

Bản quyền dịch thuật và diễn giải thuộc về justinnnguyen92© - telegram

1. Mean Reversion (Hồi quy về trung bình)

Định nghĩa: *Mean reversion* là xu hướng của chuỗi thời gian quay về mức trung bình dài hạn theo thời gian.

Ứng dụng:

- Chiến lược giao dịch *mean reversion* dựa vào việc mua khi giá thấp hơn mức trung bình và bán khi giá cao hơn mức trung bình.

Thách thức:

- Việc tìm một tài sản riêng lẻ có tính hồi quy về trung bình với hành vi có thể dự đoán được rất khó.
- Tuy nhiên, dễ dàng hơn khi tìm cặp tài sản có chung tính hồi quy về trung bình.

2. Cointegration (Đồng tích hợp)

Định nghĩa: *Cointegration* là tính chất trong đó hai hoặc nhiều tài sản, dù không hồi quy về trung bình riêng lẻ, lại có mối quan hệ hồi quy về trung bình khi xét sự chênh lệch (*spread*) giữa chúng."

Ví dụ phổ biến: "Một người đàn ông say rượu (*random walk*) đi lang thang với một con chó (*another random walk*).

- Tuy cả hai đi theo những hướng không thể đoán trước (*non-stationary*), nhưng khoảng cách giữa người và chó vẫn ổn định trong dài hạn (***distance between them is mean reverting and stationary***) → ***Cointegration.***"



Ý nghĩa thực tế:

Mặc dù rất khó để dự đoán chuyển động của từng tài sản một cách độc lập (vì chúng có thể là chuỗi không dừng (*non-stationary series*)), nhưng có thể dự đoán mối quan hệ tương đối giữa chúng (như *spread* giữa hai

tài sản) nếu có đồng tích hợp.

- Trong các hệ thống kinh tế, *cointegration* phản ánh sự tồn tại của trạng thái cân bằng dài hạn giữa các thành phần, nơi hệ thống có xu hướng hội tụ theo thời gian.

Về mặt toán học: Một chuỗi thời gian đa biến y_1, y_2, y_3, \dots được gọi là đồng tích hợp (*cointegration*) nếu có một tổ hợp tuyến tính nào đó của các chuỗi có thể trở thành chuỗi dừng (*stationary*). Ví dụ, nếu y_t là chuỗi không dừng, nhưng tổ hợp tuyến tính $w^T y_t$ trở thành chuỗi dừng, thì ta nói rằng chuỗi này có tính đồng tích hợp.

- Cointegration* có thể được hiểu là một phiên bản nâng cao của việc lấy sai phân (*differencing*) chuỗi để đạt được tính dừng.
- Thay vì phải lấy sai phân, ta tìm một tổ hợp tuyến tính phù hợp giữa các chuỗi dữ liệu để tạo thành một chuỗi dừng.

Một cách phổ biến để mô hình hóa đồng tích hợp của hai chuỗi thời gian là:

$$y_{1t} = \gamma x_t + w_{1t} y_{2t} = x_t + w_{2t}$$

Trong đó:

- x_t là một chuỗi bước ngẫu nhiên "random walk" với phương trình: $x_t = x_{t-1} + w_t$
- w_{1t}, w_{2t}, w_t là các sai số ngẫu nhiên độc lập, có phương sai $\sigma_1^2, \sigma_2^2, \sigma^2$.
- Hệ số γ quyết định mối quan hệ đồng tích hợp giữa hai chuỗi dữ liệu. Tuy mỗi chuỗi riêng lẻ y_{1t} và y_{2t} đều không dừng, nhưng do cả hai chia sẻ một xu hướng chung x_t , sự kết hợp tuyến tính giữa chúng có thể loại bỏ xu hướng này.

Chuỗi "spread" giữa y_{1t} và y_{2t} được biểu diễn là:

$$z_t = y_{1t} - \gamma y_{2t} = w_{1t} - \gamma w_{2t}$$

Chuỗi z_t là một chuỗi dừng và có tính hồi quy về trung bình (*mean-reversion*).

Giải thích chi tiết

1. Đồng tích hợp (*Cointegration*) là gì?

- Mặc dù từng chuỗi thời gian riêng lẻ y_{1t} và y_{2t} có thể không dừng (*random walk*), sự kết hợp tuyến tính giữa chúng có thể trở thành một chuỗi có tính dừng.
- Trong bối cảnh *pairs trading*, *cointegration* được sử dụng để tìm kiếm cặp tài sản có sự liên kết dài hạn nhưng tạm thời phân kỳ, cho phép khai thác lợi nhuận khi chúng quay trở lại trạng thái cân bằng.

2. Ví dụ minh họa:

- Giả sử chuỗi x_t đại diện cho "xu hướng chung" của hai tài sản.
- Chuỗi y_{1t} và y_{2t} là giá của hai tài sản có mối quan hệ với xu hướng chung x_t , cộng thêm thành phần nhiễu w_{1t} và w_{2t} .
- Spread $z_t = y_{1t} - \gamma y_{2t}$ loại bỏ xu hướng x_t và chỉ còn lại phần nhiễu w_{1t} và w_{2t} .

3. Ý nghĩa trong giao dịch:

- Nếu chuỗi *spread* z_t là chuỗi dừng, điều này cho thấy sự chênh lệch giữa hai tài sản có xu hướng hồi quy về mức trung bình.
- Khi spread lớn hơn mức trung bình, có thể thực hiện chiến lược **SHORT** tài sản tăng giá và **LONG** tài sản giảm giá, kỳ vọng rằng giá sẽ hồi tụ về mức cân bằng.

Tóm tắt:

Cointegration là công cụ quan trọng trong việc tìm kiếm cặp tài sản có quan hệ dài hạn. Dù từng chuỗi dữ liệu riêng lẻ có thể là *random walk* (không dự đoán được), nhưng sự kết hợp giữa chúng có thể tạo ra một chuỗi dừng, mang lại cơ hội thực hiện chiến lược *pairs trading* dựa trên tính hồi quy về trung bình của *spread*.

Ví dụ code:

Tính *spread* chuỗi *time-series* của EURUSD với giá đóng cửa *close* và đoạn chuỗi *time-series* của GBPUSD với giá đóng cửa *close*, kiểm tra tính dừng của chuỗi *spread* - kết luận *Cointegration* giữa 2 chuỗi

```
#
=====
# VÍ DỤ COINTEGRATION GIỮA HAI CHUỖI TIME SERIES
# cho tôi ví dụ đoạn code tính spread chuỗi time series của EURUSD với giá đóng
# của [close] và đoạn chuỗi time series của GBPUSD
# với giá đóng cửa [close], kiểm tra tính dừng của chuỗi spread - kết luận
# Cointegration giữa 2 chuỗi
#
=====
# Import các thư viện cần thiết
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import coint, adfuller

# Tạo chuỗi thời gian giả lập giá đóng cửa của EUR/USD và GBP/USD
np.random.seed(42)
T = 200 # số lượng điểm dữ liệu
initial_price_eur = 1.1 # Giá ban đầu EUR/USD
initial_price_gbp = 1.3 # Giá ban đầu GBP/USD

returns_eur = np.random.normal(0, 0.001, T) # Lợi nhuận ngẫu nhiên EUR/USD
returns_gbp = np.random.normal(0, 0.001, T) # Lợi nhuận ngẫu nhiên GBP/USD

prices_eur = initial_price_eur * np.exp(np.cumsum(returns_eur)) # Chuỗi giá EUR/USD
prices_gbp = initial_price_gbp * np.exp(np.cumsum(returns_gbp)) # Chuỗi giá GBP/USD

# Tạo DataFrame chứa giá đóng cửa
df = pd.DataFrame({
    'EURUSD_Close': prices_eur,
    'GBPUSD_Close': prices_gbp
```

```

}, index=pd.date_range(start='2022-01-01', periods=T))

# Vẽ biểu đồ giá EUR/USD và GBP/USD
plt.figure(figsize=(12, 5))
plt.plot(df.index, df['EURUSD_Close'], label='EUR/USD Close', color='blue')
plt.plot(df.index, df['GBPUSD_Close'], label='GBP/USD Close', color='orange')
plt.title('Giá đóng cửa EUR/USD và GBP/USD')
plt.xlabel('Time')
plt.ylabel('Close Price')
plt.legend()
plt.show()

# Tính toán chuỗi spread: spread = EUR/USD - gamma * GBP/USD
gamma = np.polyfit(df['GBPUSD_Close'], df['EURUSD_Close'], 1)[0] # Tìm gamma qua
hồi quy tuyến tính
df['Spread'] = df['EURUSD_Close'] - gamma * df['GBPUSD_Close']

# Kiểm tra tính dừng của chuỗi spread bằng Augmented Dickey-Fuller (ADF) Test
adf_result = adfuller(df['Spread'].dropna())
p_value = adf_result[1]

# Vẽ biểu đồ chuỗi spread
plt.figure(figsize=(12, 5))
plt.plot(df.index, df['Spread'], label='Spread (EUR/USD - gamma * GBP/USD)',
color='purple')
plt.axhline(df['Spread'].mean(), color='red', linestyle='--', label='Mean
(Spread)')
plt.title('Spread giữa EUR/USD và GBP/USD')
plt.xlabel('Time')
plt.ylabel('Spread')
plt.legend()
plt.show()

# Kiểm tra Cointegration bằng kiểm định Engle-Granger
coint_score, coint_p_value, _ = coint(df['EURUSD_Close'], df['GBPUSD_Close'])

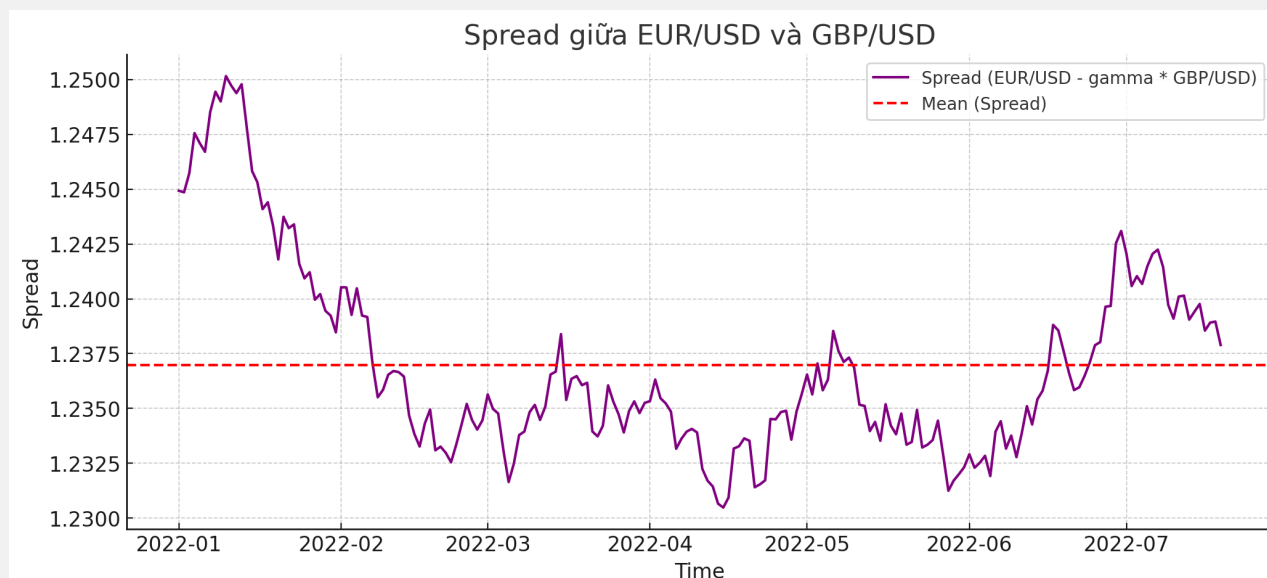
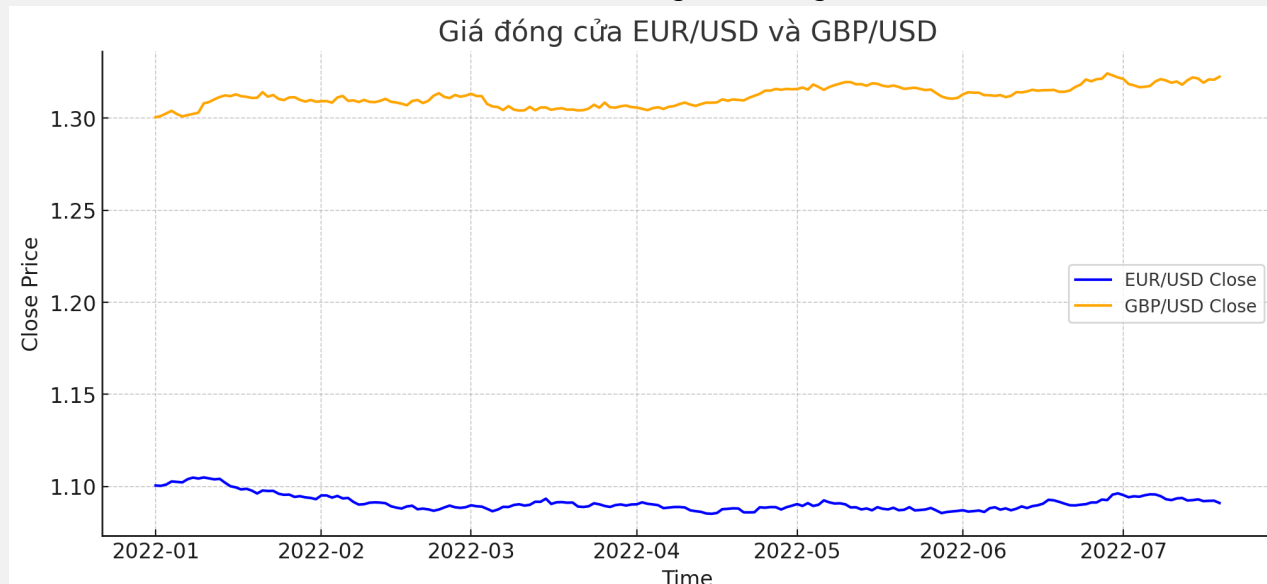
# Kết luận
conclusion = "Có Cointegration" if coint_p_value < 0.05 else "Không có
Cointegration"

p_value, coint_p_value, conclusion

```

Result

(0.21767844943807863, 0.43847298625304515, 'Không có Cointegration')



Kết quả kiểm tra Cointegration giữa EUR/USD và GBP/USD

1. Biểu đồ:

- **Biểu đồ 1 (giá đóng cửa):** Hiển thị diễn biến giá đóng cửa của EUR/USD (xanh dương) và GBP/USD (cam).
- **Biểu đồ 2 (spread):** Hiển thị chuỗi spread giữa EUR/USD và GBP/USD. Đường màu đỏ là mức trung bình của spread.

2. Kết quả kiểm tra:

- **ADF Test cho chuỗi spread:**
 - **P-value:** 0.2177 (lớn hơn 0.05) → Không bác bỏ giả thuyết H_0 → Chuỗi spread không có tính dừng.
- **Kiểm định Engle-Granger Cointegration:**

- **P-value:** 0.4385 (lớn hơn 0.05) → Không có bằng chứng cho thấy có mối quan hệ đồng tích hợp giữa hai chuỗi.

3. Kết luận:

Với dữ liệu giả lập này, **không có Cointegration** giữa **EUR/USD** và **GBP/USD**, cho thấy rằng không tồn tại mối liên hệ dài hạn ổn định giữa hai chuỗi dữ liệu. Vì vậy, chiến lược *pairs trading* có thể không hiệu quả trong trường hợp này.