.lars(form, data=longley2)
plot(rxucobj)
```

---

plot.YonX   *Plot method for YonX objects*

Description

Graphics for Shrinkage in "Simple" Linear Regression: Models with only p=1 X-variable. The default is to [1] display four TRACES in one plot and then to [2] display the "Y on X" scatterplot with Three Fitted Lines: the OLS fit is BLUE, the optimally Shrunken fit is "purple", and the largest m-Extent of shrinkage with Relative MSE at most that of OLS is "red". Note: Whenever model lack-of-fit is small, these 3 m-Extents can be quite close to each other.

Usage

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

Arguments

x	Output list object of class "YonX".
trace	One of EIGHT possible options: "all" to display 4 traces on the first plot, then the "YonX" scatter plot; "seq" to display 5 full-sized plots in sequence (in response to user prompts); "coef" to display only the estimated beta coefficient trace (a straight line); "rmse" to display only the (quadratic) estimated relative MSE risk trace; "spat" to display only the delta-factor trace (a straight line); "lglk" to display only the "-2 log(Likelihood Ratio)" trace; "YonX" to display only the Y-vs-X scatter plot with 3 fitted-lines: the OLS fitted line (BLUE), the "purple" Maximum-Likelihood (optimally biased) line, and the "red" line marking the most shrinkage with estimated MSE Relative Risk less than or equal to that of OLS; or "exev" to display only the Excess Eigenvalue trace that is redundant with the "rmse" trace.
...	Optional argument(s) passed on to plot().

Details

The effects of Shrinkage on Simple Linear Regression models (p = 1) are, in reality, NOT easier to illustrate than the corresponding effects on Multiple Linear Regressiom models (p > 1). In both situations, alternative estimates of effects and risks abound. For example, the estimate chosen can be [1] Maximum Likelihood under Normal-theory, [2] Unbiased under Normal-theory or [3] have "Correct Range". [See Obenchain (1978), equations (3.3) to (3.5), and corresponding text.] When a graphic contains only a single curve, a "reality" is that the general "shape" of the curve (plus any highlighted "points" on that curve) should "look right" or, at the very least, "reasonable".

My choices among alternative estimates of (nonlinear) MSE risk were initially made roughly 30 years ago ...and have remained mostly unchanged (primarily) for consistency with earlier versions of RXshrink.

The NEW "rmse" TRACE for class "YonX" displays MSE Relative Risk estimates from the "qrsk" vector rather than the (traditional) "rmse" estimates for all p = 1 models. This allows MSE Relative Risk estimates to *satisfy a Quadratic equation* and give the CORRECT visual impression that Relative Risk is MINIMIZED at the ML "purple" point and dotted-line at m = (1-dMSE). Note that the Relative Risk then starts to increase for m > (1-dMSE) and returns to its initial starting level at m = 0 ["blue" point and dotted-line] when the m-Extent reaches m = 2*(1-dMSE) ["red" point and dotted-line.] Finally, when 0.5 < dMSE < 1, the Relative Risk then continues to increase, reaching its Maximum at m = 1. As argued in Obenchain(1978), the "Good" Shrinkage Range is 0 < m < 2*(1-dMSE), between the "blue" and "red" vertical dotted-lines on the "rmse" TRACE.

Value

NULL

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

Examples

```
data(haldport)
form <- heat ~ p4caf
YXobj <- YonX(form, data=haldport)
plot(YXobj)
```

---

qm.ridge   *Restricted (2-parameter) Maximum Likelihood Shrinkage in Regression*

Description

Computes and displays TRACEs for a **q**-Shaped shrinkage PATH, including the **m**-Extent of shrinkage along that path, most likely under Normal-distribution theory to yield optimal reducions in MSE Risk. When rank(X-matrix) = p > 2, the most likely solution on the qm.ridge() path will be less likely to have minimal MSE risk than the optimal solution on the p-parameter eff.ridge() path. The Hoerl-Kennard "ordinary" ridge path has Shape q=0 within the qm.ridge() family.

Usage

```
qm.ridge(form, data, rscale = 1, Q = "qmse", steps = 20, 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 (after being "centered") to remove all "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 = 20.)
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 value for one-minus-delta (default = 9.9e-013.)

Details

Traditional qm.ridge() paths cannot be overall-optimal when $p > 2$ because they are restricted to using strictly "monotone" (increasing or decreasing) shrinkage factors. Still, the "best" m-Extent of qm-shrinkage is marked by a vertical dashed-line on all 5-types of qm.ridge() TRACE displays. Compared to OLS estimates, these shrunken estimates have higher likelihood of reduced MSE risk and 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 qm.ridge:

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.
mx	Matrix containing mean values of X-predictors.
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.
mClk	Most-Likely Extent of Shrinkage Observed: best multiple of (1/steps) <= p.

minC	Minimum Observed Value of Normal-theory -2*log(Likelihood-Ratio).
QS	Was a Mesh-Search for Best q-Shape requested? : 1 => Yes, 0 => No.
qML	Computable only when p=rank=2: True Most-Likely q-Shape.
kML	Computable only when p=rank=2: True Most-Likely k-Factor.
dML1	Computable only when p=rank=2: True Most-Likely Delta[1]-Factor.
dML2	Computable only when p=rank=2: True Most-Likely Delta[2]-Factor.
mML	Computable only when p=rank=2: True Most-Likely m-Extent.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Burr TL, Fry HA. (2005) Biased Regression: The Case for Cautious Application. *Technometrics* **47**, 284-296.

Goldstein M, Smith AFM. (1974) Ridge-type estimators for regression analysis. *J. Roy. Stat. Soc. B* **36**, 284-291. (The 2-parameter shrinkage family.)

Obenchain RL. (1975) Ridge Analysis Following a Preliminary Test of the Shrunken Hypothesis. *Technometrics* **17**, 431-441. doi:10.1080/00401706.1975.10489369

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. <doi:10.1214/aos/1176344314>

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108 [Best q-Shape when p = 2.]

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.3. http://localcontrolstatistics.org

See Also

eff.ridge, correct.signs, MLtrue and RXpredict.

Examples

```
data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxrobj <- qm.ridge(form, data=longley2)
rxrobj
plot(rxrobj)
str(rxrobj)
```

RXpredict	*Predictions from Models fit using RXshrink Generalized Ridge Estimation Methods.*

Description

RXpredict() makes in-sample predictions (i.e. computes "fitted.values") for all 6 forms of RXshrink estimation either at some user-specified m-Extent of Shrinkage, such as m=0.963, or at the Normal distribution-theory m-Extent most likely to achieve minimum Risk (minMSE).

Usage

```
RXpredict(x, data, m="minMSE", rscale=1)
```

Arguments

x	An object output by one of the 6 RXshrink estimation functions. Thus class(x) must be "qm.ridge", "eff.ridge", "aug.lars", "uc.lars", "MLcalc" or "correct.signs".
data	Existing data.frame containing observations on all variables used by the RXshrink function for estimation of regression coefficients.
m	The m argument can be either [i] a single "numeric" value that is non-negative and does not exceed rank(X) or [ii] the (default) string "minMSE" to request use of the observed m-Extent of shrinkage most likely to be MSE optimal under Normal distribution-theory. For example, m="0.0" requests use of the (unbaised) OLS estimate [BLUE].
rscale	One of two possible choices (0 or 1) for "rescaling" of variables (after being "centered") to remove all "non-essential" ill-conditioning. Use "rscale=0" only when the RXshrink estimation function that computed the x-object also used "rscale=0". The default of "rscale=1" should be used in all other cases.

Value

An output list object of class RXpredict:

cryprd	Predicted values for the "centered" and POSSIBLY "rescaled" outcome y-vector, cry. These values correspond, for example, to the default "predicted.values" from lm().
cry	This the "centered" and POSSIBLY "rescaled" outcome y-vector from the input data.frame.
yvecprd	Predicted values for the Y-outcome variable, yvec.
yvec	The Y-outcome vector from the input data.frame specified by the "data" argument.
m	"numeric" Value of m-Extent implied by the call to RXpredict(), possibly via a default call with m="minMSE". Restriction: 0 <= m <= rank(X).
mobs	Observed m-Extent most close to the requested m-Extent AND is on the lattice of observed m-Extents stored within the given x-object.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2005) *Shrinkage Regression: ridge, BLUP, Bayes, spline and Stein.* Electronic book-in-progress (185+ pages.) http://localcontrolstatistics.org

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.2. http://localcontrolstatistics.org

See Also

qm.ridge, eff.ridge and MLtrue.

Examples

```
data(tycobb)
form <- batavg~atbats+seasons+CMspl
rxefobj <- eff.ridge(form, data=tycobb)
tycfit <- RXpredict(rxefobj, tycobb, m="minMSE")
plot(tycfit)
tycobb$batavg[18]  # Ty Cobb's batavg = 0.401 in 1922
abline(h=tycfit$cry[18], lty=2, lwd=3, col="red")
```

---

syxi   *Linear and GAM Spline Predictions from a Single x-Variable*

Description

Compute and display (x,y) plots with their linear and gam() spline y-predictions.

Usage

```
syxi(form, data, i = 1)
```

Arguments

form	A "simple" regression formula [y~x] suitable for use with lm().
data	data.frame containing at least 10 observations on both variables in the formula.
i	A single integer "index" within 1:25.

Details

The gam() functon from the *mgcv* R-package is used to compute and, subsequently, to generate plots that visually compare the "linear" fit from lm(y~x) with a potentially "nonlinear" fit using smoothing parameters. The horizontal axis on type = "sy" plots gives potentially "straightened out" x numerical values.

Value

An output list object of class syxi:

dfname	Name of the data.frame object specified as the second argument.
xname	"xi" as Two or Three Characters.
sxname	"si" as Two or Three Characters.
dfsxf	A data.frame containing 3 variables: "yvec", "xvec", and "sxfit".
yxcor	Pearson correlation between "yvec" and "xvec".
yscor	Pearson correlation between "yvec" and "sxfit".
xscor	Pearson correlation between "xvec" and "sxfit".
lmyxc	lm() Coefficients (intercept and slope) for y ~ x.
lmysc	lm() Coefficients (intercept and slope) for y ~ sxfit.
adjR2	Adjusted R2 value from gam.sum$r.sq.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2023) Nonlinear Generalized Ridge Regression. *arXiv preprint* https://arxiv.org/abs/2103.05161

Examples

```
library(mgcv)
data(longley2)
form = GNP ~ Year
GNPpred = syxi(form, data=longley2, i = 1)
plot(GNPpred, type="xy")
title(main="y = GNP on x1 = Year")
plot(GNPpred, type="sy")
title(main="y = GNP on Spline for Year")
```

---

tycobb	*Ty Cobb batting statistics for 1905–1928 with Carl Morris' 2-piece Spline term.*

Description

Linear Regression models can be used to predict Ty Cobb's Expected true yearly batting averages from his observed yearly "batavg" and the 5 other variables stored in the "tycobb" data.frame. Predictions from such "models" can address the question: "Was Ty Cobb ever a TRUE .400 hitter?" Since a player's seasonal batavg is a "random variable," the fact that Cobb's batavg was 0.4196 in 1911 and 0.4105 in 1912 does not necessarily imply that his Expected Seasonal batavg was truly over .400 in either of those consecutive seasons. For example, his batavg was 0.4011 in 1922 (10 years later). However, his seasonal batavg had dipped to "only" 0.3341 in 1920.

"Cobb lived off the field as though he wished to live forever. He lived on the field as though it was his last day." – Branch Rickey, Major League Baseball Hall of Fame executive.

Usage

```
data(tycobb)
```

Format

A data frame with 24 observations (years) on the following 6 variables.

year Ty Cobb's 24 American League Seasons: 1905 - 1928.

hits Total number of Hits that season.

atbats Total number of times at Bat that season.

CMspl Carl Morris' Piecewise-Linear Spline term with "knot" in season 6 (1910).

seasons A linear "Trend" term: 1, 2, ..., 24.

batavg Cobb's Seasonal Batting Average ... 7 decimal places.

References

Carl Morris. (1982). "Was Ty Cobb ever a TRUE .400 hitter?" One-page Handout for his JSM Lecture on August 18 in Cincinnati, Ohio.

---

uc.lars

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

```
uc.lars(form, data, rscale = 1, type = "lar", trace = FALSE,
    eps = .Machine$double.eps, 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 (after being "centered") to remove all "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 uc.lars() 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.)

Details

uc.lars() 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/abs[PC(i)]), where 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]. MCAL=0 occurs at k=0, while MCAL = p results when k is the maximum absolute principal correlation.

Value

An output list object of class uc.lars:

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.
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.
mClk	Most Likely Extent of Shrinkage Observed: best multiple of (1/steps) <= p.
minC	Minimum Observed Value of Normal-theory -2*log(Likelihood).

Author(s)

Bob Obenchain <wizbob@att.net>

References

Hastie T, Efron, B. (2013) lars: Least Angle Regression, Lasso and Forward Stagewise. ver 1.2, <https://CRAN.R-project.org/package=lars>

Obenchain RL. (1994-2005) Shrinkage Regression: ridge, BLUP, Bayes, spline and Stein. <http://localcontrolstatistics.org>

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. <http://localcontrolstatistics.org>

See Also

aug.lars.

Examples

```
data(longley2)
form <- GNP~GNP.deflator+Unemployed+Armed.Forces+Population+Year+Employed
rxucobj <- uc.lars(form, data=longley2)
rxucobj
plot(rxucobj)
str(rxucobj)
```

YonX   *Maximum Likelihood (ML) Shrinkage in Simple Linear Regression*

Description

Compute and display Normal-theory ML Shrinkage statistics when a y-Outcome Variable is re-
gressed upon a SINGLE x-Variable (i.e. $p = 1$). This illustration is usefull in regression pedagogy.
The OLS (BLUE) estimate is a scalar in these simple cases, so the MSE optimal Shrinkage factor,
dMSE, is also a scalar less than +1 and greater than 0 when cor(y,x) differs from Zero. The cor-
responding m-Extent of Optimal Shrinkage is marked by the "purple" vertical dashed-line on all
YonX() TRACE Diagnostics.

Usage

```
    YonX(form, data, delmax = 0.999999)
```

Arguments

form   A regression formula [y ~ x] suitable for use with lm().

data   Data frame containing observations on both variables in the formula.

delmax   Maximum allowed value for Shrinkage delta-factor that is strictly less than 1.
                (default = 0.999999, which prints as 1 when rounded to fewer than 6 decimal
                places.)

Details

Since only a single x-Variable is being used, these "simple" models are (technically) NOT "Ill-
conditioned". Of course, the y-Outcome may be nearly multi-collinear with the given x-Variable,
but this simply means that the model then has low "lack-of-fit". In fact, the OLS estimate can never
have the "wrong" numerical sign in these simple $p = 1$ models! Furthermore, since "risk" estimates
are scalar-valued, no "exev" TRACE is routinely displayed; its content duplicates information in the
"rmse" TRACE. Similarly, no "infd" TRACE is displayed because any "inferior direction" COSINE
would be either: +1 ("upwards") when an estimate is decreasing, or -1 ("downwards") when an
estimate is increasing. The m-Extent of shrinkage is varied from 0.000 to 1.000 in 1000 "steps" of
size 0.001.

Value

An output list object of class YonX:

data   Name of the data.frame object specified as the second argument.

form   The regression formula specified as the first argument to YonX() must have only
                ONE right-hand-side X-variable in calls to YonX().

p   Number of X-variables MUST be $p = 1$ in YonX().

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 for error.
prinstat	Vector of five Principal Statistics: eigval, sv, b0, rho & tstat.
yxnam	Character Names of "Y" and "X" data vectors.
yvec	"Y" vector of data values.
xvec	"X" vector of data values.
coef	Vector of Shrinkage regression Beta-coefficient estimates: delta * B0.
rmse	Vector of Relative MSE Risk estimates starting with the rmse of the OLS estimate.
spat	Vector of Shrinkage (multiplicative) delta-factors: 1.000 to 0.000 by -0.001.
qrsk	Vector of Quatratic Relative MSE Risk estimates with minimum at delta = dMSE.
exev	Vector of Excess Eigenvalues = Difference in MSE Risk: OLS minus GRR.
mlik	Normal-theory Likelihood ...for Maximum Likelihood estimation of Shrinkage m-Extent.
sext	Listing of summary statistics for all M-extents-of-shrinkage.
mUnr	Unrestricted optimal m-Extent of Shrinkage from the dMSE estimate; mUnr = 1 - dMSE.
mClk	Most Likely Observed m-Extent of Shrinkage: best multiple of (1/steps) <= 1.
minC	Minimum Observed Value of CLIK Normal-theory -2*log(Likelihood-Ratio).
minE	Minimum Observed Value of EBAY (Empirical Bayes) criterion.
minR	Minimum Observed Value of RCOF (Random Coefficients) criterion.
minRR	Minimum Relative Risk estimate.
mRRm	m-Extent of the Minimum Relative Risk estimate.
mReql	m-Extent where the "qrsk" estimate is first >= the observed OLS RR at m = 0.
Phi2ML	Maximum Likelihood estimate of the Phi-Squared noncentrality parameter of the F-ratio for testing H: true beta-coefficient = zero.
Phi2UB	Unbiased Phi-Squared noncentrality estimate. This estimate can be negative.
dALT	This Maximim Likelihood estimate of Optimal Shrinkage has serious Downward Bias.
dMSE	Best Estimate of Optimal Shrinkage Delta-factor from the "Correct Range" adjustment to the Unbiased Estimate of the NonCentrality of the F-ratio for testing Beta = 0.

Author(s)

Bob Obenchain <wizbob@att.net>

References

Obenchain RL. (1978) Good and Optimal Ridge Estimators. *Annals of Statistics* **6**, 1111-1121. doi:10.1214/aos/1176344314

Obenchain RL. (2022) Efficient Generalized Ridge Regression. *Open Statistics* **3**: 1-18. doi:10.1515/stat20220108

Obenchain RL. (2022) *RXshrink_in_R.PDF* RXshrink package vignette-like document, Version 2.1. http://localcontrolstatistics.org

See Also

[correct.signs](#) and [MLtrue](#)

Examples

```
data(haldport)
form <- heat ~ p4caf
YXobj <- YonX(form, data=haldport)
YXobj
plot(YXobj)
```

Index