Linear Time Series

Christian Francq

CREST-ENSAE

Chapter 4: Unit roots and stationarity tests

Outline

Unit Root (UR) or Trend Stationary (TS) ?

To model a series with trend, 2 possibilities:

1. A trend-stationary model

$$\mathrm{TS}_t = \mathrm{DT}(t) + \mathrm{MA}_t,$$

   where
   - $\mathrm{DT}(t)$ is a deterministic trend
   - $\mathrm{MA}_t = \sum_{i\geq 0} c_i \epsilon_{t-i}$, where (ϵ_t) is i.i.d. $(0,1)$, with $\sum_{i\geq 0} |c_i| < \infty$.

2. A UR model:

$$(1-B)\mathrm{UR}_t = b + \mathrm{MA}_t,$$

   with initial value $\mathrm{UR}_0 = a$.

Examples of UR processes:

- the **random walk** (RW): $(1-B)\mathrm{UR}_t = b + \epsilon_t$,
- AR process with one root equal to 1 and the other roots outside the unit disk

TS deviates less from the trend than UR

Suppose that

$$\mathrm{TS}_t = a + bt + \mathrm{MA}_t \quad \text{and} \quad (1-B)\mathrm{UR}_t = b + \epsilon_t$$

with $\mathrm{UR}_0 = a$.

$$\begin{aligned} \mathrm{UR}_t &= b + \epsilon_t + \mathrm{UR}_{t-1} = 2b + \epsilon_t + \epsilon_{t-1} + \mathrm{UR}_{t-2} = bt + a + \sum_{i=1}^{t} \epsilon_i, \\ \Rightarrow E(\mathrm{TS}_t) &= E(\mathrm{UR}_t) = a + bt. \end{aligned}$$

However,

- $\mathrm{Var}\,\mathrm{TS}_t = \mathrm{Var}\,\mathrm{MA}_t = \sigma^2 \sum_{i\geq 0} c_i^2$ independent of t,
- $\mathrm{Var}\,\mathrm{UR}_t = \mathrm{Var} \sum_{i=1}^{t} \epsilon_i = t\sigma^2 \to \infty$ when $t \to \infty$.

Long term predictions with TS models

$$\mathrm{TS}_{t+h|t} = \mathrm{DT}(t+h) + \mathrm{MA}_{t+h|t} = \mathrm{DT}(t+h) + \sum_{i\geq h} c_i\epsilon_{t+h-i} + \sum_{0\leq i<h} c_i\epsilon_{t+h-i}.$$

If the TS model is invertible, then the prediction at horizon h is

$$\widehat{\mathrm{TS}}_{t+h|t} = \mathrm{DT}(t+h) + \widehat{\mathrm{MA}}_{t+h|t} = \mathrm{DT}(t+h) + \sum_{i\geq h} c_i\epsilon_{t+h-i}.$$

- convergence to the deterministic trend ("mean reversion").
- the prediction interval is bounded, even at the infinite horizon:

$$\begin{aligned} \mathsf{MSE}_{\mathrm{TS}}(h) &:= E(\mathrm{TS}_{t+h} - \widehat{\mathrm{TS}}_{t+h|t})^2 \\ &= \mathrm{Var}\left(\sum_{i=0}^{h-1} c_i\epsilon_{t+h-i}\right) \\ &= \sigma^2\sum_{i=0}^{h-1} c_i^2 \quad \xrightarrow{h\to\infty} \quad \sigma^2\sum_{i=0}^{\infty} c_i^2 = \mathrm{Var}\,\mathrm{TS}_t \end{aligned}$$

Long-term predictions with the UR model

$UR_t = UR_{t-1} + b + \epsilon_t$

Since

$$\mathrm{UR}_{t+h} = \epsilon_{t+h} + \cdots + \epsilon_{t+1} + bh + \mathrm{UR}_t,$$

we have

$$\widehat{\mathrm{UR}}_{t+h|t} = bh + \mathrm{UR}_t.$$

That is, one can predict a steady increase at rate b starting from the current value, and

$$\mathrm{MSE}_{\mathrm{UR}}(h) = E(\mathrm{UR}_{t+h} - \widehat{\mathrm{UR}}_{t+h|t})^2 = E(\epsilon_{t+h} + \cdots + \epsilon_{t+1})^2 = h\sigma^2 \to \infty$$

as $h \to \infty$.

Thus, when $h \to \infty$,

$$\mathrm{MSE}_{\mathrm{TS}}(h) \to \mathrm{cst}, \qquad \mathrm{MSE}_{\mathrm{UR}}(h) \to \infty$$

Effects of shocks

One can interpret the innovation as an economic shock at time t.

For instance: many studies exist on the effects of gas price shocks on the Gross Domestic Product (GDP).

$\epsilon_t < 0$, can be interpreted as an exogenous "bad news" (entails a smaller value for X_t than $\hat{X}_{t|t-1}$).

Interpreting X_{t+h} as the resultant of all shocks $\epsilon_u, u \leq t+h$, the effect of a small variation of the shock at time t on X_{t+h} is $\frac{\partial X_{t+h}}{\partial \epsilon_t}$:

$$X_{t+h} \text{ increases by } \frac{\partial X_{t+h}}{\partial \epsilon_t}\delta \text{ if } \epsilon_t \text{ increases by } \delta, \text{ with } \delta \text{ small.}$$

Effects of shocks on UR and TS

- The effect of a shock on TS is temporary:

$$\frac{\partial \mathrm{TS}_{t+h}}{\partial \epsilon_t} = \frac{\partial}{\partial \epsilon_t}\left(\mathrm{DT}(t+h) + \sum_{i\geq 0} c_i \epsilon_{t+h-i}\right) = c_h \overset{h\to\infty}{\to} 0$$

- The effect of a shock on UR is persistent: $(1-B)\mathrm{UR}_t = b + \mathrm{MA}_t$, $\mathrm{UR}_0 = a$ with $\mathrm{MA}_t = \sum_{i\geq 0} c_i \epsilon_{t-i}$,

$$\frac{\partial \mathrm{UR}_{t+h}}{\partial \epsilon_t} = \frac{\partial}{\partial \epsilon_t}\left(a + b(t+h) + \sum_{j=1}^{t+h} \mathrm{MA}_j\right) = c_0 + c_1 + \cdots + c_h \to \sum_{i\geq 0} c_i.$$

Indeed, we have:

$$\begin{aligned}
\mathrm{MA}_{t+h} &= c_0\epsilon_{t+h} + c_1\epsilon_{t+h-1} + \cdots + c_h\epsilon_t + \cdots \\
\mathrm{MA}_{t+h-1} &= c_0\epsilon_{t+h-1} + c_1\epsilon_{t+h-2} + \cdots + c_{h-1}\epsilon_t + \cdots \\
&\vdots \\
\mathrm{MA}_t &= c_0\epsilon_t + c_1\epsilon_{t-1} + c_2\epsilon_{t-2} + \cdots
\end{aligned}$$

Effects of shocks on UR and TS

This has **important consequences in terms of modeling**:

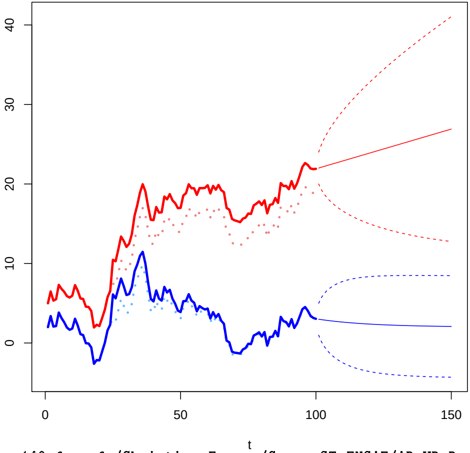
If a TS model is plausible for the GDP series, the effect a gas price shock should eventually vanish.

If a UR model is plausible, the economy should never completely recover from the effect of a recession due to negative shocks.

stationary AR (blue) and random walk (red),

with a shock at time $t = 25$.

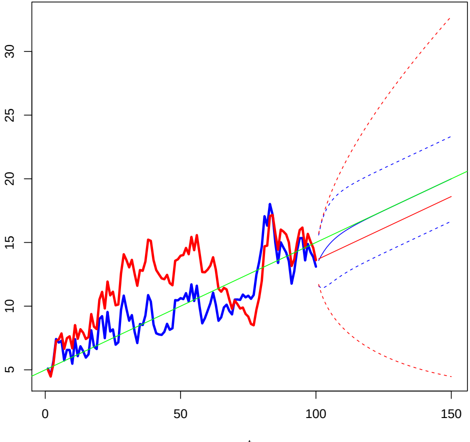
last 50 values: predictions and their confidence intervals (dotted lines)



http://christian.francq140.free.fr/Christian-Francq/Cours-ST-ENSAE/AR_UR.R

TS (blue) and UR (red)

last 50 values: predictions and their confidence intervals (dotted lines)
common trend in green (mean reversion for TS)



http://christian.francq140.free.fr/Christian-Francq/Cours-ST-ENSAE/mean_reversion.R

Estimation of an AR(1): $X_t = \rho X_{t-1} + \epsilon_t$

Suppose now that (ϵ_t) is strong WN $(0, \sigma^2)$ with $\sigma^2 > 0$.

We have seen that in the stationary case $(|\rho| < 1)$,

$$\sqrt{n}(\hat{\rho}_n - \rho) \quad \xrightarrow{\mathscr{L}} \quad \mathscr{N}\left(0, 1-\rho^2 = \frac{\sigma^2}{\gamma(0)}\right)$$

when $n \to \infty$.

At the 5% level, one rejects $H_0 : \rho = 0$ (WN) if

$$t_n^0 := \frac{\sqrt{n}\hat{\rho}_n}{\sqrt{\frac{\hat{\sigma}^2}{\hat{\gamma}(0)}}} > 1.96.$$

Estimation of a AR(1) with a unit root

When $\rho = 1$, the behaviour of $\hat{\rho}_n$ is **completely different**.

The estimator $\hat{\rho}_n$ is called super consistent because $\hat{\rho}_n \to 1$ at faster rate than $\sqrt{n}$ (the convergence to 1 may hold even if ϵ_t is not WN). More precisely, with the usual t-statistic

$$t_n = \frac{\hat{\rho}_n - 1}{\hat{\sigma}_{\hat{\rho}_n}}, \quad \hat{\sigma}^2_{\hat{\rho}_n} = \frac{\hat{\sigma}^2}{\sum_{t=1}^n X_{t-1}^2}, \quad \hat{\sigma}^2 = \frac{\sum_{t=1}^n (X_t - \hat{\rho}_n X_{t-1})^2}{(n-1)},$$

if $X_t = \rho X_{t-1} + \epsilon_t$ with $\rho = 1$ and $X_0 = 0$,

$n(\hat{\rho}_n - 1) \sim$ Dickey-Fuller distribution, $\quad t_n \sim$ other DF distributions

tabulated by Fuller (1976) in the Gaussian case (when the law of ϵ_t is non Gaussian, the distribution is only valid asymptotically).*

*To allow for $X_0 \neq 0$, the statistics are computed on $Y_t = X_{t+1} - X_1$ for $t = 1, \ldots n-1$ which satisfies also $Y_t = \rho Y_{t-1} + \epsilon_{t+1}$ with $Y_0 = 0$.

Unit root test in the AR(1) without trend

The simplest test is for

Case 1: $X_t = \rho X_{t-1} + \epsilon_t, \qquad H_0 : \rho = 1$

Null hypothesis: RW without trend,

$$H_0 : \qquad X_t = X_{t-1} + \epsilon_t$$

Alternative assumption: stationary AR(1) with zero mean,

$$H_1 : \qquad X_t = \rho X_{t-1} + \epsilon_t, \qquad |\rho| < 1.$$

The model can be written under "error correction model (ECM)" form: using the difference operator ($\nabla = 1 - B$) with $\pi = \rho - 1$,

$$\nabla X_t = \pi X_{t-1} + \epsilon_t, \qquad H_0 : \pi = 0$$

The critical region (rejection of H_0) has the form

$$\{n(\hat{\rho}_n - 1) < C\} \quad \text{or} \quad \{t_n < C^*\},$$

where the critical values C and C^* are tabulated.

Dickey-Fuller table (?adfTable for the help)

```
> library("fUnitRoots")
> adfTable(trend="nc",statistic="n")$z
0.010 0.025 0.050 0.100 0.900 0.950 0.975 0.990
25 -11.9  -9.3  -7.3  -5.3  1.01  1.40  1.79  2.28
50 -12.9  -9.9  -7.7  -5.5  0.97  1.35  1.70  2.16
100 -13.3 -10.2  -7.9  -5.6  0.95  1.31  1.65  2.09
250 -13.6 -10.3  -8.0  -5.7  0.93  1.28  1.62  2.04
500 -13.7 -10.4  -8.0  -5.7  0.93  1.28  1.61  2.04
Inf -13.8 -10.5  -8.1  -5.7  0.93  1.28  1.60  2.03
> adfTable(trend="nc",statistic="t")$z
0.010 0.025 0.050 0.100 0.900 0.950 0.975 0.990
25 -2.66 -2.26 -1.95 -1.60  0.92  1.33  1.70  2.16
50 -2.62 -2.25 -1.95 -1.61  0.91  1.31  1.66  2.08
100 -2.60 -2.24 -1.95 -1.61  0.90  1.29  1.64  2.03
250 -2.58 -2.23 -1.95 -1.62  0.89  1.29  1.63  2.01
500 -2.58 -2.23 -1.95 -1.62  0.89  1.28  1.62  2.00
Inf -2.58 -2.23 -1.95 -1.62  0.89  1.28  1.62  2.00
```

Example: CAC 40 from 01/03/1990 to 18/01/2013 (n=5795)

Letting $Y_t = CAC_{t+1} - CAC_1$, the AR coefficient is estimated by

$$\hat{\rho}_n = \frac{\sum_{t=2}^{n-1} Y_t Y_{t-1}}{\sum_{t=2}^{n-1} Y_{t-1}^2} = 0.9997686.$$

The UR hypothesis at level 5% (or even at much larger levels) cannot be rejected because

$$\hat{\pi} = 5794(0.9997686 - 1) = -1.34 >> -8.1.$$

The value of the Student statistics is $t_n = -0.1602$: same conclusion.

Doing the same test on the **returns** series $Y_t = \log(CAC_{t+1}/CAC_t)$ yields $t_n = -55.6627$, which leads to reject the UR assumption.

UR test in the AR(1) with constant

For many series, the assumption of **zero mean** in the AR(1) is not plausible.

A RW **without trend** is also not appropriate for many series.

To incorporate a trend under H_0 and H_1, different situations:

Case 1:	$X_t = \rho X_{t-1} + \epsilon_t,$	$H_0 : \rho = 1$
Case 2:	$X_t = c + \rho X_{t-1} + \epsilon_t,$	$H_0 : \rho = 1$ and $c = 0$
Case 3:	$X_t = c + \rho X_{t-1} + \epsilon_t,$	$H_0 : \rho = 1$ and $c \neq 0$
Case 4:	$X_t = c + bt + \rho X_{t-1} + \epsilon_t,$	$H_0 : \rho = 1$ and $b = 0$

- in Case 2, the asymptotic distribution of $\hat{\rho}_n$ is **not the same** as in case 1 under H_0 (a different table of critical values must be used) since an intercept c is estimated.
- in Case 3, the asymptotic distribution of $\hat{\rho}_n$ is Gaussian.

Which is the correct case to use for testing UR?

In the absence of precise idea: fit a model that is plausible under both H_0 and H_1.

If there is an obvious trend use case 4; in the absence of trend use case 2.

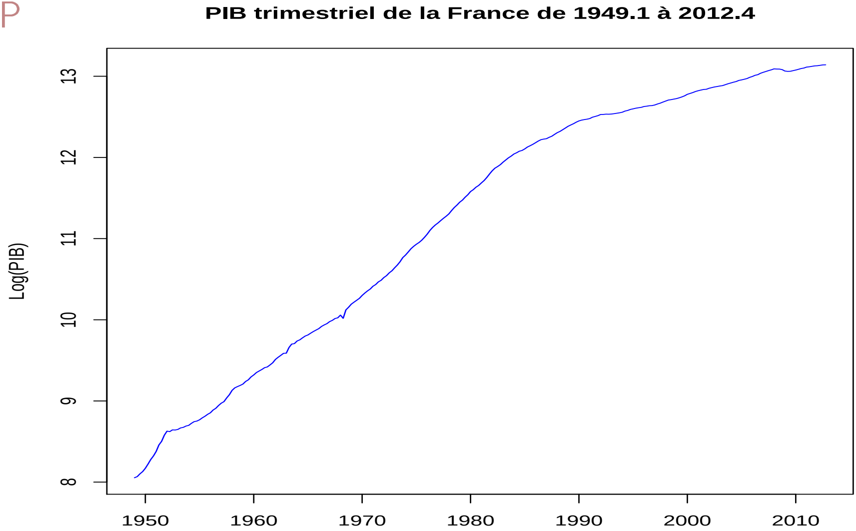
Case 3 is not much interesting (too obvious).

In cases 2 and 4, tests only depend of $\hat{\rho}_n$ (or t_n) in general, but one can use a F statistics to test the values of the 2 parameters (with a non standard Fisher Snedecor asymptotic distribution).

CAC 40

```
> adfTest(indcac, type = "nc") # cas 1
STATISTIC:
Dickey-Fuller: -0.1602
P VALUE:
0.5653
> adfTest(indcac, type = "c") # cas 2
STATISTIC:
Dickey-Fuller: -1.7264
P VALUE:
0.4155
```

GDP


PIB trimestriel de la France de 1949.1 à 2012.4
Log(PIB)
8
9
10
11
12
13
1950
1960
1970
1980
1990
2000
2010

UR test on the series log(GDP)

From 1949.1 to 1986.2

```
> adfTest(log(pib)[1:150], lags=0, type = "ct")# case 4
STATISTIC:
Dickey-Fuller: -1.3034
P VALUE:
0.8666
```

GDP, with JmulTi

```
ADF Test for series:      GDP
sample range:             [1949 Q2, 1986 Q2], T = 149
lagged differences:       0
intercept, time trend
asymptotic critical values
reference: Davidson, R. and MacKinnon, J. (1993),
"Estimation and Inference in Econometrics" p 708, table 20.1,
Oxford University Press, London
1%         5%         10%
-3.96      -3.41      -3.13
value of test statistic: -1.3034
regression results:
---------------------------------------
variable      coefficient   t-statistic
---------------------------------------
x(-1)        -0.0220       -1.3034
constant       0.2511        1.4659
trend          0.0006        1.2812
RSS            0.0347

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:             [1951 Q4, 1986 Q2], T = 139

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:    2
Final Prediction Error:   2
Hannan-Quinn Criterion:   2
Schwarz Criterion:        2
```

Error correction model

AR(1) is often too simple.

⇒ AR(p) models with $p > 1$ to account for more past values.

An AR(p) can be written under Error Correction Model (ECM) form: using the difference operator ($\nabla = 1 - B$)

$$X_t - \mu = \sum_{i=1}^{p} \phi_i (X_{t-i} - \mu) + \epsilon_t$$

⇒

$$\nabla X_t = \pi (X_{t-1} - \mu) + \sum_{i=1}^{p-1} \pi_i \nabla X_{t-i} + \epsilon_t$$

Proof

with $\pi = -1 + \sum_{i=1}^{p} \phi_i$, $\pi_i = -(\phi_{i+1} + \cdots + \phi_p)$ for $i = 1, \ldots, p-1$.

Null hypothesis on the ECM

Suppose the process is $I(1)$, that is, there is one (and only one) UR ($\pi = 0$):

$$H_0: \qquad 1 - \sum_{i=1}^{p} \phi_i z^i = (1-z)\left(1 - \sum_{i=1}^{p-1} \pi_i z^i\right)$$

with

$$1 - \sum_{i=1}^{p-1} \pi_i z^i \neq 0, \qquad \forall |z| \leq 1.$$

For a series of size n, the critical regions for H_0 are

$$\left\{(n-p)\frac{\hat{\pi}}{1-\hat{\pi}_1 - \cdots - \hat{\pi}_{p-1}} < \text{constant}\right\} \qquad \text{or} \qquad \{t_n < \text{constant}\}.$$

The lagged variables ∇X_{t-i} entail a change of the (first) test statistic, but **no change of the asymptotic distribution**
(the same table can be used but, even in the Gaussian case, it is only valid asymptotcally).

GDP

```
ADF Test for series:      GDP
sample range:             [1949 Q4, 1986 Q2], T = 147
lagged differences:       2
intercept, time trend
asymptotic critical values
1%          5%          10%
-3.96       -3.41       -3.13
value of test statistic: -2.0999
regression results:
---------------------------------------
variable       coefficient   t-statistic
---------------------------------------
x(-1)         -0.0338       -2.0999
dx(-1)          0.1207        1.5343
dx(-2)          0.3324        4.2168
constant        0.3588        2.1984
trend           0.0009        2.0756
```

Check that GDP is I(1)

```
ADF Test for series:       PIB_d1
sample range:              [1949 Q4, 1986 Q2], T = 147
lagged differences:        1
intercept, no time trend
asymptotic critical values
1%         5%         10%
-3.43      -2.86      -2.57
value of test statistic: -5.6185
regression results:
---------------------------------------
variable      coefficient   t-statistic
---------------------------------------
x(-1)        -0.5781       -5.6185
dx(-1)        -0.3130       -3.9654
constant       0.0161         5.1620
```

Perron-Phillips tests

Phillips (1987 *Econometrica*), and Phillips and Perron (1988 *Biometrika*) proposed tests of $H_0 : \rho = 1$ in semi-parametric models of the form

$$X_t = \rho X_{t-1} + u_t, \quad X_t = c + \rho X_{t-1} + u_t, \quad X_t = c + bt + \rho X_{t-1} + u_t,$$

where u_t is a very general error term (not necessarily a WN).

Remarks:

- In the stationary case ($|\rho| < 1$), the LS estimator $\hat{\rho}_n$ of ρ is generally inconsistent when u_t is autocorrelated.
- In the UR case ($\rho = 1$), it can be shown that, under appropriate assumptions, $\hat{\rho}_n$ **converges to** $\rho = 1$**, even if** u_t **is not WN.**

Perron-Phillips statistics

For UR models with very general error terms, Phillips (1987) showed that the asymptotic distribution of the LS estimator $\hat{\rho}_n$ depends on:

- the marginal variance

$$\sigma_u^2 := \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} u_t^2 \qquad \text{a.s.},$$

- the long-term variance

$$\vartheta_u^2 := \lim_{n\to\infty} \mathrm{Var}\left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \right\}.$$

Phillips (1987), and Phillips and Perron (1988) proposed non parametric estimators of these two quantities and modified the DF test statistics to reach the same asymptotic distribution in the case where u_t is iid, or when $\sigma_u^2 \neq \vartheta_u^2$.

Perron-Phillips test in Case 1

PP tests use the fact that, under general assumptions,

$$Z_\phi := n(\hat{\rho}_n - 1) - \frac{n^2 \hat{\sigma}^2_{\hat{\rho}_n}}{2\hat{\sigma}^2_u}\left(\hat{\vartheta}^2_u - \hat{\sigma}^2_u\right) \xrightarrow{\mathscr{L}} \text{first asympt. law of DF},$$

and

$$Z_t := \frac{\hat{\sigma}_u}{\hat{\vartheta}_u}\frac{\hat{\rho}_n - 1}{\hat{\sigma}_{\hat{\rho}_n}} - \frac{n\hat{\sigma}_{\hat{\rho}_n}}{2\hat{\sigma}_u\hat{\vartheta}_u}\left(\hat{\vartheta}^2_u - \hat{\sigma}^2_u\right) \xrightarrow{\mathscr{L}} \text{second asympt. law of DF}.$$

Similar results can be obtained in Cases 2 and 4.

PP tests on the GDP

```
> pib.pp <- ur.pp(log(pib)[1:150], type="Z-tau", model='trend')
> summary(pib.pp)
##################################
# Phillips-Perron Unit Root Test #
##################################

Call:
lm(formula = y ~ y.l1 + trend)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.2508426  0.1710949   1.466    0.145
y.l1        0.9779588  0.0169107  57.831   <2e-16 ***
trend       0.0005979  0.0004667   1.281    0.202

Value of test-statistic, type: Z-tau  is: -1.7231

Critical values for Z statistics:
1pct     5pct    10pct
critical values -4.021997 -3.44051 -3.14447

>
> pib.pp <- ur.pp(diff(log(pib)[1:150]), type="Z-tau", model='trend')
> summary(pib.pp)

Value of test-statistic, type: Z-tau  is: -10.7621
```

Various other tests

- Schmidt-Phillips (1992) (in Case 4, filtering of the (polynomial) deterministic trend before applying PP as in Case 2).
- Elliot, Rothenberg and Stock (1996): looking for "efficient" tests.
- UR test in presence of a break (Zivot and Andrews, 1992 and Saikkonen and Lutkepohl, 2002).
- ...

KPSS test by Kwiatkowski, Phillips, Schmidt and Shin (1992)

KPSS proposed to test the null hypothesis of a (trend) stationary model:

$$y_t = \xi t + r_t + \epsilon_t, \quad r_t = r_{t-1} + u_t,$$

for $t \geq 1$, where (ϵ_t) is a stationary process, $\xi = 0$ if there is no deterministic trend, r_0 is taken constant, and u_t iid $(0, \sigma_u^2)$. The aim is to **test stationarity** (up to a deterministic trend), that is:

$$H_0 : \qquad \sigma_u^2 = 0.$$

In the case where ϵ_t iid $\mathcal{N}(0, \sigma_\epsilon^2)$, KPSS showed that the Lagrange Multiplyer (LM) test rejects H_0 for large values of

$$LM = \frac{\sum_{k=1}^n S_k^2}{n^2 \hat{\sigma}_\epsilon^2}, \quad \text{where} \quad S_k = \sum_{i=1}^k e_i, \quad \hat{\sigma}_\epsilon^2 = n^{-1} \sum_{i=1}^n e_i^2,$$

and the e_i's are the residuals of the regression of the y_t's on 1 (and t if $\xi \neq 0$).

KPSS test

KPSS showed that the asymptotic law of LM does not depend on the law of the ϵ_t's: it can be tabulated with ϵ_t Gaussian.

However, the law depends on whether $\xi = 0$ or $\xi \neq 0$.

Moreover, the asymptotic law of LM is unchanged when (ϵ_t) satisfies the assumptions used by Perron Phillips, provided that $\hat{\sigma}_\epsilon^2$ is replaced by a consistent estimator of the long-term variance $\hat{\vartheta}_u^2$.

KPSS on the Log of the GDP and its first difference

```
KPSS test for series:  GDP
sample range:          [1949 Q2, 1986 Q2], T = 149
number of lags:        2
KPSS test based on y(t)=a+bt+e(t) (trend stationarity)
asymptotic critical values:
10%        5%         1%
0.119      0.146      0.216
value of test statistic: 0.8879

KPSS test for series:  GDP_d1
sample range:          [1949 Q2, 1986 Q2], T = 149
number of lags:        2
KPSS test based on y(t)=a+e(t) (level stationarity)
asymptotic critical values:
10%        5%         1%
0.347      0.463      0.739
value of test statistic: 0.1206
```

Testing stationarity of the Earth's global temperature

```
> kpss.test(GlobalTemp, null = "Level")# the stationarity is rejected

KPSS Test for Level Stationarity

data:  GlobalTemp
KPSS Level = 2.8864, Truncation lag parameter = 4, p-value = 0.01

Warning message:
In kpss.test(GlobalTemp, null = "Level") :
p-value smaller than printed p-value


> kpss.test(GlobalTemp, null = "Trend")# the trend stationarity is rejected

KPSS Test for Trend Stationarity

data:  GlobalTemp
KPSS Trend = 0.55127, Truncation lag parameter = 4, p-value = 0.01

Warning message:
In kpss.test(GlobalTemp, null = "Trend") :
p-value smaller than printed p-value
```

Testing stationarity of the Earth's global temperature

Using package urca instead of tseries is not more friendly

```
> summary(ur.kpss(GlobalTemp,type="tau"))#

#######################
# KPSS Unit Root Test #
#######################

Test is of type: tau with 4 lags.

Value of test-statistic is: 0.5513

Critical value for a significance level of:
10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

Testing UR model for Earth's global temperature

```
> adf.test(GlobalTemp,alternative="e")# ADF does not reject UR with trend

Augmented Dickey-Fuller Test

data:  GlobalTemp
Dickey-Fuller = -1.1538, Lag order = 5, p-value = 0.08878
alternative hypothesis: explosive

> pp.test(GlobalTemp,alternative="e")# PP does not reject UR with trend

Phillips-Perron Unit Root Test

data:  GlobalTemp
Dickey-Fuller Z(alpha) = -29.688, Truncation lag parameter = 4, p-value
= 0.99
alternative hypothesis: explosive

Warning message:
In pp.test(GlobalTemp, alternative = "e") :
p-value smaller than printed p-value
```

Stationarity tests are not magic (require an AR framework)

```
> simu<-0.1*rnorm(1000)+sin(c(1:1000)*2*pi/12)# simulation of a nonstationary series
> kpss.test(simu, null = "Level") # KPSS does not reject the stationarity in level

KPSS Test for Level Stationarity

data:  simu
KPSS Level = 0.0051126, Truncation lag parameter = 7, p-value = 0.1

Warning message:
In kpss.test(simu, null = "Level") : p-value greater than printed p-value

> adf.test(simu,alternative="stationary")# ADF rejects the non stationary model

Augmented Dickey-Fuller Test

data:  simu
Dickey-Fuller = -14.995, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(simu, alternative = "stationary") :
p-value smaller than printed p-value

> pp.test(simu,alternative="stationary")# PP rejects the non stationary model

Phillips-Perron Unit Root Test

data:  simu
Dickey-Fuller Z(alpha) = -185.02, Truncation lag parameter = 7, p-value
= 0.01
alternative hypothesis: stationary
```

References

- Dickey, D.A., Fuller W.A. (1979); Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74, 427-431.
- Hamilton. J. (1994); Time Series Analysis, Princeton University Press.
- Elliot G., Rothenberg T.J., Stock J.H.(1996); Efficient Tests for an Autoregressive Time Series with a Unit Root, Econometrica 64, 813-836.
- Kwiatkowski D., Phillips P.C.B, Schmidt P., Shin Y. (1992); Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, Journal of Econometrics 54, 159-178.
- MacKinnon, J.G. (1991); Critical Values for Cointegration Tests, Long-Run Economic Relationships, eds. R.F. Engle and C.W.J. Granger, London, Oxford, 267-276.

References (continued)

- Perron P. (1988); Trends and Random Walks in Macroeconomic Time Series, Journal of Economic Dynamics and Control 12, 297-332.
- Phillips, P.C.B. (1987); Time Series Regression with a Unit Root, Econometrica 55, 227-301.
- Phillips P.C.B., Perron P. (1988); Testing for a unit root in time series regression, Biometrika 75, 335-346.
- Said S.E., Dickey D.A. (1984); Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, Biometrika 71, 599Ű607.
- Saikkonen P., Lutkepohl H. (2002); Testing for a unit root in a time series with a level shift at unknown time, Econometric theory 18, 313-348.
- Schwert G.W. (1989); Tests for Unit Roots: A Monte Carlo Investigation, Journal of Business and Economic Statistics 2, 147-159.
- Schmidt P., Phillips P.C.B. (1992); LM Test for a Unit Root in the Presence of Deterministic Trends, Oxford Bulletin of Economics and Statistics, 54, 257-287.
- Zivot, E., Andrews D.W.K. (1992); Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis, Journal of Business & Economic Statistics 10 251-270.

End of Chapter 4 ☺ !

Proof of the ECM representation

It is not restrictive to assume $\nu = 0$ and to replace $X_t - \mu$ by X_t. We then have

$$\begin{aligned}
\nabla X_t &= (\phi_1 - 1)X_{t-1} + \sum_{i=2}^{p} \phi_i X_{t-i} + \epsilon_t \\
&= (\sum_{i=1}^{p} \phi_i - 1)X_{t-1} + \sum_{i=2}^{p} \phi_i (X_{t-i} - X_{t-1}) + \epsilon_t \\
&= \pi X_{t-1} + \sum_{i=2}^{p} \phi_i (X_{t-2} - X_{t-1}) + \sum_{i=3}^{p} \phi_i (X_{t-i} - X_{t-2}) + \epsilon_t \\
&= \pi X_{t-1} + \pi_1 \nabla X_{t-1} + \cdots + \pi_{p-2} \nabla X_{t-p-2} + \sum_{i=p}^{p} \phi_i (X_{t-i} - X_{t-p+1}) + \epsilon_t
\end{aligned}$$

Return