

Package ‘forsearch’

July 13, 2024

Title Diagnostic Analysis Using Forward Search Procedure for Various Models

Version 6.2.0

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: 0-387-95017-6> to prepare descriptive statistics of a dataset that is to be analyzed by `stats::lm()`, `stats::glm()`, `stats::nls()`, `nlme::lme()` or `survival::coxph()`. Includes graphics functions to display the descriptive statistics.

Depends R (>= 4.2)

License GPL (>= 3)

SystemRequirements gmp (>= 4.1)

Encoding UTF-8

RoxygenNote 7.2.3

Imports Hmisc(>= 4.7-0), Cairo(>= 1.6-0), formula.tools(>= 1.7.0), ggplot2(>= 3.4.0), nlme(>= 3.1-157), survival(>= 3.4), tibble(>= 3.1.8)

Suggests rmarkdown, knitr

VignetteBuilder knitr

NeedsCompilation no

Author William Fairweather [aut, cre]

Maintainer William Fairweather <wrf343@flowervalleyconsulting.com>

Repository CRAN

Date/Publication 2024-07-13 15:00:02 UTC

Contents

forsearch-package	2
aStep1	5
aStep2	6
bStep1	7

bStep2	8
cStep1	9
cStep2	10
forsearch_cph	10
forsearch_glm	12
forsearch_lm	14
forsearch_lme	17
forsearch_nls	19
identifyCoeffs	21
identifyFixedCoeffs	23
picksome	24
plotdiag.AICX	25
plotdiag.allgraphs	26
plotdiag.ANOX2	26
plotdiag.blind.fixed	27
plotdiag.Cook	28
plotdiag.deviance.residuals	29
plotdiag.deviations	30
plotdiag.fit3	31
plotdiag.leverage	32
plotdiag.loglik	33
plotdiag.lrt	34
plotdiag.params.fixed	35
plotdiag.params.random	36
plotdiag.phihatx	37
plotdiag.residuals	38
plotdiag.s2	39
plotdiag.tstats	40
plotdiag.Wald	41
search.history	42
showme	43
variablelist	44
Index	45

forsearch-package	<i>Diagnostic Analysis Using Forward Search Procedure for Various Models Diagnostic Analysis Using Forward Search Procedure for Various Models</i>
-------------------	--

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: 0-387-95017-6> to prepare descriptive statistics of a dataset that is to be analyzed by `stats::lm()`, `stats::glm()`, `stats::nls()`, `nlme::lme()` or `survival::coxph()`. 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:         6.2.0
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 procedure.
Depends:         R (>= 4.2)
License:         GPL (>= 3)
LazyData:        true
SystemRequirements: gmp (>= 4.1)
Encoding:        UTF-8
Roxygen:         list(markdown = TRUE)
RoxygenNote:     7.2.3
Imports:         Hmisc(>= 4.7-0), Cairo(>= 1.6-0), formula.tools(>= 1.7.0), ggplot2(>= 3.4.0), nlme(>= 3.1-157), sur
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 Set 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
cStep1           Create Set of Observation Numbers in Step 1 for
                  Cox Proportional Hazards Model Analysis
cStep2           Update Observation Set in Step 2
forsearch-package Diagnostic Analysis Using Forward Search
                  Procedure for Various Models Diagnostic
                  Analysis Using Forward Search Procedure for
                  Various Models
forsearch_cph    Create Statistics Of Forward Search in a Cox
                  Proportional Hazard Database
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
forsearch_nls    Create Statistics Of Forward Search in a
                  Nonlinear Model 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.ANOX2	Plot Diagnostic Statistics of Analysis of Variance Tables
plotdiag.Cook	Plot Diagnostic Statistics of Modified Cook's Distance
plotdiag.Wald	Plot Diagnostic Statistics of Wald Test Output of COXPH Function
plotdiag.allgraphs	Execute All Plotting Functions For a Select Forsearch Object
plotdiag.blind.fixed	Plot Diagnostic Statistics of Fixed Coefficients for Blinded Dataset
plotdiag.deviance.residuals	Plot Diagnostic Statistics Of Deviance Residuals
plotdiag.deviiances	Plot Diagnostic Deviance Statistics
plotdiag.fit3	Plot Diagnostic Statistics of AIC, BIC, and Log Likelihood
plotdiag.leverage	Plot Diagnostic Statistics Of Leverage
plotdiag.loglik	Plot Diagnostic Statistics of LOGLIK Output of COXPH Function
plotdiag.lrt	Plot Diagnostic Statistics of Likelihood Ratio Test of COXPH Function
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_xxx 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 the utility function `showme` displays 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	<i>Create Set of Observation Numbers in Step 1 for Linear Model Analysis</i>
--------	--

Description

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

Usage

```
aStep1(yesfactor, df1, df1.ls, inner.rank, initial.sample, formula, ycol, nop1, b.d)
```

Arguments

<code>yesfactor</code>	Logical. TRUE if there are factors in the X matrix
<code>df1</code>	Data frame being analyzed by forward search.
<code>df1.ls</code>	List, each element of which is a factor subset of <code>df1</code>
<code>inner.rank</code>	Rank of X matrix of <code>lm</code> analysis on entire database
<code>initial.sample</code>	Number of random samples from which to take set of initial observations
<code>formula</code>	Fixed parameter formula of <code>lm</code> function
<code>ycol</code>	Response column number
<code>nop1</code>	Number of observations per level of combined factor variables
<code>b.d</code>	Index of point to begin diagnostic listings

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 Set in Step 2

Description

Derives the set of observation numbers for forsearch in Step 2 for linear models

Usage

aStep2(yesfactor, form.A2, finalm, rimbs, dfa2, ycol, mstart, rnk, b.d)

Arguments

yesfactor	True or False for presence of factors
form.A2	Formula for analysis of entire dataset
finalm	See VALUE above. finalm argument is the same but only for Step 1 values
rimbs	List, each element is a matrix of obs numbers and corresponding subset codes
dfa2	Data frame being analyzed by forward search. Presence of Observation column has no effect on output
ycol	Response column number, including 1 for Observation
mstart	Number of first subset to be defined in Step 2
rnk	Rank of X matrix. For factors, this is rank with factors removed.
b.d	Number at which to begin diagnostic listings

Details

Support function, usually not called independently

Value

Vector of integers corresponding to observation numbers

Author(s)

William R. Fairweather

bStep1	<i>Create Set of Observation Numbers in Step 1 for Linear Mixed Effects Model Analysis</i>
--------	--

Description

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

Usage

```
bStep1(yesfactor, df1, df1.ls, inner.rank, initial.sample, formula, randform,
       ycol, nop1, b.d)
```

Arguments

yesfactor	Logical. TRUE if there are factors in the X matrix
df1	Data frame being analyzed by forward search.
df1.ls	List, each element of which is a factor subset of df1
inner.rank	Rank of X matrix of lme analysis on entire database
initial.sample	Number of random samples from which to take set of initial observations
formula	Two-sided fixed parameter formula of lme function
randform	One-sided random effects formula
ycol	Response column number
nop1	Number of observations per level of combined factor variables
b.d	Index of point to begin diagnostic listings

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

`bStep2`*Update Observation Numbers in Step 2*

Description

Derives the set of Step 2 observation numbers for forsearch in linear mixed effects models

Usage

```
bStep2(f2, dfa2, randm2, ms, finalm, fbg, b.d, rnk2, ycol)
```

Arguments

<code>f2</code>	Fixed parameter formula
<code>dfa2</code>	Complete data set with factor subset identification codes
<code>randm2</code>	Random parameter formula
<code>ms</code>	Number of observations beginning Step 2
<code>finalm</code>	List of expanding subset observation numbers
<code>fbg</code>	List of observation numbers by factor subgroup
<code>b.d</code>	Indicator of place in code to begin diagnostic printouts
<code>rnk2</code>	Rank of linear regression with factor variables eliminated
<code>ycol</code>	Column number of response variable

Details

Support function, usually not called independently

Value

List of expanding number sets corresponding to observation numbers

Author(s)

William R. Fairweather

cStep1	<i>Create Set of Observation Numbers in Step 1 for Cox Proportional Hazards Model Analysis</i>
--------	--

Description

Derives the first set of observation numbers for forsearch in Cox Proportional Hazards models

Usage

```
cStep1(yesfactor, df1, df1.ls, inner.rank, initial.sample, formula, f.e, ycol, nop1, b.d)
```

Arguments

yesfactor	Logical. TRUE if there are factors in the X matrix
df1	Data frame being analyzed by forward search.
df1.ls	List, each element of which is a factor subset of df1
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
f.e	Right-hand side of formula for Surv function
ycol	Response column number
nop1	Number of observations per level of combined factor variables
b.d	Index of point to begin diagnostic listings

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

cStep2	<i>Update Observation Set in Step 2</i>
--------	---

Description

Derives the set of observation numbers for step 2 for forsearch in Cox proportional hazard models

Usage

```
cStep2(f.e, finalm, dfa2, ms, rnk2, ss, b.d)
```

Arguments

f.e	Right hand side of formula
finalm	List of rows in model at each stage
dfa2	Complete data frame with factor subset indicator codes
ms	Number of observations in first stage of Step 2
rnk2	Rank of linear analysis with factor variables removed
ss	NULL or vector of observation numbers manually entered into Step 1
b.d	Indicator of starting point for diagnostic listings

Details

Support function, usually not called independently

Value

Vector of expanding number sets corresponding to observation numbers

Author(s)

William R. Fairweather

forsearch_cph	<i>Create Statistics Of Forward Search in a Cox Proportional Hazard Database</i>
---------------	--

Description

Prepares summary statistics at each stage of forward search for subsequent plotting.

Usage

```
forsearch_cph(alldata, formula.rhs, initial.sample=1000,
  n.obs.per.level=1, skip.step1=NULL, ties = "efron", maxdisturb=0.01,
  proportion=TRUE, unblinded=TRUE, begin.diagnose= 100, verbose=TRUE)
```

Arguments

<code>alldata</code>	Data frame containing variables 'Observation', 'event.time', 'status', and independent variables, in that order
<code>formula.rhs</code>	Character vector of names of independent variables in model
<code>initial.sample</code>	Number of observations in Step 1 of forward search
<code>n.obs.per.level</code>	Number of observations per level of (possibly crossed) factor levels to include in Step 1
<code>skip.step1</code>	NULL or a vector of integers for observations to be included in Step 1
<code>ties</code>	Method for handling ties in event time; = "efron", "breslow", or "exact"; see <code>survival::coxph</code>
<code>maxdisturb</code>	Maximum amount to add randomly to event.time to prevent ties.
<code>proportion</code>	TRUE causes evaluation of proportionality of Cox regression
<code>unblinded</code>	TRUE causes printing of presumed analysis structure
<code>begin.diagnose</code>	Numeric. Indicates where in code to begin printing diagnostics. 0 prints all; 100 prints none
<code>verbose</code>	TRUE causes function identifier display before and after run

Value

LIST	
Rows in stage	Observation numbers of rows included at each stage
Number of model parameters	Number of fixed coefficients in Cox model
Fixed parameter estimates	Vector of parameter estimates at each stage
Wald Test	Vector of Wald tests at each stage
Proportionality Test	Result of Cox proportionality test, if run
LogLikelihood	Vector of null and overall coefficients log likelihoods at each stage
Likelihood ratio test	Vector of LRTs at each stage
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

```
## Not run:
{# Forsearch for Cox Proportional Survival, including Step 1
veteran <- survival::veteran
veteran <- veteran[order(veteran$celltype),]
veteranx <- veteran[,c(3,4,1,2)]
veteranx$trt <- as.factor(veteranx$trt)
dimv <- dim(veteran)[1]
Observation <- 1:dimv
veteranx <- data.frame(Observation,veteranx)
names(veteranx)[2] <- "event.time"
form.1 <- "trt + celltype"
forskip <- NULL
# forskip <- c(12, 23, 38, 71, 91, 104, 116, 130, 31, 73, 62, 76)
cphptest1a.out <- forsearch_cph(alldata=veteranx, formula.rhs=form.1,
                               n.obs.per.level=2, skip.step1=forskip, ties="efron", unblinded=TRUE,
                               initial.sample=467, begin.diagnose = 100, verbose = TRUE)
}
{# Same, but skipping Step 1.
forskip <- c(12, 6, 31, 23, 38, 62, 71, 73, 91, 84, 104, 101, 116, 125,128,76)
cphptest1b.out <- forsearch_cph(alldata=veteranx, formula.rhs=form.1,
                               n.obs.per.level=2, skip.step1=forskip, ties="efron", unblinded=TRUE,
                               initial.sample=467, begin.diagnose = 100, verbose = TRUE)
}

## End(Not run)
```

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, response.cols, indep.cols, family,
              formula=NULL, binomialrhs=NULL, formula.cont.rhs, data, n.obs.per.level = 1,
              estimate.phi = TRUE, skip.step1=NULL, unblinded=TRUE, begin.diagnose=100,
              verbose=TRUE)
```

Arguments

`initial.sample` Number of random sets of observations in Step 1 of forward search
`response.cols` Vector of column numbers (1 or 2) of responses and nonresponses (if binomial)

<code>indep.cols</code>	Column number(s) of independent variables
<code>family</code>	Error distribution and link
<code>formula</code>	Formula relating response to independent variables. Required except for family=binomial
<code>binomialrhs</code>	Quoted character.Right-hand side of formula. Required for family=binomial
<code>formula.cont.rhs</code>	Quoted character.Right-hand side of formula, omitting factor variables. Required for all families
<code>data</code>	Name of database
<code>n.obs.per.level</code>	Number of observations per level of (possibly crossed) factor levels
<code>estimate.phi</code>	TRUE causes phi to be estimated; FALSE causes phi to be set = 1
<code>skip.step1</code>	NULL, or vector of observation numbers to include at end of Step 1
<code>unblinded</code>	TRUE allows print of formula of analysis function
<code>begin.diagnose</code>	Numeric. Indicates where in code to begin printing diagnostics. 0 prints all; 100 prints none
<code>verbose</code>	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. Current version assumes independent variables are all continuous.

Value

LIST

<code>Rows in stage</code>	Observation numbers of rows included at each stage
<code>Family</code>	Family and link
<code>Number of model parameters</code>	Number of fixed effect parameters
<code>Fixed parameter estimates</code>	Matrix of parameter estimates at each stage
<code>Residual deviance</code>	Vector of deviances
<code>Null deviance</code>	Vector of null deviances
<code>PhiHat</code>	Vector of values of phi parameter
<code>Deviance residuals and augments</code>	Deviance residuals with indication of whether each is included in fit
<code>AIC</code>	Vector of AIC values
<code>Leverage</code>	Matrix of leverage of each observation at each stage
<code>Call</code>	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) with Rolling Stock as a factor
Observation<-1:67
Month<-c(9,8,3,1,10,6,7,1,8,4,3,3,12,11,10,9,9,4,3,12,12,10,7,2,12,2,12,12,12,
  11,3,10,4,2,12,12,9,11,1,10,8,6,1,10,6,12,8,4,9,6,12,10,7,2,5,12,5,5,4,3,1,
  9,11,9,7,3,2)
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)
RollingStock <- as.factor(RollingStock)
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)
train2022 <- data.frame(Observation, Year, RollingStock, Traffic, Deaths)
forsearch_glm(initial.sample = 100, response.cols = 5,
  indep.cols = 2:4, formula=Deaths~Year + RollingStock + Traffic,
  formula.cont.rhs="Year + Traffic",
  family = poisson("log"), data = train2022,
  n.obs.per.level = 1, estimate.phi = TRUE, skip.step1 = NULL,
  unblinded = TRUE, begin.diagnose=100)
```

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, begin.diagnose = 100,
             verbose = TRUE)
```

Arguments

formula	Fixed effects formula as described in help(lm). The only permitted operators are +, :, and * . Terms must be found in data or as constructed by I(xxx) where xxx is found in data
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. Set to rank of X'X if model contains constructed variables such as I(x^3), for example in polynomial regression
skip.step1	NULL or a vector of integers for observations to be included in Step 1
unblinded	TRUE causes printing of presumed analysis structure
begin.diagnose	Numeric. Indicates where in code to begin printing diagnostics. 0 prints all; 100 prints none
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

Vector 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=200,begin.diagnose=100)
## 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=200,begin.diagnose=100)

# 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=200,begin.diagnose=100)

# Polynomial regression
C1 <- 7*runif(60) + 1
y <- 4 + C1 - 6*C1^2 + 9*C1^3 + rnorm(60)
Observation <- 1:60
dfpoly <- data.frame(Observation,C1,y)
forsearch_lm(formula = y ~ C1 + I(C1^2) + I(C1^3), data = dfpoly, initial.sample = 200,
begin.diagnose=100)

## End(Not run)
```

forsearch_lme	<i>Create Statistics Of Forward Search For a Linear Mixed Effects Database</i>
---------------	--

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(fixedform, alldata, randomform, initial.sample=1000, n.obs.per.level=1,
  skip.step1=NULL, unblinded=TRUE, begin.diagnose = 100, verbose = TRUE)
```

Arguments

fixedform	2-sided formula for fixed effects
alldata	data frame, first column of which must be "Observation"
randomform	1-sided formula for random effects
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
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 layout. 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. Variables in the randomform formula must be character variables, but *not* factors

Value

LIST

- Number of observations in Step 1
 - Number of observations included in Step 1
- Step 1 observation numbers
 - Observation numbers useful in skipping step 1

Rows by outer subgroup	List of row numbers, by outer subgroup
Rows by outer-inner subgroups	List of row numbers, by outer-inner 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 <- as.factor(rep(c("G1", "G2"), each=8))
lmetest1 <- data.frame(Observation, y, x1, x2, x3, group)
forsearch_lme(fixedform=y~x1+x2+x3, alldata=lmetest1, randomform= ~1|group,
  n.obs.per.level=1, initial.sample=200)
## 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 <- as.factor(rep(c("G1","G2"),each=30))
lmetest2 <- data.frame(Observation,y,AN1,AN2,group)
forsearch_lme(fixedform=y~AN1*AN2, alldata=lmetest2, randomform= ~1|group,
              initial.sample=500)

# 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)
group <- as.factor(rep(c("G1","G2"),each=30))
group <- c(group,group)
lmetest3 <- data.frame(Observation,y,AN1,AN2,COV,group)
forsearch_lme(fixedform=y~AN1*AN2+COV, alldata=lmetest3, randomform= ~ 1 | group,
              initial.sample=500)

## End(Not run)

```

forsearch_nls

Create Statistics Of Forward Search in a Nonlinear 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_nls(phaselist, data, poolstart, poolformula, algorithm=
  "default", controlarg=NULL, initial.sample=1000, skip.step1=NULL,
  begin.diagnose=100, verbose=TRUE)

```

Arguments

phaselist	LIST of formula, formulacont, start, nopl for each phase
data	Name of database. First 2 variables are Observation and Phases (both mandatory)
poolstart	List Start values for Step 2

poolformula	Formula for pooled data from all phases for Step 2
algorithm	algorithm for nls function.
controlarg	nls control. Default is NULL to use preset nls.control
initial.sample	Number of observation sets in Step 1 of forward search
skip.step1	NULL or a vector of integers for observations to be included in Step 1
begin.diagnose	Numeric. Indicates where in code to begin printing diagnostics. 0 prints all; 100 prints none
verbose	TRUE causes function identifier to display before and after run

Details

All datasets are considered to be in phases. See vignette for definition and discussion. There is a phaselist for each phase and an element for each phaselist input variable. In addition, there is a (pool)start and a (pool)formula input variable for the pooled dataset.

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	Same as number of levels of poolstart input variable
Sigma	Estimate of random error at final stage; used to standardize all residuals
Fixed parameter estimates	Vector of parameter estimates at each stage
s^2	Estimate of random error 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.
 Pinheiro, JC and DM Bates. Mixed Effects Models in S and S-PLUS, Springer, New York, 2000.
 Example from nlstools package

Examples

```
## Not run:
t<-(0:35)/3
V02<-c(377.1111,333.3333,352.1429,328.7500,369.8750,394.4000,352.6667,337.3333,
366.4286,364.0000,293.8889,387.0000,364.8889,342.2222,400.3000,375.1111,
320.5556,385.1667,527.0714,688.6364,890.8182,1145.1538,1254.9091,1327.5000,
```

```

1463.9000,1487.8333,1586.6667,1619.1000,1494.4167,1640.4545,1643.3750,
1583.6364,1610.8000,1568.5000,1464.5833,1652.8000)
Observation <- 1:36
Phases <- as.factor(c(rep("REST",18), rep("EXERCISE",18)))
test01 <- data.frame(Observation,Phases,t,V02)

formula.1 <-as.formula(V02~V02rest)
formulacont.1 <- as.formula(V02~V02rest)
start.1 <- list(V02rest = 400)
nopl.1 <- 1

formula.2<-
  as.formula(V02~(V02rest+(V02peak-V02rest)*(1-exp(-(t-5.883)*I(1/mu))))))
formulacont.2<-
  as.formula(V02~(V02rest+(V02peak-V02rest)*(1-exp(-(t-5.883)*I(1/mu))))))
start.2 <- list(V02rest = 400, V02peak = 1600, mu = 1)
nopl.2 <- 6

phaselist <- list(
  REST=
  list(formula=formula.1,formulacont=formulacont.1,start=start.1,nopp=nopl.1),
  EXERCISE=
  list(formula=formula.2,formulacont=formulacont.2,start=start.2,nopp=nopl.2))

pstart <- list(V02rest=400, V02peak = 1600, mu = 1)
pformula <- as.formula(V02~(t<=5.883)*(V02rest)+
  (t>5.883)*(V02rest+(V02peak-V02rest)*
  (1-exp(-(t-5.883)*I(1/mu))))))
forsearch_nls(phaselist=phaselist, data=test01,
  poolstart=pstart, poolformula=pformula, algorithm="default",
  controlarg=nls.control(maxiter=50,warnOnly=TRUE), initial.sample = 155)

## 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", 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
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

```
info3 <- system.file("extdata", "Machines.0.R", package="forsearch");
info3 <- source(info3);
info3 <- info3[[1]];
identifyCoeffs(fixed=score~1, data=info3, random= ~1 | Worker)
```

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, verbose = TRUE)
```

Arguments

formula	2-sided formula for fixed effects
data	Name of file (to be) run by forsearch_lm
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

```
info3 <- system.file("extdata", "crossdata.R", package="forsearch");  
crossdata <- source(info3);  
crossdata <- crossdata[[1]];  
identifyFixedCoeffs(formula=y~x1*x2, data=crossdata)
```

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)
```

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.

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

`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="none",
  verbose = TRUE)
```

Arguments

<code>forn</code>	Name of output file from <code>forsearch_glm</code>
<code>maintitle</code>	Main title of plot
<code>subtitle</code>	Subtitle of plot
<code>caption</code>	Content of caption
<code>wmf</code>	File name of stored plot; omit ".wmf"
<code>Cairo</code>	TRUE causes use of Cairo graphics
<code>printgraph</code>	TRUE causes graph to print to file and closes device
<code>addline</code>	add a line to the graph; "none", "loess", or "straight"); abbreviation allowed
<code>verbose</code>	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.

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=" ", blind.label=FALSE, 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
blind.label	TRUE causes 'blind' to be added to graph and to file name for fixed parameters
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

plotdiag.ANOX2 *Plot Diagnostic Statistics of Analysis of Variance Tables*

Description

Plot output from forsearch_XXX to show change in anova p-values as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.ANOX2(forn, anova.rows=NULL, ylab.extend=c("proportionality","variance"),
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",
verbose = TRUE)
```

Arguments

forn	Name of output file from forsearch_XXX
anova.rows	Row numbers of p values to include together on the plot
ylab.extend	Type of anova table. "proportionality" is a test of proportionality for a coxph analysis; "variance" is a test of null hypothesis of a lm or lme test
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
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot anova test p values 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.

plotdiag.blind.fixed *Plot Diagnostic Statistics of Fixed Coefficients for Blinded Dataset*

Description

Plot output from forsearch_XXX to show change in fixed coefficients as the number of observations in the forward search procedure increases. Save plot in folder containing working directory. Run on blinded data only.

Usage

```
plotdiag.blind.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",
  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
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.

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 = "none", 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	Character variable to add a line to the graph; options: "none", "loess", and "straight"; abbreviation allowed
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.

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", verbose = TRUE)
```

Arguments

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

`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="none",
  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; "none", "loess", or "straight"
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.

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",
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
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.

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, 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
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.

plotdiag.loglik *Plot Diagnostic Statistics of LOGLIK Output of COXPH Function*

Description

Plot output from `forsearch_cph` to show change in loglik pairs as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.loglik(forn,
  maintitle= "Put main title here",
  subtitle= "Put subtitle here" ,
  caption="Put caption here",
  wmf = "Put_stored_name_here",
  Cairo=TRUE,
  printgraph = TRUE,
  verbose=TRUE)
```

Arguments

forn	Name of output file from forsearch_cph
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
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot Wald Test statistics from forsearch_cph

Author(s)

William R. Fairweather

References

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

plotdiag.lrt

Plot Diagnostic Statistics of Likelihood Ratio Test of COXPH Function

Description

Plot output from forsearch_cph to show change in likelihood ratio test as the number of observations in the forward search procedure increases. Save plot in folder containing working directory.

Usage

```
plotdiag.lrt(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"),
  verbose=TRUE)
```

Arguments

forn	Name of output file from forsearch_cph
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. Default none
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot likelihood ratio test statistics from forsearch_cph

Author(s)

William R. Fairweather

References

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

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

Description

Plot output from forsearch_xxx to show change in fixed 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",
  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
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.

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", 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
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.

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="none",
  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; "none", "loess", or "straight"
printgraph	TRUE causes graph to print to file and closes device
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.

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", 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
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.

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"),
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
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.

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",
  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
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.

plotdiag.Wald

Plot Diagnostic Statistics of Wald Test Output of COXPH Function

Description

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

Usage

```
plotdiag.Wald(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"),
verbose=TRUE)
```

Arguments

form	Name of output file from forsearch_cph
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. Default none
verbose	If TRUE, indicates beginning and end of function

Value

Process and plot Wald Test statistics from forsearch_cph

Author(s)

William R. Fairweather

References

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

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, verbose = TRUE)
```

Arguments

list1	Name of a forsearch_xxx output file
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

```
info3 <- system.file("extdata", "crossdata.for1.R", package="forsearch");  
info3 <- source(info3);  
info3 <- info3[[1]];  
search.history(list1=info3, verbose=TRUE)
```

showme

Display Abbreviated Output of FORSEARCH_XXX Function

Description

Output of forsearch_XXX 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_XXX 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

variablelist	<i>Identify Level(s) to Which Each Factor Observation Belongs</i>
--------------	---

Description

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

Usage

```
variablelist(datadf, prank)
```

Arguments

datadf	Data frame of independent variables in analysis. First column of data frame is Observation number
prank	Number of continuous variables among independent variables

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

Index

* datagen

- forsearch_cph, 10
- forsearch_glm, 12
- forsearch_lm, 14
- forsearch_lme, 17
- forsearch_nls, 19

* hplot

- plotdiag.AICX, 25
- plotdiag.allgraphs, 26
- plotdiag.ANOX2, 26
- plotdiag.blind.fixed, 27
- plotdiag.Cook, 28
- plotdiag.deviance.residuals, 29
- plotdiag.deviations, 30
- plotdiag.fit3, 31
- plotdiag.leverage, 32
- plotdiag.loglik, 33
- plotdiag.lrt, 34
- plotdiag.params.fixed, 35
- plotdiag.params.random, 36
- plotdiag.phihatx, 37
- plotdiag.residuals, 38
- plotdiag.s2, 39
- plotdiag.tstats, 40
- plotdiag.Wald, 41

* manip

- aStep1, 5
- aStep2, 6
- bStep1, 7
- bStep2, 8
- cStep1, 9
- cStep2, 10
- identifyCoeffs, 21
- identifyFixedCoeffs, 23
- picksome, 24
- search.history, 42
- showme, 43
- variablelist, 44

* package

- forsearch-package, 2

- aStep1, 5
- aStep2, 6

- bStep1, 7
- bStep2, 8

- cStep1, 9
- cStep2, 10

- forsearch (forsearch-package), 2
- forsearch-package, 2
- forsearch_cph, 10
- forsearch_glm, 12
- forsearch_lm, 14
- forsearch_lme, 17
- forsearch_nls, 19

- identifyCoeffs, 21
- identifyFixedCoeffs, 23

- picksome, 24
- plotdiag.AICX, 25
- plotdiag.allgraphs, 26
- plotdiag.ANOX2, 26
- plotdiag.blind.fixed, 27
- plotdiag.Cook, 28
- plotdiag.deviance.residuals, 29
- plotdiag.deviations, 30
- plotdiag.fit3, 31
- plotdiag.leverage, 32
- plotdiag.loglik, 33
- plotdiag.lrt, 34
- plotdiag.params.fixed, 35
- plotdiag.params.random, 36
- plotdiag.phihatx, 37
- plotdiag.residuals, 38
- plotdiag.s2, 39
- plotdiag.tstats, 40
- plotdiag.Wald, 41

search.history, [42](#)
showme, [43](#)
variablelist, [44](#)