Abstract

Pair trading is a classic approach to algorithmically trading financial assets such as equities and commodities. Traditional pair trading uses only 2 non-stationary time series to form cointegrated pair with stationarity, where mean-reversion method can be applied. This study extends the pair trading strategy to multi-dimensional cointegrated set of financial time series and test the effectiveness of cointegration on China’s stock market, particularly, CSI 300 Index component stocks.

Introduction

Linear regression is a commonly used method to find the relationship between several time series processes. However, Granger and Newbold [1] argued that linear regression analysis might lead to spurious regression, a situation in which the regression relationship appears to be statistically significant when variables are unrelated. Many economic time series exhibit random walk and other non-stationary behavior, leading to a higher possibility of finding spurious correlation under regression analysis. On the contrast, cointegration theory provides an alternative approach to study non-stationary time series. When a time series is not stationary, there is no tendency for its level to return to a constant mean over time; moreover, the volatility of the process is expected to grow boundlessly in the long run and any prediction based on the historical observations is not possible. Cointegration theory allows the identification of integration among time series that have similar dynamics in the long run and the estimation of their relationships. In this study, we will be using Johansen’s approach to exam cointegration relationship among multiple time series of stock prices. For cointegrated stock groups, we expect the stock with least 7-day return rate to rebound and thus, we buy in and hold the stock until its bounce.

Methodology

Integration: $y_t$ is $I(d)$ if its $(d-1)^{th}$ difference is $I(0)$. That is, $Δ^d y_t$ is stationary.

Augmented Dickey-Fuller Test: ADF test is a common way to assess the integration of a time series. Assume time series $y_t$ follows autoregressive process of order $p (AR(p))$:

$$Δy_t = C + ρ y_{t-1} + \sum_{i=2}^{p} \phi_i Δy_{t-i} + ε_t$$

where:

$$\begin{cases} p = \sum_{i=1}^{n} \pi_i - 1 \\ \alpha = \sum_{i=1}^{n} \pi_i \end{cases}$$

The null and alternative hypothesis are:

$$H_0: \rho = 0 \quad H_1: \rho < 0$$

ADF test statistic is:

$$F_p = \frac{\hat{\rho}}{SE(\hat{\rho})}$$

Cointegration: An $m$ 1 vector time series $y_t$ is said to be cointegrated of order $(d, b)$, $CI(d, b)$ where $0 < b \leq d$, if each of its component series $y_{ij}$ is $I(d)$ but some linear combination $y^0$ is $I(d-b)$ for some constant vector $a \neq 0$.

Johansen’s Cointegration Test: Estimate $y_t$ by vector autoregressive (VAR) process with order $p$:

$$y_t = \Phi y_{t-1} + \cdots + \Phi_0 y_{t-p} + u_t$$

The error correction model is:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=0}^{k-1} \beta_i \Delta y_{t-i} + u_t$$

where:

$$\sum_{i=0}^{k-1} \beta_i = 1$$

Let the rank of $\Pi$ be $r$:

if $r = 0$: there is no cointegration relationships

if $0 < r < k$: there are $r$ cointegration relationships

Gregory-Hansen Cointegration Test with Structural Breaks: structural breaks or regime shifts are common in financial time series due to market emergencies or major changes in policies and international relationships. Normal cointegration test assumes the cointegration to be both stationary in long-run mean and volatility. Gregory-Hansen’s Cointegration Test allows one unknown structural breaks in:

(i) Level shift;

(ii) Level shift with trend;

(iii) Regime shift (both level and slope coefficient can change)

Portfolio Pool: 300 stocks are categorized into 62 industrial classes and 37 of them have more than 2 stocks within each class. Among those 37 groups, 18 groups have passed either Johansen’s Cointegration Test or Gregory-Hansen’s Cointegration Test.

Johansen’s Cointegration Test: the cointegration test result of industry group “Insurance”

Gregory-Hansen’s Cointegration Test: the test result of industry group “Insurance”

Investment Result: The annual return of investment for year 2021 is 43.4%, compared to 11.27% of investment return for non-cointegrated group and -7% for CSI 300 Index.

Discussion

- The cointegration filtration exhibits outstanding performance for gaining excess return compared to benchmark (CSI 300 Index)
- Overrejection of Johansen’s Cointegration Test when dimensionality is high: it has been pointed out that the Johansen’s cointegration test tends to overreject the null hypothesis (no cointegration relationship) when either the cross-sectional dimension (N) or time duration dimension (T) is high. Our study result generally supports this, but further research is expected.

<table>
<thead>
<tr>
<th>N</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equally Weighted Return (no-cointegrated class)</td>
<td>16.69%</td>
<td>20.86%</td>
<td>17.92%</td>
<td>5.90%</td>
<td>23.42%</td>
<td>5.61%</td>
<td>-</td>
<td>12.20%</td>
</tr>
<tr>
<td>Equally Weighted Return (cointegrated class)</td>
<td>-17.89%</td>
<td>35.87%</td>
<td>-15.21%</td>
<td>25.31%</td>
<td>50.84%</td>
<td>79.55%</td>
<td>4.63%</td>
<td>2.01%</td>
</tr>
</tbody>
</table>