Package 'feVAR'

June 19, 2022

Type Package

Title Fixed Effects Vector Autoregressive Models

Version 0.1.1

Date 2022-06-19

Author Alessandro Magrini

Maintainer Alessandro Magrini <alessandro.magrini@unifi.it>

Description Estimation and analysis of fixed effects vector autoregressive models.

Depends R (>= 3.5.0)

Imports graphics, stats

License GPL-2

NeedsCompilation no

R topics documented:

`gammadlm-package` *Fixed Effects Vector Autoregressive Models*

Description

Estimation and analysis of fixed effects vector autoregressive models.

Details

Package:	feVAR
Type:	Package
Version:	0.1.1
Date:	2022-06-19
License:	GPL-2

Fixed effects Vector Autoregressive Models are Vector Autoregressive Models (VAR) for panel data, firstly introduced by Holtz-Eakin et al. (1988). Let $Y_1, \ldots, Y_m$ be m endogenous variables and $X_1, \ldots, X_q$ be q exogenous variables. Also, let $\boldsymbol{y}_{i,t}$ be an m-dimensional vector including the values of endogenous variables observed on unit i at time t, and $\boldsymbol{x}_{i,t}$ be a q-dimensional vector including the values of exogenous variables observed on unit i at time t. A fixed effect Vector Autoregressive Model (feVAR) of order $p \in \mathrm{N}_0$ is defined as:

$$\boldsymbol{y}_{i,t} = \boldsymbol{\alpha}_i + \sum_{j=1}^{p} \boldsymbol{B}_j \boldsymbol{y}_{i,t-j} + \boldsymbol{G}\boldsymbol{x}_{i,t} + \boldsymbol{\varepsilon}_{i,t}$$

where:

- $\boldsymbol{\alpha}_i$ is a p-dimensional vector including the intercepts of the endogenous variables for unit i;
- $\boldsymbol{B}_j$ is a $p \times p$ matrix including the regression coefficients among the endogenous variables at time lag $j = 1, \ldots, p$;
- $\boldsymbol{G}$ is a $q \times q$ matrix including the regression coefficients of exogenous variables on endogenous variables;
- $\boldsymbol{\varepsilon}_{i,t}$ is a p-dimensional vector of random errors for unit i at time t such that $\mathrm{E}(\boldsymbol{\varepsilon}_{i,t}) = 0$ and $\mathrm{E}(\boldsymbol{\varepsilon}_{i,t}\, \boldsymbol{\varepsilon}_{k,t}) = 0 \; \forall i, k, t$.

The order p represents the number of lags up to which endogenous variables are supposed to influence each other, which can be determined based on information criteria.

The main functions of the package are:

- unirootTest, to check stationarity of the time series;
- feVAR, to estimate a fixed effect autoregressive model through bias-corrected ordinary least squares (Dhaene & Jochmans, 2016);
- residualPlot, to display residual diagnostics.
- stabilityCheck, to check stationarity (stability) of an estimated model;
- predict.feVAR and predictPlot, to perform and display predictions.

Author(s)

Alessandro Magrini <alessandro.magrini@unifi.it>

References

G. Dhaene, and K. Jochmans (2016). Bias-corrected estimation of panel vector autoregressions. *Economics Letters*, 145: 98-103. DOI: 10.1016/j.econlet.2016.06.010

D. Holtz-Eakin, W. Newey, and H. S. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6), 1371-1395. DOI: 10.2307/1913103

agrisus2020 *EU agricultural sustainability data*

Description

Data on several indicators covering the economic, social and environmental dimensions of agricultural sustainability for 27 EU countries plus United Kingdom in the period 2004-2020.

Usage

```
data(agrisus2020)
```

Format

Object of class `data.frame` with a total of 476 observations on the following 16 variables:

`Country` Country name.

`Country_code` Country code.

`Year` Time of measurement (year).

`TFP_2015` Total Factor Productivity (TFP) index of agriculture (2015=100). Source: USDA, Economic Research Service (ERS).

`NetCapital_GVA` Net capital stocks in agriculture (2015 US dollars) to gross value added of agriculture (2015 US dollars). Source: Faostat.

`Manager_ratio` Ratio young/elderly for farm managers (number of managers with less than 35 years by 100 managers with 55 years and over). Source: CMEF.

`FactorIncome_paid_2010` Real income of agricultural factors per paid annual work unit (index 2010=100). Source: Eurostat.

`EntrIncome_unpaid_2010` Net entrepreneurial income of agriculture per unpaid annual work unit (index 2010=100). Source: Eurostat.

`Income_rur` Median equivalised net income in rural areas (power purchasing standard). Source: Eurostat.

`Unempl_rur` At-risk-of-poverty rate in rural areas (%). Source: Eurostat.

`Poverty_rur` Unemployment rate in rural areas (%). Source: Eurostat.

`RenewProd_UAA` Production of renewable energy from agriculture (share of utilized agricultural area, %). Source: CMEF.

`Organic_p` Area under organic cultivation (share of utilized agricultural area, %). Source: Faostat.

`GHG_UAA` Greenhouse gas emissions due to agriculture (million CO2 equivalent grams per hectare of utilized agricultural area). Source: Faostat.

`GNB_UAA` Gross nitrogen balance (tonnes of nutrient per hectare of utilized agricultural area). Source: Eurostat.

`GDP` Gross domestic product (million 2015 international US dollars). Source: World Bank.

Note

This dataset was employed in Magrini & Giambona (2022) and in Magrini (2022).

References

United States Department of Agriculture (2022). International agricultural productivity. USDA, Economic Research Service (ERS), Washington, DC, US. https://www.ers.usda.gov/data-products/international-agricultural-productivity

European Commission (2022). Eurostat database. https://ec.europa.eu/eurostat/data/database

European Commission (2020). Common Monitoring and Evaluation Framework (CMEF) for the CAP 2014-2020. https://agridata.ec.europa.eu/extensions/DataPortal/cmef_indicators.html

Food and Agriculture Organization (2022). Faostat statistical database. https://www.fao.org/faostat/en/#home

A. Magrini, F. Giambona (2022). A composite indicator to assess sustainability of agriculture in European Union countries. *Social Indicators Research*, published online: April 2022. DOI: 10.1007/s11205-022-02925-6

A. Magrini (2022). Assessment of agricultural sustainability in European Union countries: A group-based multivariate trajectory approach. *Advances in Statistical Analysis*, published online: March 2022. DOI: 10.1007/s10182-022-00437-9

`djstock` *Data on Dow Jones stock*

Description

Closing price of Dow Jones companies from 20 March 2019 to 9 June 2022.

Usage

```
data(djstock)
```

Format

Object of class `data.frame` with a total of 814 observations on 31 variables. The first variable represents the date of observation, while the subsequent variables represent the closing price of each Dow Jones company.

`feVAR` *Estimation of a fixed effect vector autoregressive model*

Description

Estimation of a fixed effect vector autoregressive model through bias-corrected ordinary least squares (Dhaene & Jochmans, 2016). Both panel and non-panel data are allowed.

Usage

```
feVAR(var.names, unit=NULL, time=NULL, exogenous=NULL, data, max.nlags=NULL, nlags=NULL,
  add.intercept=TRUE, box.cox=1, ndiff=0, local.adapt=FALSE, ic="bic", quiet=FALSE)
```

Arguments

var.names	Character vector of length 2 or greater including the names of the endogenous variables, which must be quantitative variables.
unit	Character containing the name of the variable that identifies the units of observation. If NULL (the default), a single unit of observation is assumed.
time	Character containing the name of the time variable, which must be in numeric or date format. If NULL (the default), data are assumed to be temporally ordered.
exogenous	Character vector including the names of the exogenous variables, which can be either quantitative or qualitative variables. If NULL (the default), no exogenous variable is included in the model.
data	Object of class data.frame containing the variables in var.names, exogenous, unit and time.
max.nlags	Non-negative value indicating the maximum lag order p of the model. If NULL (the default) and argument nlags is not NULL, the maximum lag order is determined automatically and the best one is selected based on information criteria.
nlags	The lag order p of the model. If NULL (the default), the best lag order is selected based on information criteria (see argument max.nlags).
add.intercept	Logical value indicating whether the intercept should be included in the model. Default is TRUE. See 'Note'.
box.cox	Named vector with non-negative real values indicating the parameters of the Box-Cox transformation (Box & Cox, 1964) for variables in var.names and in exogenous. If box.cox has no names and length greater than one, the same ordering as in c(var.names,exogenous) is assumed. If box.cox has no names and length equal to one, the same parameter is used for all variables in var.names and in exogenous. Value 0 of the parameter equates to the logarithmic transformation, while value 1 means no transformation. Default is 1 for all variables in var.names and in exogenous. See 'Note'.
ndiff	Named vector with non-negative integer values indicating the number of differences for variables in var.names and in exogenous. If ndiff has no names and length greater than one, the same ordering c(var.names,exogenous) is assumed. If ndiff has no names and length equal to one, the same number of differences is used for all variables in var.names and in exogenous. Value 0 means no differencing. Default is 0 for all variables in var.names and in exogenous. See 'Note'.
local.adapt	Logical value indicating whether the selection of the lag order should be performed locally (i.e., for each variable in each equation). Default is FALSE.
ic	The information criterion to be used in the automated selction of the lag order p of the model, which should be one among 'bic' (Bayesian information criterion), 'hqic' (Hannan & Quinn information criterion), and 'aic' (Akaike information criterion). Ignored if nlags is not NULL. Default is 'bic'.
quiet	Logical value indicating whether prompt messages should be suppressed. Default is FALSE.

Details

S3 methods available for class feVAR include:

- summary provides a summary of parameter estimation;

- coef returns point estimation of parameters for each equation;

- `residuals` returns the residuals;

- `fitted.values` returns the fitted values;

- `confint` returns confidence intervals of parameters for each equation;

- predict.feVAR performs predictions.

Also, the function `stabilityCheck` can be used to check whether an estimated vector autoregressive model is stationary (stable).

Value

An object of class `feVAR` including the following components:

- `models`: list of objects of class `lm`, including ordinary least squares estimation of each equation;
- `call`: list including the main arguments passed to the function `feVAR`;
- `nlags`: the lag order `p` of the model;
- `intercepts`: matrix including the estimated intercepts for each unit (by row) and endogenous variable (by column);
- `Beta`: point estimation of regression coefficients among endogenous variables (matrix B);
- `Gamma`: point estimation of regression coefficients of exogenous variables on endogenous variables (matrix G);
- `Sigma`: estimated covariance matrix of random errors;
- `ic`: value of information criteria for the estimated model;
- `data`: original data (before Box-Cox transformation and differencing);
- `data.used`: data used for parameter estimation (after Box-Cox transformation and differencing).

Note

Second-order stationarity (expected value and autocorrelation function independent of time) of all time series is a basic assumption of the model, that is guaranteed if no time series contains unit roots. Before calling the function feVAR, the user is strongly recommended to check the absence of unit roots in each time series through the function unirootTest. Stationarity of an estimated model (stability) can be check through the function stabilityCheck.

The first order difference of logarithmic values (`box.cox=0` and `ndiff=1`) approximates the relative changes with respect to the previous time point.

If a variable contains negative values, the Box-Cox transformation will be not applied and a warning is returned.

If the number of differencing exceeds `n-5`, where `n` is the sample size, differencing will be not applied and a warning is returned.

If an endogenous variable is not differenced, the intercepts capture the effect of time-invariant factors besides exogenous variables. Otherwise, the intercept represents the drift (i.e., the coefficient of a linear deterministic trend). To avoid the drift, argument `add.intercept` should be set to `FALSE` when differencing is applied to all endogenous variables.

References

H. Akaike (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6): 716-723. DOI: 10.1109/TAC.1974.1100705

G. Dhaene, and K. Jochmans (2016). Bias-corrected estimation of panel vector autoregressions. *Economics Letters*, 145: 98-103. DOI: 10.1016/j.econlet.2016.06.010

E. J. Hannan, and B. G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, Series B, 41: 190-195.

G. E. Schwarz (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2): 461-464. DOI: 10.1214/aos/1176344136

See Also

unirootTest; stabilityCheck; predict.feVAR; predictPlot; residualPlot.

Examples

```
##  EXAMPLE WITH A SINGLE UNIT OF OBSERVATION  ##

data(djstock)
varNames <- colnames(djstock)[-1]

# fit a model with p=1 on data in logarithmic differences ('box.cox'=0 and 'ndiff'=1)
m_dj <- feVAR(varNames, time="Date", data=djstock,
  box.cox=0, ndiff=1, nlags=1)
summary(m_dj)  ## summary of parameter estimation

# automated selection of the lag order
m_dj_auto <- feVAR(varNames, time="Date", data=djstock,
  box.cox=0, ndiff=1, ic="bic", max.nlags=5)
summary(m_dj_auto)  ## summary of parameter estimation


##  EXAMPLE WITH PANEL DATA  ##

data(agrisus2020)
varNames <- colnames(agrisus2020)[4:15]

# fit a model with p=1 on data in logarithmic differences ('box.cox'=0 and 'ndiff'=1)
m_agr <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, nlags=1)
summary(m_agr)  ## summary of parameter estimation

# automated selection of the lag order
m_agr_auto <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, ic="bic", max.nlags=4)
summary(m_agr_auto)  ## summary of parameter estimation
```

LAG *Matrix of lagged values*

Description

Generate the matrix of lagged values for a quantitative variable.

Usage

```
LAG(x, p, unit=NULL, cut=0)
```

Arguments

x	Numerical vector representing temporally ordered data.
p	Numerical value indicating the number of lags. If p is not integer, it is rounded to the nearest integer. If p is lower than 1, x is returned with no changes.
unit	Character containing the name of the units of observation. If NULL (the default), a single unit of observation is assumed.
cut	Numerical value indicating the number of initial time points to be set to missing. Default is 0.

Value

A matrix with length(x) rows and p columns provided that p is greater or equal to 1. Otherwise, the vector x with no changes.

Examples

```
# single unit of observation
x <- rnorm(10)  ## simulate a white noise on 10 time points
LAG(x, 3)       ## 3 lags

# panel data
x <- rnorm(30)
u <- rep(1:3, each=10)  ## 3 units, each with 10 time points
LAG(x, 3, unit=u)       ## 3 lags
```

---

predict.feVAR *Predict method for class feVAR*

Description

Perform predictions based on a fixed effects vector autoregressive model.

Usage

```
## S3 method for class 'feVAR'
predict(object, newdata=NULL, n.ahead=0, unit.id=NULL, level=0.95, ...)
```

Arguments

object	Object of class feVAR.
newdata	Object of class data.frame on which predictions should be based. Ignored if n.ahead is equal to 0. If NULL (the default) and n.ahead is equal to 0, predictions are based on data used for parameter estimation (in-sample predictions).
n.ahead	Integer non-negative number indicating the number of steps ahead of predictions. If equal to 0, in-sample predictions are returned (see argument newdata).

unit.id	Name or identificative number of the units of observation for which predictions should be made. Ignored in case of a single unit of observation. If NULL (the default), predictions are performed for all units of observation.
level	Numerical value between 0 and 1 indicating the level of prediction intervals. Default is 0.95.
...	Further parameters to be passed to the generic method predict.

Value

A named list of data frames with one component for each endogenous variable. Each data frame has several columns including the unit name (only in case of panel data), the time of observation, point predictions, lower and upper bound of prediction intervals.

See Also

feVAR.

Examples

```
## Not run:
data(agrisus2020)
varNames <- colnames(agrisus2020)[4:15]
m_agr <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, nlags=1)

predict(mod_agr)             ## in-sample prediction
predict(mod_agr, n.ahead=3)  ## 3 steps ahead

## End(Not run)
```

---

predictPlot *Graphics of predictions*

Description

Graphic of in-sample or out-of-sample predictions for a fixed effects vector autoregressive models.

Usage

```
predictPlot(model, unit.id, var.names=NULL, n.ahead=0, newdata=NULL,
  level=0.95, xlim=NULL, ylim=NULL, add.grid=TRUE, cex.axis=NULL, cex.points=0.6,
  xlab=NULL, ylab=NULL, titles=NULL, fit.col="dodgerblue", out.col="red",
  obs.col="grey40", new.col="grey40", fit.lty=1, out.lty=1, obs.lty=1, new.lty=2,
interval.col="grey70", las=NULL, mar=c(3.5,3.5,2,2), mgp=c(2.3,0.8,0), ...)
```

Arguments

model	An object of class feVAR.
unit.id	Name or identificative number of one unit of observation for which predictions should be displayed. Ignored in case of a single unit of observation.
var.names	Charcater vector indicating the name of endogenous variables for which prediction should be displayed.

n.ahead	Integer non-negative number indicating the number of steps ahead of predictions. If equal to 0, in-sample predictions are returned.
newdata	Object of class data.frame containing new observations to be added to the graphics. If NULL (the default), no further observations are added. Ignored if n.ahead is equal to 0.
level	Numeric value between 0 and 1 indicating the level of confidence intervals. If equal to 0, confidence intervals are not provided. Default is 0.95.
xlim	Numeric vector of length 2 indicating the limits of the x-axis.
ylim	Numeric vector of length 2 indicating the limits of the y-axis.
add.grid	Logical value indicating whether the grid should be added. Default is TRUE.
cex.axis	Numeric vector of length 2 indicating the expansion factor for x-axis and y-axis labels, respectively. Default is (1,1).
cex.points	Numeric value indicating the expansion factor for points. If 0, points are suppressed. Default is 0.6.
xlab	Label for the x-axis.
ylab	Label for the y-axis.
titles	Character vector containing the title for each graphic. If NULL (the default), the names of endogenous variables are used as titles.
fit.col	Color for the time series of fitted values. Default is black.
out.col	Color for the time series of predictions. Default is red.
obs.col	Color for the time series of observed data. Default is black.
new.col	Color for the time series of new data. Default is black.
fit.lty	Line type for the time series of fitted values. Default is 1.
out.lty	Line type for the time series of predictions. Default is 1.
obs.lty	Line type for the time series of observed data. Default is 1.
new.lty	Line type for the time series of new data. Default is 1.
interval.col	Color of prediction intervals. Default is grey40.
las	Numeric value indicating the orientation of the tick mark labels. Possible values are: parallel to the axis (0), horizontal (1), perpendicular to the axis (2), vertical (3). If NULL (the default), it is set automatically based on the length of axis labels.
mar	Numeric vector of length 4 indicating the margin sizes in the following order: bottom, left, top, right. Default is c(3.5,3.5,2,2).
mgp	Numeric vector of length 3 indicating the location of axis labels, tick mark labels and tick marks, respectively, relative to the plot window. Default is c(2.3,0.8,0).
...	Further parameters to be passed to the generic method plot.

See Also

feVAR; predict.feVAR.

Examples

```
## Not run:
data(agrisus2020)
varNames <- colnames(agrisus2020)[4:15]
m_agr <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, nlags=1)

# predictions for the first four variables on unit 'Italy'
predictPlot(m_agr, unit.id="Italy", var.names=varNames[1:4])  ## in-sample
predictPlot(m_agr, unit.id=15, var.names=1:4)                 ## same as before
predictPlot(m_agr, unit.id=15, var.names=1:4, n.ahead=3)      ## 3 steps ahead

## End(Not run)
```

---

residualPlot *Graphics of residual diagnostics*

Description

Display graphics of residual diagnostics for a fixed effect vector autoregressive model.

Usage

```
residualPlot(model, type="ts", var.names=NULL, max.nlags=NULL, signif=0.05,
 acf.return=FALSE, ylim=NULL, cex.points=0.6, add.grid=TRUE, xlab=NULL, ylab=NULL,
  titles=NULL, las=0, mar=c(3.5,3.5,2,2), mgp=c(2.3,0.8,0), ...)
```

Arguments

model	An object of class feVAR.
type	Type of diagnostic, which should be one among 'ts' (1): time series plot of residuals, 'acf' (2): autocorrelograms of residuals, 'qq' (3): normal quantile plot of residuals, and 'fitVSres' (4): fitted values versus residuals.
var.names	Charcater vector indicating the name of endogenous variables for which prediction should be displayed.
max.nlags	Integer non-negative number indicating the number of lags to be displayed in the ACF. If NULL (the default), it is automatically determined. Ignored if type is not equal to 'acf'.
signif	Numeric value between 0 and 1 indicating the significance level of the autocorrelation test. Ignored if type is not equal to 'acf'. Default is 0.05.
acf.return	Logical value indicating whether autocorrelations should be returned. Ignored if type is not equal to 'acf'. Default is FALSE.
ylim	Numeric vector of length 2 indicating the limits of the y-axis.
cex.points	Numeric value indicating the expansion factor for points. If equal to 0, points are suppressed. Ignored if type is not equal to 'qq' or 'fitVSres'. Default is 0.6.
add.grid	Logical value indicating whether the grid should be added. Ignored if type is equal to 'acf'. Default is TRUE.
xlab	Label for the x-axis.

ylab	Label for the y-axis.
titles	Character vector containing the title for each graphic. If NULL (the default), the names of endogenous variables are used as titles.
las	Numeric value indicating the orientation of the tick mark labels. Possible values are: parallel to the axis (0), horizontal (1), perpendicular to the axis (2), vertical (3). If NULL (the default), it is set automatically based on the length of axis labels.
mar	Numeric vector of length 4 indicating the margin sizes in the following order: bottom, left, top, right. Default is c(3.5,3.5,2,2).
mgp	Numeric vector of length 3 indicating the location of axis labels, tick mark labels and tick marks, respectively, relative to the plot window. Default is c(2.3,0.8,0).
...	Further parameters to be passed to the generic method plot.

See Also

feVAR; predict.feVAR.

Examples

```
## Not run:
data(agrisus2020)
varNames <- colnames(agrisus2020)[4:15]
m_agr <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, nlags=1)

# time series plot of residuals
residualPlot(m_agr, type="ts", cex.main=1.1)

# autocorrelogram of residuals
residualPlot(m_agr, type="acf", cex.main=1.1)

# normal quantile plot of residuals
residualPlot(m_agr, type="qq", cex.main=1.1)

# fitted versus residuals
residualPlot(m_agr, type="fitVSres", cex.main=1.1)

## End(Not run)
```

---

stabilityCheck *Stability check for a fixed effect vector autoregressive model*

Description

Check whether an estimated fixed effect vector autoregressive model is stationary (stable).

Usage

```
stabilityCheck(model)
```

Arguments

model	Object of class feVAR.

Details

A vector autoregressive model is stationary (stable) if all eigenvalues of the companion matrix have modulus less than 1.

Value

Numerical vector including the modulus of each eigenvector of the companion matrix.

See Also

feVAR.

Examples

```
data(agrisus2020)
varNames <- colnames(agrisus2020)[4:15]
m_agr <- feVAR(varNames, unit="Country", time="Year", exogenous="GDP",
  data=agrisus2020, box.cox=0, ndiff=1, nlags=1)

stabilityCheck(m_agr)  ## all moduli are lower than 1 -> stable
```

---

unirootTest *Unit root and stationarity tests for panel data*

Description

Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests for a multivariate time series, that may have a panel structure.

Usage

```
unirootTest(var.names, unit=NULL, time=NULL, data, box.cox=1, ndiff=0, max.nlags=NULL)
```

Arguments

var.names	Character vector including the name of the variables to be differenced.
unit	Character containing the name of the variable that identifies the units of observation. If NULL (the default), a single unit of observation is assumed.
time	Character containing the name of the time variable, which must be in numeric or date format. If NULL (the default), data are assumed to be temporally ordered.
data	Object of class data.frame containing the variables in var.names, unit and time.

box.cox — Named vector with non-negative real values indicating the parameters of the Box-Cox transformation (Box & Cox, 1964) for variables in x.names. If box.cox has no names and length greater than one, the same ordering as in x.names is assumed. If box.cox has no names and length equal to one, the same parameter is used for all variables in x.names. Value 0 of the parameter equates to the logarithmic transformation, while value 1 means no transformation. Default is 1 for all variables in x.names.

ndiff — Named vector with non-negative integer values indicating the number of differences for variables in x.names. If ndiff has no names and length greater than one, the same ordering as in x.names is assumed. If ndiff has no names and length equal to one, the same number of differences is used for all variables in x.names. Value 0 means no differencing. Default is 0 for all variables in x.names.

max.nlags — Non-negative integer value representing the maximum lag length at which to perform the tests (see 'Details'). If NULL (the default), it is taken as the squared root of the length of the time series.

.

Details

The variables subjected to the tests must be quantitative. Missing values internal to the time series are imputed through linear interpolation, otherwise they are deleted out.

The lag length at which to perform the tests is selected through AIC-based backward elimination starting from the lag length specified in argument max.lag.

The null hypothesis of the ADF test (Dickey & Fuller, 1981) is the presence of unit roots, while the null hypothesis of the KPSS test (Kwiatkowski *et al.*, 1992) is the absence of unit roots. Therefore, p-value higher than 0.05 for the ADF test or p-value lower than 0.05 for the KPSS test suggest the presence of unit roots and the need of further differencing.

In case of panel data, p-values are combined according to the method by Demetrescu *et al.* (2006).

Value

One list for each variable in var.names, each with three components:

- statistic: test statistic for each test;
- lag.selected: lag length selected for each test;
- p.value: p-value of each test, which is a single value if unit is NULL, otherwise one value for each unit of observation plus another one indicating the combined p-value;
- box.cox: parameter of the Box-Cox transformation for each variable subjected to the tests;
- ndiff: order of differencing for each variable subjected to the tests.

Note

The first order difference of logarithmic values (box.cox=0 and ndiff=1) approximates the relative changes with respect to the previous time point.

If a variable contains negative values, the Box-Cox transformation will be not applied and a warning is returned.

If the number of differencing exceeds n-5, where n is the sample size, differencing will be not applied and a warning is returned.

References

M. Demetrescu, U. Hassler, and A. Tarcolea (2006). Combining significance of correlated statistics with application to panel data. *Oxford Bulletin of Economics and Statistics*, 68(5), 647-663. DOI: 10.1111/j.1468-0084.2006.00181.x

D. A. Dickey, and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4): 1057-1072. DOI: 10.2307/1912517

D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3): 159-178. DOI: 10.1016/0304-4076(92)90104-Y

Examples

```
data(agrisus2020)
varNames <- names(agrisus2020)[4:15]

# tests on variables in level
unirootTest(varNames, unit="Country", time="Year", data=agrisus2020)

# tests on variables in logarithmic differences
unirootTest(varNames, unit="Country", time="Year", data=agrisus2020,
  box.cox=0, ndiff=1)
```

Index