

# Cointegration Analysis in Time Series

Kai Chien Chen

# Table of contents

# 簡介

- **共整合 ( Cointegration )** 用於分析非平穩時間序列之間的長期關係。
- 常見於 **經濟學** 和 **金融學** 中。
- 範例：**購買力平價 ( PPP )** 與匯率。

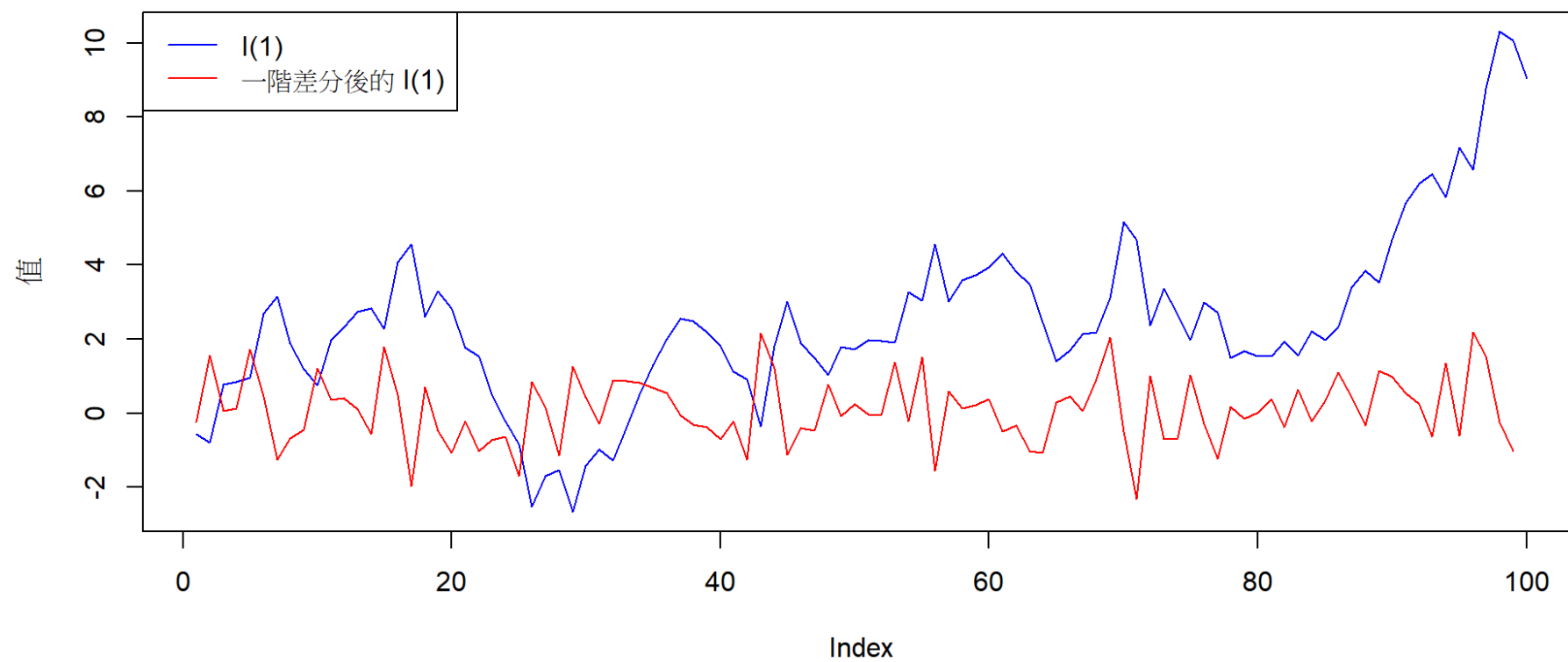
# 時間序列與 I(1) 序列

- 時間序列：隨時間收集的數據點序列。
- I(1) 序列：一階差分後成為定態的序列。
  - 例如：股票價格、GDP。

# I(1) 序列繪圖

## ► Code

I(1) 序列（非平穩）



# 範例

購買力平價（PPP）PPP 令  $q_t$  為實際匯率的對數， $s_t$  表示名義匯率的對數， $p_t$  和  $p_t^*$  分別表示國內與國外物價水平的對數。則有

$$q_t = s_t + p_t^* - p_t$$

# 什麼是共整合？

- 共整合：當兩個或多個  $I(1)$  序列有一個穩定的長期關係時。
- 數學上，若存在一個向量  $A$  使得： $A'y_t \sim I(0)$ ，其中  $y_t$  是  $I(1)$  變量的向量。

# 誤差修正模型

- ECM 描述了共整合序列的短期效果和長期效應。
- 模型為： $\Delta y_t = \beta_0 \Delta Z_t + (\varphi - 1)[y_{t-1} - (\frac{\beta_0 + \beta_1}{1 - \varphi})Z_{t-1}] + \epsilon_t$
- 其中：

$$y_t = \varphi y_{t-1} + \beta_0 Z_t + \beta_1 Z_{t-1} + \epsilon_t$$

$$Z_t = Z_{t-1} + u_t \text{ 其中}$$

$$u_t \sim N(0, \sigma^2)$$



# Engle-Granger 兩步驟法

## (步驟 1)

1. 假設我們有兩個時間序列  $y_t$  和  $z_t$  。
2. 估計共整合：

$$y_t = \beta_0 + \beta_1 z_t + e_t$$

3. 對  $\hat{e}_t$  進行 ADF 檢驗：

$$\Delta \hat{e} = a_0 + a_1 e_{t-1} + \sum_{i=1}^n a_{i+1} \Delta e_{t-i} + \epsilon_t$$

# Engle-Granger 兩步驟法 (步驟 2)

1. 零假設： $H_0 : a_1 = 0$   $H_1 : a_1 < 0$
2. 如果拒絕  $H_0$ ，則兩者存在共整合。

# Johansen 檢驗

- Johansen 檢驗 用於檢測多個共整合關係。

# VAR(p) 模型與 VECM 轉換

假設  $\mathbf{X}_t$  為  $k$  個變數的時間序列向量：

$$\mathbf{X}_t = \mathbf{A}_1 \mathbf{X}_{t-1} + \mathbf{A}_2 \mathbf{X}_{t-2} + \cdots + \mathbf{A}_p \mathbf{X}_{t-p} + \epsilon_t$$

將其改寫為差分形式的向量誤差修正模型（VECM）：

$$\Delta \mathbf{X}_t = \Pi \mathbf{X}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{X}_{t-i} + \epsilon_t$$

According to (Beveridge and Nelson JME 1981)

# 共整合矩陣 $\Pi$ 與長期關係

- 共整合矩陣  $\Pi = \alpha\beta'$ 
  - $\beta$ : 共整合向量矩陣，描述長期平穩關係
  - $\alpha$ : 調整速度矩陣，表示變數回到均衡的速度

# II 的秩與共整合關係

根據 II 的秩，我們可以確定共整合向量的數量：

1. 秩為 0 (  $\text{rank}(\Pi) = 0$  )：

- 無共整合向量，即無長期均衡關係。

2. 秩介於 0 和  $k$  之間 (  $0 < \text{rank}(\Pi) < k$  )：

- 存在  $r$  個共整合向量，代表  $r$  個長期均衡關係。

3. 滿秩  $k$  (  $\text{rank}(\Pi) = k$  )：

- 所有變數平穩，無需進行共整合檢定。

# $\Pi$ 的eigenvalue

1. 如果

$$rank(\Pi) = 0 \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

$$\Rightarrow \log(1 - \lambda_i) = 0 \quad \forall i$$

$$rank(\Pi) = k \Rightarrow \log(1 - \lambda_i) \neq 0 \quad \forall i$$

則  $x_t$  不存在共整合關係。



2. 如果  $rank(\Pi) = r$  , 且假設

$$\begin{cases} \lambda_1, \lambda_2, \dots, \lambda_r \neq 0 \\ \lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_k = 0 \end{cases}$$

亦即 ,

$$\begin{cases} \log(1 - \lambda_i) \neq 0 & \text{for } i = 1, 2, \dots, r \\ \log(1 - \lambda_i) = 0 & \text{for } i = r + 1, r + 2, \dots, k \end{cases}$$

則  $x_t$  存在共整合關係。

# Johansen 檢定步驟

## 1. 設定虛無假設與對立假設

- $H_0$ : 最大共整合階次為  $r$  (共整合關係數量)
- $H_1$ : 最大共整合階次為  $k$  or  $r + 1$

## 2. 計算兩種統計量：

- 跡檢定統計量：檢查  $r$  是否足夠
- 最大特徵值檢定統計量：檢查  $r + 1$  個共整合向量的可能性

# 跡檢定統計量(Trace Test)

跡檢定的公式為：

$$\text{Trace Statistic} = -T \sum_{i=r+1}^k \ln(1 - \lambda_i)$$

- 若統計量 > 臨界值，拒絕  $H_0$ ，表示存在多於  $r$  個共整合關係

# 最大特徵值檢定統計量(Max Test)

最大特徵值檢定的公式為：

$$\text{Max-Eigen Statistic} = -T \ln(1 - \lambda_{r+1})$$

# Cointegration in R

- R package: [urca](#) for conducting the Johansen test.

# R Code Example

# 1. Load Necessary Packages and data

```
1 # 安裝必要packages
2 # install.packages("vars")
3 library(urca)
4 library(tidyverse)
5 library(tseries)
6 library(quantmod)
```

```
1 # 下載資料
2 getSymbols("EWA", from="2006-04-26", to="2012-04-09")
```

```
[1] "EWA"
```

```
1 getSymbols("EWC", from="2006-04-26", to="2012-04-09")
```

```
[1] "EWC"
```

```
1 getSymbols("IGE", from="2006-04-26", to="2012-04-09")
```

```
[1] "IGE"
```

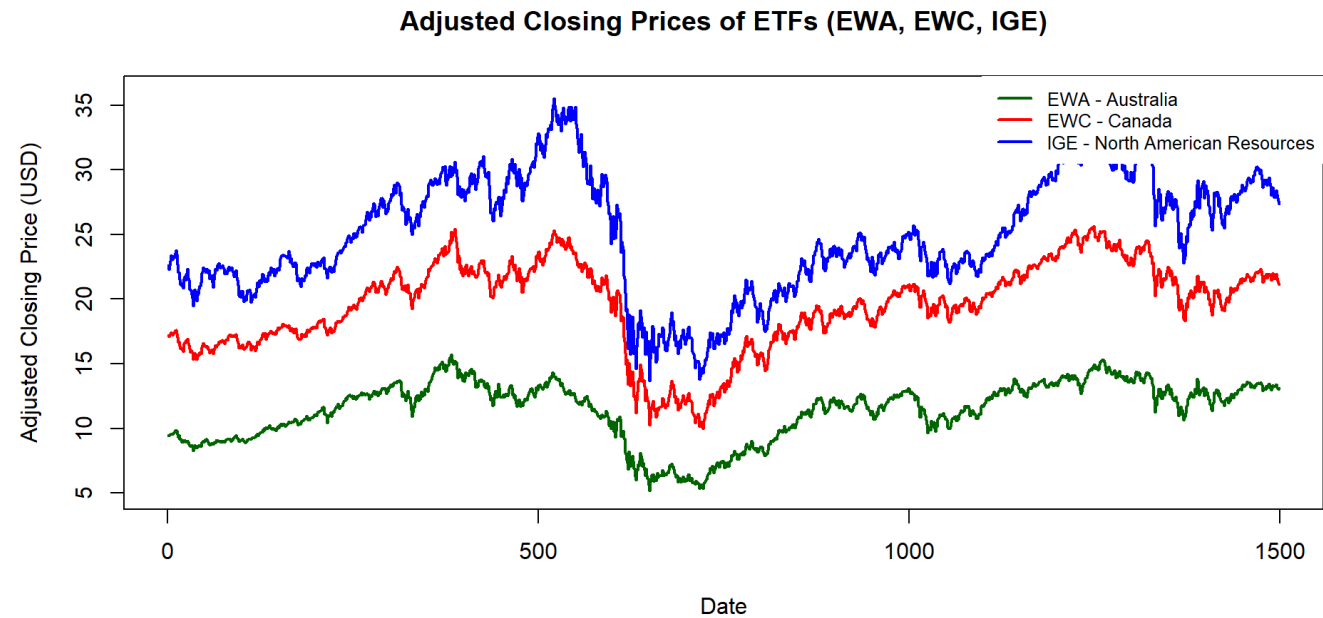
```
1 # 提取調整後的價格
2 ewaAdj = unclass(EWA$EWA.Adjusted)
3 ewcAdj = unclass(EWC$EWC.Adjusted)
4 igeAdj = unclass(IGE$IGE.Adjusted)
```

這三個資料為：

1. **EWA - iShares MSCI Australia ETF** 這是一個追蹤澳大利亞股票市場的 **ETF**，反映澳大利亞股票的整體表現。
2. **EWC - iShares MSCI Canada ETF** 這是一個追蹤加拿大股票市場的 **ETF**，反映加拿大市場的整體走勢。
3. **IGE - iShares North American Natural Resources ETF** 這是一個追蹤北美自然資源公司的 **ETF**，通常包括能源和材料類股，反映北美自然資源行業的表現。



# 2. Plot and describe the Data



Overall (N=1499)	
EWA.Adjusted	
Mean (SD)	11.4 (2.28)
Median [Min, Max]	12.1 [5.18, 15.7]
EWC.Adjusted	
Mean (SD)	19.5 (3.41)
Median [Min, Max]	20.1 [9.99, 25.6]
IGE.Adjusted	
Mean (SD)	25.0 (4.57)
Median [Min, Max]	24.8 [13.7, 35.5]

# 5. Augmented Dickey-Fuller Test (ADF)

```
1 # ADF検定
2 adf.test(ewaAdj)
```

Augmented Dickey-Fuller Test

```
data: ewaAdj
Dickey-Fuller = -1.732, Lag order = 11, p-value = 0.6918
alternative hypothesis: stationary
```

```
1 adf.test(ewcAdj)
```

Augmented Dickey-Fuller Test

```
data: ewcAdj
Dickey-Fuller = -1.8042, Lag order = 11, p-value = 0.6612
alternative hypothesis: stationary
```

```
1 adf.test(igeAdj)
```

Augmented Dickey-Fuller Test

```
data: igeAdj
```

# 6. Select the Optimal Lag for VAR

```
1 # 使用 VARselect 函數選擇最佳滯後階數
2 library(vars)
3 var_select <- VARselect(data.frame(ewaAdj, ewcAdj, igeAdj), lag.max = 10, t
4 var_select$selection
```

AIC (n)	HQ (n)	SC (n)	FPE (n)
4	2	2	4

# 7. Johansen Cointegration Test

## ► Code

```
#####  
# Johansen-Procedure #  
#####
```

```
Test type: trace statistic , with linear trend
```

```
Eigenvalues (lambda):
```

```
[1] 0.010707118 0.007278267 0.003526434
```

```
Values of teststatistic and critical values of test:
```

	test	10pct	5pct	1pct
r <= 2	5.28	6.50	8.18	11.65
r <= 1	16.21	15.66	17.95	23.52
r = 0	22.22	22.71	24.50	27.22

## ► Code

```
#####  
# Johansen-Procedure #  
#####
```

```
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
```

```
Eigenvalues (lambda):
```

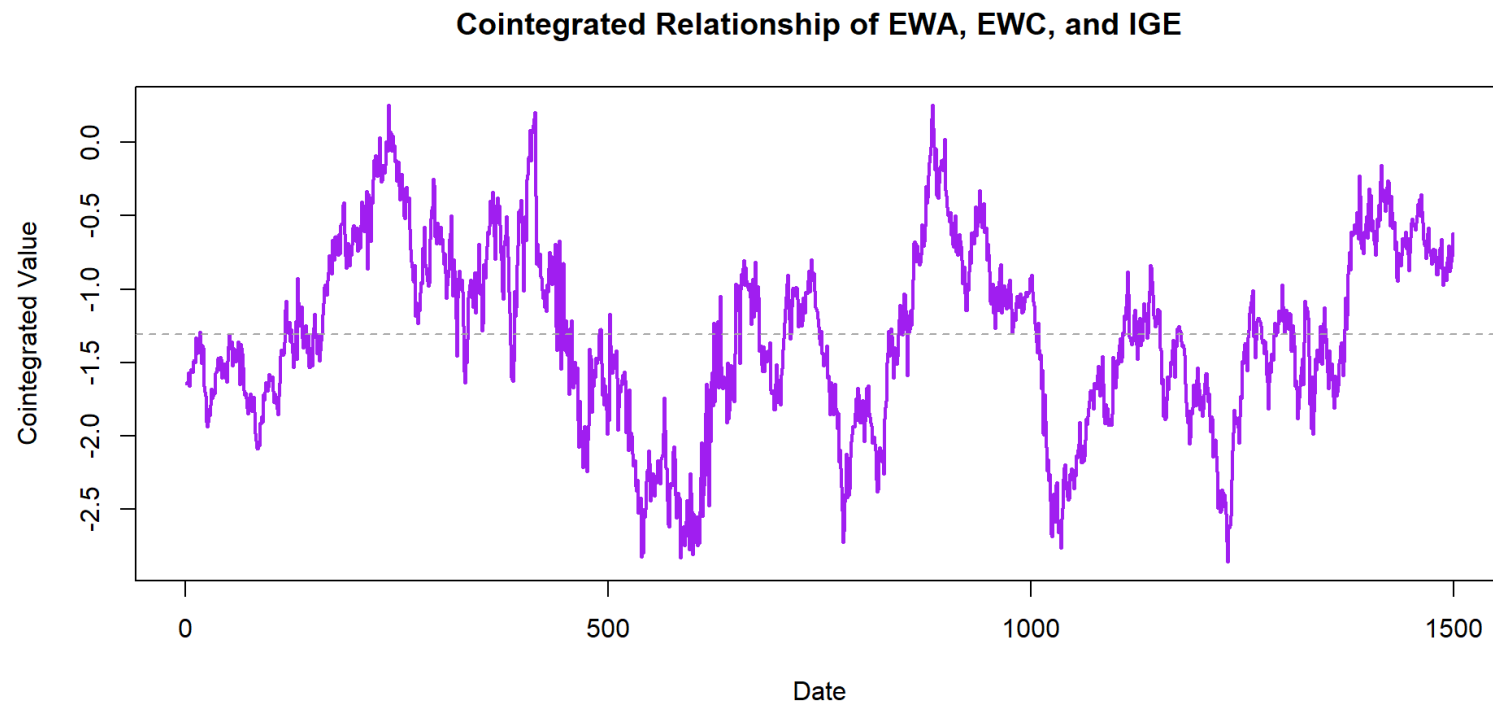
```
[1] 0.010707118 0.007278267 0.003526434
```

```
Values of teststatistic and critical values of test:
```

	test	10pct	5pct	1pct
r <= 2	5.28	6.50	8.18	11.65
r <= 1	10.93	12.91	14.90	19.19
r = 0	16.10	18.00	21.07	25.75

# 8. Calculate and Plot Cointegrated Relationship

► Code



# 9. ADF Test for Cointegrated Series

```
1 # ADF検定共整合時間序列  
2 adf.test(CI)
```

Augmented Dickey-Fuller Test

```
data: CI  
Dickey-Fuller = -3.1871, Lag order = 11, p-value = 0.09014  
alternative hypothesis: stationary
```

```
1 kpss.test(CI)
```

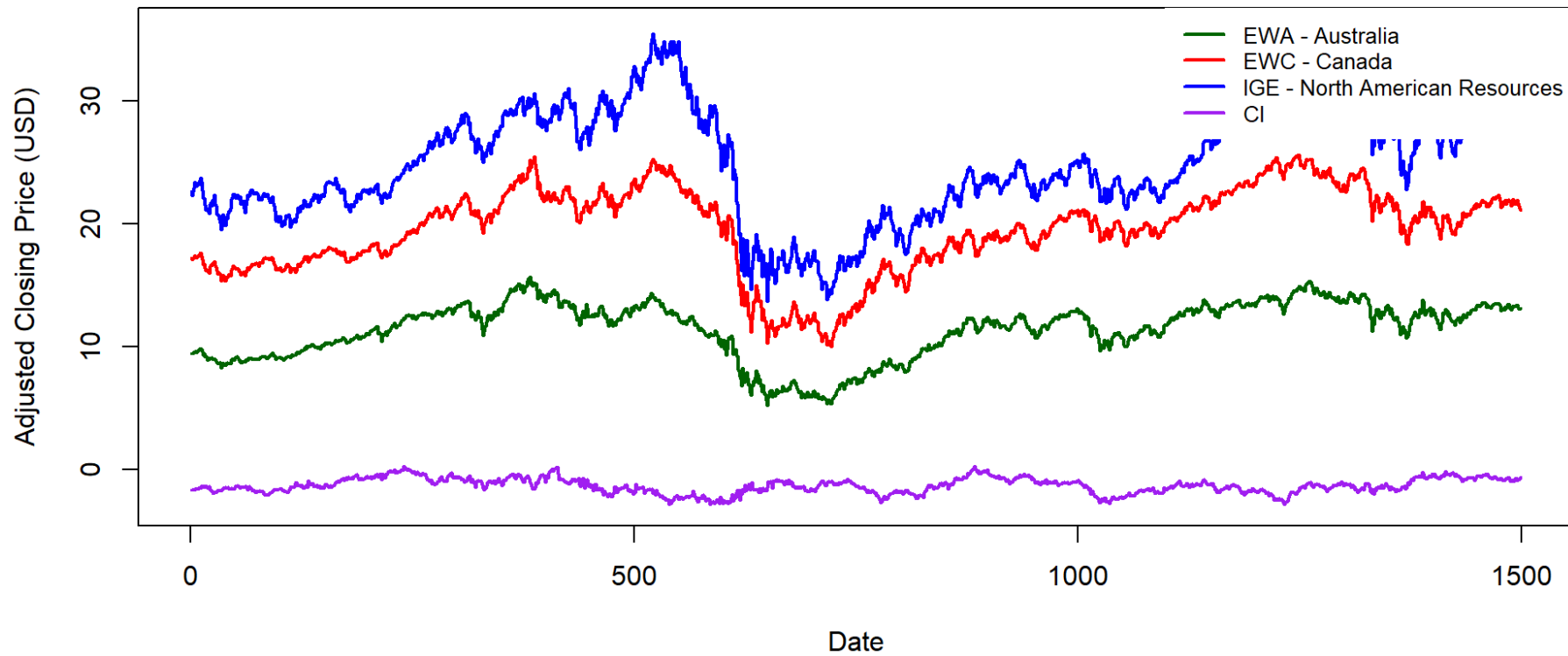
KPSS Test for Level Stationarity

```
data: CI  
KPSS Level = 0.75332, Truncation lag parameter = 7, p-value = 0.01
```

# 10. Plot All Series and Cointegration

## ► Code

Adjusted Closing Prices of ETFs (EWA, EWC, IGE)





# 結果差異的原因

- trace 檢定統計量傾向於更保守，因為它是對所有特徵值的累積檢定，對所有的共整合向量的數量進行評估，檢查是否有額外共整合的可能性。
- 最大特徵值檢定更集中於每一個特徵值的增量效果，因此可能更敏感於檢測出一個特定的共整合關係。

# 為何結果不同

# 共整合的應用

- 廣泛應用於：
  - **經濟學**：消費與收入之間的長期關係。
  - **金融學**：股價與利率之間的關係。

# 共整合的限制

- 共整合假設 **長期穩定性**，但這在現實中不一定成立。
- 對於複雜數據，可能需要其他方法，例如 **非線性模型**。

# 結論

- 共整合在分析非平穩時間序列中至關重要。
- 應用範圍包括 **宏觀經濟模型**、**金融時間序列**和 **預測**。
- 未來的研究可以探索 **非線性共整合**模型。

# 參考資料和markdown

- 林常青教授計量經濟(二)講義:Time Series Analysis (3): Spurious regression and Cointegration
- 時間序列分析: 總體經濟與財務金融之應用 / 陳旭昇 著
- <https://www.quantstart.com/articles/Johansen-Test-for-Cointegrating-Time-Series-Analysis-in-R/>

**END**