

Cointegration

Rob Hayward

June 6, 2013

1 Introduction

This paper will examine time series with a focus on stationarity, integration and cointegration. The first part looks at univariate series, the second looks at multivariate and the final section assesses cointegration. The final section concludes.

2 Stationary data

The standard series to be investigated can take the form of

$$y_t = TD_t + z_t \quad (1)$$

It is possible to differentiate between *trend stationary* and *difference stationary* processes.

$$y_t = y_{t-1} + \mu = y_0 + \mu t \quad (2)$$

and

$$y_t = y_{t-1} + \varepsilon = y_0 + \sum_{i=0}^t \varepsilon_i \quad (3)$$

If all the roots of the autoregressive polynomial $\phi_p(z)$ lie outside the unit circle and there is a trend stationary process; if at least one of the roots lies on the unit circle and there is a difference stationary process.

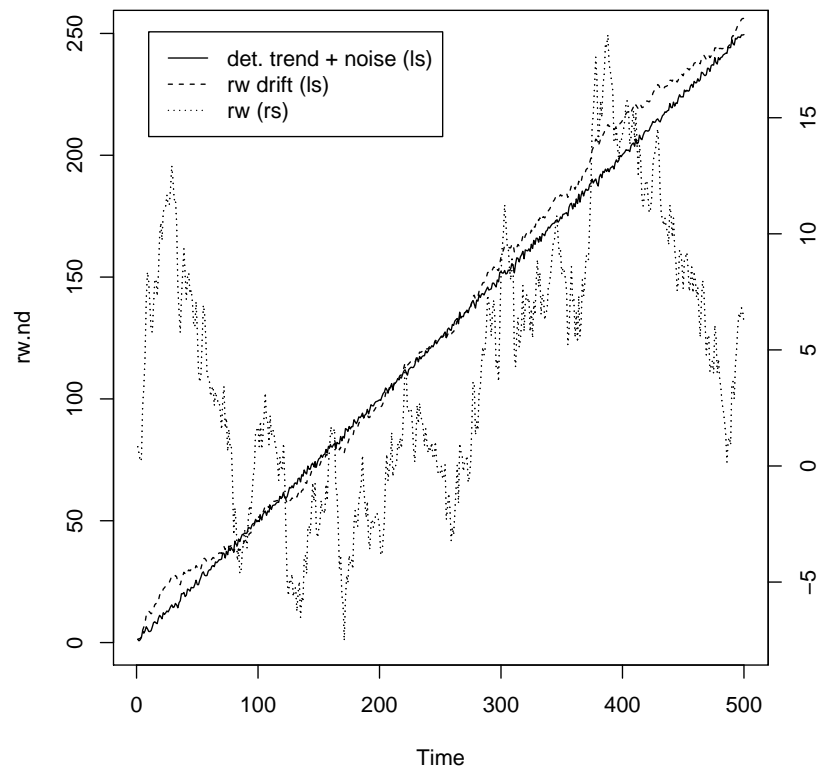
$$\phi_p(z) = 1 - \phi_1(z) - \phi_2(z)^2 - \phi_3(z)^3 \dots \phi_p(z)^p \quad (4)$$

It is possible to create and plot these different types of time series.

```
set.seed(123456)
e <- rnorm(500)
rw.nd <- cumsum(e)
```

```
trd <- 1:500
# random walk with drift
rw.wd <- 0.5 * trd + cumsum(e)
# deterministic trend and noise
dt <- e + 0.5 * trd
```

```
par(mar = rep(5, 4))
plot.ts(dt, lty = 1, ylab = "", xlab = "")
lines(rw.wd, lty = 2)
par(new = T)
plot.ts(rw.nd, lty = 3, axes = FALSE)
axis(4, pretty(range(rw.nd)))
lines(rw.nd, lty = 3)
legend(10, 18.7, legend = c("det. trend + noise (ls)", "rw drift (ls)", "rw (rs)"),
      lty = c(1, 2, 3))
```



There are also a series of tests that can be put in place to assess the nature of the time series. There are three types of stationary series: *trend stationary*, *difference stationary* and *difference stationary with a constant*

Rather than estimating

$$y_t = \beta_1 + \beta_2 t + \rho y_{t-1} + \sum_{j=1}^k \gamma_j \Delta y_i + u_{1t} \quad (5)$$

to test whether ρ is equal to unity, y_{t-1} is taken from each side to produce.

$$\Delta y_t = \beta_1 + \beta_2 t + \pi y_{t-1} + \sum_{j=1}^k \gamma_j \Delta y_i + u_{1t} \quad (6)$$

where π is equal to $1 - \rho$ and therefore if π is significantly different from zero, ρ cannot be one and there is no unit root.

Information Criteria used to assess the appropriate lags of the dependent variable to introduce to remove any serial correlation in the residuals.

Using the *usca* package and the *ur.df* function on UK real consumer spending data (*lc*). Set up the data as a timeseries.

```
library(urca)
library(xtable)
data(Raotbl3)
lc <- ts(Raotbl3$lc, start = c(1966, 4), end = c(1991, 2), frequency = 4)
```

Test the trend (*lc.ct*), drift (*lc.co*) and first difference (*lc2.ct*).

```
lc.ct <- ur.df(lc, lags = 3, type = "trend")
lc.co <- ur.df(lc, lags = 3, type = "drift")
lc2 <- diff(lc)
lc2.ct <- ur.df(lc2, type = "trend", lags = 3)
```

The three different equations are tested by 'trend', 'drift' or 'none' and there are two tests that take place. The first tests whether π is equal to zero. The test is the usual t-value on the lagged dependent variable. Critical values have been provided by Dickey and Fuller. There is also an F-test of the coefficient on the

```
a <- cbind(t(lc.ct@teststat), lc.ct@cval)
print(xtable(a, digits = 2))
```

3 Cointegration

This is the overview of cointegration and the methods use to analyse cointegrated relationships. Non-stationary data may exhibit *spurious regression*. If

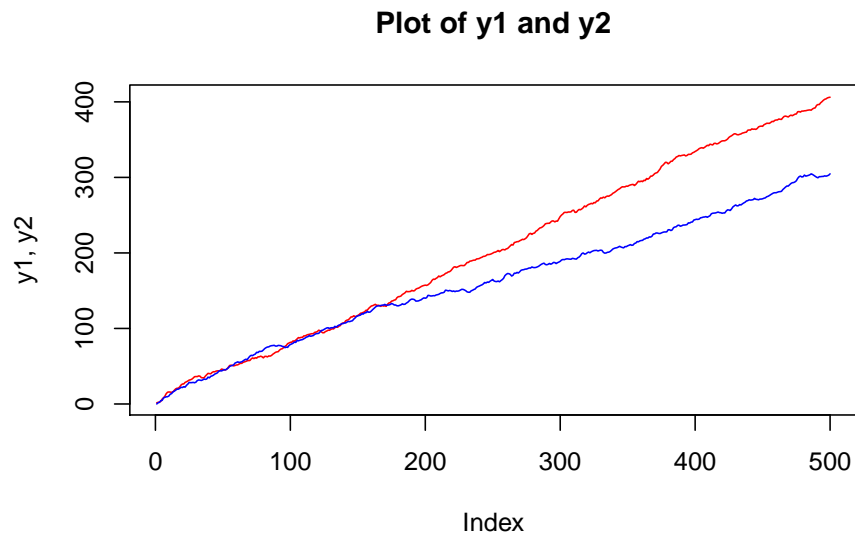
	statistic	1pct	5pct	10pct
tau3	-2.24	-4.04	-3.45	-3.15
phi2	3.74	6.50	4.88	4.16
phi3	2.60	8.73	6.49	5.47

two normal random variables are created (e1 and e2) and two series (y1 and y2) have a trend plus a random shock.

```
library(lmtest)
library(xtable)
set.seed(123456)
e1 <- rnorm(500)
e2 <- rnorm(500)
trd <- 1:500
y1 <- 0.8 * trd + cumsum(e1)
y2 <- 0.6 * trd + cumsum(e2)
```

Now plot the two series

```
plot(y1, type = "l", main = "Plot of y1 and y2", col = "red", ylab = "y1, y2")
lines(y2, col = "blue")
```



Run a regression of $y1$ on $y2$ and it appears that there is a strong relationship.

```
sr.reg <- lm(y1 ~ y2)
print(xtable(sr.reg))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-29.3270	1.3672	-21.45	0.0000
y2	1.4408	0.0075	191.62	0.0000

However, the Durbin-Watson statistics shows there is a large amount of auto-correlation in the residuals.

```
sr.dw <- dwtest(sr.reg)$statistic
sr.dw

##      DW
## 0.01715
```

The statistic will be around 2 if there is no autocorrelation.