Package 'forsearch'

January 5, 2023

Title Diagnostic Analysis Using Forward Search Procedure for Various Models

Version 3.0.1

Description Identifies potential data outliers and their impact on estimates and analyses. Uses the forward search approach of Atkinson and Riani, ``Robust Diagnostic Regression Analysis", 2000,<ISBN: o-387-95017-6>) to prepare descriptive statistics of a dataset that is to be analyzed by stats::lm(), stats::glm(), or nlme::lme(). Includes graphics functions to display the descriptive statistics.

Depends R (>= 4.2)

License GPL (>= 3)

SystemRequirements gmp (>= 4.1)

Encoding UTF-8

RoxygenNote 7.2.1

Imports Hmisc(>= 4.6-0), Cairo(>= 1.5-14), formula.tools(>= 1.7.0), ggplot2(>= 3.3.5), nlme(>= 3.1-152), tibble(>= 3.1.6)

Suggests rmarkdown, knitr

VignetteBuilder knitr

NeedsCompilation no

Author William Fairweather [aut, cre]

Maintainer William Fairweather <wrf343@flowervalleyconsulting.com>

Repository CRAN

Date/Publication 2023-01-05 17:50:08 UTC

R topics documented:

forsearch-package 2
aStep1 4
aStep2 5
bStep1 6
bStep2 7
forsearch_glm 8

forsearch_lm . . . 10
forsearch_lme . . . 12
identifyCoeffs . . . 15
identifyFixedCoeffs . . . 16
picksome . . . 17
plotdiag.AICX . . . 18
plotdiag.allgraphs . . . 19
plotdiag.Cook . . . 20
plotdiag.deviance.residuals . . . 21
plotdiag.deviances . . . 22
plotdiag.fit3 . . . 23
plotdiag.leverage . . . 24
plotdiag.params.fixed . . . 25
plotdiag.params.random . . . 26
plotdiag.phihatx . . . 28
plotdiag.residuals . . . 29
plotdiag.s2 . . . 30
plotdiag.tstats . . . 31
search.history . . . 32
showme . . . 32
showmegl . . . 33
showmelme . . . 34
variablelist . . . 34

Index **36**

`forsearch-package` *Diagnostic Analysis Using Forward Search Procedure for Various Models Diagnostic Analysis Using Forward Search Procedure for Various Models*

Description

Identifies potential data outliers and their impact on estimates and analyses. Uses the forward search approach of Atkinson and Riani, "Robust Diagnostic Regression Analysis", 2000,<ISBN: o-387-95017-6>) to prepare descriptive statistics of a dataset that is to be analyzed by stats::lm(), stats::glm(), or nlme::lme(). Includes graphics functions to display the descriptive statistics.

Details

The DESCRIPTION file:

Package:	forsearch
Title:	Diagnostic Analysis Using Forward Search Procedure for Various Models
Version:	3.0.1
Authors@R:	person("William","Fairweather", email = "wrf343@flowervalleyconsulting.com", role = c("aut", "cre"
Description:	Identifies potential data outliers and their impact on estimates and analyses. Uses the forward search a
Depends:	R (>= 4.2)

License:	GPL (>= 3)
LazyData:	true
SystemRequirements:	gmp (>= 4.1)
Encoding:	UTF-8
Roxygen:	list(markdown = TRUE)
RoxygenNote:	7.2.1
Imports:	Hmisc(>= 4.6-0), Cairo(>= 1.5-14), formula.tools(>= 1.7.0), ggplot2(>= 3.3.5), nlme(>= 3.1-152), tib
Suggests:	rmarkdown, knitr
VignetteBuilder:	knitr
Author:	William Fairweather [aut, cre]
Maintainer:	William Fairweather <wrf343@flowervalleyconsulting.com>

Index of help topics:

aStep1	Create Set of Observation Numbers in Step 1 for Linear Model Analysis
aStep2	Update Observation Number in Step 2
bStep1	Create Set of Observation Numbers in Step 1 for Linear Mixed Effects 'Model Analysis
bStep2	Update Observation Numbers in Step 2
forsearch-package	Diagnostic Analysis Using Forward Search Procedure for Various Models Diagnostic Analysis Using Forward Search Procedure for Various Models
forsearch_glm	Create Statistics of Forward Search in a Generalized Linear Model Database
forsearch_lm	Create Statistics Of Forward Search in a Linear Model Database
forsearch_lme	Create Statistics Of Forward Search For a Linear Mixed Effects Database
identifyCoeffs	Index To Identify Fixed and Random Coefficients To Appear Together on Plot
identifyFixedCoeffs	Index To Identify Fixed Coefficients To Appear Together on Plot
picksome	Structured Sets of Random Samples of Observations
plotdiag.AICX	Plot Diagnostic AIC Statistics
plotdiag.Cook	Plot Diagnostic Statistics of Modified Cook's Distance
plotdiag.allgraphs	Execute All Plotting Functions For a Select Forsearch Object
plotdiag.deviance.residuals	Plot Diagnostic Statistics Of Deviance Residuals
plotdiag.deviances	Plot Diagnostic Deviance Statistics
plotdiag.fit3	Plot Diagnostic Statistics of AIC, BIC, and Log Likelihood

plotdiag.leverage	Plot Diagnostic Statistics Of Leverage
plotdiag.params.fixed	Plot Diagnostic Statistics of Fixed Coefficients
plotdiag.params.random	Plot Diagnostic Statistics Of Random Coefficients
plotdiag.phihatx	Plot Diagnostic PhiHat Statistics
plotdiag.residuals	Plot Diagnostic Statistics Of Residuals Or Squared Residuals
plotdiag.s2	Plot Diagnostic Statistics Of Residual Variation
plotdiag.tstats	Plot Diagnostic T Statistics
search.history	Create Tabular History Of Forward Search
showme	Display Abbreviated Output Of FORSEARCH_LM Function
showmegl	Display Abbreviated Output Of FORSEARCH_GLM Function
showmelme	Display Abbreviated Output Of FORSEARCH_LME Function
variablelist	Identify Level(s) to Which Each Factor Observation Belongs

Ensure that data frame has a leading column of observation numbers. Run forsearch_xxx to create a file of diagnostic statistics to be used as input to such plotting functions as plotdiag.residuals, plotdiag.params.fixed, plotdiag.params.random, plotdiag.s2, plotdiag,leverage, and plotdiag.Cook. The file of diagnostic statistics can be voluminous, and utility functions of showme, showmelme, and showmegl (for lm, lme and glm analyses, respectively) display the output more succinctly. Plotting of statistics for fixed and for random coefficients is limited by graphical restraints in some cases. The function identifyCoeffs provides a set of indexing codes so that plotdiag.params.random can display diagnostics for selected fixed or random model parameters. The function identifyFixedCoeffs does the same for lm models.

Author(s)

William R. Fairweather, Flower Valley Consulting, Inc., Silver Spring MD USA NA

Maintainer: NA William R. Fairweather <wrf343 at flowervalleyconsulting.com>

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000. Pinheiro, JC and DM Bates. Mixed-Effects Models in S and S-Plus, Springer, New York, 2000.

---

aStep1 *Create Set of Observation Numbers in Step 1 for Linear Model Analysis*

Description

Derives the first set of observation numbers for forsearch in linear models

Usage

```
aStep1(yesfactor, data, inner.rank, initial.sample, formula, ycol, nopl)
```

Arguments

yesfactor	Logical. TRUE if there are factors in the X matrix
data	Data frame being analyzed by forward search.
inner.rank	Rank of X matrix of lm analysis on entire database
initial.sample	Number of random samples from which to take set of initial observations
formula	Fixed parameter formula of lm function
ycol	Response column number
nopl	Number of observations per level of combined factor variables

Details

Support function, usually not called independently

Value

Produces set of observation numbers for Step 1. Accounts for presence of factors in the dataset

Note

Presence of Observation column has no effect on outcome

Author(s)

William R. Fairweather

---

aStep2 *Update Observation Number in Step 2*

Description

Derives the next set of observation numbers for forsearch in linear models

Usage

```
aStep2(thislm, data, ycol, thisi)
```

Arguments

thislm	A lm object
data	Data frame being analyzed by forward search
ycol	Column number of response variable
thisi	Iteration number

Details

Support function, usually not called independently

Value

Vector of integers corresponding to observation numbers

Author(s)

William R. Fairweather

---

bStep1 *Create Set of Observation Numbers in Step 1 for Linear Mixed Effects 'Model Analysis*

Description

Derives the first set of observation numbers for forsearch_lme in linear mixed effects models

Usage

```
bStep1(fixed, yf, mnf, nOuter, yobs, s.o, nopl, nobs, i.s, fbg, b.d, verbose=TRUE)
```

Arguments

fixed	Fixed parameter formula of lm function
yf	Indicator of presence of factors in the X matrix
mnf	Maximum number of observations in any outer subgroup
nOuter	Number of outer subgroups
yobs	Column number of response variable
s.o	Vector of original observation numbers
nopl	Number of observations to take for each subgroup. Default is 1
nobs	Total number of observations in database
i.s	Number of random sets of observations in Step 1
fbg	Observation numbers for each outer subgroup
b.d	Numeric indicator of position in code to begin diagnostic prints
verbose	TRUE causes printing of function ID at beginning and end of run

Details

Support function, usually not called independently

Value

Produces set of observation numbers for Step 1. Accounts for presence of factors in the dataset

Author(s)

William R. Fairweather

`bStep2` *Update Observation Numbers in Step 2*

Description

Derives the next set of observation numbers for forsearch in linear mixed effects models

Usage

```
bStep2(fixed, nOuter, mnf, mstart, nobs, yobs, fbg, n.f, s.o, ras, b.d, verbose)
```

Arguments

`fixed`	Fixed parameter formula of lm function
`nOuter`	Number of outer subgroups
`mnf`	Maximum number of observations in an outer subgroup
`mstart`	Number of observations in each outer subgroup
`nobs`	Number of observations in entire database
`yobs`	Column number of response variable
`fbg`	List of observation numbers by outer subgroup
`n.f`	Vector of number of observation in each outer subgroup
`s.o`	Original observation numbers prior to renumbering in each outer subgroup
`ras`	List of observation numbers in each outer subgroup
`b.d`	Indicator of place in code to begin diagnostic printouts
`verbose`	TRUE causes printing of function ID at beginning and end of run

Details

Support function, usually not called independently

Value

List of expanding number sets corresponding to observation numbers

Author(s)

William R. Fairweather

`forsearch_glm` *Create Statistics of Forward Search in a Generalized Linear Model Database*

Description

Prepares summary statistics at each stage of forward search for subsequent plotting. Forward search is conducted in three steps: Step 1 to identify minimal set of observations to estimate unknown parameters, and Step 2 to add one observation at each stage such that observations in the set are best fitting at that stage. A preliminary step (Step 0) contains code for pre-processing of the data.

Usage

```
forsearch_glm(initial.sample=1000, cobs, response.cols, indep.cols, family, data,
  n.obs.per.level = 1, estimate.phi = TRUE, skip.step1 = NULL, unblinded = TRUE,
  diagnose = FALSE, verbose = TRUE)
```

Arguments

`initial.sample`	Number of random sets of observations in Step 1 of forward search
`cobs`	Number of observations to include in each inner subgroup of Step 1.
`response.cols`	Vector of column numbers (2) of responses and nonresponses
`indep.cols`	Column number(s) of independent variables
`family`	Error distribution and link
`data`	Name of database
`n.obs.per.level`	Number of observations per level of (possibly crossed) factor levels
`estimate.phi`	TRUE causes phi to be estimated; FALSE causes phi to be set = 1
`skip.step1`	NULL, or vector of observation numbers to include at end of Step 1
`unblinded`	TRUE allows print of formula of analysis function
`diagnose`	TRUE causes printing of intermediate steps of function
`verbose`	TRUE causes function identifier to display before and after run

Details

Step 2 is determined by the results of Step 1, which itself is random. So, it is possible to reproduce the entire run by using the skip.step1 argument. Inner subgroups are produced by presence of categorical variables.

Value

LIST

`Rows in stage`	Observation numbers of rows included at each stage
`Family`	Family and link
`Number of model parameters`	Number of fixed effect parameters
`Fixed parameter estimates`	Matrix of parameter estimates at each stage
`Residual deviance`	Vector of deviances
`Null deviance`	Vector of null deviances
`PhiHat`	Vector of values of phi parameter
`Deviance residuals and augments`	Deviance residuals with indication of whether each is included in fit
`AIC`	Vector of AIC values
`Leverage`	Matrix of leverage of each observation at each stage
`Call`	Call to this function

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

```
# Train deaths (Atkinson and Riani, 2000)

Observation<-1:67




Year<-c(97,96,96,95,94,94,91,91,90,89,89,89,88,88,87,86,86,86,86,84,84,84
,84,84,83,83,82,81,81,80
,80,79,79,79,78,78,77,76,76,75,75,75,75,74,74,73,73,73,72,72,71,71,71
,71,70,69,69,69,69,69
,69,68,67,67,67,67,67)

RollingStock<-c(2,2,3,2,1,1,1,1,2,3,1,1,1,2,1,2,1,3,2,2,1,2,2,3,1,2
,1,1,2,3,1,1,1,1,1,1
,1,3,3,2,3,1,2,3,1,1,1,3,3,1,3,3,1,1,1,2,1,1,2
```

```
,1,1,1,1,1,1,1,1)

Traffic<-c(0.436,0.424,0.424,0.426,0.419,0.419,0.439,0.439,0.431,0.436,0.436,0.436,
0.443,0.443,0.397,0.414,0.414,0.414,0.414,0.389,0.389,0.389,0.389,0.389,0.401,
0.401,0.372,0.417,0.417,0.43,0.43,0.426,0.426,0.426,0.43,0.43,0.425,0.426,0.426,
0.436,0.436,0.436,0.436,0.452,0.452,0.433,0.433,0.433,0.431,0.431,0.444,0.444,
0.444,0.444,0.452,0.447,0.447,0.447,0.447,0.447,0.447,0.449,0.459,0.459,0.459,
0.459,0.459)

Deaths<-c(7,1,1,1,5,2,4,2,1,1,2,5,35,1,4,1,2,1,1,3,1,3,13,2,1,1,1
,4,1,2
,1,5,7,1,1,3,2,1,2,1,2,6,1,1,1,10,5,1,1,6,3,1,2,1,2,1,1,6,2,2
,4,2,49,1,7,5,9)

train <- data.frame(Observation,Year,RollingStock,Traffic,Deaths)
train<-train[1:25,]
train.out <- forsearch_glm(initial.sample=100, cobs=4, response.cols=5, indep.cols=2:4,
   family=poisson("log"),  data=train,  n.obs.per.level = 1,
   estimate.phi = TRUE, skip.step1 = NULL, unblinded = TRUE)
## Not run:
# Recode RollingStock as categorical variable. Note change in n.obs.per.level
#      argument
train2 <- train
train2[,3] <- as.factor(train2[,3])
train2.out <- forsearch_glm(initial.sample=100, cobs=4, response.cols=5, indep.cols=2:4,
   family=poisson("log"),  data=train2,  n.obs.per.level = 2,
   estimate.phi = TRUE, skip.step1 = NULL, unblinded = TRUE)

## End(Not run)
```

`forsearch_lm` *Create Statistics Of Forward Search in a Linear Model Database*

Description

Prepares summary statistics at each stage of forward search for subsequent plotting. Forward search is conducted in two steps: Step 1 to identify minimal set of observations to estimate unknown parameters, and Step 2 to add one observation at each stage such that observations in the set are best fitting at that stage.

Usage

```
forsearch_lm(formula, data, initial.sample=1000, n.obs.per.level = 1,
                   skip.step1 = NULL, unblinded = TRUE, diagnose = FALSE,
                   verbose = TRUE)
```

Arguments

`formula`	Fixed effects formula as described in stats::lm

data	Name of database
initial.sample	Number of observations in Step 1 of forward search
n.obs.per.level	Number of observations per level of (possibly crossed) factor levels
skip.step1	NULL or a vector of integers for observations to be included in Step 1
unblinded	TRUE causes printing of presumed analysis structure
diagnose	TRUE causes printing of intermediate steps of function
verbose	TRUE causes function identifier to display before and after run

Details

Step 2 is determined by the results of Step 1, which itself is random. So, it is possible to reproduce the entire run by using the skip.step1 argument.

Value

LIST

Rows in stage	Observation numbers of rows included at each stage
Standardized residuals	Matrix of errors at each stage
Number of model parameters	Rank of model
Sigma	Estimate of random error at final stage; used to standardize all residuals
Fixed parameter estimates	Matrix of parameter estimates at each stage
s^2	Estimate of random error at each stage
Leverage	Matrix of leverage of each observation at each stage
Modified Cook distance	Estimate of sum of squared changes in parameter estimates at each stage
Call	Call to this function

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

```
# Multiple regression
Observation <- 1:16
y <- runif(16)
x1 <- runif(16)
x2 <- runif(16)
x3 <- runif(16)
lmtest1 <- data.frame(Observation,y,x1,x2,x3)
forsearch_lm(formula=y~x1+x2+x3, data=lmtest1, initial.sample=100,diagnose=FALSE)
## Not run:
# Analysis of variance
Observation <- 1:30
y <- runif(30)
AN1 <- as.factor(c(rep("A1",5),rep("A2",5),rep("A3",5)))
AN1 <- c(AN1,AN1)
AN2 <- as.factor(c(rep("B1",15),rep("B2",15)))
lmtest2 <- data.frame(Observation,y,AN1,AN2)
forsearch_lm(formula=y~AN1*AN2, data=lmtest2, initial.sample=100,diagnose=TRUE)

# Analysis of covariance
Observation <- 1:60
y <- runif(60)
AN1 <- as.factor(c(rep("A1",10),rep("A2",10),rep("A3",10)))
AN1 <- c(AN1,AN1)
AN2 <- as.factor(c(rep("B1",30),rep("B2",30)))
COV <- runif(60)
lmtest3 <- data.frame(Observation,y,AN1,AN2,COV)
forsearch_lm(formula=y~AN1*AN2+COV, data=lmtest3, initial.sample=100,diagnose=FALSE)

## End(Not run)
```

`forsearch_lme` *Create Statistics Of Forward Search For a Linear Mixed Effects Database*

Description

Prepares summary statistics at each stage of forward search for subsequent plotting. Forward search is conducted in four steps: Step 0 to set up accounting for group structure, Step 1 to identify minimal set of observations to estimate unknown fixed parameters, Step 2 to identify the order of the remaining observations, and a final stage to extract the intermediate statistics based on increasing sample size.

Usage

```
forsearch_lme(fixed, data, random, formula, initial.sample=1000, n.obs.per.level=1,
   skip.step1=NULL,
   XmaxIter = 1000, XmsMaxIter = 1000, Xtolerance = 0.01, XniterEM = 10,
   XmsMaxEval = 400, XmsTol = 1e-05, Xopt = "optim", unblinded=TRUE,
   begin.diagnose = 1000, verbose = TRUE)
```

Arguments

fixed	2-sided formula for fixed effects
data	data frame, first column of which must be "Observation"
random	1-sided formula for random effects
formula	a simplified formula of the form resp ~ cov | group where resp is the response, cov is the primary covariate, and group is the grouping factor, as in nlme::groupedData
initial.sample	Number of observations in Step 1 of forward search
n.obs.per.level	Number of observations per level of (possibly crossed) factor levels
skip.step1	NULL or a vector of integers for observations to be included in Step 1
XmaxIter	lme control parameter
XmsMaxIter	lme control parameter
Xtolerance	lme control parameter
XniterEM	lme control parameter
XmsMaxEval	lme control parameter
XmsTol	lme control parameter
Xopt	lme control parameter
unblinded	TRUE causes printing of presumed analysis structure
begin.diagnose	Numeric indicator of place in coding to begin printing diagnostic information. 0 prints all information, 100 prints none.
verbose	TRUE causes function identifier to display before and after run

Details

data will be grouped within the function, regardless of initial condition. Steps 1 and 2 are based on formula argument and use of lm() function on each outer subgroup defined in formula. Intermediate data is extracted based on fixed and random arguments. Step 2 is determined by the results of Step 1, which itself is random. So, it is possible to reproduce the entire run by using the skip.step1 argument.

Value

LIST

Number of rows included in Step 1	Number of observations included in Step 1
Rows by subgroup	List of row numbers, by subgroup
Rows in stage	Observation numbers of rows included at each stage
Sigma	Estimate of random error at final stage; used to standardize all residuals
Standardized residuals	Matrix of errors at each stage

Fixed parameter estimates	Matrix of parameter estimates at each stage
Random parameter estimates	Matrix of parameter estimates at each stage
Leverage	Matrix of leverage of each observation at each stage
Modified Cook distance	Estimate of sum of squared changes in parameter estimates at each stage
Dims	Dims from fit of lme function
t statistics	t statistics for each fixed parameter
Fit statistics	AIC, BIC, and log likelihood
Call	Call to this function

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000. Pinheiro, JC and DM Bates. Mixed-Effects Models in S and S-Plus, Springer, New York, 2000. https://CRAN.R-project.org/package=nlme

Examples

```
# Multiple regression in grouped data
Observation <- 1:16
y <- runif(16)
x1 <- runif(16)
x2 <- runif(16)
x3 <- runif(16)
group <- rep(c("G1","G2"),each=8)
lmetest1 <- data.frame(Observation,y,x1,x2,x3,group)
lmetest1 <- nlme::groupedData(formula = y ~ 1 | group, data = lmetest1)
forsearch_lme(fixed=y~x1+x2+x3,data=lmetest1,random= ~ 1 | group,
   formula = y ~ 1 | group, initial.sample=100)
## Not run:
# Analysis of variance in grouped data
Observation <- 1:60
y <- runif(60)
AN1 <- as.factor(c(rep("A1",5),rep("A2",5),rep("A3",5)))
AN1 <- c(AN1,AN1,AN1,AN1)
AN2 <- as.factor(c(rep("B1",15),rep("B2",15)))
AN2 <- c(AN2,AN2)
group <- rep(c("G1","G2"),each=30)
lmetest2 <- data.frame(Observation,y,AN1,AN2,group)
lmetest2 <- nlme::groupedData(formula = y ~ 1 | group, data = lmetest2)
forsearch_lme(fixed=y~AN1*AN2,data=lmetest2,random= ~1 | group,
             formula= y ~ 1 | group,initial.sample=100)

# Analysis of covariance in grouped data
```

```
Observation <- 1:120
y <- runif(120)
AN1 <- as.factor(c(rep("A1",10),rep("A2",10),rep("A3",10)))
AN1 <- c(AN1,AN1,AN1,AN1)
AN2 <- as.factor(c(rep("B1",10),rep("B2",10)))
AN2 <- c(AN2,AN2,AN2,AN2,AN2,AN2)
COV <- runif(120)
lmetest3 <- data.frame(Observation,y,AN1,AN2,COV,group)
lmetest3 <- nlme::groupedData(formula = y ~ 1 | group, data = lmetest3)
forsearch_lme(fixed=y~AN1*AN2+COV,data=lmetest3,random= ~ 1 | group,
        formula= y ~ 1 | group, initial.sample=100)

## End(Not run)
```

`identifyCoeffs` *Index To Identify Fixed and Random Coefficients To Appear Together on Plot*

Description

Runs the defined, grouped linear mixed effects (lme) model. Displays the resulting fixed and random coefficients. Attaches codes for identifying them to the plotting functions of this package.

Usage

```
identifyCoeffs(fixed, data, random,
    XmaxIter = 1000, XmsMaxIter = 1000,
    Xtolerance = 0.01, XniterEM = 1000, XmsMaxEval = 400, XmsTol = 1e-05,
    Xopt = "optim",
    diagnose = FALSE, verbose = TRUE)
```

Arguments

`fixed`	2-sided formula for fixed effects
`data`	Name of file (to be) run by forsearch_lme
`random`	1-sided formula for random effects
`XmaxIter`	lme control parameter
`XmsMaxIter`	lme control parameter
`Xtolerance`	lme control parameter
`XniterEM`	lme control parameter
`XmsMaxEval`	lme control parameter
`XmsTol`	lme control parameter
`Xopt`	lme control parameter
`diagnose`	If TRUE, displays code to help diagnose main function errors
`verbose`	If TRUE, indicates beginning and end of function

Details

Plotting functions cannot plot more than a few coefficients on one graph. This function prepares an index of the coefficients so that the user can more easily identify which ones should appear together in a plot.

Value

Index of fixed and random coefficients from forsearch_lme.

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

`identifyFixedCoeffs` *Index To Identify Fixed Coefficients To Appear Together on Plot*

Description

Runs the defined linear (lm) model. Displays the resulting coefficients. Attaches codes for identifying them to the plotting functions of this package.

Usage

```
identifyFixedCoeffs(formula, data, diagnose = FALSE, verbose = TRUE)
```

Arguments

`formula`	2-sided formula for fixed effects
`data`	Name of file (to be) run by forsearch_lm
`diagnose`	If TRUE, displays code to help diagnose main function errors
`verbose`	If TRUE, indicates beginning and end of function

Details

Plotting functions cannot plot more than a few coefficients on one graph. This function prepares an index of the coefficients so that the user can more easily identify which ones should appear together in a plot.

Value

Index of coefficients from forsearch_lm.

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

`picksome` *Structured Sets of Random Samples of Observations*

Description

Restricts Step 1 of forward search procedures to ensure that every possible combination of levels is included for observations containing factors

Usage

```
picksome(subsetlist, nobs, initial.sample, n.obs.per.level, rank,
     verbose)
```

Arguments

`subsetlist`	List, each element is a data frame of 2 columns with code indicating the highest possible level of interaction to which each observation can belong. Usually, output from variablelist function.
`nobs`	Number of observations in data frame containing observations of forward search.
`initial.sample`	Number of randomized sets of observations in Step 1 of forward search.
`n.obs.per.level`	Number of observations to pull from each level.
`rank`	Rank of the X matrix of the analytical function to be used on data frame.
`verbose`	If TRUE, indicates beginning and end of function run.

Details

Support function, usually not called independently. Argument n.obs.per.level is set by user in forsearch_xxx function call.

Value

Matrix, each row of which identifies observations in each set of random sample of observations.

Author(s)

William R. Fairweather

Examples

```
V1 <- rep(c(3,7,9,11),times=4);
V2 <- V1+3;
V3 <- V1+90;
V1 <- as.factor(V1);
V2 <- as.factor(V2);
V3 <- as.factor(V3);
Observation <- 1:16;
testdf <- data.frame(Observation, V1,V2,V3);

testoutdf <- variablelist(testdf);
picksome(testoutdf,nobs=16,initial.sample=11,n.obs.per.level=1,rank=4,
   verbose=FALSE)
```

---

plotdiag.AICX *Plot Diagnostic AIC Statistics*

Description

Plot output from forsearch_glm to show change in AIC statistics as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.AICX(forn, maintitle = "Put main title here",
   subtitle = "Put subtitle here", caption="Put caption title here",
   wmf = "Put_plot_file_title_here",
   Cairo=TRUE, printgraph=TRUE,addline=c("none","loess","straight"),
   diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_glm
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit '.wmf'
Cairo	TRUE causes use of Cairo graphics

printgraph	TRUE causes graph to print to file and closes device
addline	add a line to the graph; abbreviation allowed
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot AIC statistics from forsearch_glm

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.allgraphs *Execute All Plotting Functions For a Select Forsearch Object*

Description

Executes all the plotting functions for a select analytical function such as lm or glm; default omits titles and subtitles and attempts to plot all fixed and random coefficients.

Usage

```
plotdiag.allgraphs(object, mt=" ", st=" ", cpt=" ", cc=NULL, ccrand = NULL,Cairo=TRUE)
```

Arguments

object	Name of forsearch object file
mt	Maintitle of graph
st	Subtitle of graph
cpt	Caption on the graph
cc	Fixed variable code numbers of coefficients to be included in graph
ccrand	Random variable code numbers of parameters to be included in graph
Cairo	TRUE causes use of Cairo graphics

Value

Prints search history and creates graphical files in current subdirectory

Author(s)

William R. Fairweather

Examples

```
## Not run:
info3 <- system.file("extdata", "train.for3.R", package="forsearch");

info3 <- source(info3);

info3 <- info3[[1]];

plotdiag.allgraphs(object=info3, mt=" ", st=" ", cpt=" ", cc=NULL, ccrand = NULL
   Cairo=FALSE)

## End(Not run)
```

---

plotdiag.Cook *Plot Diagnostic Statistics of Modified Cook's Distance*

Description

Plot output from forsearch_lm or forsearch_lme to show change in Modified Cook's distance as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.Cook(forn, maintitle = "Put main title here", subtitle = "Put subtitle here",
caption = "Put caption here", wmf = "Put_plot_file_title_here",
Cairo=TRUE,printgraph=TRUE, addline = c("none","loess","straight"),
diagnose = FALSE, verbose = TRUE)
```

Arguments

forn	Name of forward search output file
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
addline	add a line to the graph; abbreviation allowed
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot Cook distance statistics from forsearch_lm or forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.deviance.residuals

Plot Diagnostic Statistics Of Deviance Residuals

Description

Plot output from forsearch_glm to show change in deviance residuals or augmented deviance residuals, either of which can be squared, as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.deviance.residuals(forn, squared = FALSE, augmented=TRUE, hilos = c(1, 0),
maintitle="Put main title here", subtitle="Put subtitle here", caption="Put caption here",
wmf= "Put_graph_title_here", Cairo=TRUE,printgraph=TRUE,
legend = "Dummy legend name", diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of forward search output file
squared	TRUE causes residuals to be squared before plotting
augmented	TRUE causes graphing of augmented deviance residuals, see Details
hilos	Number of observations having high and number having low values of residuals to identify. No low values are identified for squared residual plot
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Caption of plot
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics

printgraph	TRUE causes graph to print to file and closes device
legend	Legend title
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Details

We reserve the use of the term 'Deviance residuals' to deviance residuals of the observations that were used to create the model fit, and use the term 'Augmented deviance residuals' to refer to deviance residuals of all available observations. The latter are created by predicting the fit of the model to all observations.

Value

Process and plot changes in deviance residuals or squared deviance residuals from forsearch_glm

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.deviances *Plot Diagnostic Deviance Statistics*

Description

Plot output from forsearch_glm to show change in deviances as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.deviances(forn, devtype, maintitle = "Put main title here",
subtitle = "Put subtitle here", caption="Put caption here",
wmf = "Put_plot_file_title_here",
Cairo=TRUE, printgraph=TRUE,addline=c("none","loess","straight"),
diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_glm
devtype	Type of deviance: "R" or "N" for Residual deviance or Null deviance
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
addline	add a line to the graph; abbreviation allowed
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot deviances from forsearch_glm

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.fit3 *Plot Diagnostic Statistics of AIC, BIC, and Log Likelihood*

Description

Plot output from forsearch_lme to show change in AIC, BIC, and log likelihood as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.fit3(forn, maintitle = "Put main title here", subtitle = "Put subtitle here",
    caption = "Put caption here", wmf = "Put_stored_name_here",
    Cairo=TRUE,printgraph=TRUE, legend="Dummy legend name",
    diagnose = FALSE, verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_lm
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
legend	Legend name
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot trends of AIC, BIC, and log likelihood statistics from forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.leverage *Plot Diagnostic Statistics Of Leverage*

Description

Plot output from forsearch_lm or forsearch_lme to show change in leverage of each observation as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.leverage(forn, hilos = c(1, 0), maintitle = "Put main title here",
subtitle = "Put subtitle here", caption="Put caption here",wmf = "Put_graph_title_here",
Cairo=TRUE, printgraph = TRUE,diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of forward search output file
hilos	Vector with number of highest observations and number of lowest observations on graph to identify
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot Cook distance statistics from forsearch_lm or forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.params.fixed *Plot Diagnostic Statistics of Fixed Coefficients*

Description

Plot output from forsearch_xxx to show change in random coefficients as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.params.fixed(forn, coeff.codenums=NULL, maintitle = "Put main title here",
subtitle = "Put subtitle here", caption="Put caption here",wmf = "Put_stored_name_here",
Cairo=TRUE, printgraph=TRUE,legend = "Dummy legend name",
diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_xxx
coeff.codenums	Numeric vector of coefficients to include together on the plot. Codes are output by identifyFixedCoeffs (for lm files) or by identifyCoeffs function (for lme files)
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
legend	Name of legend
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot fixed coefficient statistics from forsearch_lm or forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

---

plotdiag.params.random

Plot Diagnostic Statistics Of Random Coefficients

Description

Plot output from forsearch_lme to show change in root mean squares of random coefficients as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.params.random(forn, coeff.codenums=NULL,
asfacets=FALSE, facetdir=c("h","v"),
maintitle = "Put maintitle here", subtitle = "Put subtitle here",
caption = "Put caption here",wmf = "Put_stored_name_here",
Cairo=TRUE,printgraph = TRUE,
legend = "Dummy legend name", diagnose = FALSE, verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_lme
coeff.codenums	columns of output file to be included in graph
asfacets	TRUE causes printing in facets
facetdir	"v" lays out the facets vertically, "h" lays them out horizontally
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
legend	Name of legend
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot RMS of random coefficients from forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

plotdiag.phihatx *Plot Diagnostic PhiHat Statistics*

Description

Plot output from forsearch_glm to show change in phiHat statistics as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.phihatx(forn, maintitle = "Put main title here",
subtitle = "Put subtitle here", caption="Put caption here",
wmf = "Put_graph_filename_here",
Cairo=TRUE, printgraph=TRUE, addline=c("none","loess","straight"),
diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_glm
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
addline	add a line to the graph; abbreviation allowed
printgraph	TRUE causes graph to print to file and closes device
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot phiHat statistics from forsearch_glm

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

`plotdiag.residuals` *Plot Diagnostic Statistics Of Residuals Or Squared Residuals*

Description

Plot output from forsearch_lm or forsearch_lme to show change in residuals or squared residuals as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.residuals(forn, squared = FALSE, hilos = c(1, 0), maintitle, subtitle,
caption, wmf, Cairo=TRUE,printgraph=TRUE,
legend = "Dummy legend name", diagnose = FALSE, verbose = TRUE)
```

Arguments

`forn`	Name of forward search output file
`squared`	TRUE causes residuals to be squared before plotting
`hilos`	Number of observations having high and number having low values of residuals to identify. No low values are identified for squared residual plot.
`maintitle`	Main title of plot
`subtitle`	Subtitle of plot
`caption`	Caption of plot
`wmf`	File name of stored plot; omit ".wmf"
`Cairo`	TRUE causes use of Cairo graphics
`printgraph`	TRUE causes graph to print to file and closes device
`legend`	Legend title
`diagnose`	If TRUE, displays code to help diagnose main function errors
`verbose`	If TRUE, indicates beginning and end of function

Value

Process and plot changes in residuals or squared residuals from forsearch_lm or forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

plotdiag.s2 *Plot Diagnostic Statistics Of Residual Variation*

Description

Plot output from forsearch_lm to show change in residual variation as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.s2(forn, maintitle = "Put main title here", subtitle = "Put subtitle here",
    caption = "Put caption here", wmf = "Put_graph_filename_here",
    Cairo=TRUE,printgraph=TRUE, addline = c("none","loess","straight"),
    diagnose = FALSE, verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_lm
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
addline	add a line to the graph; abbreviation allowed
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot residual variation statistics from forsearch_lm

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

plotdiag.tstats *Plot Diagnostic T Statistics*

Description

Plot output from forsearch_lm or forsearch_lme to show change in t statistics as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.tstats(forn, coeff.codenums=NULL, maintitle = "Put main title here",
subtitle = "Put subtitle here", caption="Put caption here", wmf = "Put_stored_name_here",
Cairo=TRUE, printgraph=TRUE,legend = "Dummy legend name",
diagnose = FALSE,verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_lm or forsearch_lme
coeff.codenums	Numeric vector of coefficients to include together on the plot. Codes are output by identifyFixedCoeffs (for lm files) or by identifyCoeffs function (for lme files)
maintitle	Main title of plot
subtitle	Subtitle of plot
caption	Content of caption
wmf	File name of stored plot; omit ".wmf"
Cairo	TRUE causes use of Cairo graphics
printgraph	TRUE causes graph to print to file and closes device
legend	Name of legend
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot t statistics of fixed coefficients from forsearch_lm or forsearch_lme

Author(s)

William R. Fairweather

References

Atkinson, A and M Riani. Robust Diagnostic Regression Analysis, Springer, New York, 2000.

Examples

search.history *Create Tabular History Of Forward Search*

Description

The forward search functions output a list of vectors, each of which indicates which observations are in the model at each stage of the search. This function processes that list to create a more easily understood matrix of the observation numbers that are newly entered into the model and any that were temporarily removed from the model over the course of the search.

Usage

```
search.history(list1, diagnose = FALSE, verbose = TRUE)
```

Arguments

list1	Name of a forsearch_xxx output file
diagnose	If TRUE, displays code to help diagnose main function errors
verbose	If TRUE, indicates beginning and end of function

Value

Printout of matrix showing evolution of observations to enter or leave the model during the course of the forward search

Author(s)

William R. Fairweather

Examples

showme *Display Abbreviated Output Of FORSEARCH_LM Function*

Description

Output of forsearch_lm function can be voluminous. This function displays the output in an abbreviated format. Primarily for programmer use.

Usage

```
showme(x, verbose = TRUE)
```

Arguments

x	Name of forsearch_lm output file
verbose	If TRUE, indicates the beginning and end of function run

Value

Abbreviated printout of output of forsearch_lm function

Author(s)

William R. Fairweather

Examples

---

showmegl *Display Abbreviated Output Of FORSEARCH_GLM Function*

Description

Output of forsearch_glm function can be voluminous. This function displays the output in an abbreviated format. Primarily for programmer use.

Usage

```
showmegl(x, verbose = TRUE)
```

Arguments

x	Name of forsearch_glm output file
verbose	If TRUE, indicates the beginning and end of function run

Value

Abbreviated printout of output of forsearch_glm function

Author(s)

William R. Fairweather

Examples

showmelme *Display Abbreviated Output Of FORSEARCH_LME Function*

Description

Output of forsearch_lme function can be voluminous. This function displays the output in an abbreviated format. Primarily for programmer use.

Usage

```
showmelme(x, verbose = TRUE)
```

Arguments

x	Name of forsearch_lme output file
verbose	If TRUE, indicates the beginning and end of function run

Value

Abbreviated printout of output of forsearch_lme function

Author(s)

William R. Fairweather

Examples

variablelist *Identify Level(s) to Which Each Factor Observation Belongs*

Description

For a data frame with factor variables V1, V2, V3, etc having levels n1, n2, n3, etc, lists the n1*n2*n3*... possible interaction levels and identifies which of the observations of the data frame belong in which of these interaction levels.

Usage

```
variablelist(datadf, verbose = TRUE)
```

Arguments

datadf	Data frame of independent variables in analysis. First column of data frame is Observation number
verbose	If TRUE, indicates beginning and end of function run

Details

Support function, usually not called independently

Value

List, each element is a data frame of 2 columns with code indicating the highest possible level of interaction to which each observation can belong

Author(s)

William R. Fairweather

Examples

```
V1 <- rep(c(3,7,9,11),times=4)
V2 <- V1+3
V3 <- V1+90
V1 <- as.factor(V1)
V2 <- as.factor(V2)
V3 <- as.factor(V3)
Observation<-1:16
testdf <- data.frame(Observation, V1,V2,V3)
variablelist(testdf)
```

Index

∗ **datagen**
forsearch_glm, 8
forsearch_lm, 10
forsearch_lme, 12
∗ **hplot**
plotdiag.AICX, 18
plotdiag.allgraphs, 19
plotdiag.Cook, 20
plotdiag.deviance.residuals, 21
plotdiag.deviances, 22
plotdiag.fit3, 23
plotdiag.leverage, 24
plotdiag.params.fixed, 25
plotdiag.params.random, 26
plotdiag.phihatx, 28
plotdiag.residuals, 29
plotdiag.s2, 30
plotdiag.tstats, 31
∗ **manip**
aStep1, 4
aStep2, 5
bStep1, 6
bStep2, 7
identifyCoeffs, 15
identifyFixedCoeffs, 16
picksome, 17
search.history, 32
showme, 32
showmegl, 33
showmelme, 34
variablelist, 34
∗ **package**
forsearch-package, 2

aStep1, 4
aStep2, 5

bStep1, 6
bStep2, 7

forsearch *(*forsearch-package*)*, 2
forsearch-package, 2
forsearch_glm, 8
forsearch_lm, 10
forsearch_lme, 12

identifyCoeffs, 15
identifyFixedCoeffs, 16

picksome, 17
plotdiag.AICX, 18
plotdiag.allgraphs, 19
plotdiag.Cook, 20
plotdiag.deviance.residuals, 21
plotdiag.deviances, 22
plotdiag.fit3, 23
plotdiag.leverage, 24
plotdiag.params.fixed, 25
plotdiag.params.random, 26
plotdiag.phihatx, 28
plotdiag.residuals, 29
plotdiag.s2, 30
plotdiag.tstats, 31

search.history, 32
showme, 32
showmegl, 33
showmelme, 34

variablelist, 34