Nonstationary time series models

Susan Thomas

13 November, 2009

Goals

- Trends in economic data.
- Alternative models of time series trends:
  - deterministic trend, and
  - stochastic trend.
- Comparison of deterministic and stochastic trend models

The statistical models

Two models for a a series that shows a growth through time:

1. A model with a deterministic trend

$$x_t = \alpha + \beta t + (1 - \theta L)\epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

2. A model with a stochastic trend, ie, an integrated process of some order. The simplest is an I(1) process like a random walk:

$$y_t = y_{t-1} + (1 - \theta L)\nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2)$$

Deterministic trend model

- In the model: $y_t = \alpha + \beta t + (1 - \theta L)\epsilon_t$, the unconditional mean converges to

$$E(\hat{y}_t) = \alpha + \beta t$$

  This is a linear function of time, t.
- Implication of the model: the main dynamic in the model is the long term “trend”. All deviations from the trend is eventually washed out and the trend is the long term expected value.
- To obtain the innovations,
  1. Estimate the deterministic trend
  2. Detrend the series by subtracting the deterministic trend as $y_t - \alpha - \beta t$.
  3. The resulting residuals are stationary.
- Therefore, such a process is also called a “trend-stationary” process.

The stochastic trend model

- If the variable is x_t,

$$x_t = x_{t-1} + (1 - \theta L)\epsilon_t$$

- This is called an $I(1)$ or "integrated to the order of 1" process since the first difference

$$\Delta x_t = x_t - x_{t-1} = (1 - \theta L)\epsilon_t$$

  is stationary.
- Here the expected value of the variable at any time in point is the value of the previous period.

Types of stochastic trend models used

Two popular forms of stochastic trend models (with white noise innovations) are:

1. Pure random walk: $z_t = z_{t-1} + \epsilon_t$.
   When z_t is rolled back to understand the information it contains, it becomes
   $$z_t = z_0 + \sum_{i=0}^{t} \epsilon_i$$
2. Random walk with drift: $z_t = \delta + z_{t-1} + \epsilon_t$, where each z_t is:
   $$z_t = z_0 + \delta t + \sum_{i=0}^{t} \epsilon_i$$
   Here, we see that there is still a deterministic time trend, but the deviations of z_t from the trend is not stationary.
3. In both cases, if z_t is differenced once to give $\tilde{z}_t = z_t - z_{t-1}$, it becomes stationary.
   1. $z_t - z_{t-1} = \tilde{z}_t = \epsilon_t$
   2. $z_t - z_{t-1} = \tilde{z}_t = \delta + \epsilon_t$

   These are called "difference-stationary" process.

Terminology: general stochastic trend model

- Recall the vocabulary of ARMA(p, q) models.
- Suppose we have a time–series X_t where the first differences ΔX_t are ARMA(p, q). Then X_t is said to be *integrated* of order 1, generally referred to as:

$$\text{ARIMA}(p, 1, q).$$

- For example, if X_t is ARIMA($p, 1, q$) process "with drift", it would be specified as:

$$\begin{aligned} X_t &= \alpha + X_{t-1} + \Xi(L)\epsilon_t, \\ \text{where } \Xi(L) & \quad \text{is such that} \\ \Phi(L)\epsilon_t &= \Theta(L)\omega_t \\ \omega_t &\sim N(0, \sigma^2_\omega) \end{aligned}$$

Terminology: higher order general stochastic trend model

- More generally, we may have ARIMA(p, d, q) processes.
  For example, a random walk with drift model having ARMA(p,q) innovation, which is integrated to the order of d is written as:

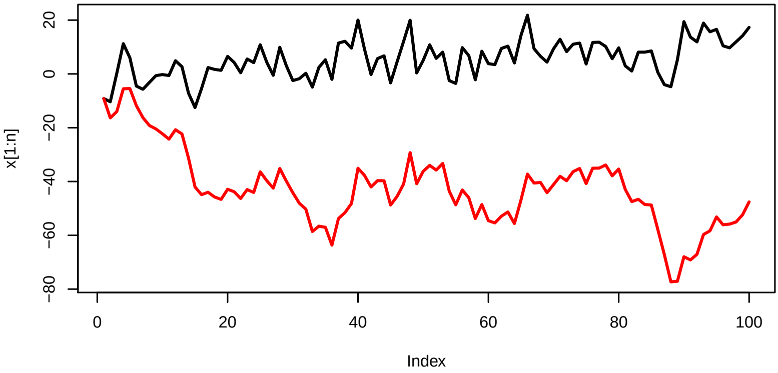
$$X_t(1-L)^d = \alpha + \Xi(L)\epsilon_t$$

- However, several economic time series appear to be integrated to the order of one, ($I(1)$).
  Processes with higher orders of integration are rare.

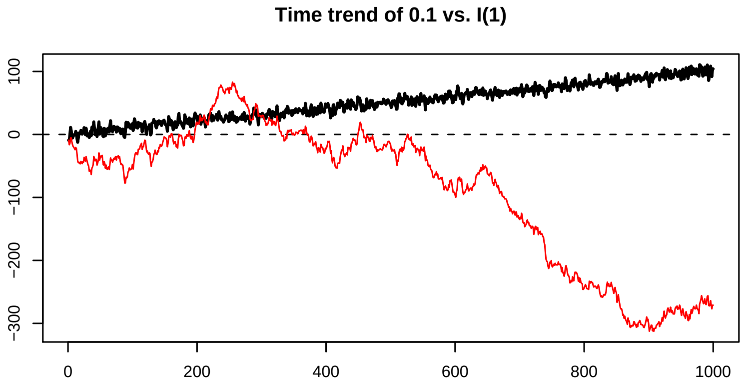
A Monte Carlo simulation of a time trend vs. I(1)

```
e <- rnorm(5000, sd=5)
x <- xi <- NULL
x[1] <- xi[1] <- e[1]
for (i in 2:5000) {
  x[i] <- 1.2 + 0.1*i + e[i] + 0.5*e[i-1];
  xi[i] <- xi[i-1] + e[i] + 0.5*e[i-1]
}

hilo <- range(rbind(x[1:100], xi[1:100]))
plot(x[1:100], type="l", lwd=2, ylim=hilo)
lines(xi[1:100], type="l", col="red", lwd=2)

N <- c(100, 200, 500)
par(mfrow=c(3,1))
for (i in 1:3) {
  n <- N[i]
  hilo <- range(rbind(x[1:n], xi[1:n]))
  plot(1:n, x[1:n], type="l", ylim=hilo,
       xlab="", ylab="", lwd=2,
       main="Time trend of 0.25 vs. I(1)")
  abline(h=0, lty=2)
  lines(1:n, xi[1:n], col="red")
}
```

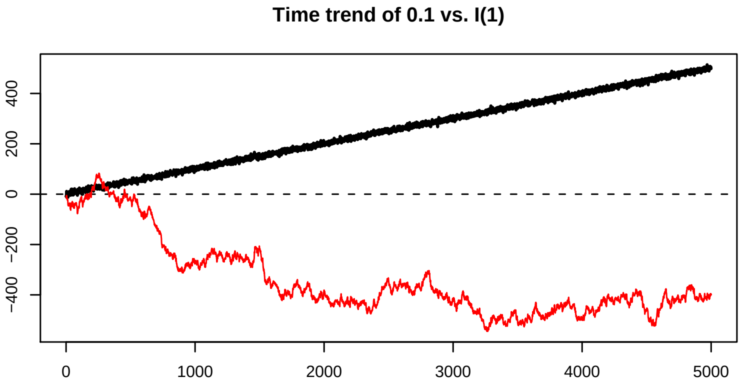
First 100 obs



First 1000 obs



First 1000 obs



Part I

Motivation to worry about trend stationary vs. difference stationary processes

Spurious regressions in econometrics

Clive W. J. Granger and P. Newbold, "Spurious regressions in econometrics." Journal of Econometrics, 2:pg 111 – 120, (1974)

What do you think about this regression?

```
> lm(formula = y ~ x)
>
> Residuals:
>      Min        1Q    Median        3Q       Max
> -25.61685  -4.50486  -0.07608   4.56791  27.31734
>
> Coefficients:
>              Estimate Std. Error  t value Pr(>|t|)
> (Intercept) -3.927e+00  1.347e-01   -29.16   <2e-16 ***
> x            6.667e+00  7.767e-05 85829.09   <2e-16 ***
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> Residual standard error: 6.727 on 9998 degrees of freedom
> Multiple R-squared:     1, Adjusted R-squared:     1
> F-statistic: 7.367e+09 on 1 and 9998 DF,  p-value: < 2.2e-16
```

Where this came from

```
> obs <- 10000
>
> tind <- seq(1, obs, 1)
> x <- 0.2 + 0.3*tind + rnorm(obs,2)
> y <- 7 + 2*tind + rnorm(obs,4)
>
> m1 <- lm(y ~ x)
```

We have a problem

- The DGPs of X_t and Y_t were totally unrelated.
- Yet, OLS appears to establish a strong correlation between them.
- This seems to happen when both X and Y have time trends – e.g. a regression of sugarcane production in Maharashtra on spark plug production in Detroit.
- Economists had indulged in a vast number of such regressions in the post–war years, so this was a really damaging problem - it invalidated a great deal of existing empirical papers.

The setting

- Suppose we have two series, x_t, y_t, which are uncorrelated random walks:

$$\begin{aligned} x_t &= x_{t-1} + \epsilon_t \\ y_t &= y_{t-1} + \nu_t, \text{ where} \\ \epsilon_t &\sim iid(0, \sigma^2_\epsilon) \\ \nu_t &\sim iid(0, \sigma^2_\nu) \\ E(\epsilon_t, \nu_t) &= 0 \\ E(\epsilon_t, \epsilon_{t-i}) &= 0 \forall i \\ E(\nu_t, \nu_{t-i}) &= 0 \forall i \end{aligned}$$

- Model for how x_t, y_t are related: $y_t = \alpha + \beta x_t + \gamma_t$
- Null hypothesis:

$$\begin{aligned} H_0 : \alpha &= 0; H_0 : \beta = 0 \\ H_0 : \alpha &= \beta = 0 \end{aligned}$$

- Statistical tests:
  1. t-statistic for $\beta = 0, \alpha = 0$
  2. F-statistic for the joint of $\alpha = \beta = 0$

Fundamental problem in testing the null

- The above framework follows the classical framework to test one model vs. another model.
- However, following this framework leads to fundamental flaws.

Problem in testing $H_0 : \alpha = 0$

- Under $H_0 : \alpha = 0$, the model becomes:

$$y_t = \beta x_t + \gamma_t$$

  This is a projection of a random walk on a random walk. However, $\hat{\rho}_{xy}$ converge to 1, even though the two are unrelated.
- Monte Carlo simulations show that
  - When the series are I(0) and uncorrelated, $\hat{\rho}_{xy}$ is distributed as nearly normal with a mean of 0.
  - When the series are I(1) with uncorrelated innovations, $\hat{\rho}_{xy}$ tends more towards a value of 1.
  - When the series have higher order of integration also, $\hat{\rho}_{xy}$ is very likely to have values of $+1, -1$.
- **G. U. Yule.**, "Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series." Journal of the Royal Statistical Society, 89:pages 164 (1926)

Problems in testing $H_0 : \beta = 0; \alpha = \beta = 0$

- Under $H_0 : \beta = 0$, the model is:

$$y_t = \alpha + \gamma_t$$

  This is a model of white noise with a level of α.

- Under $H_0 : \alpha = \beta = 0$, the model is:

$$y_t = \gamma_t$$

  This also implies that y_t is a white noise series.

- However, the null that we want to test is **not** the white noise model.

Distributions of OLS β with I(1) variables

- Some of the documented problems if dealing with I(1) variables:
  - t-statistics on α and β are not t-distributed; in fact, they do not have limiting distributions.
  - t-statistics on α and β diverge in distribution as $T \rightarrow \infty$.
  - F-statistic diverges in distribution as $T \rightarrow \infty$.
  - The autocorrelation coefficients of the OLS residuals, $\hat{\rho}_{\gamma} \rightarrow 1$
- Therefore, any inference made using conventional tests of hypothesis in an OLS setting will be false when the variables regressed are non–stationary.

  P. C. B. Phillips. "Understanding Spurious Regressions in Econometrics." Journal of Econometrics, 33:pages 311340 (1986)

Some aspects documented for nonstationary timeseries in a multivariate setting

- Very high correlations between variables.
- Very high autocorrelation in OLS residuals even if $H_0 : \beta = 0$ is not rejected.
- An increasing rejection rate for $H_0 : \beta = 0$ as $T \to \infty$.

Part II

Sample size T and convergance issues for trend stationary vs. difference stationary models

OLS estimates for trend stationary models

- Model: $y_t = \alpha + \beta t + \epsilon_t$.
- Ordinary OLS:
  $\sqrt{N}(\hat{\beta}_{ols} - \beta) = [(1/N)\sum_{i=1}^{N} x_i x_i']^{-1}[(1/\sqrt{N}\sum_{i=1}^{1} x_i \epsilon_i]$
- For the deterministic time trend model:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} 1 & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} \epsilon_t \\ \sum_{t=1}^{T} t\epsilon_t \end{bmatrix}$$

Convergance properties for the OLS estimates of a deterministic trend

- $\hat{\alpha}$ converges at the rate of $T^{1/2}$, and $\hat{\beta}$ at the rate of $T^{3/2}$. Therefore,

$$\begin{bmatrix} T^{1/2}\hat{\alpha} \\ T^{3/2}\hat{\beta} \end{bmatrix} = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{T} 1 & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} \epsilon_t \\ \sum_{t=1}^{T} t\epsilon_t \end{bmatrix}$$

$$\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix}^{-1} \right)$$

- If the trend is truly deterministic, then all the standard tests of OLS estimates are asymptotically valid, given that the correct rates of convergance are used for each parameter.
- However, identifying whether the trend observed in a given data set arises from a deterministic trend vs. stochastic trend, gets complicated.

Problems with β inference with I(1) data

- Typically, the deterministic trend model is estimated as $y_t = \alpha + \beta t + \epsilon_t$, and tested for the null:

$$H_0 : \alpha = \beta = 0$$

- If the underlying data is I(1), the estimate of β does converge to 0.
  But the standard error for β does not converge to 0 under large samples.
  The behaviour of the t-statistic for the estimated β is not well understood with non-stationary data as T increases.
- If the underlying data is I(1), the estimate for α is divergent. As $T \rightarrow \infty$, $\sigma_\alpha \rightarrow \infty$.
- **S. N. Durlauf and P. C. B. Phillips** , "Trends vs. random walks in time series analysis", Econometrica 56, pgs 1333-54 (1988).

Detecting I(1) when ϕ_i is very close to 1

- The stationary process –

$$x_t = \rho x_{t-1} + \epsilon_t$$

  Innovations move the process away from 0, but they tend to die away over time.
  $\mathrm{Var}(x_t) = \sigma^2_\epsilon/(1-\rho)$, a finite and tame quantity.

- The process

$$x_t = x_{t-1} + \epsilon_t$$

  is profoundly different.

- Difficulty in identification:
  - If $\rho = 1$, it is an $I(1)$ process.
  - If $\rho = 0.999$ it is $I(0)$.

  It proves to be difficult to tell the two apart.

The statistical problem

- Model: $y_t = \rho y_{t-1} + \epsilon_t$.
  What happens at the boundary of $\rho < 1$ and $\rho = 1$? Say $\rho = 1 - \nu$, where ν is very small.
- For finite T, the variance of the series is:

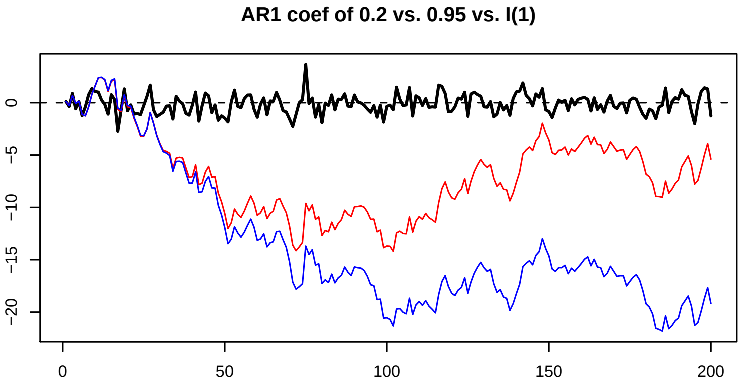
$$\begin{aligned}\sigma_y^2 &= \sigma^2(1 + \rho^2 + \rho^4 + \ldots + \rho^{2(t-1)}) \\ &= \sigma^2 t + \sigma^2 \sum_{i=1}^{(t-1)} \rho^{2i}\end{aligned}$$

- When T is finite, σ_y^2 behaves as if it has a trend – ie, y_t behaves like an I(1) process.
- Asymptotically, it behaves like a I(0) process.
- This parameterisation is useful when deriving asymptotics for the tests of unit root.

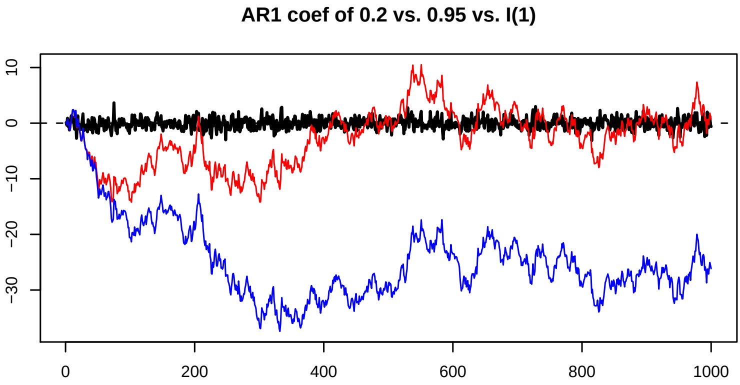
A Monte Carlo simulation of a time trend vs. I(1)

```
> e <- rnorm(25000)
>
> x <- xni <- xi <- NULL
> x[1] <- xni[1] <- xi[1] <- e[1]
> for (i in 2:25000) {
>   x[i] <- 0.2*x[i-1] + e[i];
>   x[i] <- 0.99*x[i-1] + e[i];
>   xi[i] <- xi[i-1] + e[i]
> }
>
> pdf("pix/i1sim-1.pdf", pointsize=8, width=5.6, height=3.2)
>
> n <- 100
>  hilo <- range(rbind(x[1:n], xni[1:n], xi[1:n]))
>  plot(1:n, x[1:n], type="l", ylim=hilo,
>       xlab="", ylab="", lwd=2,
>       main="AR1 coef of 0.5 vs. I(1)")
>  abline(h=0, lty=2)
>  lines(1:n, xni[1:n], col="red")
>  lines(1:n, xi[1:n], col="blue")
>}
```

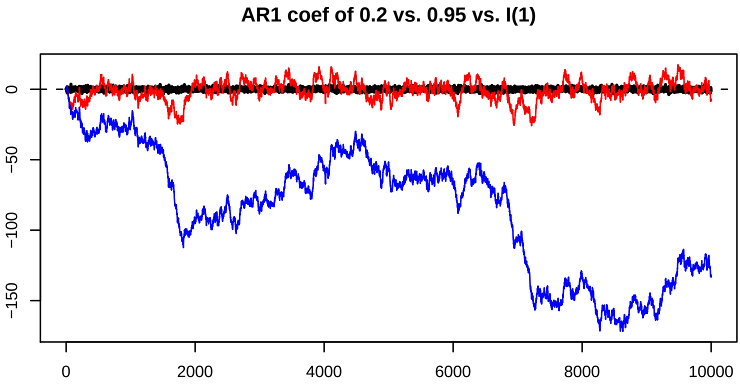
First 100 obs



First 1000 obs



First 10000 obs



Data demands

- 100 observations is roughly half a year of daily data, 8 years of monthly data, 25 years of quarterly data.
- 1000 obs is roughly 4 years of daily data, 83 years of monthly data, 250 years of quarterly data.
- Typically, economic time series are available in the range of 50–60 years of data, typically quarterly (120 observations) or monthly (720 observations).

HW:

1. If the DGP is truly a stationary series, what is the joint test of the hypothesis that $\alpha = \beta = 0$ in the deterministic trend model?
2. Write a Monte Carlo experiment that will test that the asymptotic distribution of the $(\hat{\alpha}_{ols}, \hat{\beta}_{ols})$ for a deterministic trend model is really the same as an ordinary OLS estimator.