

# Operación por Pares

## Cointegración

Emanuel Mejia

```
library(quantmod)
library(tseries)
```

## 1 Pruebas de Cointegración

### 1.1 Prueba de Engle-Granger

La hipótesis nula de la prueba de Engle-Granger es que no existe cointegración, y concluimos que hay cointegración solo si rechazamos esta hipótesis.

$H_0$  : No existe cointegración

$H_1$  : Existe cointegración

#### Pasos de esta prueba:

1. Realizar una prueba de raíz unitaria (como Dickey-Fuller aumentada) para verificar si  $y$  y  $x$  tienen raíces unitarias. Si ambas son integradas del mismo orden (usualmente  $I(1)$ ), ir al siguiente paso.
2. Ejecutar una regresión de  $y$  sobre  $x$  y guardar los residuos.
3. Realizar una prueba de raíz unitaria sobre los residuos (sin incluir constante ni tendencia).
4. Si se rechaza la hipótesis de raíz unitaria, concluimos que  $y$  y  $x$  están cointegradas. Si no se rechaza, concluimos que no hay cointegración.

Por lo tanto, si  $y_t$  y  $x_t$  están cointegradas, en  $y_t = \alpha + \beta x_t + \epsilon_t$ , el término de error  $\epsilon_t$  será  $I(0)$ . Si no,  $\epsilon_t$  será  $I(1)$ . Así, se puede probar la existencia de cointegración evaluando la raíz unitaria en los residuos OLS.

### 1.2 Prueba de Phillips-Ouliaris

La prueba de Engle-Granger asume que los errores de regresión son independientes y con varianza constante, lo cual raramente ocurre en la práctica. La prueba de Phillips-Ouliaris mejora la prueba de Engle-Granger al considerar errores que no son ruido blanco.

$H_0$ : No existe cointegración

$H_1$ : Existe cointegración

### 1.3 Prueba de Johansen

Otra mejora sobre la prueba de Engle-Granger es la prueba desarrollada por Johansen, la cual permite detectar múltiples vectores de cointegración (útil en modelos con más de dos variables).

## 2 Pairs Trading (Operación por Pares)

Recordemos que si dos series de tiempo están cointegradas, permanecen “cerca” una de la otra en el largo plazo. En otras palabras, el spread (residuo OLS) entre ellas,  $z_t = x_t - \beta y_t$  es media-revertiente.

Esta propiedad de reversión a la media del spread puede aprovecharse para hacer operaciones de trading, y comúnmente se conoce como “pairs trading” o “arbitraje estadístico”.

### ¿Cómo funciona?

Se asume que el spread  $z_t = x_t - \beta y_t$  es estacionario o media-revertiente con media cero.

#### 2.1 Escenario 1: El spread es bajo $z_t < -s_0$

Esto sugiere que la acción  $x$  está infravalorada y la acción  $y$  sobrevalorada.

**Estrategia - Comprar el spread:**

**Comprar la acción  $x$ . Vender en corto la acción  $y$ .**

Cerrar las posiciones cuando el spread vuelva a cero después de  $n$  periodos (es decir  $z_{t+n} = 0$ ).

#### 2.2 Escenario 2: El spread es alto $z_t > s_0$

Esto sugiere que la acción  $x$  está sobrevalorada y la acción  $y$  infravalorada.

**Estrategia - Vender el spread:**

**Vender en corto la acción  $x$ . Comprar la acción  $y$ .**

Cerrar las posiciones cuando el spread vuelva a cero después de  $n$  periodos (es decir  $z_{t+n} = 0$ ).

### 2.3 Diseño del Pairs Trading

#### 2.3.1 Selección de Pares

Identificar pares de acciones que potencialmente estén cointegradas.

Criterios comunes para seleccionar los pares:

- Pertenencia al mismo sector o industria.
- Comportamiento histórico similar en precios.
- Alta correlación como punto de partida (aunque correlación  $\neq$  cointegración).

### 2.3.2 Prueba de Cointegración

Verificar si los pares seleccionados están efectivamente cointegrados.

Aplicar pruebas estadísticas como:

- Engle-Granger
- Phillips-Ouliaris
- Johansen

Si la cointegración se confirma, se puede considerar que los precios tienen una relación de largo plazo estable.

### 2.3.3 Diseño de la Estrategia de Trading

Estudiar la dinámica del spread entre las dos acciones:

$$z_t = x_t - \beta y_t$$

Establecer reglas claras de trading:

- **Umbral de entrada** (por ejemplo, cuando  $z_t$  supera  $\pm 2$  o  $\pm 3$  desviaciones estándar).
- **Condiciones de salida** (cuando  $z_t$  revierte a la media o a un valor objetivo).
- **Gestión del riesgo** (stop-loss, tamaño de la posición, etc.).

El objetivo es explotar las desviaciones temporales del equilibrio para obtener beneficios cuando el spread vuelve a su valor esperado.

## 3 Netflix VS Amazon

Centrémonos en el par NFLX vs. AMZN, que corresponde a las acciones de Netflix y Amazon.

### 3.1 Obtener precios de acciones

Con el siguiente código podemos obtener los precios de las acciones usando el paquete `quantmod`.

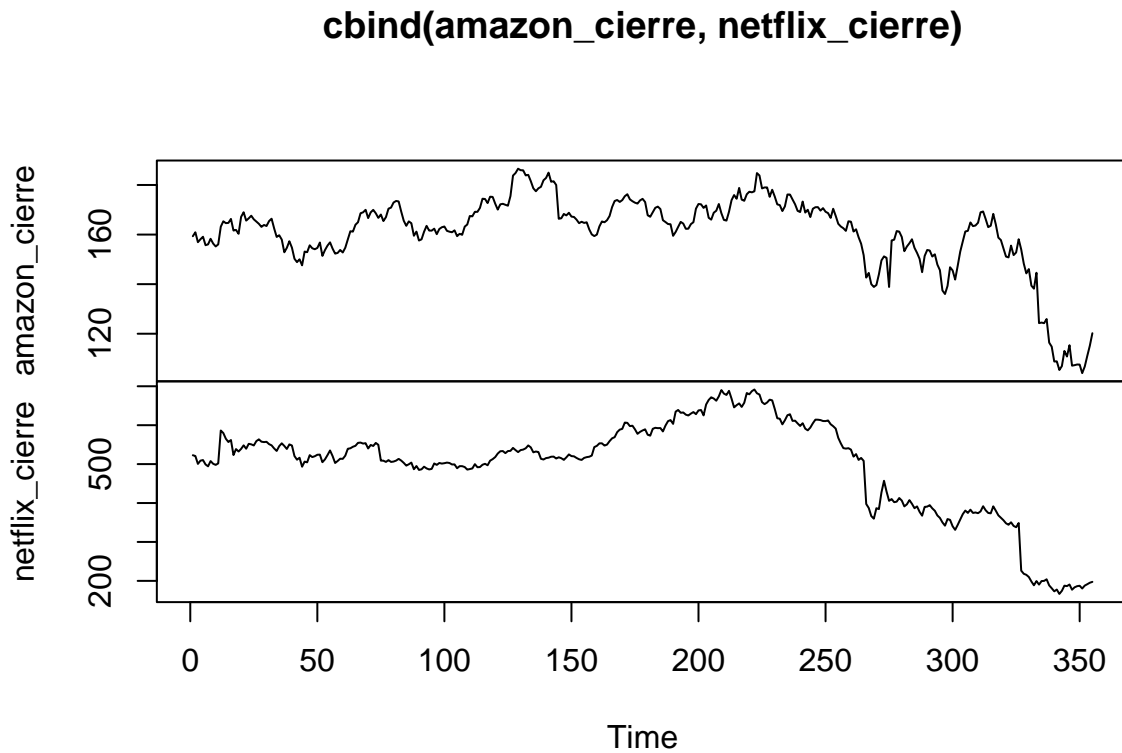
```
inicio <- as.Date("2021-01-01")
fin <- as.Date("2022-06-01")
stocks <- c("AMZN", "NFLX")

# Cargamos datos de dos acciones
# auto.assign=FALSE para permitir que el objeto se guarde
# en variable local y no en la sesión de R
amazon <- getSymbols("AMZN", src = "yahoo", from = inicio, to = fin,
                     auto.assign = FALSE, return.class = "ts")
netflix <- getSymbols("NFLX", src = "yahoo", from = inicio, to = fin,
                     auto.assign = FALSE, return.class = "ts")
```

### 3.2 Análisis Inicial

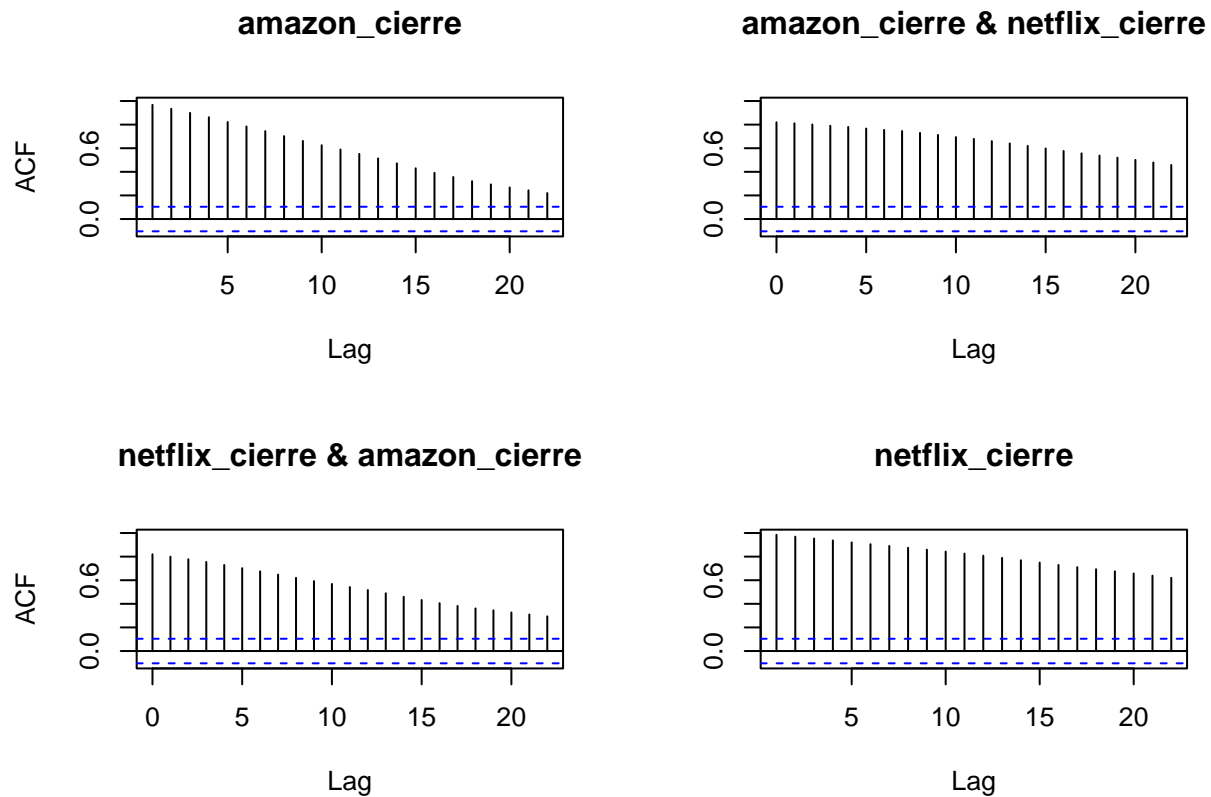
Graficamos los precios de cierre de NFLX y AMZN, las funciones de autocorrelación, y examinamos estacionariedad.

```
amazon_cierre <- amazon[,4]  
netflix_cierre <- netflix[,4]  
plot(cbind(amazon_cierre, netflix_cierre))
```



Los gráficos de series de tiempo de Netflix y Amazon no muestran estacionariedad, hay una disminución general en los precios desde 2021.

```
forecast::Acf(cbind(amazon_cierre, netflix_cierre))
```



Las funciones de autocorrelación (ACF) decaen muy lentamente, lo cual sugiere que estas series podrían tener una raíz unitaria.

Además, a partir de los gráficos de correlación cruzada, podemos ver que la correlación entre los precios de cierre de Amazon y Netflix es positiva.

Pero debemos tener cuidado. Esta correlación positiva podría deberse a tendencias estocásticas “ocultas” en estas variables. Debemos verificar si están cointegradas o no.

Parece que la varianza cambia con el tiempo; utilizaremos una **transformación logarítmica**.

### 3.3 Pruebas de Raíz Unitaria

Realizamos pruebas de raíz unitaria para los precios de cierre de NFLX y AMZN. ¿Son estacionarias? ¿Están ambas integradas del mismo orden?

```
tseries::adf.test(log(amazon_cierre), alternative = "stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data: log(amazon_cierre)
## Dickey-Fuller = -2.3473, Lag order = 7, p-value = 0.4303
## alternative hypothesis: stationary
```

```
tseries::adf.test(log(netflix_cierre), alternative = "stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data: log(netflix_cierre)
## Dickey-Fuller = -0.45515, Lag order = 7, p-value = 0.9836
## alternative hypothesis: stationary
```

```
tseries::adf.test(diff(log(amazon_cierre)), alternative = "stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data: diff(log(amazon_cierre))
## Dickey-Fuller = -5.9655, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

```
tseries::adf.test(diff(log(netflix_cierre)), alternative = "stationary")
```

```
##
## Augmented Dickey-Fuller Test
##
## data: diff(log(netflix_cierre))
## Dickey-Fuller = -6.6909, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

Para ambas series logarítmicas de los precios de cierre de Netflix y Amazon, el *p-value* es mucho mayor al 5%, por lo que no rechazamos la hipótesis nula de no estacionariedad.

En cambio, para sus primeras diferencias, ambos *p-values* son menores al 5%, y rechazamos la hipótesis nula de no estacionariedad (y por lo tanto otra raíz unitaria).

Por lo tanto, tanto los logaritmos de los precios de cierre de Amazon como de Netflix tienen una raíz unitaria y son  $I(1)$ . Podemos proceder al siguiente paso del análisis.

### 3.4 Pruebas de Cointegración

Realizar la prueba de Engle-Granger para cointegración

```
coint_reg <- lm(log(netflix_cierre) ~ log(amazon_cierre))
coef(summary(coint_reg))
```

```
##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept)    -6.779378  0.39766372 -17.04802 5.705160e-48
## log(amazon_cierre)  2.550723  0.07833995  32.55967 2.315878e-108
```

```
coint_res <- ts(coint_reg$res)
pp.test(coint_res)
```

```
##
## Phillips-Perron Unit Root Test
##
## data: coint_res
## Dickey-Fuller Z(alpha) = -22.031, Truncation lag parameter = 5, p-value
## = 0.04491
## alternative hypothesis: stationary
```

Para probar la cointegración usando la prueba de Engle-Granger, primero realizamos una regresión del logaritmo del precio de cierre de Netflix sobre el logaritmo del precio de cierre de Amazon, y luego aplicamos una prueba de raíz unitaria sobre los residuos.

Con base en la prueba de raíz unitaria de Phillips-Perron, el *p-value* es menor al 5%, por lo que rechazamos la hipótesis nula de raíz unitaria en los residuos. Esto indica que la prueba de Engle-Granger proporciona evidencia de que los logaritmos de los precios de cierre de Netflix y Amazon están cointegrados, ya que los residuos son estacionarios.

```
po.test(cbind(log(netflix_cierre), log(amazon_cierre)))
```

```
##
## Phillips-Ouliaris Cointegration Test
##
## data: cbind(log(netflix_cierre), log(amazon_cierre))
## Phillips-Ouliaris demeaned = -17.893, Truncation lag parameter = 3,
## p-value = 0.0852
```

También podemos utilizar una prueba de cointegración más robusta, como la prueba de cointegración de Phillips-Ouliaris. Según esta prueba, no se rechaza la hipótesis nula de que no existe cointegración al 5% de significancia, pero sí se rechaza al 10%.

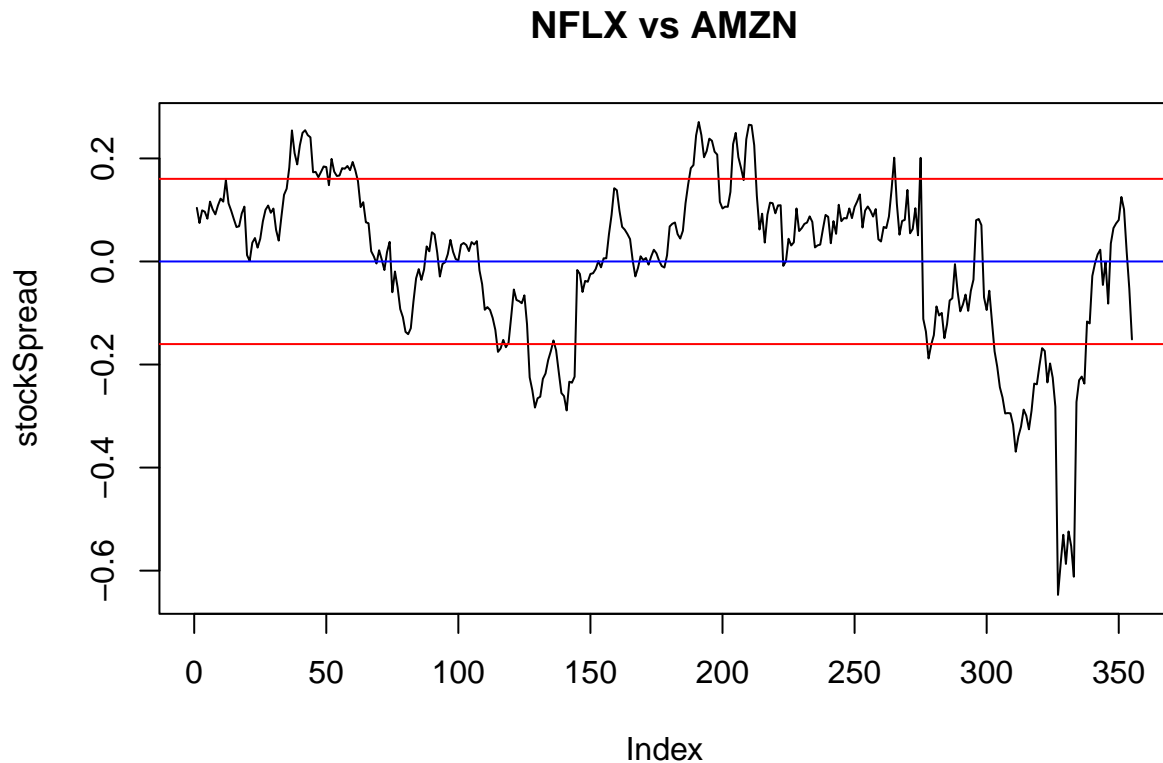
### 3.5 Estrategia

Graficar el spread (residuo) y discutir una posible estrategia de trading.

```
stockSpread<- coint_reg$residuals
plot(stockSpread, main = "NFLX vs AMZN", type = "l")
sd(stockSpread)
```

```
## [1] 0.1603209
```

```
abline(a=mean(stockSpread), b=0, col= "blue")
abline(a=sd(stockSpread), b=0, col= "red")
abline(a=-sd(stockSpread), b=0, col = "red")
```



Al analizar el spread de los residuos, podemos definir señales de trading para abrir y cerrar posiciones. Podemos usar las desviaciones estándar del spread como umbrales.

Por ejemplo, consideremos que nuestras señales de trading se basan en una desviación estándar de los residuos, es decir, 0.1603 y -0.1603, respectivamente.

Entonces, nuestra estrategia sería la siguiente:

- Cuando el spread está por encima de 0.1603, esto significa que Netflix está sobrevaluado en comparación con Amazon, por lo tanto, vendemos NFLX y compramos AMZN.
- Cuando el spread está por debajo de -0.1603, esto indica que Netflix está subvaluado respecto a Amazon, por lo tanto, compramos NFLX y vendemos AMZN.

Y siempre que el spread vuelva a converger hacia 0, cerramos la posición.