Cointegration

Rob Hayward

May 31, 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 conintegration. The final section concludes.

2 Stationary data

The standard series to be investigated can take the form of

$$y_t = TD_t + z_t \tag{1}$$

y_t is the series to be considered, TD_t is the *deterministic trend* where $TD_t = \beta_0 + \beta_1 t$ and $_t$ is a *stochastic component* with $\Phi(L)z_t = \Theta(L)\varepsilon_t$. If the auto-regressive compoents sum to more than zero (or the roots of the ploynominal lie inside the unit circle) the series is explosive and difficult to analyse; if the components sum to lesss than zero (or the roots lie outside the unit circle) the series tends to return to a constant level (there is a stationary mean); if any of the roots lie on the unit circle (and the others are outside) the series has a unit root. A series with a unit root could be a random walk.

$$y_t = y_{t-1} + \varepsilon \tag{2}$$

where

$$\varepsilon \sim N(0, \sigma^2)$$

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

$$y_t = y_{t-1} + \mu = y_0 + \mu t \tag{3}$$

and

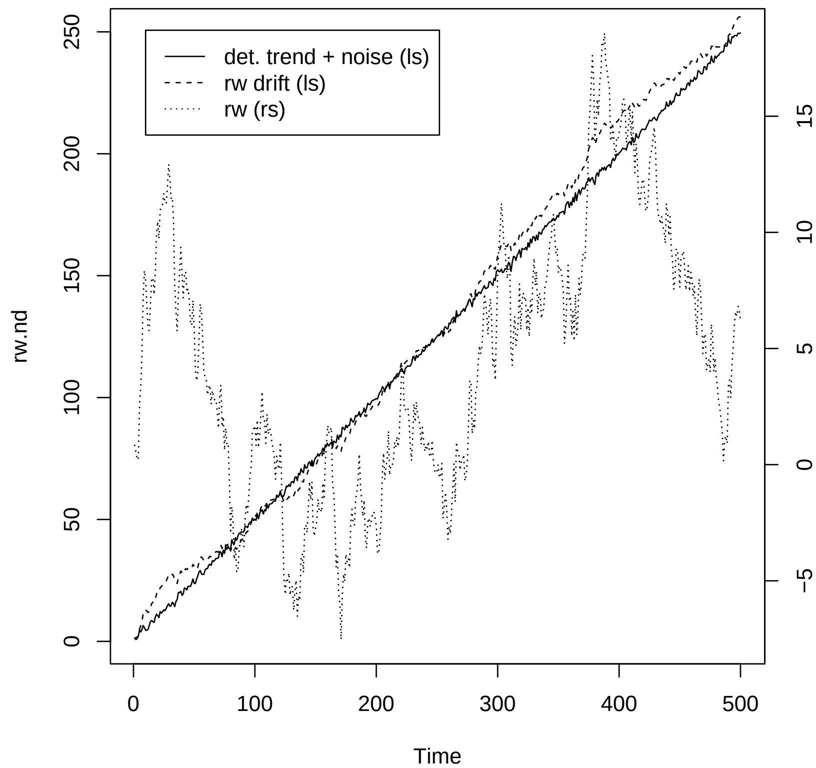
$$y_t = y_{t-1} + \varepsilon = y_0 + \sum_{i=0}^{t} \varepsilon_t \tag{4}$$

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))
```



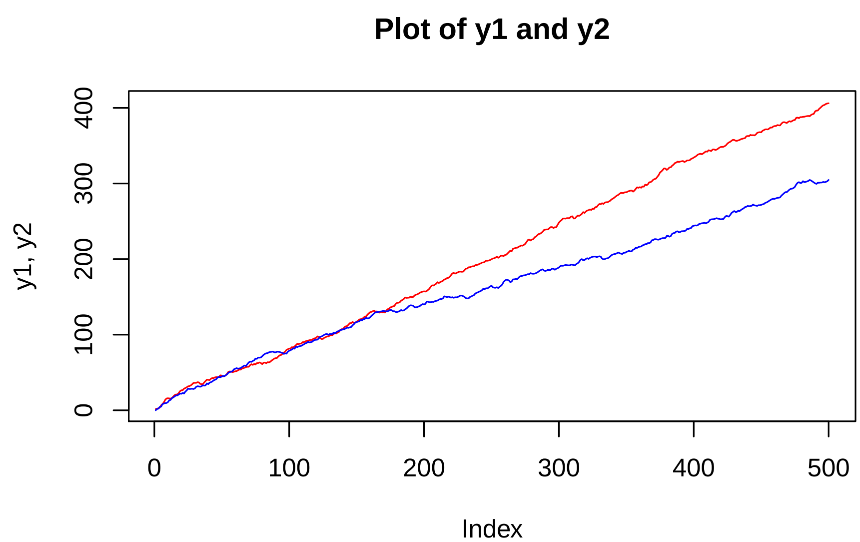
3 Cointegration

This is the overview of conintegration and the methods use to analyse conintegrated relationships. Non-stationary data may exhibit *spurious regression*. If two norman 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.