Package ‘pdR’

August 18, 2024

Type Package

Title Threshold Model and Unit Root Tests in Cross-Section and Time Series Data

Version 1.9.3

Date 2024-08-18

Author Ho Tsung-wu [aut, cre]

Maintainer Ho Tsung-wu <tsungwu@ntnu.edu.tw>

Description
Threshold model, panel version of Hylleberg et al. (1990) <DOI:10.1016/0304-4076(90)90080-D> seasonal unit root tests, and panel unit root test of Chang (2002) <DOI:10.1016/S0304-4076(02)00095-7>.

LazyData TRUE

LazyLoad yes

Depends R (>= 3.5.0)

Imports plm, pglm, lme4, glmmTMB

NeedsCompilation no

Repository CRAN

Date/Publication 2024-08-18 13:10:01 UTC

Contents

`pdR-package` *Panel Data Regression: Threshold Model and Unit Root Tests*

Description

Functions for analysis of panel data, including the panel threshold model of Hansen (1999,JE), panel unit root test of Chang(2002,JE) based upon instuments generating functions (IGF), and panel seasonal unit root test based upon Hylleberg et al.(1990,JE).

Details

This version offers formatted output. This package designs a specification function ptm() to estimate the panel threshold model of Hansen(1999). The key feature of ptm() is to generalize Hansen's original code to allow multiple (more-than-one) regime-dependent right-hand-side independent variables; Dr. Hansen's original code admits only 1 regime-dependent right-hand-side independent variable. This version also includes panel unit root tests based on the instrument generating functions(IGF), proposed by Chang (2002, J. of Econometrics), and the panel version of Hylleberg et al.(1990) seasonal unit root test, proposed by Otero, et al. (2005, 2007).

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.
Maintainer: Ho Tsung-wu <tsungwu@ntnu.edu.tw>

References

Chang, Y. (2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. Journal of Econometrics, 110, 261-292.
Hansen B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.
Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. Journal of Econometrics,44, 215-238.
Otero, J., Smith, J., and Giulietti, M. (2005) Testing for seasonal unit roots in heterogeneous panels. Economics Letters, 86, 229-235.
Otero, J., Smith, J., and Giulietti, M. (2007) Testing for seasonal unit roots in heterogeneous panels in the presence of cross section dependence. Economics Letters, 86, 179-184.
Pesaran M. Hashem (2007) A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics, 22, 265-312.

`bank_income` *Panel data of bank,2001Q1~2010Q1*

Description

A quarterly panel data frame with 1000 observations on the following 7 variables, unbalanced panel data ranges from 2001Q1~2010Q1.

Usage

```
data("bank_income")
```

Format

`ID` a numeric vector

`Qtr` a numeric vector

`preTax_Income` a numeric vector

`shortRatio` a numeric vector

`longRatio` a numeric vector

`Current_ratio` a numeric vector

`LoanDeposit_ratio` a numeric vector

Examples

```
data(bank_income)
```

cigaretts *Cigaretts consumption of US states*

Description

Cigaretts consumption of US states

Usage

```
data(cigaretts)
```

Format

A data frame of 48 US states' cigaretts consumption

`State` State abbreviation, N

`Year` Year, t

`Y_SALES` Cigarette sales in packs per capita, deflated by population

`X1_PRICE` P=Real price per pack of cigarettes, deflated by 1983 CPI.

`X2_PIMIN` Real minimum price in adjoining states per pack of cigarettes, deflated by CPI

`X3_NDI` Per capita disposable income

References

Baltagi Badi H. (2005) Econometric Analysis of Panel Data. John Wiley.

Examples

```
data(cigaretts)
head(cigaretts)
```

contts *Function for extracting components from a lm object*

Description

Extract the standard error and t-stat of the a-th parameter estimate of a lm object

Usage

```
contts(lm, a)
```

Arguments

`lm`	lm object
`a`	The a-th parameter estimate of a linear model regression

Value

se.coef	The standard error of the selected coefficient
t.stat	The t-stat of the selected coefficient

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle.

References

Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
x=rnorm(100)
y=1+0.2*x+rnorm(100)
LMout=lm(y~x)
contts(LMout,1)

#$se.coef
#[1] 0.1081023

#$t.stat
#(Intercept)
#   10.60401
```

`crime` *Annual crime dataset of US counties*

Description

Annual crime dataset of US counties

Usage

```
data(crime)
```

Format

A data frame of US counties

`county` counties index, N

`year` Year, t

`crmrte` crime rate(crime/population)

`prbarr` probability of arrest (arrests/offenses)

`prbconv` probability of conviction, given arrest `prbpris` probability of a prison, given conviction

`avgsen` sanction severity(average prison sentence in days)

`polpc` ability of police force to detect crime(# of police per capita)

`density` population density(POP/area)

`taxpc` Taxpayment per capita

`region` region index of county

`smsa` =1 if SAMA, POP>50000; =0 else

`pctmin` See Baltagi(2006) for details

`wcon` See Baltagi(2006) for details

`wtuc` See Baltagi(2006) for details

`wtrd` See Baltagi(2006) for details

`wfir` See Baltagi(2006) for details

`wser` See Baltagi(2006) for details

`wmfg` See Baltagi(2006) for details

`wfed` See Baltagi(2006) for details

`wsta` See Baltagi(2006) for details

`wloc` See Baltagi(2006) for details

`mix` See Baltagi(2006) for details

`pctymle` See Baltagi(2006) for details

References

Baltagi Badi H. (2005) Econometric Analysis of Panel Data. John Wiley. Baltagi Badi H. (2006) Estimating an Economic Model of Crime Using Panel Data from North Carolina. J.of Applied Econometrics 21: 543¡V547.

`dur_john` *The cross-country growth data in Durlauf and Johnson(1995)*

Description

The Durlauf-Johnson data manupipulated by Hansen(2000),excluding missing variables and oil states

Usage

```
data(dur_john)
```

Format

A data frame with 19 countries

- `gdpGrowth` Economic growth measured by GDP of 1960 and 1985
- `logGDP60` log Per capita GDP in 1960
- `Inv_GDP` Average ratio of investment (including Government Investment) to GDP from 1960 to 1985
- `popGrowth` Average growth rate of working-age population 1960 to 1985
- `School` Average fraction of working-age population enrolled in secondary school from 1960 to 1985
- `GDP60` Per capita GDP in 1960
- `Literacy` fraction of the population over 15 years old that is able to read and write in 1960

Details

Steven N. Durlauf and Paul A. Johnson, "Multiple Regimes and Cross-Country Growth Behavior," Journal of Applied Econometrics, Vol. 10, No. 4, 1995, 365-384.

Examples

```
data(dur_john)
head(dur_john)
```

`hegy.reg` *Generate the HEGY regressors.*

Description

This function generates the level regresors in HEGY regression, without differenced lag terms.

Usage

```
hegy.reg(wts)
```

Arguments

`wts`	Univariate time series, with a possibly seasonal stochastic trend

Details

This function automatically identifies the frequency of time series data, and generate necessary level components as described in Eq.(3.7) of Hylleberg et. al (1990).

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. Journal of Econometrics,44, 215-238.
Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y=inf_Q[,1]
hegy.reg(y)
```

`HEGY.test` *Seasonal unit root test based on Hylleberg et al. (1990)*

Description

The function performs seasonal unit root test based on Eq.(3.6) of Hylleberg et al. (1990), univariate time series.

Usage

```
HEGY.test(wts, itsd, regvar = 0, selectlags = list(mode = "signf", Pmax = NULL))
```

Arguments

`wts`	Univariate time series
`itsd`	Options for c(i,t,sd) i=1, intercept;=0 no intercept t=1, trend;=0 no deterministic trend sd=1, season dummy 1:(s-1);=0 no.
`regvar`	Additional regressors
`selectlags`	Selection of lags mode, Criteria for selection, having three options: "signf","bic","aic". Pmax, maximum number of lags.

Details

Mode for selectlags has three options, AIC and BIC use R built-in functions for linear model and their meanings are popular and straightforward. They include only lags that meet specific criterion, others are dropped from regressors. That is, lag orders of your model may not be a regular sequence. See also selPsignf() and selPabic().

Value

stats	Tests statistics for HEGY regression coefficients.
hegycoefs	HEGY regression coefficients.
lagsorder	Lags order. "aic" or "bic" returns a scalar; "signf" returns a sequence of numbers
lagcoefs	Coefficients of lag terms.
regvarcoefs	Coefficient(s) of additional regressor(s).

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. Journal of Econometrics,44, 215-238.
Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y<-inf_Q[,1]
hegy.out<-HEGY.test(wts=y, itsd=c(1,0,c(1:3)),regvar=0, selectlags=list(mode="aic", Pmax=12))

hegy.out$stats #HEGY test statistics
names(hegy.out) # HEGY objects, which can be called by using $, see below.
hegy.out$hegycoefs
hegy.out$regvarcoefs
```

---

htest_pglm   *Specification test for panel glm models*

Description

This function performs Hausman specification test for panel glm.

Usage

```
htest_pglm(RE, FE, re.method)
```

Arguments

RE	Random effect objects. Support pglm, glmer, glmmTMB
FE	Fixed effect objects.
re.method	Method that used to estimate the random effect estimation, in addition to "pglm", it also supports "glmmTMB" of package glmmTMB, and "glmer" of package lme4.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from phtest() of plm

References

Hausman J.A. (1978). Specification Tests in Econometrics. Econometrica, 46, 1251-1271.

Examples

```
data(ship)
library(pglm)
Eq1="accident  ~  op_75_79+co_65_69+co_70_74+co_75_79"
FE.pois <- pglm(Eq1,data=ship,family = "poisson",model = "within",index = 'ship',R=10)

RE.pois <- pglm::pglm(Eq1,data=ship,family = "poisson", model = "random", index = 'ship')

## Hausman test
htest_pglm(RE=RE.pois,  FE=FE.pois, re.method="pglm")

Eq2=accident  ~  op_75_79+co_65_69+co_70_74+co_75_79 + (1 | ship)
re.glmmTMB=glmmTMB::glmmTMB(Eq2,data=ship, family="poisson")

## Hausman test
htest_pglm(RE=re.glmmTMB,  FE=FE.pois, re.method="glmmTMB")
```

IGF *Unit root test based on Change(2002)*

Description

This function estimates the unit root regression based on instrument generating function of Change(2002) and returns useful outputs.

Usage

```
IGF(y, maxp, ic, spec)
```

Arguments

y	A univariate time series data
maxp	the max number of lags
ic	Information criteria, either "AIC" or "BIC"
spec	regression model specification. =0, no intercept and trend. =1, intercept only. =2, intercept and trend.

Details

Estimate univariate unit root test of Chang(2002).

Value

`tstat.IGF`	IGF unit root test
`beta`	regression coefficients. The first one is the AR(1) coefficient of unit root, and the last one is the intercept or trend
`sdev`	The IGF standard error for unit root coefficient
`cV`	The scalar C in IGF equation
`p`	The optimal number of lag

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Chang, Y. (2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. Journal of Econometrics, 110, 261-292.

Examples

```
data(inf19)
y <- inf19[,1]
IGF(y,maxp=35,ic="BIC",spec=2)$tstat.IGF
```

`inf19` *Monthly inflation time series of 19 countries*

Description

Monthly inflation time series of 19 countries, 1984.1~2011.3

Usage

```
data(inf19)
```

Format

A data frame with 19 countries

`AUSTRIA` inflation of Austria

`BELGIUM` inflation of Belgium

`CANADA` inflation of Canada

`DENMARK` inflation of Denmark

`FINLAND` inflation of Finland

`FRANCE` inflation of France

`GREECE` inflation of Greece

`ICELAND` inflation of Iceland

`ITALY` inflation of Italy

`JAPAN` inflation of Japan

`LUXEMBOURG` inflation of Luxembourg

`NETHERLANDS` inflation of Netherlands

`NORWAY` inflation of Norway

`PORTUGAL` inflation of Portugal

`SPAIN` inflation of Spain

`SWEDEN` inflation of Sweden

`SWITZERLAND` inflation of Switzerland

`UK` inflation of UK

`USA` inflation of USA

Details

Monthly CIP, seasonaly differenced of log CPI of 19 countries

Examples

```
data(inf19)
head(inf19)
```

`inf_M` *Monthly inflation time series of 20 countries*

Description

Monthly inflation time series of 19 countries, 1971.1~2011.12

Usage

```
data(inf_M)
```

Format

A data frame with 20 countries

`AUSTRALIA` inflation of Austrlia

`AUSTRIA` inflation of Austria

`BELGIUM` inflation of Belgium

`CANADA` inflation of Canada

`DENMARK` inflation of Denmark

`FINLAND` inflation of Finland

`FRANCE` inflation of France

`GREECE` inflation of Greece

`ICELAND` inflation of Iceland

`ITALY` inflation of Italy

`JAPAN` inflation of Japan

`LUXEMBOURG` inflation of Luxembourg

`NETHERLANDS` inflation of Netherlands

`NORWAY` inflation of Norway

`PORTUGAL` inflation of Portugal

`SPAIN` inflation of Spain

`SWEDEN` inflation of Sweden

`SWITZERLAND` inflation of Switzerland

`UK` inflation of UK

`USA` inflation of USA

Details

Monthly CIP, seasonaly differenced of log CPI of 20 countries

Examples

```
data(inf_M)
head(inf_M)
```

`inf_Q` *Quarterly inflation time series of 20 countries*

Description

Quarterly inflation time series of 19 countries, 1971Q1~2014Q4

Usage

```
data(inf_Q)
```

Format

A data frame with 19 countries

- `AUSTRLIA` inflation of Austrlia
- `AUSTRIA` inflation of Austria
- `BELGIUM` inflation of Belgium
- `CANADA` inflation of Canada
- `DENMARK` inflation of Denmark
- `FINLAND` inflation of Finland
- `FRANCE` inflation of France
- `GREECE` inflation of Greece
- `ICELAND` inflation of Iceland
- `ITALY` inflation of Italy
- `JAPAN` inflation of Japan
- `LUXEMBOURG` inflation of Luxembourg
- `NETHERLANDS` inflation of Netherlands
- `NORWAY` inflation of Norway
- `PORTUGAL` inflation of Portugal
- `SPAIN` inflation of Spain
- `SWEDEN` inflation of Sweden
- `SWITZERLAND` inflation of Switzerland
- `UK` inflation of UK
- `USA` inflation of USA

Details

Quarterly CIP, seasonaly differenced of log CPI of 20 countries

Examples

```
data(inf_Q)
head(inf_Q)
```

`interpolpval` *Extracting critical value and p-value from Table 1 of Hylleberg et. al (1990)*

Description

Hylleberg et. al (1990,pp.226-227) offer simulated critical values for seasonal unitr to test. interpolpval() is an internal call and should not be used independently.

Usage

```
interpolpval(code, stat, N, swarn = TRUE)
```

Arguments

`code`	Type of HEGY model, this will be automatically identified.
`stat`	Empirical test statistics.
`N`	Sample size calculating stat above.
`swarn`	Logical. Whether the warning message for negative p-value will be returned? The default is TRUE.

Value

`table`	Table for critical value and p-value.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. Journal of Econometrics,44, 215-238.
Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

`invest` *investment data of 565 listed companies, 1973-1987*

Description

investment data of 565 listed companies, 1973-1987, from Hansen's example

Usage

```
data(invest)
```

Format

A pooled data frame

`invest[,1]` investment/assets

`invest[,2]` Tobin's Q

`invest[,3]` cash-flow/assets

`invest[,4]` debt/assets

Details

This is a pooled data frame, without date (T) and cross-section(N) ID columes

Examples

```
#data(invest)
#head(invest)
```

`ipsHEGY` *IPS-HEGY seasonal unit root test in panel data, Otero et al.(2007).*

Description

This function performs panel data-based HEGY seasonal unit root test, the asymptotics is based upon Otero et al.(2007).

Usage

```
ipsHEGY(data, itsd, Sel, pmax, CIPS = TRUE)
```

Arguments

`data`	Panel data, T by N
`itsd`	Options for c(i,t,sd). i=1, intercept;=0 no intercept. t=1, trend;=0 no deterministic trend. sd=1, season dummy 1:(s-1);=0 no.
`Sel`	Selection of lags, having three options: "signf","bic","aic".
`pmax`	Maximum number of lags for searching optimal criteria.
`CIPS`	Logical. If TRUE, using Pesaran(2007) to account for cross-section correlation. The default is TRUE.

Details

Mode for selectlags has three options, AIC and BIC use R built-in functions for linear model and their meanings are popular and straightforward. "signf" includes only statistically significant lags, and statistically insignificant lags are dropped from regressors. That is, once you select this option, lags of your model may not be continuous.
The critical values for panel HEGY are standard normal for individual t-ratios, however,you need to perform simulation for the critical values of F joint test, at pdR 1.3. To this end, you are encouraged to work this out for yourself: using arima.sim() to sample seasonal time series with unit root (1-order difference) and obtain their statistics under the null using ipsHEGY(), then it is streightforward to obtain critical values.
Otero et al. (2007) provide critical values for quarterly frequency.

Value

P_HEGY	Panel HEGY statistics.
U_HEGY	Idividual HEGY statistics of N units.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Otero, J., Smith, J., and Giulietti, M. (2005) Testing for seasonal unit roots in heterogeneous panels. Economics Letters, 86, 229-235.
Otero, J., Smith, J., and Giulietti, M. (2007) Testing for seasonal unit roots in heterogeneous panels in the presence of cross section dependence. Economics Letters, 86, 179-184.
Pesaran M. Hashem (2007) A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics, 22, 265-312.

Examples

```
data(inf_Q)
dataz<-inf_Q
itsd<-c(1,0,c(1:3))
#Seasonal dummy only takes quarters 1:3,
#becasue of the presence of common intercept.
Sel<-"bic" # "aic","bic", "signf".
pmax<-12

OUT<-ipsHEGY(dataz,itsd,Sel,pmax,CIPS=FALSE)
OUT$P_HEGY
OUT$U_HEGY

# Simulation of critical values
```

lagSelect *Select the optimal number of lags, given criteria*

Description

Determine the optimal number of lags for dynamic regression

Usage

```
lagSelect(y, maxp, ic)
```

Arguments

y	A univariate time series data
maxp	the max number of lags
ic	Information criteria, either "AIC" or "BIC"

Details

Information criteria "AIC" and "BIC" use the R built-in functions.

Value

It returns an integer, indicating the optimal lags

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

Examples

```
#library(pdR)
#data(inf19)
#y<-inf19[,1]
#lagSelect(y,maxp=25,ic="BIC")
```

`lookupCVtable` *Function for looking up tabulated critical values and associated p-values of HEGY test.*

Description

Function for looking up tabulated critical values and associated p-values,Hylleberg et. al (1990, Table 1a and Table 1b).

Usage

```
lookupCVtable(code)
```

Arguments

`code`	Type of HEGY model, this will be automatically identified.

Value

`table`	Table for critical value and p-value.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S.(1990) Seasonal integration and cointegration. Journal of Econometrics,44, 215-238.
Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

`model` *Estimate specified panel threshold model*

Description

This function is the main function estimating threshold regression for function ptm()

Usage

```
model(r, trim, rep, it, qq1, cf, xt, ct, thresh, tt, qn1, n, qn, cc, yt, ty, k)
```

Arguments

r	vector of threshold estimate(s).
trim	value of trimmed percentage.
rep	number bootstrap repetition.
it	number of regime during computation, used in a for loop.
qq1	defined parameter.
cf	special declaration, e.g. lag().
xt	regime independent variables.
ct	trace of regime dependent variables.
thresh	threshold variable.
tt	length of time period.
qn1	as defined by nrow(qq1).
n	number of cross-section units.
qn	number of quantiles to examine.
cc	as defined by 2*log(1-sqrt(conf_lev)).
yt	vectorized dependent variable.
ty	trace of yt.
k	number of regime-independent independent variables.

Note

Original code offered by Dr. B. E.Hansen (http://www.ssc.wisc.edu/~bhansen/).

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

---

pIGF — *Panel unit root test of Chang(2002)*

Description

Compute the panel unit root test statistic of Chang(2002).

Usage

```
pIGF(datamat, maxp, ic, spec)
```

Arguments

`datamat`	T by N panel data.T is the time length,N is the number of cross-section units.
`maxp`	the max number of lags
`ic`	Information criteria, either "AIC" or "BIC".
`spec`	model specification. =0, no intercept and trend. =1, intercept only. =2, intercept and trend.

Details

This function estimates the panel unit root test based on univariate instrument generating function of (Chang,2002).

Value

`panel.tstat`	panel IGF test statistics
`pvalue`	P-value of the panel.tstat

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Chang,Y.(2002) Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. Journal of Econometrics, 110, 261-292.

Examples

```
data(inf19)
datam <- inf19
pIGF(datam,maxp=25,ic="BIC",spec=2)
```

`productivity` *Productivity data of 48 US state,1970-1986*

Description

Gross state production data

Usage

```
data(productivity)
```

Format

A data frame with US production

state US state index, 1-48

year Year index

y_gsp Gross state product

x1_hwy Expediture of public utility- highway construction

x2_water Expediture of public utility- water

x3_other Expediture of others

x4_private Private consumption of each state

x5_emp Employment rate of each state

x6_unemp Unemployment rate of each state

Examples

```
data(productivity)
head(productivity)
```

---

ptm — *Threshold specification of panel data*

Description

A generalized specification for estimating panel threshold model.

Usage

```
ptm(dep, ind1, ind2, d, bootn, trimn, qn, conf_lev, t, n)
```

Arguments

dep	Dependent variable
ind1	Independent variables: regime dependent
ind2	Independent variables:regime independent
d	Threshold variable
bootn	Vector of bootstrap repetition
trimn	Vector of trimmed percentage
qn	Number of quantiles to examine
conf_lev	Confidence level
t	Length of time period
n	Number of cross-section units

Details

This code fits only balanced panel data. It generalizes the simple code of Dr. Hansen (http://www.ssc.wisc.edu/~bhansen/), allowing multiple (more-than-one) regime-dependent (ind1) variables. We generalize the original code to better fit general need of threshold modeling in panel data.
bootn and trimn are vector of 3 by 1, indicating numbers of three corresponding regimes.
This version corrects a slight error incurred by argument max_lag, which is used by Hansen to arrange investment data via lags. In this package, users manipulate data to fit personal research to ptm(), hence this argument is omitted lest should degree of freedom will suffer a loss of N.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, College of Management, National Taiwan Normal University.

References

Hansen B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

Examples

```
# library(pdR)
#data(invest)
#dat<-invest[1:1500,]    # subsetting the first 1500 obs., #for simplicity
#t <- 15            #Length of time period
#nt <- nrow(dat)
#n <- nt/t           # number of cross-section units

#dep<- as.matrix(dat[,1])       # investment/assets
#th1<- as.matrix(dat[,2])  #Tobin's Q
#th2<- as.matrix(dat[,3])  # cash-flow/assets
#ind1<- cbind(th1,th2)  #regime-dep covariates
#d <- as.matrix(dat[,4])      # Threshold variable
#ind2 <- cbind((th1^2),(th1^3),(th1*d)) # regime-indep covariates:
#bootn<-c(100,200,300)  # bootstrapping replications for each  threshold esitmation
#trimn<-c(0.05,0.05,0.05)   #trimmed percentage for each threshold  esitmation

#qn<-400
#conf_lev<-0.95

#Output=ptm(dep,ind1,ind2,d,bootn,trimn,qn,conf_lev,t,n)
#Output[[1]] #Formatted output of 1st threshold, 2 regimes
#Output[[2]] #Formatted output of 2nd threshold, 3 regimes
#Output[[3]] #Formatted output of 3rd threshold, 4 regimes

# In the output, the Regime-dependent Coefficients matrix
# is, from top to bottom, regime-wise.
```

`ret` *Returns a data.frame of sequential lag matrix.*

Description

ret() is similar to embed(), but returns a data.frame specified with colnames, not matrix.

Usage

```
ret(wts, k)
```

Arguments

`wts`	Univariate time series.
`k`	k-1 lagged terms.

Details

ret() is similar to embed(), but returns a data.frame with colnames, not matrix. Moreover, unlike embed(),ret() fills lagged cells with NA, instead of trimming them.

Value

A T by k dataframe returns. If you need 2 lags, you have to specify k=3, then it returns a dataframe with T by 3 dataframe, the first column is lag0.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y=inf_Q[,2]
ret(y,3)
```

`r_est` *A subroutine for model()*

Description

This function is a subroutine for model(), estimation procedure.

Usage

```
r_est(y, r, trim, tt, qq1, qn1, qn, n, cf, xt, ct, thresh)
```

Arguments

`y`	vector of dependent variable.
`r`	numer of regime.
`trim`	value of trimmed percentage.
`tt`	length of time period.
`qq1`	parameter defined by as.matrix(unique(thresh)[floor(sq*nrow(as.matrix(sort(unique(thresh)))))]).
`qn1`	as defined by nrow(qq1).
`qn`	number of quantiles to examine.
`n`	parameter of cross-section units.
`cf`	special declaration, e.g. lag().
`xt`	regime independent variables.
`ct`	trace of regime dependent variables.
`thresh`	threshold variable.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

Original code from Dr. Hansen (http://www.ssc.wisc.edu/~bhansen/).

SeasComponent *Generate a data matrix of seasonal components*

Description

Generate a data matrix of seasonal components, having two pattern cycles.

Usage

```
SeasComponent(wts, type)
```

Arguments

`wts`	A univeriate time series with monthly or quarterly frequency.
`type`	Types of patterns of seasonal cycle. ="dummyCycle", generating dummy variables for the pattern of seasonal cycle, Barsky & Miron (1989) ="trgCycle", generating trigonometric variables for the pattern of seasonal cycle, Granger & Newbold (1986).

Details

This function generates data matrix for controlling the pattern of seasonal cycles. type="dummyCycle" generates DUMMY variables with season frequency. type="trgCycle" generates trigonometric pattern.

Value

A dataframe returns. Number of columns is determined by frequency.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Barsky, Robert B. and Jeffrey A. Miron (1989) The Seasonal Cycle and the Business Cycle. Journal of Political Economy, 97 (3): 503-32.
Granger, Clive William John and Newbold, Paul (1986) Forecasting Economic Time Series. 2nd edition. Published by New Milford, Connecticut, U.S.A.: Emerald Group Pub Ltd.
Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y=inf_Q[,2]
SeasComponent(y,type="dummyCycle")
SeasComponent(y,type="trgCycle")
```

selPabic *Selection of lags.*

Description

Lagged coefficient estimates are kept if they meet the inequality condition of AIC or BIC.

Usage

```
selPabic(lmdet, type, Pmax = NULL)
```

Arguments

`lmdet`	Object of lm()
`type`	Take the value of "aic" or "bic".
`Pmax`	The maximum number of lag orders.

Details

This is an internal function used for HEGY.test(). Beginning with pamx, the lag order will be drop if its inclusion worsens the minimum condition. Hence, they may not be a regular sequence. For example, for pmax=10, the selected lags may look like (1,4,5,8,9), rather than 1,2,3,...10.

Value

This function returns the lag orders.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y=inf_Q[,1]
hegy.out<-HEGY.test(wts=y, itsd=c(1,0,c(1:3)),regvar=0, selectlags=list(mode="aic", Pmax=12))
hegy.out$lagsorder
hegy.out$lagcoefs
```

selPsignf *Selection of lags.*

Description

Lagged coefficient estimates are kept if they are statistically significant

Usage

```
selPsignf(lmdet, cvref = 1.65, Pmax = NULL)
```

Arguments

`lmdet`	Object of lm()
`cvref`	Reference of critical values, the default is 1.65.
`Pmax`	The maximum number of lag orders.

Details

This is an internal function used for HEGY.test(). Beginning with pamx, the lag order will be kept if it is statistically significant. Hence, the lag orders may not be a regular sequence. For example, for pmax=10, the seelcted lags may look like (1,4,5,8,9), rather than 1,2,3,...10.

Value

This function returns the lag orders.

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>, modifed from Javier Lopez-de-Lacalle

References

Javier Lopez-de-Lacalle in https://github.com/cran/uroot/blob/master/R/hegy.R

Examples

```
data(inf_Q)
y=inf_Q[,1]
hegy.out<-HEGY.test(wts=y, itsd=c(1,0,c(1:3)),regvar=0, selectlags=list(mode="signf", Pmax=12))
hegy.out$lagsorder
hegy.out$lagcoefs
```

ship — *Panel data on the number of ship accidents*

Description

Panel data on the number of ship accidents, McCullagh and Nelder(1989)

Usage

```
data("ship")
```

Format

- `accident` the number of ship accidents
- `ship` Ship iD
- `service` the number of months in service
- `op_75_79` the operating period between 1975 and 1979
- `co_65_69` consecutive construction periods of 5 years
- `co_70_74` consecutive construction periods of 5 years
- `co_75_79` consecutive construction periods of 5 years
- `yr_con` years of construction
- `yr_op` years of operation

References

McCullagh, P., and J. A. Nelder (1989) Generalized Linear Models. 2nd ed. London: Chapman and Hall/CRC.

Examples

```
data(ship)
```

SMPLSplit_est — *Estimation of sub-sampled data*

Description

A function for estimating the subsampled data.

Usage

```
SMPLSplit_est(data,dep,indep,th,plot,h=1,nonpar=2)
```

Arguments

`data`	the data in either data.frame or matrix.
`dep`	the name of dependent variable.
`indep`	the name(s) of independent variable(s).
`th`	the name of threshold variable.
`plot`	=1, plot; =0, do not plot.
`h`	h=1, heteroskedasticity-consistent covariance; h=0, homoskedastic case.
`nonpar`	Indicator for non-parametric method used to estimate nuisance scale in the presence of heteroskedasticity (only relevant if h=1).Set nonpar=1 to estimate regressions using a quadratic.Set nonpar=2 to estimate regressions using an Epanechnikov kernel with automatic bandwidth.

Details

This code estimates the parameters of sub-sampled data. It generalizes the simple code of Dr. Hansen, allowing White Corrected Heteroskedastic Errors.

Value

`threshold`	values of threshold estimates.
`est0`	coefficient estimates of global data.
`est.low`	coefficient estimates of low regime.
`est.high`	coefficient estimates of high regime.
`est0.info`	additional information of global data.
`est.joint.info`	additional information of joint threshods.
`est.low.info`	additional information of est.low.
`est.high.info`	additional information of est.high.

Note

Original code offered by Dr. B. E.Hansen (http://www.ssc.wisc.edu/~bhansen/).

References

Hanse B. E. (2000) Sample Splitting and Threshold Estimation. Econometrica, 68, 575-603.

Examples

```
## Not run, becasue of bootstrap replicaiton takes time. Users may unmark # and run.
data("dur_john")
rep <- 500
trim_per <- 0.15
dep <- "gdpGrowth"
indep <- colnames(dur_john)[c(2,3,4,5)]

SMPLSplit_est(data=dur_john,dep,indep,th="GDP60",plot=0,h=1,nonpar=2)
```

`SMPLSplit_example` *Example code for sample splitting*

Description

A sample code for learning sample splitting.

Usage

```
SMPLSplit_example(data,dep,indep,th1,th2,trim_per,rep,plot)
```

Arguments

`data`	the data in either data.frame or matrix.
`dep`	the name of dependent variable.
`indep`	the name(s) of independent variable(s)
`th1`	the first threshold variable.
`th2`	the second threshold variable.
`trim_per`	trimmed percentage.
`rep`	nNumber of bootstrap repetitions.
`plot`	=1, plot; =0, do not plot.

Details

This code is the learning example for learning Hansen's econometric sample splitting. I detailed the description of each threshold stage.

Note

Original code offered by Dr. B. E.Hansen (http://www.ssc.wisc.edu/~bhansen/).

References

Hanse B. E. (2000) Sample Splitting and Threshold Estimation. Econometrica, 68, 575-603.

Examples

```
## Not run, becasue of bootstrap replicaiton takes time. Users may unmark # and run.
data("dur_john")
#rep <- 500
#trim_per <- 0.15
#dep <- "gdpGrowth"
#indep <- colnames(dur_john)[c(2,3,4,5)]
#th1 <- "GDP60"
#th2 <- "Literacy"
#OUT=SMPLSplit_est(data=dur_john,dep,indep,th=th1,plot=0,h=1,nonpar=2)
```

```
#OUT$TEST
#OUT$Hypothesis
#OUT$Threshold
#stat=matrix(as.numeric(OUT$TEST),byrow = TRUE,8,2)
#colnames(stat)=c("F-Stat","P-value")
#rownames(stat)=OUT$Hypothesis
#stat
```

SMPLSplit_het *Testing for sample splitting*

Description

A function for testing sample split given subsampled data.

Usage

```
SMPLSplit_het(data,dep,indep,th,trim_per,rep,plot)
```

Arguments

data	the data in either data.frame or matrix
dep	the name of dependent variable.
indep	the name(s) of independent variable(s).
th	the name of threshold variable.
trim_per	trimmed percentage.
rep	number of bootstrap repetition.
plot	=1, plot; =0, do not plot.

Details

This code tests for the presence of threshold. It generalizes the simple code of Dr. Hansen, allowing Heteroskedastic Errors (White Corrected).

Value

fstat	LM-test for no threshold.
pvalue	bootstrap P-Value.

Note

Original code offered by Dr. B. E.Hansen (http://www.ssc.wisc.edu/~bhansen/).

References

Hanse B. E. (2000) Sample Splitting and Threshold Estimation. Econometrica, 68, 575-603.

sse_calc — *a subroutine of model()*

Description

SSE calculation

Usage

```
sse_calc(y, x)
```

Arguments

This function is a sub-routine for model(), calculating SSE of each regression

y	vector of dependent variable.
x	matrix of independent variables.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

Original code from Dr. Hansen (http://www.ssc.wisc.edu/~bhansen/).

tbar — *Compute the resursive mean*

Description

Compute the resursive mean of each series

Usage

```
tbar(x)
```

Arguments

x	A univariate time series data

Details

This function computes the resursive mean

Author(s)

Ho Tsung-wu <tsungwu@ntnu.edu.tw>

Examples

```
data(inf19)
y <- inf19[,1]
tbar(y)
```

---

thr_sse *a subroutine calculating SSE*

Description

This function is a sub-routine for model(), calculating SSE of each threshold regression.

Usage

```
thr_sse(y, q, r, cf, xt, ct, thresh, tt, n)
```

Arguments

y	parameter.
q	qq1 in model().
r	parameter.
cf	as defined in model().
xt	as defined in model().
ct	as defined in model().
thresh	as defined in model().
tt	as defined in model().
n	as defined in model().

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

Original code from Dr. Hansen (http://www.ssc.wisc.edu/~bhansen/).

`tr` *A sub-routine calculating trace*

Description

Estimation of trace.

Usage

```
tr(y, tt, n)
```

Arguments

This function is a sub-routine for model(), calculating trace of matrix

`y`	data vector.
`tt`	time period length.
`n`	number of cross-section units.

References

Hanse B. E. (1999) Threshold effects in non-dynamic panels: Estimation, testing and inference. Journal of Econometrics,93, 345-368.

Original code from Dr. Hansen (http://www.ssc.wisc.edu/~bhansen/).

Index