Lecture 12: Cointegration and Statistical Arbitrage

Pair trading

Brian Thelen
258 West Hall
bjthelen@umich.edu

Reference: Chapter 15 of Ruppert/Matteson

Overview

- Background and Concept of Cointegration
- Theory and statistical inference/estimation
- Examples and Utility
- Summary

Basic Ideas

如果k阶差分平稳,就可以用I(k)来描述

Recall: Backward operator B: $BX_n = X_{n-1}$

Notation/Definition. Stochastic process X_t is $I(k)$ if the k^{th} order difference, $Y_t = (I - B)^k X_t$, is stationary.

Stationary Model. Suppose an asset's price or log-price, X_n, was a stationary process (denote as $I(0)$) – what are some investment strategies 均值和协方差不变

- Buy when stock price is below mean, $\mu - \boxed{\beta_b}$
- Sell when stock price is above mean, $\mu + \boxed{\beta_s}$

Implications

Example. Our primary model has been that the log-price at time n, $\log(X_n)$ is random walk type of process with stationary differences, i.e., - denote by $I(1)$, and log-returns $\log(X_{n+1}) - \log(X_n)$ are stationary. 我们以前处理的数据只有1阶差分(收益率)平稳

Basic Ideas - Alternative 1

Random Walk Model. Suppose log-price data X_n is a true random walk with starting mean μ_o and at each time n,

$$\begin{aligned} X_0 &= \mu_0 + \epsilon_0 \\ X_1 &= X_0 + \triangle\mu + \epsilon_1 \\ &\vdots \\ X_n &= X_{n-1} + \triangle\mu + \epsilon_n \end{aligned}$$

Example. State an alternative/similar investment strategy for buying low and selling high.

Answer. Note that this is a random walk with drift –

$$X_n = \mu_o + n\triangle\mu + \sum_{i=1}^{n} \epsilon_i$$

Could look at some rule where

Buy when $X_n \leq \boxed{n\triangle\mu - \beta_{b,n}}$ and sell when $X_n \geq n\triangle\mu + \beta_{s,n}$

Differences between Random Walk and Stationary

- Have presented similar strategies for random walk and stationary models
- What are differences?
- Let T be time it takes to go from "buy" threshold to hit the "sell" threshold

For random walk model, $E(T) = \infty$

有限时间内可以找到做多或者卖出的机会

For stationary model, $E(T) =$ finite constant

Implications: For stationary model, there is always a reversion to the mean, where-as for random walk there may not be

价格平稳就说明存在均值回归

Co-Integration - Two time series

Definition. Two time series $Y_{1,t}$ and $Y_{2,t}$ are said to co-integrated if each is $I(1)$, but there exists a λ such that

$$Y_{1,t} - \lambda Y_{2,t} \text{ is stationary, i.e., } I(0)$$ 相差是平稳的

Example. Suppose W_t follows a random walk model, and

$$\begin{aligned} Y_{1,t} &= \beta_1 W_t + \epsilon_{1,t} \\ Y_{2,t} &= \beta_2 W_t + \epsilon_{2,t} \end{aligned}$$

Show that these time series are co-integrated.

Answer. Let $\lambda = \frac{\beta_1}{\beta_2}$ – then

$$\begin{aligned} Y_{1,t} - \lambda Y_{2,t} &= \beta_1 W_t + \epsilon_{1,t} - \frac{\beta_1}{\beta_2}\left[\beta_2 W_t + \epsilon_{2,t}\right] \\ &= \epsilon_{1,t} - \frac{\beta_1}{\beta_2}\epsilon_{2,t} \end{aligned}$$

and this is a stationary white noise process.

Co-Integration - More than two time series

d个资产套利

Definition. Suppose have d time series $Y_{1,t}, Y_{2,t}, \ldots, Y_{d,t}$. They are said to be co-integrated with order $1 \le r \le d$, if there exists r linearly independent linear combinations of the time series that are stationary, i.e., there is an $r \times d$ matrix $\mathbf{A}$ which is of rank r, and

$$\begin{bmatrix} X_{1,t} \\ \vdots \\ X_{r,t} \end{bmatrix} = \mathbf{A} \begin{bmatrix} Y_{1,t} \\ Y_{2,t} \\ \vdots \\ Y_{d,t} \end{bmatrix}$$

are stationary time series.

The of vector X is constant

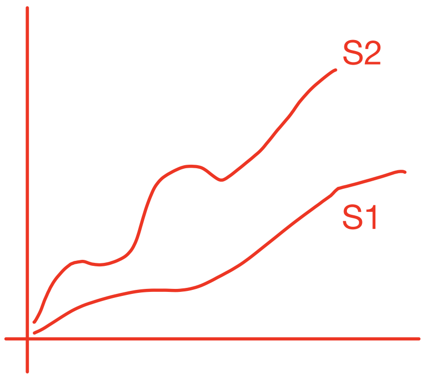
Implications

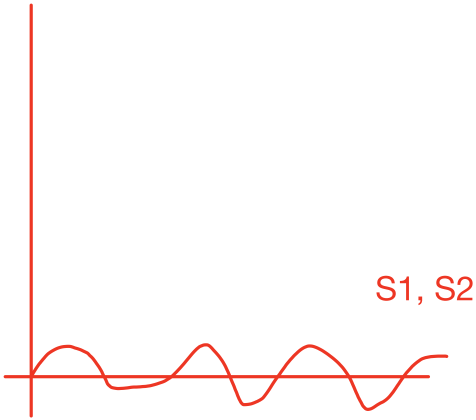
Cross auto covariance Cov(X_t,X_{t+h}) only depends on h

Pairs Trading

- Pairs trading is focused on using pairs of co-integrated asset prices
  - Similar assets
  - Assuming convergence of the prices – highly correlated

Pictures





Determining Stationarity vs. Non-Stationarity

Suppose Y_t follows the ARMA(p,q) model of

$$Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

Conditions for Stationarity Process is stationary if all (possibly) complex roots of the polynomial equation for the AR coefficients,

$$\phi(x) = 1 - \alpha_1 x - \cdots - \alpha_p x^p$$

have absolute value larger than 1. ARMA单位根检验,只有AR决定平稳性,
可以用来检验数据是否平稳

Example. Suppose have AR(1) model,

$$Y_t = \alpha_1 Y_{t-1} + \epsilon_t$$

Derive the conditions for stationarity.

Answer. Requires that the root of

$$1 - \alpha_1 x$$

have absolute value larger than 1. But the root of this equation is clearly $x = \frac{1}{\alpha_1}$, and so this is equivalent to $|\alpha_1| < 1$.

Problem. Suppose Y_t follows the ARMA(p,q) model of

$$Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

(a) Show that $\sum_{i=1}^{p} |\alpha_i| < 1$ implies stationarity.

(b) Give an example of a stationary AR(2) model that does not satisfy condition in part **(a)** . 参数小于1是充分不必要条件

Answer. (a) Suppose x is such that $|x| \leq 1$. Then

$$1 - \sum_{i=1}^{p} \alpha_i x^p \geq 1 - \sum_{i=1}^{p} |\alpha_i||x|^p \geq 1 - \sum_{i=1}^{p} |\alpha_i| > 0$$

and so x cannot be a root.

(b) The polynomial of

$$(1 + .8x)(1 + .7x) = 1 + 1.5x + .56x^2$$

has roots that have absolute value strictly larger than 1. This corresponds to the AR model of

$$Y_t = -1.5Y_{t-1} - .56Y_{t-2} + \epsilon_t$$

Testing Stationarity

Background: In ARMA model estimate AR coefficients $\{\hat{\alpha}_i\}_1^p$ and want to test stationarity. Testing whether all roots x_i of

$$\phi(x) = 1 - \sum_{i=1}^{p} \hat{\alpha}_i x^i$$

satisfy that $|x_i| > 1$ for $i = 1, 2, \ldots, p$.

One Approach: Generate AR coefficient estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_p$, do a root-finding on the polynomial, and assess evidence.

Example. Suppose $\tilde{\alpha}_1 = 1.207$ and $\hat{\alpha}_2 = -.224$. Using R's polyroot, find that roots have absolute value of 1.022 and 4.377. What can you say?? 获得参数后,直接计算单位根

Answer. Since 1.022≈ 1, the randomness of data/estimates does not preclude that the real root has unit magnitude.

Alternative Tests: Dickey-Fuller and Phillips-Perron

- Null hypothesis is that there is a unit magnitude root (non-stationary) vs. alternative hypothesis there is not
- Small p-value suggests stationarity

P-value越小越好

Regression Approach - Cointegration

Example. Suppose have two time series Y_{1t} and Y_{2t} as in example, of

$$\begin{aligned} Y_{1,t} &= \beta_1 W_t + \epsilon_{1,t} \\ Y_{2,t} &= \beta_2 W_2 + \epsilon_{2,t} \end{aligned}$$

1. 用回归估计lambda
2. 使用PPtest对残差进行平稳性检验

- For the λ which makes $Y_{1,t} - \lambda Y_{2,t}$ stationary, also makes the differences **small**
- Suggests estimating λ via regression model/analysis of

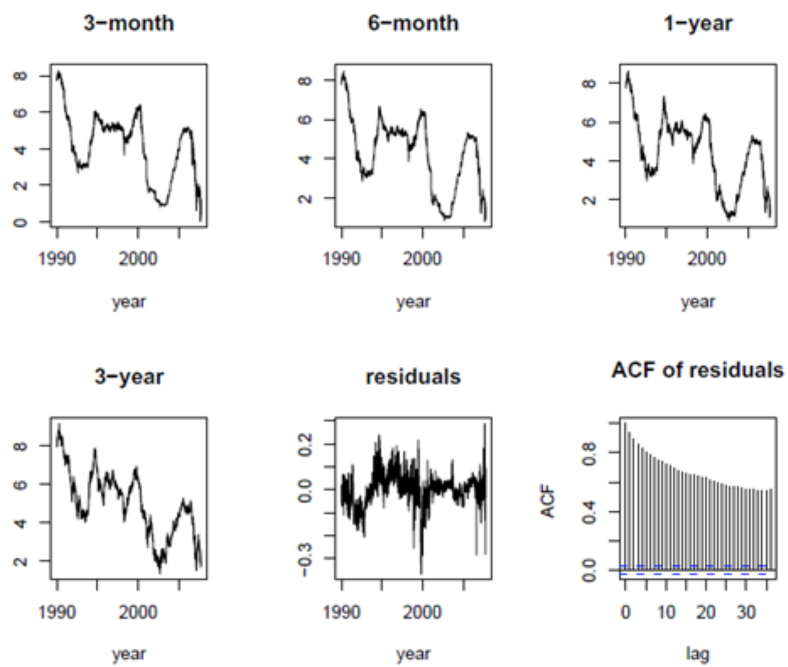
$$Y_{1t} = \lambda Y_{2t} + \delta_t$$

  and then carry out Unit Root tests on the residuals.
- Taking this approach with Phillips-Perron test on the resulting residuals corresponds to the so-called Phillips-Ouliarus test (po.test in R package tseries)
- Can generalize to more than two time series

Example (from Ruppert/Matteson).

Background. Have 3-month, 6-month, 1-year, 2-year, and 3-year bond yield data (daily) from Jan 1990 to Oct 2008.

- Track each other closely – suggests possibility of cointegration



在债券之间进行套利

Fig. 15.1. *Time series plots of the five yields and the residuals from a regression of the 1-year yields on the other four yields. Also, a ACF plot of the residuals.*

Phillips-Ouliaris test

R-package: tseries

```
Usage: po.test(x, demean = TRUE, lshort = TRUE)
Arguments
x -- a matrix or multivariate time series.
demean -- a logical indicating whether an intercept is included
 in the cointegration regression or not.
lshort -- a logical indicating whether the short or long version
 of the truncation lag parameter is used.

Details:
The Phillips-Perron Z(alpha) statistic for a unit root in the
residuals of the cointegration regression is computed, see also
pp.test. The unit root is estimated from a regression of the
first variable (column) of x on the remaining variables of x
without a constant and a linear trend. To estimate sigma^2 the
Newey-West estimator is used. If lshort is TRUE, then the
truncation lag parameter is set to trunc(n/100), otherwise
trunc(n/30) is used.

Value.
statistic -- the value of the test statistic.
parameter -- the truncation lag parameter.
p.value -- the p-value of the test.
method -- a character string indicating what type of test was performed.
data.name -- a character string giving the name of the data.
```

Example (from Ruppert/Matteson).

Run Phillips-Ouliaris Test

```
        Phillips-Ouliaris Cointegration Test

data:  dat[, c(3, 1, 2, 4, 5)]
Phillips-Ouliaris demeaned = -323.546, Truncation lag
parameter = 47, p-value = 0.01

Warning message:
In po.test(dat[, c(3, 1, 2, 4, 5)]) : p-value smaller
than printed p-value
```

PO的p-value很小,说明残差平稳

- Interpretation
- ACF of residuals - what does this say about "reversion time" for exploiting this in statistical arbitrage

Market-Neutrality

Example. Suppose have SML model and two assets where the log-return data satisfies that

$$R_{1,t} = \beta_1 R_{M,t} + \eta_{1,t}$$

$$R_{2,t} = \beta_2 R_{M,t} + \eta_{2,t}$$

If you buy S_1 worth shares of β_1 and short S_2 worth of shares of β_2, get the return is

$$R_t = S_1 R_{1,t} - S_2 R_{2,t}$$

Note that if $S_1\beta_1 - S_2\beta_2 = 0$, this represents market-neutral strategy, and it's profitability/gain depends on the process $\underline{S_1\epsilon_{1,t} - S_2\epsilon_{2,t}}$.

- How much and when (near/long term)

Pairs-Trading

Suppose have model

$$Y_{1,t} = \beta_1 Y_{M,t} + \epsilon_{1,t}$$

$$Y_{2,t} = \beta_2 Y_{M,t} + \epsilon_{2,t}$$

$Y_{1,t}, Y_{2,t}, Y_{M,t}$ are log-prices of the assets 1,2, and market. The cointegration model would assume/hope that if the assets are closely related, that the resulting process of

$$Y_{1,t} - \frac{\beta_1}{\beta_2} Y_{2,t} = \epsilon_{1,t} - \frac{\beta_1}{\beta_2} \epsilon_{2,t}$$

is stationary.

- It is automatically market-neutral
- Exploit differences between the log-prices

Summary

- Pairs-trading trying to exploit tight correlations/relationships between stocks/assets
  - Example of Market-neutral strategies
- One approach relies on theory of cointegration
  - Much more on this in a few books/references in Ruppert/Matteson
- Part of the success depends on the resulting stationary process
  - What type of ARMA model – relates to "speed of convergence"
  - Can even have long-memory types of stationary processes – important to assess (topic not covered in class, but is in Ruppert)