Time Series Analysis

Econometrics

Notes draw heavily from Tsay's 3rd edition Analysis of Financial Time Series textbook, Hamilton's 1994 Time Series Analysis textbook, and Greene's 7th edition Econometric Analysis textbook.

This Lecture

- Random walk models and forecasts
- Interpreting trend and intercept terms
- Unit root testing — Augmented Dickey Fuller (ADF) test
- Some R commands

Nonstationary Processes

- Price levels, interest rates, exchange rates and many other time series can be nonstationary.
- Key points about unit root/random walk processes:
  - A stationary series can usually obtained with appropriate differencing
  - Pay careful attention to drift and trend
  - Typical properties of econometric estimators will not apply
- Random walk is just one form of nonstationarity - structural breaks/mean shifts, for example.

Random Walk

Let $p_t = log(P_t)$, a_t be white noise and

$$p_t = p_{t-1} + a_t$$

- Each period p_t has a 50% chance of going up or down.
- With logged stock prices, if a_t is Normal, returns are Normal, prices are Lognormal.
- Lag polynomial $(1 - \phi_1 L)$ with $\phi_1 = 1$, a "unit root".
- By repeated substitution

$$p_t = a_t + a_{t-1} + a_{t-2} + ...$$

- All past shocks are "permanent"

$$Var(p_t) = \sigma_a^2 + \sigma_a^2 + \sigma_a^2 + ... \to \infty$$

Forecasting a Random Walk

A random walk is not forecastable and not mean-reverting.

$$\hat{p}_h(1) = E(p_{h+1}|p_h, p_{h-1}, ...) = E(p_{h+1}|p_h) = p_h$$

$$\hat{p}_h(2) = E(p_{h+2}|p_h, p_{h-1}, ...) = E(p_{h+1} + a_{h+2}|p_h) = p_h$$

$$\implies \hat{p}_h(l) = p_h$$

Forecast errors grow linearly rather than approaching the variance of the series (which is undefined):

$$e_h(l) = a_{h+l} + a_{h+l-1} + ... + a_{h+1}$$

$$Var(e_h(l)) = l\sigma_a^2$$

Problem for modeling prices and other economic variables: they are usually not negative.

Random walk with drift

Suppose log returns have a non-zero mean:

$$r_t = p_t - p_{t-1} = \mu + a_t$$

$$\implies p_t = \mu + p_{t-1} + a_t$$

$\mu =$ average price growth rate, often called “drift” or “time trend” of the log prices.

$$\begin{aligned} p_t &= \mu + p_{t-1} + a_t \\ &= \mu + \mu + p_{t-2} + a_t + a_{t-1} \\ &= t\mu + p_0 + a_t + \ldots + a_1 \end{aligned}$$

This is a trend plus an MA representation of the pure random walk.

Interpreting the intercept

- In MA(q), the intercept is the mean
- In stationary AR(p) or ARMA(p,q), intercept is the mean multiplied by the lag polynomial.
- In a random walk with drift, intercept is the slope of the drift - the average growth rate of the series.

Time trends and unit roots

Alternative model: suppose the series has a deterministic trend, but it is stationary around the trend.

$$p_t = \beta_0 + \beta_1 t + r_t$$

where r_t is stationary.
It can be very difficult, if not impossible, to tell a random walk with drift apart from a trend stationary series.
Trend-stationary, 2nd moment does not depend on t:

- $E(p_t) = \beta_0 + \beta_1 t$
- $Var(p_t) = Var(r_t)$

Do we really believe progress is deterministic?

Time trends and unit roots

Do we really believe uncertainty in predicting GDP in 2100 is the same as predicting GDP in the year 3100?
Random Walk, moments depend on t. This is more believable:

- $E(p_t) = p_0 + \mu t$
- $Var(p_t) = t\sigma_a^t$

Our forecast is the drift (same as with deterministic trend), but uncertainty grows with forecast horizon.
If we KNOW it's trend-stationary, could transform trend-stationary to stationary by running regression on a trend and calculating

$$p_t^\star = p_t - \hat{\beta}_0 - \hat{\beta}_1 t$$

General Nonstationary Models

Consider an ARMA(p,q) model

$$\phi(B)r_t = \phi_0 + \theta(B)a_t$$

- If $\phi(B)$ has a unit root, we really have ARIMA (Autoregressive, Integrated, Moving Average).
- y_t is ARIMA$(p, 1, q)$ process if the differenced series $y_t - y_{t-1} = \Delta y_t$ is an ARMA(p, q) process.
- Very few economics & finance examples require more differencing.
  - Technically, y_t is an ARIMA$(p, 2, q)$ process if the *difference* of the differenced series $\Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2}$ is ARMA(p, q).

Distinguishing between unit roots and time trends

- Is it $y_t \sim I(1)$ or is it $y_t = \alpha + \delta t + a_t$ with $a_t \sim I(0)$?
- Most conclude it is impossible to definitively know the difference.
- It is rarely legitimate to claim that y_t is one or the other given the precision (or lack thereof) with which we can estimate δ.
- Instead:

1. State that I am restricting myself to evaluate a subset of possible processes, for example AR(p) with fixed p.
   - e.g., $y_t = \rho y_{t-1} + \epsilon_t$
2. Reasonable conclusions:
   - The data are consistent/inconsistent with the claim that $\rho = 1$
   - I can reject or fail to reject the hypothesis that $\rho = 1$.
   - $\rho = 0.99999$ may be possible even when I fail to reject, and technically this is stationary. Must recognize that we are always taking about an approximation to the right process.

Inference with unit roots

- The estimated coefficients are not Normally distributed even as the sample size gets big.
- WHY? Central Limit Theorem requires stationary data (conditionally independent observations). If the data has a unit root, the CLT doesn't apply to coefficients calculated from that data.
- Different distributions for hypothesis tests depending on the case.
- Software package or various textbooks can tell you the critical value for specific model cases.
- *Want null and alternative hypotheses that are reasonable competing models of the data.*

Inference with unit roots: Case 1

AR(1) with a unit root

- Case 1: no constant (no drift):

$$y_t = \rho y_{t-1} + \epsilon_t, \ \epsilon_t \sim iid(0, \sigma^2)$$

- $H_0 : \rho = 1, \quad H_a : \rho < 1$, IMPOSE $\phi_0 = 0$.
- Under the null, random walk without drift
- Alternative: AR(1) with mean ZERO
- Unrealistically constraining alternative hypothesis
- $\hat{\rho}$ not Normally distributed, unusual critical values.
- Rarely applicable, but important for understanding, and built-in to most stats packages.
- ur.df(y,type=c(”none”)) or CADFtest(y,type=”none”)

Inference with unit roots: Case 1

AR(1) with a unit root

- Case 1: no constant (no drift):

$$y_t = \rho y_{t-1} + \epsilon_t, \ \epsilon_t \sim iid(0, \sigma^2)$$

- $H_0 : \rho = 1$, Estimated $\hat{\rho} = \frac{\sum y_{t-1} y_t}{\sum y_{t-1}^2}$
- OLS estimator converges to a ratio of two types of Brownian motion, rather than to a Normal.
- distribution is similar to a Normal, but skewed to the left.
- Different critical values.
- The limiting distribution is specific to this case with true $\rho = 1$ and no constant.
- Other cases will have different limiting distributions, different tables.

Inference with unit roots: Case 2

- Case 2: include constant, but truth is constant = 0 (misspecified drift when no drift exists).

$$y_t = \phi_0 + \rho y_{t-1} + \epsilon_t, \ \epsilon_t \sim iid(0, \sigma^2)$$

- $H_0 : \phi_0 = 0, \ \rho = 1, \ \ H_a : \phi_0 \neq 0, \ \rho < 1.$
- Again $\hat{\phi}_0$ and $\hat{\rho}$ converge to some other non-normal distribution.
- Because of inclusion of constant, limiting distribution of ρ is more negatively-skewed - leads to failure to reject null of a unit root more often.

Inference with unit roots: Case 2

- This case seems odd - we include constant when we think it doesn't belong under the null, and it makes the distribution farther from Normal.
  - **yet, you almost always want to use Case 2 NOT Case 1 if there is no obvious trend.**
  - We want a reasonable description of the data under both the null and the alternative.
  - $H_0 : y_t = y_{t-1} + \epsilon_t, \quad H_a : y_t = \phi_0 + \rho y_{t-1} + \epsilon_t$
  - e.g., your data is all positive (e.g., GDP, prices) one of these two stories is most likely - you would not have generated all positive observations with $y_t = \rho y_{t-1} + \epsilon_t$ and $\rho < 1$, so rejecting the null in Case 1 leaves you scratching your head.
  - interpretation of $\phi_0 \neq 0$ depends on ρ. If $\rho = 1$, ϕ_0 is the drift. If $\rho < 1$, ϕ_0 is the nonzero mean.

Inference with unit roots: Case 2

- *R* syntax: ur.df(y,type=c("drift")) or CADFtest(y,type="drift").
  - ur.df() from the "urca" package (Unit Root Cointegration Analysis)
  - CADFtest() from the "CADFtest" package (Covariate Augmented Dickey Fuller)
- Let's focus on ur.df() for a minute.
  - Reports a "tau2" test statistic, which is an individual test of y_{t-1} coefficient ρ.
  - Reports a "phi1" test statistic, which is a joint test of ρ and $\phi_0 = 0$.
  - If you reject "phi1", either you DON'T have a unit root, the intercept is NOT zero, OR BOTH.
  - If you reject "phi1", but NOT "tau1", you have a unit root with drift - need to investigate Case 4 below to see if a deterministic trend is driving growth or if it's drift in a random walk.

Inference with unit roots: Case 3

- Case 3: include constant and it belongs there (truth is random walk with drift).

$$y_t = \phi_0 + \rho y_{t-1} + \epsilon_t,\ \epsilon_t \sim iid(0, \sigma^2)$$

- $H_0 : \rho = 1,$ IMPOSE $\phi_0 \neq 0, \quad H_a : \rho < 1.$
- Null: random walk WITH drift. Alternative: AR(1) with NO trend, NON-ZERO mean.
- Implies $y_t = \phi_0 t + \epsilon_1 + \epsilon_2 + ... + \epsilon_t$

Inference with unit roots: Case 3

- Shouldn't we then just always pretend there's a constant, with $\phi_0 \neq 0$ but possibly arbitrarily close? No.
  - $H_0 : y_t = \phi_0 + y_{t-1} + \epsilon_t$
    - data should exhibit a trend if the null is true.
  - $H_a : y_t = \phi_0 + \rho y_{t-1} + \epsilon_t$ with $\rho < 1$
    - data here would not exhibit a trend.
    - Not really a sensible description of the most (macro/financial/commodity price) time series of interest. This case is rarely used.
- Strangely, coefficients are asymptotically Normal, use "type=c("drift")" in R, but use Normal critical values, not the ones the program spits out.

Inference with unit roots: Case 4

- Case 4: add a time trend to the model from Case 2.

$$y_t = \phi_0 + \rho y_{t-1} + \beta t + \epsilon_t,\ \epsilon_t \sim iid(0, \sigma^2)$$

- $H_0 : \phi_0 > 0,\ \rho = 1,\ \beta = 0$
  - $y_t = \phi_0 + y_{t-1} + \epsilon_t$
  - observed trend comes from random walk with drift
- $H_a : \phi_0 \neq 0,\ \rho < 1,\ \beta > 0$
  - $y_t = \phi_0 + \rho y_{t-1} + \beta t + \epsilon_t$
  - observed trend comes from deterministic component.
- All parameters converge at different rates, distribution of $\hat{\rho}$ is even more strongly negatively skewed than Case 2.

Inference with unit roots: Case 4

- *R* syntax: ur.df(y,type=c("trend")) or CADFtest(y,type="trend").
- Again, let's focus on ur.df(). It reports a few test statistics:
  - "tau3", an individual test of y_{t-1} coefficient ρ.
  - "phi3", a joint test of ρ and $\beta = 0$.
  - "phi2", a joint test of ρ, $\beta = 0$, and $\phi_0 = 0$.
  - If you reject "phi3", either you DON'T have a unit root, you HAVE a trend, OR BOTH.
  - If you reject "phi3", but NOT "tau3", you have a unit root AND a deterministic trend. Use "phi2" to tell you if some of the growth may also be coming from drift (although you have confirmed the deterministic trend part).
  - If you FAIL to reject "phi3", either you have a unit root, you DON'T have a trend, OR BOTH. Since you don't have a deterministic trend, return to Case 2 to test for the unit root and drift/intercept.

Inference with unit roots: Rules of thumb

- **Rules of thumb if you want to test for the presence of a unit root ($H_0 : \rho = 1$):**
  - Use Case 2 if your data appears not to have persistent growth (option: type=drift).
  - Use Case 4 if your data appears to have persistent growth (option: type=trend).
    - In ur.ca() output, use "phi3" and "tau3" to see if deterministic trend term explains the growth along with a unit root.
    - If not, return to Case 2 to and use "phi1" and "tau1" to see if drift explains the growth along with unit root.

Inference with unit roots: Rules of thumb

- **In each case the NULL IS A NON-STATIONARY SERIES! Rejecting the null means you have a stationary series.**
- In practice, more intuitive to test lag coefficient of zero rather than one:

$$\Delta p_t = \phi_0 + \rho p_{t-1} + e_t, \quad H_0 : \rho = 0,\ \phi_0 > 0,\ \ H_a : \rho < 1,\ \phi_0 \neq 0$$

$$\Delta p_t = \phi_0 + \rho p_{t-1} + \beta t + e_t, \quad H_0 : \rho = 0,\ \phi_0 > 0,\ \beta = 0,\ \ H_a : \rho < 1,\ \phi_0$$

Augmented Dickey Fuller Test

- Augmented to allow for AR(p).
  - Looking for lag order *p* that eliminates serial correlation in the errors (i.e., don't include MA terms, still use AIC to select *p*).
  - Step 1: Estimate optimal lag order of

$$\Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \ldots + \Delta p_{t-p+1} + \epsilon_t$$

  -
  - This is algebraic rearrangement of an AR(p) with $\alpha_0 = 0$ for Case 2, and $\alpha_0 \neq 0$ for Case 4.

Augmented Dickey Fuller Test

- Step 2:
  - =In Case 2, estimate

$$\Delta p_t = \phi_0 + \rho p_{t-1} + \alpha_1 \Delta p_{t-1} + \ldots + \Delta p_{t-p+1} + e_t,$$

    - with same hypotheses for ϕ_0, ρ as before.
    - All α's converge to Normal, ϕ's and ρ's converge to the weird distribution we already know from Case 2.
  - Likewise in Case 4, estimate

$$\Delta p_t = \phi_0 + \rho p_{t-1} + \beta t + \alpha_1 \Delta p_{t-1} + \ldots + \Delta p_{t-p+1} + e_t$$