Package 'AutoSEARCH'

March 27, 2015

Type Package

Title General-to-Specific (GETS) Modelling

Version 1.5

Depends R (>= 2.4.1), zoo, lgarch

Date 2015-03-27

Author Genaro Sucarrat

Maintainer Genaro Sucarrat <genaro.sucarrat@bi.no>

Description General-to-Specific (GETS) modelling of the mean and variance of a regression. NOTE: The package has been succeeded by gets, also available on the CRAN, which is more user-friendly, faster and easier to extend. Users are therefore encouraged to consider gets instead.

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2015-03-27 23:49:39

R topics documented:

AutoSEARCH-package 2
eqwma 3
gedestp 5
gets.mean 6
gLag 11
gLog.ep 12
info.criterion 12
jb.test 13
ols.fit1 14
regs.mean.sm 15
regs.vol.sm 16
skewness.test 17
sm 17

Index **21**

AutoSEARCH-package *General-to-Specific (GETS) Modelling*

Description

GETS modelling of the mean and variance of a regression.

NOTE: The package has been succeeded by the package gets, also available on the CRAN, which is more user-friendly, faster and easier to extend. The development focus has switched to gets, so users are therefore encouraged to consider gets instead.

Details

Package:	AutoSEARCH
Type:	Package
Version:	1.5
Date:	2015-03-27
License:	GPL-2
LazyLoad:	yes

The code was originally developed in relation with G. Sucarrat and A. Escribano (2012): 'Automated Financial Model Selection: General-to-Specific Modelling of the Mean and Volatility Specifications', Oxford Bulletin of Economics and Statistics 74, Issue 5 (October), pp. 716-735.

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

References

G. Sucarrat and A. Escribano (2012): 'Automated Financial Model Selection: General-to-Specific Modelling of the Mean and Volatility Specifications', Oxford Bulletin of Economics and Statistics 74, Issue 5 (October), pp. 716-735

See Also

AutoSEARCH package: sm, gets.mean, gets.vol
gets package: arx, getsm, getsv, isat

Examples

```
#Generate from AR(1) model:
set.seed(123)
y <- arima.sim(list(ar=0.4), 200)

#Estimate AR(2) with intercept as mean specification
#and log-ARCH(4) as log-volatility specification:
```

```
sm(y, mc=TRUE, ar=1:2, arch=1:4)

#General-to-Specific model selection of the mean:
mymodel <- gets.mean(y, mc=TRUE, ar=1:2, arch=1:4)

#General-to-Specific model selection of the
#simplified mean specification:
gets.vol(mymodel$resids, arch=1:4)
```

---

eqwma *Equally Weighted Moving Average (EqWMA) of the pth. exponentiated values*

Description

The function eqwma returns an Equally Weighted Moving Average (EqWMA) of the pth. exponentiated values lagged. Optionally, the absolute values are computed before averaging, and the log of is returned. The function leqwma is essentially a wrapper to eqwma in which the absolute values are used and the logarithm is applied.

If x is financial return (possibly mean-corrected) and p=2, then this gives the socalled 'historical' model, also known as an integrated ARCH model where the ARCH coefficients all have the same value with sum equal to one. In the log-variance specification the lag of log(EqWMA) is thus a financial volatility proxy. It may be an imperfect proxy compared with high-frequency data (which can also be included as regressors), but - in contrast to high-frequency data - is always available and easy to compute

Usage

```
eqwma(x, length = 5, lag = 1, start = 1, p = 1, log = FALSE, abs = FALSE,
  as.vector = TRUE)
leqwma(x, length = 5, lag = 1, start = 1, p = 2, as.vector=FALSE)
```

Arguments

x	numeric vector, time-series or zoo object. Missing values in the beginning and/or at the end of the series is allowed, as they are removed with the na.trim command
length	integer or vector of integers each equal to or greater than 1. The length or lengths of the moving window or windows of averages
lag	integer equal to or greater than 0. If 0, then the moving averages are not lagged
start	integer equal to or greater than 1 (default: start=1, i.e. the first observation). Where to start the moving windows of averages
p	numeric value greater than zero. The exponent p in x^p for eqwma and in abs(x)^p for leqwma
log	logical. If TRUE, then the logarithm of the moving average is returned. If FALSE (default), then the logarithm is not applied

abs	logical. If TRUE, then x is transformed to absolute values before x is exponentiated
as.vector	logical. If TRUE, then a univariate series is returned as a vector. If FALSE, then a univariate series is returnes as a matrix. Note: multivariate series are always returned as a matrix

Details

The intended primary use of eqwma is to construct mixed frequency regressors for the mean specification.

The intended primary use of leqwma is to construct volatility proxies in for the log-variance specification. The default is the lagged log of an equally weighted moving average of the squared residuals, where each average is made up of m observations. This is equivalent to an integrated ARCH(p) model where the p coefficients are all equal. For further details on the use of log(EqWMA) as a volatility proxy, see Sucarrat and Escribano (2012)

Value

numeric vector, time series or zoo object

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

References

Genaro Sucarrat and Alvaro Escribano (2012): 'Automated Financial Model Selection: General-to-Specific Modelling of the Mean and Volatility Specifications', Oxford Bulletin of Economics and Statistics 74, Issue no. 5 (October), pp. 716-735

See Also

zoo, sm, gets.mean, gets.vol

Examples

```
##generate an iid normal series:
set.seed(123)
x <- rnorm(100)

##compute lag of EqWMA(20) for x^2:
eqwma(x, p=2)

##compute lag of EqWMA(5) and lag of EqWMA(10) for x:
eqwma(x, length=c(5,10))

##compute lag of log(EqWMA(20)) for x^2:
leqwma(x)

#compute lag of log(EqWMA(5)) and lag of log(EqWMA(8))
```

```
#for abs(x)^2:
leqwma(x, length=c(4,8))
```

---

gedestp *Estimate and compute log-likelihood of the standardised Generalised Error Distribution (GED)*

Description

gedestp and gedlogl are auxiliary functions called by gets.mean and gets.vol.

The gedestp function estimates the shape parameter of a standardised (zero mean, unit variance) GED. The estimation method is based on an index of kurtosis approach, and the code is based on the estimatep function from the normalp package by Angelo M. Mineo.

The gedlogl function computes the log-likelihood of a standardised GED with shape parameter p.

Usage

```
gedestp(x, method = c("inverse", "direct"))
gedlogl(z, p = 2)
```

Arguments

x	numeric vector
z	numeric vector
method	"inverse" or "direct"
p	numeric value, the shape parameter

Value

numeric, either an estimate of the shape parameter or the log-likelihood

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

References

Mineo, A.M. (1994): 'Un nuovo metodo di stima di p per una corretta valutazione dei parametri diintensita e di scala di una curva normale di ordine p', Atti della XXXVII Riunione Scientifica della Societa Italiana di Statistica, San Remo, Vol. 2, pp. 147-154

See Also

gets.mean, gets.vol

Examples

```
#estimate p of a standard normal:
set.seed(123)
x <- rnorm(200)
gedestp(x)

#log-likelihood of the standard normal series:
gedlogl(x, p=2)
```

`gets.mean` *General-to-Specific (GETS) Modelling of an AR-X model with log-ARCH-X errors*

Description

The starting model is referred to as the General Unrestricted Model (GUM). The `gets.mean` function undertakes multi-path GETS model selection of the mean specification, whereas `gets.vol` does the same for the log-variance specification.

Usage

```
gets.mean(y, mc = NULL, ar = NULL, ewma = NULL, mx = NULL,
  arch = NULL, asym = NULL, log.ewma = NULL, vx = NULL,
  keep = NULL, p = 2, varcov.mat = c("ordinary", "white"),
  t.pval = 0.05, do.pet = TRUE, wald.pval = 0.05,
  ar.LjungB = c(2, 0.025), arch.LjungB = c(2, 0.025),
  tau = 2, info.method = c("sc", "aic", "hq"),
  info.resids = c("mean", "standardised"),
  include.empty = FALSE, zero.adj = 0.1, vc.adj = TRUE,
  tol = 1e-07, LAPACK = FALSE, max.regs = 1000,
  verbose = TRUE, smpl = NULL, alarm = FALSE)
gets.vol(e, arch=NULL, asym=NULL, log.ewma=NULL, vx=NULL,
  p=2, keep=c(1), t.pval=0.05, wald.pval=0.05, do.pet=TRUE,
  ar.LjungB=c(1, 0.025), arch.LjungB=c(1, 0.025), tau=2,
  info.method=c("sc", "aic", "hq"),
  info.resids=c("standardised", "log-sigma"), include.empty=FALSE,
  zero.adj=0.1, vc.adj=TRUE, tol=1e-07, LAPACK=FALSE, max.regs=1000,
  verbose=TRUE, alarm=FALSE, smpl=NULL)
```

Arguments

`y`	numeric vector, time-series or zoo object. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
`e`	numeric vector, time-series or zoo object. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package

mc	logical, TRUE or FALSE (default). TRUE includes intercept in the mean specification, FALSE does not
ar	integer vector, say, c(2,4) or 1:4. The AR-lags to include in the specification
ewma	either NULL (default) or a list with arguments sent to the eqwma function. In the latter case a lagged moving average of y is included as a regressor
mx	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
arch	integer vector, say, c(1,3) or 2:5. The log-ARCH terms to include in the log-volatility specification
asym	integer vector, say, c(1) or 1:3. The asymmetry or leverage terms to include in the log-volatility specification
log.ewma	NULL (default) or a list. If NULL then log(EWMA) is not included as volatility proxy. If a list, then log(EWMA) is included as a volatility proxy.
vx	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
keep	NULL (default) or an integer vector. If keep = NULL, then no regressors are excluded from removal. Otherwise, the regressors associated with the numbers in keep are excluded from the removal space. For example, keep=c(1) excludes the constant from removal. The regressor numbering is contained in the reg.no column of the gum.mean data frame (see below)
p	numeric value greater than zero. The power of the log-volatility specification.
varcov.mat	character vector, "ordinary" or "white". If "ordinary" then the ordinary variance-covariance matrix is used for inference. Otherwise the White (1980) heteroscedasticity robust matrix is used
t.pval	numeric value between 0 and 1. The significance level used for the two-sided regressor significance tests
do.pet	logical, TRUE (default) or FALSE. If TRUE then a Parsimonious Encompassing Test (PET) against the GUM is undertaken at each regressor removal for the joint significance of all the deleted regressors along the current path
wald.pval	numeric value between 0 and 1. The significance level used for the PETs
ar.LjungB	NULL or a two-element vector where the first element contains the order of a Ljung and Box (1979) test for serial correlation in the standardised residuals, and where the second element contains the significance level. If NULL, then the standardised residuals are not checked for serial correlation after each removal. The default is c(2, 0.025)
arch.LjungB	NULL or a two-element vector where the first element contains the order of a Ljung and Box (1979) test for ARCH (serial correlation in the squared standardised residuals), and where the second element contains the significance level. If NULL, then the standardised residuals are not checked for ARCH after each removal. The default is c(2, 0.025)

tau	NULL or a numeric value greater than 1. If NULL, then the shape parameter in a Generalised Error Distribution (GED) of the standardised residuals is estimated for the log-likelihood used in the calculation of the information criterion. If tau is equal to a numeric value, a GED(tau) is used. Default: tau=2 (i.e. the standard normal density)
info.method	character string, "sc" (default), "aic" or "hq", which determines the information criterion used to select among terminal models. The abbreviations are short for the Schwarz or Bayesian information criterion (sc), the Akaike information criterion (aic) and the Hannan-Quinn (hq) information criterion
info.resids	character string, "mean" (default) or "standardised" which sets the residuals to be used in the computation of the information criterion
include.empty	logical, TRUE or FALSE (default). If TRUE then an empty model is included among the terminal models, if it passes the diagnostic tests, even if it is not equal to one of the terminals
zero.adj	numeric value between 0 and 1. The quantile adjustment for zero values. The default 0.1 means that the zero residuals are replaced by means of the 10 percent quantile of the absolute residuals before taking the logarithm
vc.adj	logical, TRUE (default) or FALSE. If true then the log-volatility constant is adjusted by means of the estimate of E[log(z^2)]. This adjustment is needed for the standardised residuals to have unit variance. If FALSE then the log-volatility constant is not adjusted
tol	numeric value (default = 1e-07). The tolerance for detecting linear dependencies in the columns of the regressors (see qr() function). Only used if LAPACK is FALSE
LAPACK	logical, TRUE or FALSE (default). If true use LAPACK otherwise use LINPACK (see qr() function)
max.regs	integer value, sets the maximum number of regressions along a deletion path. Default: max.regs=1000
verbose	logical, TRUE (default) or FALSE. FALSE returns less output and is therefore faster
smpl	Either NULL (default; the whole sample is used for estimation) or a two-element vector of dates with the start and end dates of the sample to be used in estimation. For example, smpl=c("2001-01-01", "2009-12-31")
alarm	Logical, either TRUE or FALSE (default). If TRUE, then a sound or beep is emitted when the specification search terminates in order to alert the user

Details

See Sucarrat and Escribano (2012)

Value

A list with a subset of the following:

volatility.fit	zoo-object with the fitted values of the volatility (sigma^p) of the final log-volatility specification

resids.ustar	zoo-object with the residuals of the AR-representation of the final log-volatility specification
resids	zoo-object with the residuals of the final mean specification
resids.std	zoo-object with the standardised residuals
Elogzp	estimate of E[log(z^p)]
call	the function call
gum.mean	a data frame with the estimation results of the GUM
gum.volatility	a data frame with the estimation results of the log-volatility GUM
gum.diagnostics	data frame with selected diagnostics of the GUM
keep	if any, the regressors that are excluded from deletion
insigs.in.gum	a numeric integer vector with the insignificant regressors of the GUM
paths	a list containing the simplification paths, that is, the sequences of deleted regressors
terminals	the distinct terminal models
terminals.results	the value and type of the information criterion (info) used in selecting among terminal specifications, and the number of observations (T) and parameters (k) used in the calculation of the information criterion
specific.mean	data frame with the estimation results of the final mean specification
specific.volatility	data frame with the estimation results of the final log-volatility specification
specific.diagnostics	data frame with selected diagnostics of the standardised residuals

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

References

Genaro Sucarrat and Alvaro Escribano (2012): 'Automated Financial Model Selection: General-to-Specific Modelling of the Mean and Volatility Specifications', Oxford Bulletin of Economics and Statistics 74, Issue no. 5 (October), pp. 716-735

G. Ljung and G. Box (1979): 'On a Measure of Lack of Fit in Time Series Models'. Biometrika 66, pp. 265-270

See Also

AutoSEARCH package: sm
gets package: arx, getsm, getsv, isat

Examples

```
#Generate AR(1) model and four independent normal regressors:
set.seed(123)
y <- arima.sim(list(ar=0.4), 200)
xregs <- matrix(rnorm(4*200), 200, 4)

#General-to-Specific model selection of the mean:
mymodel <- gets.mean(y, mc=TRUE, ar=1:5, mx=xregs)

#General-to-Specific model selection of the mean
#with the intercept excluded from removal:
mymodel <- gets.mean(y, mc=TRUE, ar=1:5, mx=xregs, keep=1)

#General-to-Specific model selection of the mean
#with no intercept and with a log-ARCH(4) specification
#in the log-volatility using the standardised residuals
#when computing the log-likelihood for the information
#criterion:
mymodel <- gets.mean(y, mc=FALSE, ar=1:5, mx=xregs, arch=1:4,
  info.resids="standardised")

#General-to-Specific model selection of the mean with
#non-default serial-correlation diagnostics settings:
mymodel <- gets.mean(y, mc=TRUE, ar=1:5, mx=xregs,
  ar.LjungB=c(6, 0.05))

#General-to-Specific model selection of the mean with
#very liberal (i.e. 20 percent) significance levels (20 percent):
mymodel <- gets.mean(y, mc=TRUE, ar=1:5, mx=xregs, t.pval=0.2,
  wald.pval=0.2)

#Generate iid normal residuals and a matrix of independent
#normals:
set.seed(123)
e <- rnorm(200)
xregs <- matrix(rnorm(4*200), 200, 4)

#General-to-Specific model selection of log-volatility:
mymodel <- gets.vol(e, arch=1:5, vx=log(xregs^2))

#General-to-Specific model selection of log-volatility
#with the log-ARCH(1) term excluded from removal:
mymodel <- gets.vol(e, arch=1:5, vx=log(xregs^2), keep=2)

#General-to-Specific model selection of log-volatility
#with all the log-ARCH terms excluded from removal:
mymodel <- gets.vol(e, arch=1:5, vx=log(xregs^2), asym=1:2,
  log.ewma=list(length=5), keep=2:6)

#If e is a daily (weekends excluded) financial return series,
#then the following specification includes a lagged volatility
#proxy both for the week (5-day average of squared return) and
```

```
#for the month (20-day average of squared returns), in addition
#to five log-ARCH terms:
mymodel <- gets.vol(e, arch=1:5, log.ewma=list(length=c(5,20)) )

#General-to-Specific model selection with very liberal
#(20 percent) significance levels:
mymodel <- gets.vol(e, arch=1:5, vx=log(xregs^2), t.pval=0.2,
  wald.pval=0.2)
```

---

gLag *Lag a series*

Description

A wrapper to the glag function in the lgarch package.

Usage

```
gLag(y, k=1, na.value=NA)
```

Arguments

y	numeric vector, time-series or zoo object
k	integer equal to or greater than 1. Default: k=1
na.value	the value to replace the lost values with. Default: na.replace=NA

Value

the lagged vector, time series or zoo object

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

See Also

glag

gLog.ep *Adjust for zero values and compute log(abs(e)^p)*

Description

Adjusts a series called e - typically a series of residuals or financial returns - for zero values, so that the logarithm can be applied on the absolute pth. exponentiated values. Next, log(abs(e)^p) is computed

Usage

```
gLog.ep(e, zero.adj=0.1, p=2, na.replace=NA)
```

Arguments

e	numeric vector, time series or zoo object
zero.adj	numeric value between 0 and 1 (the quantile adjustment for zero values). The default 0.1 means zeros are replaced by the 10 percent quantile of abs(e) before taking the logarithm
p	numeric value greater than zero. The power of the log-volatility specification
na.replace	the value to replace NA values with. Default: na.replace=NA

Value

log(abs(e)^p), a numeric, where the zeros in e have been adjusted

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

info.criterion *Computes the Value of an Information Criterion*

Description

Given a log-likelihood, the number of observations and the number of estimated parameters, the value of a chosen information criterion is computed

Usage

```
info.criterion(logl, n=NULL, k=NULL, method=c("sc", "aic", "aicc", "hq"))
```

Arguments

`logl`	numeric, the value of the log-likelihood
`n`	integer, number of observations
`k`	integer, number of parameters
`method`	character, either "sc" (default), "aic", "aicc" or "hq"

Value

a list with elements:

`method`	type of information criterion
`n`	number of observations
`k`	number of parameters
`value`	the value on the information criterion

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

`jb.test` *Jarque-Bera test for normality*

Description

Jarque-Bera test for normality

Usage

```
jb.test(x)
```

Arguments

`x`	a numeric vector

Value

A list with elements:

`skewness`	a numeric, an estimate of the excess skewness relative to the normal
`kurtosis`	a numeric, an estimate of the excess kurtosis relative to the normal
`statistic`	a numeric, the Jarque-Bera test statistic
`df`	2, the degrees of freedom
`p.value`	numeric between 0 and 1, the p-value of the test under the null of normality

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

`ols.fit1` *Fast and accurate OLS estimation by means of QR decomposition*

Description

ols.fit1 and ols.fit2 are auxiliary functions called by `sm`, `gets.mean` and `gets.vol`. The ols.fit2 function returns slightly more information than ols.fit1, which makes the latter faster. However, variance-covariance are to be needed in a second step, then ols.fit2 is faster due to the additional information provided by it.

Usage

```
ols.fit1(y, x, tol=1e-07, LAPACK=FALSE)
ols.fit2(y, x, tol=1e-07, LAPACK=FALSE)
```

Arguments

`y`	numeric vector, the regressand
`x`	numeric matrix, the regressors
`tol`	numeric value (default = 1e-07). The tolerance for detecting linear dependencies in the columns of the regressors (see qr() function). Only used if LAPACK is FALSE
`LAPACK`	logical, TRUE or FALSE (default). If true use LAPACK otherwise use LINPACK (see qr() function)

Value

A list that contains some or all of the following elements:

`qr`

`rank`

`qraux`

`pivot`

`xtxinv`

`xtx`

`xty`

`coefficients`

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

See Also

`qr`, `solve`

regs.mean.sm *Create the regressors of an AR-X model*

Description

Creates the regressors of an AR-X model, see sm.

Usage

```
regs.mean.sm(y, mc=NULL, ar=NULL, ewma=NULL, mx=NULL)
```

Arguments

y	numeric vector, time-series or zoo object. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
mc	logical, TRUE or FALSE (default). TRUE includes intercept in the specification, FALSE does not
ar	integer vector, say, c(2,4) or 1:4. The AR-lags to include in the specification
ewma	NULL or a list of arguments sent to the eqwma function
mx	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package

Value

Matrix with regressors

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

See Also

sm, regs.vol.sm

`regs.vol.sm` *Create the regressors of a log-ARCH-X model*

Description

Creates the regressors of a log-ARCH-X model, see `sm` and `gets.vol`

Usage

```
regs.vol.sm(e, vc=TRUE, arch=NULL, asym=NULL, log.ewma=NULL, vx=NULL, p=2,
  zero.adj=0.1)
```

Arguments

`e`	numeric vector, time-series or zoo object. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
`vc`	logical, TRUE (default) or FALSE. TRUE creates an intercept, FALSE does not
`arch`	integer vector, say, c(1,3) or 2:5. The ARCH-lags to include in the log-volatility specification
`asym`	integer vector, say, c(1) or 1:3. The asymmetry or leverage terms to include in the log-volatility specification
`log.ewma`	NULL (default) or a list. If NULL then log(EWMA) is not included as volatility proxy. If a list, then log(EWMA) is included as a volatility proxy.
`vx`	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
`p`	numeric value greater than zero. The power of the log-volatility specification.
`zero.adj`	numeric value between 0 and 1. The quantile adjustment for zero values. The default 0.1 means that the zero residuals are replaced by means of the 10 percent quantile of the absolute residuals before taking the logarithm

Value

Matrix with regressors

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

See Also

`sm`, `regs.mean.sm`

skewness.test — *Chi-squared test for skewness in the standardised residuals*

Description

Chi-squared test for skewness in the standardised residuals

Usage

```
skewness.test(x)
```

Arguments

x	numeric vector

Value

A list with elements:

statistic	the test statistic
p.value	the p-value of the test under the null of no-skewness

Author(s)

Genaro Sucarrat, http://www.sucarrat.net/

See Also

jb.test

sm — *Estimate an AR-X Model with Log-ARCH-X Errors*

Description

Estimation is by OLS in two stages. In the first the AR-X mean specification is estimated, whereas in the second stage the residuals from the first are used to fit a log-ARCH-X model to the log-variance. The natural logarithm of the squared residuals constitutes the regressand in the second step.

The AR-X mean specification can contain an intercept, AR-terms, lagged moving averages of the regressand and other conditioning covariates ('X'). The log-variance specification can contain log-ARCH terms, asymmetry or 'leverage' terms, log(EqWMA) where EqWMA is a lagged equally weighted moving average of past squared residuals (a volatility proxy) and other conditioning covariates ('X').

Usage

```
sm(y, mc=NULL, ar=NULL, ewma=NULL, mx=NULL, arch=NULL, asym=NULL,
  log.ewma=NULL, vx=NULL, p=2, zero.adj=0.1, vc.adj=TRUE,
  varcov.mat=c("ordinary", "white"), qstat.options=NULL,
  tol=1e-07, LAPACK=FALSE, verbose=TRUE, smpl=NULL)
```

Arguments

y	numeric vector, time-series or zoo object. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
mc	logical, TRUE or FALSE (default). TRUE includes intercept in the specification, FALSE does not
ar	integer vector, say, c(2,4) or 1:4. The AR-lags to include in the specification
ewma	list of arguments sent to the leqwma function
mx	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
arch	integer vector, say, c(1,3) or 2:5. The ARCH-lags to include in the log-volatility specification
asym	integer vector, say, c(1) or 1:3. The asymmetry or leverage terms to include in the log-volatility specification
log.ewma	NULL (default) or a list. If NULL then log(EWMA) is not included as volatility proxy. If a list, then log(EWMA) is included as a volatility proxy.
vx	numeric matrix, time-series or zoo object of conditioning covariates. Note that missing values in the beginning or at the end of the series is allowed, as they are removed with the na.trim command from the zoo package
p	numeric value greater than zero. The power of the log-volatility specification.
zero.adj	numeric value between 0 and 1. The quantile adjustment for zero values. The default 0.1 means that the zero residuals are replaced by means of the 10 percent quantile of the absolute residuals before taking the logarithm
vc.adj	logical, TRUE (default) or FALSE. If true then the log-volatility constant is adjusted by means of the estimate of E[log(z^2)]. This adjustment is needed for the standardised residuals to have unit variance. If FALSE then the log-volatility constant is not adjusted
varcov.mat	character vector, "ordinary" or "white". If "ordinary" then the ordinary variance-covariance matrix is used for inference. Otherwise the White (1980) heteroscedasticity robust matrix is used
qstat.options	NULL or an integer vector of length two, say, c(2,5). The first value sets the order of the AR diagnostic test, whereas the second value sets the order of the ARCH diagnostic test. NULL (default) sets the vector to c(1,1)
tol	numeric value (default = 1e-07). The tolerance for detecting linear dependencies in the columns of the regressors (see qr() function). Only used if LAPACK is FALSE

LAPACK	logical, TRUE or FALSE (default). If true use LAPACK otherwise use LINPACK (see qr() function)
verbose	logical, TRUE (default) or FALSE. FALSE returns less output and is therefore faster
smpl	Either NULL (default; the whole sample is used for estimation) or a two-element vector of dates with the start and end dates of the sample to be used in estimation. For example, smpl=c("2001-01-01", "2009-12-31")

Details

See Sucarrat and Escribano (2012)

Value

A list with the following elements:

call	the function call
mean.fit	zoo-object with the fitted values of the mean specification
resids	zoo-object with the residuals of the mean specification
volatility.fit	zoo-object with the fitted values of the volatility (sigma^p) specification
resids.ustar	zoo-object with the residuals of the AR-representation of the log-volatility specification
resids.std	zoo-object with the standardised residuals
Elogzp	estimate of E(log(z^p))
mean.results	data frame with the estimation results of the mean specification
volatility.results	data frame with the estimation results of the log-volatility specification
diagnostics	data frame with selected diagnostics of the standardised residuals

Author(s)

Genaro Sucarrat (http://www.sucarrat.net/)

References

Genaro Sucarrat and Alvaro Escribano (2012): 'Automated Financial Model Selection: General-to-Specific Modelling of the Mean and Volatility Specifications', Oxford Bulletin of Economics and Statistics 74, Issue no. 5 (October), pp. 716-735

Halbert White (1980): 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity', Econometrica 48, pp. 817-838

See Also

AutoSEARCH package: gets.mean, gets.vol
gets package: arx, getsm, getsv, isat

Examples

```
#Generate AR(1) model and independent normal regressors:
set.seed(123)
y <- arima.sim(list(ar=0.4), 200)
xregs <- matrix(rnorm(4*200), 200, 4)

#estimate AR(2) with intercept:
sm(y, mc=TRUE, ar=1:2)

#estimate AR(2) with intercept and four conditioning regressors
#in the mean:
sm(y, mc=TRUE, ar=1:2, mx=xregs)

#estimate a log-volatility specification with a log-ARCH(4)
#structure:
sm(y, arch=1:4)

#estimate a log-volatility specification with a log-ARCH(4)
#structure and an asymmetry or leverage term:
sm(y, arch=1:4, asym=1)

#estimate a log-volatility specification with a log-ARCH(4)
#structure, an asymmetry or leverage term, a 30-period log(EWMA) as
#volatility proxy, and the squareds of the conditioning regressors
#in the log-volatility specification:
sm(y, arch=1:4, asym=1, log.ewma=list(length=30), vx=log(xregs^2))

#estimate AR(2) with intercept and four conditioning regressors
#in the mean, and a log-volatility specification with a log-ARCH(4)
#structure, an asymmetry or leverage term, a 30-period log(EWMA) as
#volatility proxy, and the squareds of the conditioning regressors
#in the log-volatility specification:
sm(y, mc=TRUE, ar=1:2, mx=xregs, arch=1:4, asym=1,
  log.ewma=list(length=30), vx=log(xregs^2))
```

Index

∗Topic **Financial Econometrics**
eqwma, 3
gedestp, 5
gets.mean, 6
gLag, 11
gLog.ep, 12
ols.fit1, 14
regs.mean.sm, 15
regs.vol.sm, 16
sm, 17
∗Topic **Statistical Models**
eqwma, 3
gedestp, 5
gets.mean, 6
gLag, 11
gLog.ep, 12
info.criterion, 12
jb.test, 13
ols.fit1, 14
regs.mean.sm, 15
regs.vol.sm, 16
skewness.test, 17
sm, 17
∗Topic **Time Series**
eqwma, 3
gedestp, 5
gets.mean, 6
gLag, 11
gLog.ep, 12
ols.fit1, 14
regs.mean.sm, 15
regs.vol.sm, 16
sm, 17

AutoSEARCH *(*AutoSEARCH-package*)*, 2
AutoSEARCH-package, 2

eqwma, 3

gedestp, 5
gedlogl *(*gedestp*)*, 5
gets.mean, 2, *4, 5*, 6, *14*, *19*
gets.vol, 2, *4, 5*, *14*, *16*, *19*
gets.vol *(*gets.mean*)*, 6
gLag, 11
glag, *11*
gLog.ep, 12

info.criterion, 12

jb.test, 13, *17*

leqwma *(*eqwma*)*, 3
lgarch, *11*

na.trim, *3*

ols.fit1, 14
ols.fit2 *(*ols.fit1*)*, 14

qr, *14*

regs.mean.sm, 15, *16*
regs.vol.sm, *15*, 16

skewness.test, 17
sm, 2, *4*, *9*, *14–16*, 17
solve, *14*

zoo, *4*, *12*