Additional stuff you need to know about time-series analysis

Jason Qiang Guo, New York

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Weak stationarity

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Weak stationarity

- Mean stationarity: $E(Y_t) = \mu \ \forall t$

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Weak stationarity

- Mean stationarity: $E(Y_t) = \mu \ \forall t$
- Variance stationarity: $Var(Y_t) = E[(Y_t - \mu)^2] = \sigma^2 \ \forall t$

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Weak stationarity

- Mean stationarity: $E(Y_t) = \mu \ \forall t$
- Variance stationarity: $Var(Y_t) = E[(Y_t - \mu)^2] = \sigma^2 \ \forall t$
- Covariance stationarity: $Cov(Y_t, Y_{t-s}) = E[(Y_t - \mu)(Y_{t-s} - \mu)] = \gamma_s \ \forall s$, meaning the serial correlation of two observations $\{Y_t, Y_{t-s}\}$ depends only on the lag s and not on "where" in the series they fall.

Stationarity

Stationarity means that the fundamental form of the data generating process remains the same over time. This is indicated in the moments of the process.

Weak stationarity

- Mean stationarity: $E(Y_t) = \mu \; \forall t$
- Variance stationarity: $Var(Y_t) = E[(Y_t - \mu)^2] = \sigma^2 \; \forall t$
- Covariance stationarity: $Cov(Y_t, Y_{t-s}) = E[(Y_t - \mu)(Y_{t-s} - \mu)] = \gamma_s \; \forall s$, meaning the serial correlation of two observations $\{Y_t, Y_{t-s}\}$ depends only on the lag s and not on "where" in the series they fall.

Strong stationarity:

$f_{Y_1, Y_2, \ldots, Y_t}(Y_1, Y_2, \ldots, Y_t) = f_{Y_{1+s}, Y_{2+s}, \ldots, Y_{t+s}}(Y_{1+s}, Y_{2+s}, \ldots, Y_{t+s})$ (joint probability distribution) for all t and s.

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim$ *i.i.d.* (white noise)

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim$ *i.i.d.* (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim$ *i.i.d.* (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Check conditions for stationarity:

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim i.i.d.$ (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Check conditions for stationarity:

- $E(Y_t) = E(Y_0 + \sum_{i=1}^{t} \nu_i) = Y_0$, but it can drift away from its mean and have no tendency to return to the mean.

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim i.i.d.$ (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Check conditions for stationarity:

- $E(Y_t) = E(Y_0 + \sum_{i=1}^{t} \nu_i) = Y_0$, but it can drift away from its mean and have no tendency to return to the mean.
- $Var(Y_t) = Var(Y_0 + \sum_{i=1}^{t} \nu_i) = t\sigma^2$

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim i.i.d.$ (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Check conditions for stationarity:

- $E(Y_t) = E(Y_0 + \sum_{i=1}^{t} \nu_i) = Y_0$, but it can drift away from its mean and have no tendency to return to the mean.
- $Var(Y_t) = Var(Y_0 + \sum_{i=1}^{t} \nu_i) = t\sigma^2$
- $Cov(Y_t, Y_{t-s}) =$

Integrated processes

Integrated process is simply that the value of Y_t is the sum of random "shocks".

I(1) series

- $Y_t = Y_{t-1} + \nu_t$, $\nu_t \sim i.i.d.$ (white noise)
- Recursively, $Y_t = Y_{t-2} + \nu_{t-1} + \nu_t = Y_0 + \sum_{i=1}^{t} \nu_i$

Check conditions for stationarity:

- $E(Y_t) = E(Y_0 + \sum_{i=1}^{t} \nu_i) = Y_0$, but it can drift away from its mean and have no tendency to return to the mean.
- $Var(Y_t) = Var(Y_0 + \sum_{i=1}^{t} \nu_i) = t\sigma^2$
- $Cov(Y_t, Y_{t-s}) =$
  $Cov(Y_0 + \sum_{i=1}^{t} \nu_i, Y_0 + \sum_{i=1}^{t-s} \nu_i) = Cov(\sum_{i=1}^{t-s} \nu_i, \sum_{i=1}^{t-s} \nu_i) +$
  $Cov(\sum_{i=t-s+1}^{t} \nu_i, \sum_{i=1}^{t-s} \nu_i) = Var(\sum_{i=1}^{t-s} \nu_i) = (t-s)\sigma^2$

Integrated processes

- Typically, when we want to look at the stationarity of a series we might have the following equation:

$$Y_t = \rho Y_{t-1} + \nu_t,$$

Y_t is stationary if $\rho < 1$ and non-stationary if $\rho \geq 1$. But $\Delta Y_t = Y_t - Y_{t-1} = \nu_t$ is stationary, and in a more general form, $\Delta Y_t = Y_t - Y_{t-1} = c + \nu_t$ is unit root stationary.

Integrated processes

- Typically, when we want to look at the stationarity of a series we might have the following equation:

$$Y_t = \rho Y_{t-1} + \nu_t,$$

  Y_t is stationary if $\rho < 1$ and non-stationary if $\rho \geq 1$. But $\Delta Y_t = Y_t - Y_{t-1} = \nu_t$ is stationary, and in a more general form, $\Delta Y_t = Y_t - Y_{t-1} = c + \nu_t$ is unit root stationary.

- Given the potential existence of non-stationarity, how do we test it?

Integrated processes

- Typically, when we want to look at the stationarity of a series we might have the following equation:

$$Y_t = \rho Y_{t-1} + \nu_t,$$

Y_t is stationary if $\rho < 1$ and non-stationary if $\rho \geq 1$. But $\Delta Y_t = Y_t - Y_{t-1} = \nu_t$ is stationary, and in a more general form, $\Delta Y_t = Y_t - Y_{t-1} = c + \nu_t$ is unit root stationary.

- Given the potential existence of non-stationarity, how do we test it? Dickey-Fuller test: $H_0 : \rho = 1$ against $H_a : \rho < 1$, and we construct a standard-looking t-test $\hat{\tau} = \dfrac{\hat{\rho} - 1}{se(\hat{\rho})}$, and do a one-sided test. BUT!

Integrated processes

- Typically, when we want to look at the stationarity of a series we might have the following equation:

$$Y_t = \rho Y_{t-1} + \nu_t,$$

  Y_t is stationary if $\rho < 1$ and non-stationary if $\rho \geq 1$. But $\Delta Y_t = Y_t - Y_{t-1} = \nu_t$ is stationary, and in a more general form, $\Delta Y_t = Y_t - Y_{t-1} = c + \nu_t$ is unit root stationary.

- Given the potential existence of non-stationarity, how do we test it? Dickey-Fuller test: $H_0 : \rho = 1$ against $H_a : \rho < 1$, and we construct a standard-looking t-test $\hat{\tau} = \dfrac{\hat{\rho} - 1}{se(\hat{\rho})}$, and do a one-sided test. BUT! D-F distribution is more right-skewed than t-statistics, meaning we may incorrectly conclude the series is stationary when it really isn't if we use a t-test. For instance, the the 5% critical value (one-sided) is -1.95 and not -1.65.

Integrated processes

- Another way to do D-F test:

$$\Delta Y_t = (\rho - 1)Y_{t-1} + \nu_t = \delta Y_{t-1} + \nu_t,$$

and we need to test if $\delta = 0$, which is equivalent to testing if $\rho = 1$.

Integrated processes

- Another way to do D-F test:

$$\Delta Y_t = (\rho - 1)Y_{t-1} + \nu_t = \delta Y_{t-1} + \nu_t,$$

and we need to test if $\delta = 0$, which is equivalent to testing if $\rho = 1$.

Note, there are different sources of non-stationarity, like drift (add a constant) and trend (adding βt to the series), and we can use slightly different D-F tests (All different types of D-F tests have been built in R and STATA) to test stationarity.

Co-integration

When both X and Y are $I(1)$, then the hope is that they are cointegrated, i.e., they are in a long run equilibrium relationship. How do we know if that's the case? if residuals are stationary.

Co-integration

When both X and Y are $I(1)$, then the hope is that they are cointegrated, i.e., they are in a long run equilibrium relationship. How do we know if that's the case? if residuals are stationary.

- regress Y on X, and we obtain residuals $\hat{\epsilon}$ and estimator $\hat{\beta}$.

Co-integration

When both X and Y are $I(1)$, then the hope is that they are cointegrated, i.e., they are in a long run equilibrium relationship. How do we know if that's the case? if residuals are stationary.

- regress Y on X, and we obtain residuals $\hat{\epsilon}$ and estimator $\hat{\beta}$.
- Do a D-F test with residuals, and if residuals are stationary, then Y and X are co-integrated and $\hat{\beta}$ is the co-integrated value.

Co-integration

When both X and Y are $I(1)$, then the hope is that they are cointegrated, i.e., they are in a long run equilibrium relationship. How do we know if that's the case? if residuals are stationary.

- regress Y on X, and we obtain residuals $\hat{\epsilon}$ and estimator $\hat{\beta}$.
- Do a D-F test with residuals, and if residuals are stationary, then Y and X are co-integrated and $\hat{\beta}$ is the co-integrated value.

Spurious Regression: Two unrelated non-stationary series can have a strong correlation. Example:

- $X_t = X_{t-1} + \nu_t$ and $Y_t = Y_{t-1} + \epsilon_t$, where both ν_t and $\epsilon_t \sim$ i.i.d.
- We can find a strong correlation between these two series if we run an OLS regression, but it is simply spurious.

Error Correction Model

Co-integration tells us that Y and X in the long run will enter the equilibrium with adjustment. But what if the question is how fast is the adjustment when Y and X are out of equilibrium? Here comes the Error Correction Model:

$$\Delta Y_t = \beta_0 \Delta X_t + \gamma[Y_{t-1} - \beta_3 X_{t-1}] + \nu_t$$

Error Correction Model

Co-integration tells us that Y and X in the long run will enter the equilibrium with adjustment. But what if the question is how fast is the adjustment when Y and X are out of equilibrium? Here comes the Error Correction Model:

$$\Delta Y_t = \beta_0 \Delta X_t + \gamma [Y_{t-1} - \beta_3 X_{t-1}] + \nu_t$$

- β_0 captures the short-run relationship between X and Y.

Error Correction Model

Co-integration tells us that Y and X in the long run will enter the equilibrium with adjustment. But what if the question is how fast is the adjustment when Y and X are out of equilibrium? Here comes the Error Correction Model:

$$\Delta Y_t = \beta_0 \Delta X_t + \gamma[Y_{t-1} - \beta_3 X_{t-1}] + \nu_t$$

- β_0 captures the short-run relationship between X and Y.
- γ captures the rate at which the series returns to its equilibrium level. The coefficient should be negative and its absolute value should be smaller than 1 so as to indicate its re-equilibrating properties.

Error Correction Model

Co-integration tells us that Y and X in the long run will enter the equilibrium with adjustment. But what if the question is how fast is the adjustment when Y and X are out of equilibrium? Here comes the Error Correction Model:

$$\Delta Y_t = \beta_0 \Delta X_t + \gamma[Y_{t-1} - \beta_3 X_{t-1}] + \nu_t$$

- β_0 captures the short-run relationship between X and Y.
- γ captures the rate at which the series returns to its equilibrium level. The coefficient should be negative and its absolute value should be smaller than 1 so as to indicate its re-equilibrating properties.
- β_3 captures the long-run relationship between X and Y.

Error Correction Model

ECM Equivalence with ADL(1,1) Model. The following ADL(1,1) model

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \rho Y_{t-1} + \nu_t$$

can be rewritten as the following error correction model

$$\Delta Y_t = \beta_0 X_t + (\rho - 1)[Y_{t-1} - \left(\frac{-(\beta_0 + \beta_1)}{\rho - 1}\right) X_{t-1}] + \nu_t.$$

Let $\rho - 1 = \gamma$, and $-\dfrac{\beta_0 + \beta_1}{\rho - 1} = \beta_3$, we have exactly the same format as ECM. So the short term relationship between X and Y is captured by β_0, the long term relationship is captured by $\dfrac{\beta_0 + \beta_1}{1 - \rho}$, and the re-equilibrating rate is $\rho - 1$.

Error Correction Model

Estimating ECM: Engle-Granger Two Step Procedure

- Estimate the following regression: $Y_t = \alpha + \gamma X_t + \epsilon_t$

Error Correction Model

Estimating ECM: Engle-Granger Two Step Procedure

- Estimate the following regression: $Y_t = \alpha + \gamma X_t + \epsilon_t$
- From these estimates, generate the residuals, i.e. $\hat{\epsilon} = Y_t - \hat{\alpha} + \hat{\gamma} X_t$

Error Correction Model

Estimating ECM: Engle-Granger Two Step Procedure

- Estimate the following regression: $Y_t = \alpha + \gamma X_t + \epsilon_t$
- From these estimates, generate the residuals, i.e. $\hat{\epsilon} = Y_t - \hat{\alpha} + \hat{\gamma} X_t$
- Now include the lag of the residuals from the initial regression, so that we have $\Delta Y_t = \beta_0 + \beta_1 \Delta X_t + \epsilon_{\hat{t-1}}$

Short term relationship and long term relationship revisited

When we have an ADL(1,1) model, we can use Impulse Response Function and Unit Response Function to think about short term and long term effects of X on Y.

Horizon	Impulse	Response	Unit	Response
0	0	0	0	0
1	1	β_0	1	β_0
2	0	$\beta_1 + \rho\beta_0$	1	$\beta_0 + \beta_1 + \rho\beta_0$
3	0	$\rho(\beta_1 + \rho\beta_0)$	1	$\beta_0 + \beta_1 + \rho(\beta_0 + \beta_1 + \rho\beta_0)$
4	0	$\rho^2(\beta_1 + \rho\beta_0)$	1	$\beta_0 + \beta_1 + \rho(\beta_0 + \beta_1 + \rho(\beta_0 + \beta_1 + \rho\beta_0))$

Note: The shock occurs in period 1

Short term effects of X on Y is β_0, the effect in the period of impulse. Long term effects can either be interpreted as the unit response of Y after the shock with infinite number of periods, or the cumulative of the impulse response of Y and all of its delayed responses after the shock.

$$\begin{aligned}\text{Long Term Effects}\beta_3 &= (\beta_0 + \beta_1) + \rho(\beta_0 + \beta_1) \\ &+ \rho^2(\beta_0 + \beta_1), \ldots, \rho^{t-2}(\beta_0 + \beta_1) + \rho^{t-1}\beta_0 \\ &= \frac{\beta_0 + \beta_1}{1 - \rho}\end{aligned}$$

L Operator:

Assume $LY_t = Y_{t-1}$, meaning every term in the equation of $Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \rho Y_{t-1} + \nu_t$ is multiplied by L. Then

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \rho Y_{t-1} + \nu_t$$

can be rewritten as

$$(1 - \rho L) Y_t = (\beta_0 + \beta_1 L) X_t + \nu_t,$$

which leads to

$$Y_t = \left(\frac{\beta_0 + \beta_1 L}{1 - \rho L}\right) X_t + \frac{\nu_t}{1 - \rho L}$$