Non-stationary Time Series

Michael

1   The Big Picture on Time Series

I have to say it takes me a while to grasp the gist of *time series*. Fortunately, I got the big picture, which can be the contour for doing empirical research in macroeconomics. For most students, the starting point of learning time series is the stationary process like AR(1) or MA(1). However, later I realised that lots of time series from real world are not stationary, which make researchers' life difficult and fun. So, what's the big picture of time series? Here we go: at this moment, it's about data-generating process for me.

Follow the idea by Campbell and Perron (1991), we assume that a time series $\{y_t\}$ is a realization of a deterministic trend and a stochastic component:

$$y_t = TD_t + z_t, \tag{1.1}$$

where TD_t assigns a deterministic trend, which can be assumed to have the linear trend, $TD_t = \beta_1 + \beta_2 t$, and z_t represents the stochastic compoment $\phi(L)z_t = \theta(L)\varepsilon_t$ with $\varepsilon_t \sim$ i.i.d. Or expand it out, we have:

$$y_t = \beta_0 + \beta_1 t + z_t; \qquad \phi(L)z_t = \theta(L)\varepsilon_t \ \ \varepsilon_t \sim i.i.d$$

The intuition behind this assumption can be summarised by the following table:

Summary of data-generating process		
Deterministic Trend		**Stochastic Component**
Fixed effects	Trend	Stochastic pattern with random shocks
β_1	$\beta_2 t$	$\phi(L)z_t = \theta(L)\varepsilon_t$
Reflects Human Being's determination		Law of nature (or unknown process)

For the deterministic trend, things are easy as we can predict it very well if they really follow a certain trend, which is highly unlikely in real data sets. It is common sense that everything is changing. Therefore, we care more about the stochastic part[1].

Now, if all roots of the autoregressive polynomial lie outside the unit circle, then $\{y_t\}$ is stationary around a deterministic trend. In this case, one could remove the trend from the original series $\{y_t\}$ and fit an ARMA(p, q) to the residuals. In R programming, you can do it like this:

---

[1] or random part, think about normal distribution phenomena, then you will realise this part is the key.

```
1  # Detrend the time series
   detrended <- residuals(lm(y ~ seq(along = y)))
3  # fit ARMA model
   arma <- arima(detrend, ma=..., ar=...,)
```

The *trend-stationary* model is also called the integrated model of order zero or $I(0)$. But, what if some roots lie on the unit circle[2]? If it happens, things become tricky and that's why I am writing this notes now. For the unit root situation, $\Delta z_t = (1 - L)z_t$ is stationary around a constant mean, the series $\{y_t\}$ is *difference-stationary* because one has to apply the first difference filter with respect to time to obtain a stationary process. This process can be analyzed with *autoregressive integrated moving average* (ARIMA)(p, d, q), where d refers to the order of intergration.

Take away: by decomposing time series $\{y_t\}$, we realise that stochastic part is the key, and unit root situation is the aiming ring.

Before we moving to the next section, let's state the distinction between a trend- and a difference-stationary process by giving the following two processes:

$$y_t = y_{t-1} + \mu = y_0 + \mu t \tag{1.2}$$

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \sum_{s=1}^{t} \varepsilon_s \tag{1.3}$$

2 Unit Root Processes

As we have said that unit root processes are so important, it's worth to devote ourselves to this topic. Let's review the data-generating process:

$$y_t = y_0 + \mu t + z_t \tag{2.1}$$
$$= y_0 + \mu t + (\phi(L)z_t = \theta(L)\varepsilon_t) \tag{2.2}$$

Now, let's decompose the stochastic component z_t into a cyclical component c_t and a stochastic trend TS_t. It is assumed that *cyclical component c_t* is a mean-stationary process. Then, equation (2.1) can be rewritten as

$$y_t = y_0 + \mu t + c_t + TS_t \tag{2.3}$$

Equation (2.3) will be the **key function** for almost all our analysis in this document. Now, let's list all situations for y_t in equation (2.3):

- If y_t is trend stationary, then the stochastic trend (TS_t) is zero. Then, y_t can be modeled by ARMA(p, q).

- If y_t is difference stationary, then the autoregressive polynomial contains a unit root that can be factored out by doing the difference, $\phi(L) = (1 - L)\phi^*(L)$. After doing the difference, we get Δz_t which can be modeled as the moving average process.

---

[2]we don't consider the cases that roots are inside the unit circle as y_t will explore in this case.

Now, if we represent $z_t = TS_t + c_t = \psi(1)S_t + \psi^*(L)\varepsilon_t$, where $\psi(1)$ is the sum of moving average coefficients and S_t is the sume of the past and present of random shocks $\sum_{s=1}^{t} \varepsilon_s$, and $\psi^*(L)$ is the polynomial coefficients. The full equation is[3]:

$$z_t = c_t + TS_t = \psi(1) \sum_{s=1}^{t} \varepsilon_s + \frac{\psi(L) - \psi(1)}{(1 - L)} \varepsilon_t \tag{2.4}$$

Final step, substitute (2.4) into (2.3) to get:

$$y_t = y_0 + \mu t + TS_t + c_t \tag{2.5}$$

$$= y_0 + \mu t + \psi(1) \sum_{s=1}^{t} \varepsilon_s + \frac{\psi(L) - \psi(1)}{(1 - L)} \varepsilon_t \tag{2.6}$$

I am sorry to reclaim that equation (2.6) will be the **key function** for almost all our analysis of time series data. The intuition and summary on (2.6) is given in Figure 2.1.

$$y_t = y_0 + \mu t + TS_t + c_t$$

$$= \underbrace{y_0 + \mu t}_{\text{Determinist trend}} + \underbrace{\psi(1) \sum_{s=1}^{t} \varepsilon_s}_{\substack{\text{Stochastic compoent:} \\ \text{long-run impact of a} \\ \text{shock to the level of z}}} + \underbrace{\frac{\psi(L) - \psi(1)}{(1 - L)} \varepsilon_t}_{\substack{\text{Stochastic cyclical} \\ \text{compoent: short-run} \\ \text{impact on the level of z}}}$$

Figure 2.1: Intuition on y_t decomposition

In equation (2.6), if $\psi(1) > 1$, then the time series is characterised by an explosive path; if it is less than 1, we have explained that the time series can be modeled with ARMA process; if it is equal to 1, then the data-generating process is a random walk.

2.1 Random Walk

It's worth to know *random walk* process well before the unit root test. A *pure random walk* without a drift is defined as

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \sum_{s=1}^{t} \varepsilon_t, \tag{2.7}$$

where $\{\varepsilon_t\}$ is an i.i.d. process. With assuming $y_0 = 0$, it is trivial to show that (1) $E[y_t] = 0$ and $var(y_t) = t\sigma^2$. Suppose that the data-generating process for $\{y_t\}$ is

$$y_t = y_{t-1} + \varepsilon_t, \ \varepsilon_t = \rho\varepsilon_{t-1} + \xi_t, \tag{2.8}$$

---

[3]Check equation (1.3), to see what's going on.

where $|\rho| < 1$ and ξ_t is a white noise process instead. Then, $\{y_t\}$ is *not* a random walk process, but it still has a unit root and is a first-order non-stationary process. Let us now consider the case of a *random walk with drift*,

$$y_t = \mu + y_{t-1} + \varepsilon_t = y_0 + \mu t + \sum_{s=1}^{t} \varepsilon_t, \tag{2.9}$$

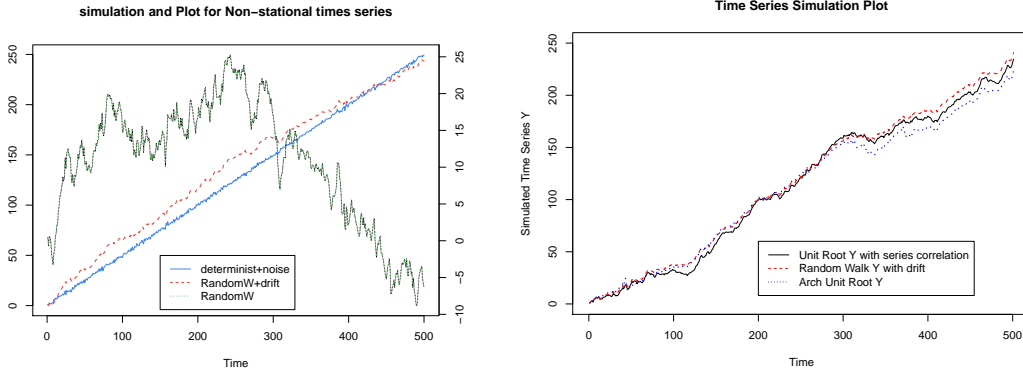where $\{\varepsilon_t\}$ is white noise.

Figure 2.2: Simulated Times Series Plot

The simulated time series paths in Figure 2.2 are very informative. By ocular econometrics, it should be evident that the statistical discrimination between a deterministic trend contaminated with noise and a random walk with drift is not easy based on the left panel of Figure 2.2. From the right panel, we can expect that time series simulated based on ARCH model has higher volatility.

2.2 Unit Root Test

Recall the decomposition of a time series $\{y_t\}$ as in equation (1.1). Now we want to investigate if the process $\{z_t\}$ does contain a unit root,

$$z_t = y_t - TD_t.$$

Hence, a deterministic trend is removed from the original series first and the residuals are tested for a unit root. Dickey and Fuller (1979) proposed the following test regression that is delineated from an assumed $AR(1)$-processes of $\{z_t\}$:

$$z_t = \theta z_{t-1} + \varepsilon_t \tag{2.10}$$

$$z_t - z_{t-1} = \theta z_{t-1} - z_{t-1} + \varepsilon_t \tag{2.11}$$

$$\Delta z_t = (\theta - 1)z_{t-1} + \varepsilon_t \tag{2.12}$$

$$\Delta z_t = \pi z_{t-1} + \varepsilon_t \tag{2.13}$$

Under the null hypothesis of a unit root, $\pi = 0$, which is equivalent to $\theta = 1$ and the alternative is a trend stationary process; i.e., $\pi < 0$ or $\theta < 1$. However, in reality, neither the existence nor the form of deterministic component is known a *priori*. Hence, we

have to choose from the set of deterministic variables DV_t the one that best suits the data-generating process.

Now we consider the equation with deterministic regressors DR_t,

$$\Delta y_t = \tau' DR_t + \pi y_{t-1} + u_t, \tag{2.14}$$

where τ is the coefficient vector of the deterministic part and $\{u_t\}$ assigns an error term. A weakness of the original DF test is that it does not take a possible serial correlation of the error process $\{u_t\}$ into account. If the noise component is an $AR(p)$-process, it can be shown that the test regression

$$\Delta y_t = \tau' DR_t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta_{t-j} + u_t, \; with \; k = p - 1 \tag{2.15}$$

ensures that the serial correlation in the error is removed. This test regression is called the *augmented Dickey-Fuller* (ADF) test.

With the help of AIC or BIC, we can decide the lag order k. To do the unit root test, we need first estimate the encompassing ADF-test equation

$$\Delta y_t = \beta_1 + \beta_2 t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t \tag{2.16}$$

Then, the further steps to be taken are dependent on this estimation until one can conclude that the series is

- stationary around a zero mean

- stationary around a non-zero mean

- stationary around a linear trend

- contains a unit root with zero drift

- contains a unit root with non-zero drift.

We now apply the ADF test to the US real GDP time series. First, we fit the following model:

$$\Delta y_t = \beta_1 + \beta_2 t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t$$

The lag orders are selected based on AIC. Table 2.1 gives the regression results

Table 2.1: ADF test: Regression Results

Coefficients	Estimate	Std.Error	t value	Pr(>|t|)
β_1	0.0924	0.0645	1.431	0.153
π	-0.0110	0.0083	-1.325	0.186
β_2	0.0000	0.0000	1.125	0.261
γ_1	0.3481	0.0569	6.119	0.0000

For the joint test, here is the result:

Table 2.2: ADF test: key statistics

Hypothesis	Test	Value	Critical Values		
			1%	5%	10%
$\pi = 0$	τ_3	-1.3248	-3.98	-3.42	-3.13
$\beta_1 = \beta_2 = \pi = 0$	ϕ_2	19.4217	6.15	4.71	4.05
$\beta_2 = \pi = 0$	ϕ_3	2.7346	8.34	6.30	5.36

Based on the result, $\phi_3 = 2.7346$ which implies that we could not reject $\beta_2 = \pi = 0$. This means that real GDP does contain a unit root. This finding is reiterated by the value of $\tau_3 = -1.3248$, which implies we could not reject null that $\pi = 0$. Therefore, a unit root cannot be rejected. Next, it is tested whether the consumption series is a random walk with or without drift:

$$\Delta y_t = \beta_1 + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t$$

The results are presented in table 2.3 and 2.4.

Table 2.3: ADF test: Regression Results

Coefficients	Estimate	Std.Error	t value	$\Pr(>|t|)$
β_1	0.0202	0.0075	2.701	0.00735
π	-0.0017	0.0008	-2.049	0.04140
γ_1	0.3403	0.0565	6.024	0000

Table 2.4: ADF test: key statistics

Hypothesis	Test	Value	Critical Values		
			1%	5%	10%
$\pi = 0$	τ_2	-2.0491	-3.44	-2.87	-2.57
$\beta_1 = \pi = 0$	ϕ_1	28.4718	6.47	4.61	3.79

In the end, we can state that US real GDP data behaves like a *ransom walk with a drift*. The study by Holden and Perman (1994) also suggests that the consumption series behaves like a *pure random walk*. The research by Harvey et al. (2003) also found that the log consumption: output and log investment:output great ratios are not, in general, stationary. The findings are thus not favourable to the neoclassical stochastic growth hypotheses (Harvey et al., 2003). Further more, their findings are also in favour of growth and real business cycle models that do not assume exogenous and permanent stochastic productivity shocks, which lead to the common trend hypothesis.

We now apply the ADF test to the US real gross private domestic investment (GPDI) by fitting the following model first,

$$\Delta y_t = \beta_1 + \beta_2 t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t$$

The lag orders are selected based on AIC. Table 2.1 gives the regression results

Table 2.5: ADF test: Regression Results - USGPDI

Coefficients	Estimate	Std. Error	t value	Pr	
(Intercept)	0.4210983	0.1244785	3.383	0.000827	***
z.lag.1	-0.0754591	0.0227653	-3.315	0.001047	**
tt	0.0007324	0.0002246	3.262	0.001254	**
z.diff.lag1	0.1855826	0.0593962	3.124	0.001981	**
z.diff.lag2	0.1413190	0.0603198	2.343	0.019887	*
z.diff.lag3	0.0664264	0.0586191	1.133	0.258172	
z.diff.lag4	-0.1627121	0.0578142	-2.814	0.005258	**

For the joint test, here is the result:

Table 2.6: ADF test: key statistics - USGPDI

			Critical Values		
Hypothesis	Test	Value	1%	5%	10%
$\pi = 0$	τ_3	-3.3146	-3.98	-3.42	-3.13
$\beta_1 = \beta_2 = \pi = 0$	ϕ_2	6.9544	6.15	4.71	4.05
$\beta_2 = \pi = 0$	ϕ_3	5.503	8.34	6.30	5.36

References

Campbell, J. Y. and Perron, P. (1991). Pitfalls and opportunities: what macroeconomists should know about unit roots. *NBER macroeconomics annual*, 6:141–201.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Harvey, D. I., Leybourne, S. J., and Newbold, P. (2003). How great are the great ratios? *Applied Economics*, 35(2):163–177.

Holden, D. and Perman, R. (1994). Unit roots and cointegration for the economist. In *Cointegration*, pages 47–112. Springer.