

國立政治大學金融學系研究所

碩士學位論文

配對交易與機器學習在台灣股票市場之應用

Applications of Pairs Trading and Machine Learning in Taiwan

Stock Market

指導教授：林士貴 博士

蔡瑞煌 博士

研究生：徐瑀暄 撰

中華民國 一百零七 年 七 月

## 謝辭

對於這篇論文，首先我要感謝我的指導老師林士貴教授，老師花了非常多的時間與我討論，對於我的研究提供許多建議與研究方向，並一次又一次幫我挑出論文不足之處，真的非常感謝老師的指導和付出。接著我要感謝我的共同指導老師蔡瑞煌教授，利用課餘時間和我們討論，讓我們能在機器學習這個領域有更好的學習。接下來我要感謝元大證券計量交易部的曾盟雅經理及郭憲聰學長，兩位在實習期間不斷的根據我的研究給我建議與方向，我的研究也因此更加豐富。

接著我要感謝我的同門陳勃文和洪子軒，除了在論文方面我們互相討論及督促彼此的進度，也一同討論學習機器學習這個新的領域，當然也要感謝班上其他同學，像賴家蔚幫我解決了很多關於程式上的技術問題，還有很多同學，這裡就不一一列舉，你們讓我在碩士班兩年的生活更加的豐富及充實。

另外，還要感謝我的女友鄭雅薰，在碩二這忙碌一年陪伴著我，和我一同努力並互相加油打氣。最後，我要感謝我的父母，謝謝你們在生活上給予我充足的資源，讓我無後顧之憂地完成我的學業，在我面對困難時也適時地給我意見及鼓勵，並無條件支持我的任何決定，讓我從工學院轉換跑道學習金融，沒有你們就沒有今日的我，未來踏入職場我會更加努力，不會讓你們失望的！

徐瑤暄 謹至于

國立政治大學 金融研究所

民國一〇七年七月

# 配對交易及機器學習在台灣股票市場之應用

學生:徐瑀暄

指導教授:林士貴 博士

蔡瑞煌 博士

國立政治大學金融系

## 摘要

本研究根據 Vidyamurthy (2004)以及後續相關文獻所提出的統計套利配對交易方法對台灣股票市場進行實證研究。本文使用的模型為 Engle and Granger (1987)提出的二階段共整合檢定。我們利用上述模型檢定台灣股票，找出具共整合性質之股票配對，利用技術指標—布林通道找出價格異常的時間點進行交易，建構配對交易投資組合；本研究進一步將類神經網路模型加入，用於預測共整合殘差走勢，建構類神經網路結合布林通道之配對交易策略並建構投資組合。實證結果顯示和 Avellaneda and Lee (2010)結果相同，市場上確實存在市場中立性的報酬，且兩個策略的投資組合皆有優於大盤的績效和穩健性；此外類神經網路確實有幫助我們減少進場次數提高勝率，並且使投資組合的最大虧損下降，但也因此降低了投資組合的總報酬。

**關鍵詞：**共整合、配對交易、布林通道、類神經網路、投資組合

# Application of Pairs Trading and Machine Learning in Taiwan Stock Market

Student: Yu-Hsuan, Hsu

Advisor : Dr. Shih-Kuei, Lin

Dr. Rua-Huan, Tsai

Department of Money and Banking, National Chengchi University

## Abstract

This paper used the statistic arbitrage pairs trading method according to Vidyamurthy (2004) and other papers based on this book. This paper followed papers to conduct empirical research on Taiwan stock market. The models used in this paper is two-steps cointegration test that proposed by Engle and Granger (1987). We tested Taiwan stocks through the above models to test cointegration, and find the investable pairs. After finding out investable pairs, we used Bollinger Band to find out abnormal stock price to trade. Then we constructed the portfolio to study its performance. This study further adds the neural network model to predict cointegral residual and constructs a strategy with Bollinger Band and neural network model. The result shows that the strategy helping us find market neutral return, which is the same as the result of Avellaneda and Lee (2010). Furthermore, our portfolio is also better than investing in benchmark. Neural network model truly helps us reduce trading frequency and decrease drawdown, but it also decreases return at the same time.

***Keywords : Cointegration 、 Pairs Trading 、 Bollinger Band 、 Neural Network 、 Portfolio***

## 目次

第一章	緒論.....	1
第一節	研究動機.....	1
第二節	研究目的.....	3
第三節	研究架構.....	4
第二章	文獻探討.....	5
第一節	共整合配對交易.....	5
第二節	機器學習建構交易策略.....	8
第三章	研究方法.....	10
第一節	投資組合理論.....	10
第二節	單根檢定.....	12
第三節	共整合檢定.....	15
第四節	布林通道技術指標.....	17
第五節	倒傳遞類神經網路.....	18
第六節	配對交易策略.....	23
第七節	配對交易投資組合的建構方法.....	26
第八節	交易策略績效評估指標.....	27
第四章	實證分析.....	28
第一節	實證資料與研究期間.....	28
第二節	台灣股票配對組合篩選結果.....	30
第三節	布林通道配對交易策略實證結果.....	34
第四節	機器學習配對交易策略實證結果.....	41
第五章	結論.....	50
參考文獻	.....	52
附錄	.....	55

## 表次

表 3-1 到傳遞類神經網路輸入變數.....	22
表 4-1 台灣證券交易市場簡介.....	29
表 4-2 宏亞與天仁布林通道交易策略績效參數.....	36
表 4-3 布林通道配對交易策略績效參數之敘述統計表.....	37
表 4-4 布林通道配對交易投資組合績效參數.....	39
表 4-5 隱藏層節點改變之穩健性分析表.....	42
表 4-6 宏亞與天仁機器學習交易策略績效參數.....	44
表 4-7 機器學習配對交易策略績效參數之敘述統計表.....	45
表 4-8 機器學習配對交易投資組合績效參數.....	47



## 圖次

圖 3-1 類神經網路模型基本架構圖 .....	18
圖 3-2 處理單元示意圖 .....	19
圖 3-3 雙曲線正切函數轉換示意圖 .....	20
圖 3-4 布林通道配對交易策略進出場方式示意圖 .....	24
圖 4-1 加高與希華兩股票之收盤價與共整合殘差圖 .....	32
圖 4-2 配對交易組合數目之產業分配統計圖 .....	33
圖 4-3 宏亞與天仁收盤價與共整合殘差圖 .....	36
圖 4-4 宏亞與天仁布林通道交易策略績效及進場訊號圖 .....	36
圖 4-5 布林通道配對交易投資組合績效圖與虧損圖 .....	38
圖 4-6 布林通道配對交易投資組合與大盤走勢圖 .....	39
圖 4-7 布林通道配對交易投資組合日損益分析圖 .....	40
圖 4-8 布林通道配對交易投資組合月損益分析圖 .....	40
圖 4-9 布林通道配對交易投資組合年損益分析圖 .....	40
圖 4-10 類神經網路模型訓練過程之誤差及預測能力圖 .....	41
圖 4-11 宏亞與天仁機器學習交易策略績效及進場訊號圖 .....	43
圖 4-12 機器學習配對交易投資組合績效圖與虧損圖 .....	46
圖 4-13 機器學習配對交易投資組合與大盤走勢圖 .....	47
圖 4-14 機器學習配對交易投資組合日損益分析圖 .....	49
圖 4-15 機器學習配對交易投資組合月損益分析圖 .....	49
圖 4-16 機器學習配對交易投資組合年損益分析圖 .....	49



# 第一章 緒論

## 第一節 研究動機

近年來由於全球金融市場開放，加上 2008 年金融海嘯以來世界各國的貨幣寬鬆政策，又由於國際金融局勢日趨複雜，政治與經濟情勢變動對資產價格產生巨大影響的可能性漸增，諸如英國脫歐、川普當選美國總統等……，皆造成全球市場商品價格大幅波動，傳統利用財報資料、歷史交易資訊來判斷長期持有投資策略，其績效穩健性開始受市場風險挑戰。

為規避此類系統性風險，最常使用的方法便是市場中立性投資策略(Market Neutral Strategy)，其概念是利用多個關聯性很高的資產，採用其價格間的變動關係，同時建立多空部位投資組合來消除大部份市場風險，降低投資組合績效與市場走勢相關性獲得穩定報酬，而在多種市場中立性投資策略中，配對交易(Pairs Trading)便是其中一種常用且建構困難度較低的方式。

配對交易(Pairs Trading)是一種利用二、三個資產配置，達到市場中性概念的投資策略，最早於 1980 年代中期華爾街著名投行 Morgan Stanley 的數量交易員 Nunzio Tartaglia 成立的數量分析團隊提出並使用，當時獲得優異績效，至今仍為許多投資機構投資人所使用，亦為許多研究的研究對象。故本研究期望能在這複雜的國際局勢中，藉由配對交易策略在台灣股票市場中的可行性分析及獲利能力分析，探討台灣股票市場是否存在建構配對交易這種市場中立性交易策略的機會，期望能從中找到能有效分散市場風險並具有穩定獲利能力的投資策略，供從事相關交易及研究工作的工作參與者參考。



此外有鑑於機器學習是近年來快速竄起的資訊處理技術，具有良好的預測能力，目前已廣泛的被運用在各個不同的領域，改變各行各業的運作模式。因此本研究嘗試運用機器學習中最廣泛應用且預測能力良好的類神經網路模型，作為預測工具，結合傳統的配對交易概念，建構機器學習配對交易策略，期望除了傳統配對交易策略，能找到新的策略建構方法，並運用於金融市場中，讓我們能夠在這複雜的國際局勢中，創造更多的獲利機會。



## 第二節 研究目的

Vidyamurthy (2004)定義配對交易為兩種類型：一類是基於統計套利的配對交易，一類是基於風險套利的配對交易，本研究便屬於統計套利，依據 Engle and Granger (1987)二階段共整合檢定法，檢定高度相關的價格時間序列間是否存在長期均衡關係，若確定其有長期的均衡關係，一旦兩者之間出現了異常的價格走勢，就可能產生套利的機會。當我們利用技術分析方法判定其價格異常，便會建構一個簡單的交易策略，買進股價相對被低估的資產，賣出股價相對被高估的資產。當未來兩者之間的價格得到修正，便可以進行相反的平倉操作來獲取利潤。儘管配對交易策略非常簡單，但卻被廣泛的應用，主要原因是：首先，配對交易的收益與市場互相獨立，即市場中立性，也就是說它與市場的上漲或者下跌無關；其次，其收益的波動性相對於一般交易較小；第三，其收益相對於一般交易穩定。

本研究使用配對交易策略概念，應用於台灣股票市場，對台灣上市上櫃股票進行檢定，找出具共整合關係之股票配對，並利用技術指標判斷股價異常時機，建構配對交易策略，探討配對交易策略的效果，最後利用篩選結果建構以配對交易策略為基礎的投資組合。此外加入機器學習進行預測期望除了傳統的技术方法，能利用機器學習優秀的預測能力，改良配對交易策略，建構出機器學習結合技術指標的配對交易策略投資組合，進行比較。希望透過本研究得到的實驗結果可以讓對沖基金做為參考，提供新的市場中立性交易策略，也希望透過這樣的方式建構出可以有效獲利的交易策略。

### 第三節 研究架構

本文主要架構如下：第二章為統計套利配對交易及機器學習模型應用於交易策略的文獻回顧，第三章為本研究應用之配對交易建構理論、共整合檢定模型、技術指標及機器學習模型介紹，第四章為本研究的實證結果，分為共整合檢定結果、配對交易策略實證研究、技術指標配對交易投資組合實證研究及機器學習配對交易投資組合實證研究等四部分，第五章為本研究之結論。



## 第二章 文獻探討

### 第一節 共整合配對交易

配對交易是廣泛被使用的交易策略，在學術界是相當重要的研究議題，相關的文獻內容也很龐大，本研究僅探討與本研究使用的配對交易方法相關之重要論文。Vidyamurthy (2004)認為在執行配對交易策略時，應先就股票基本面及歷史股價資料挑選出潛在具有共整合關係的配對，並以線性組合建構配對交易部位，Vidyamurthy (2004)中選擇使用的 Engle and Granger (1987)二階段共整合檢定法，檢定兩個高度相關的價格時間序列間是否存在長期均衡關係，二階段共整合檢定法首先以單根檢定確認配對的成分資產價格為  $I(1)$  序列(一階整合序列)，後將其中一個商品價格序列為自變數，另一商品作為應變數建立線性迴歸關係，在對殘差進行單根檢定，若結果為顯著，表示殘差為定態，則兩價格序列間具有共整合關係，可嘗試以其為標的建立配對交易策略。在確定兩資產有共整合性質後，便需判斷交易的進出場訊號，在交易訊號方面，Vidyamurthy (2004)則建議以無母數方法估算出能使與其報酬最大化的門檻值，令殘差序列偏差超過此最佳門檻值為啟動交易條件。

Avellaneda and Lee (2010)同樣使用共整合的概念，認為利用多個風險因子，能夠有效解釋資產的報酬，並使得剩下無法解釋的市場中立報酬為穩定的，因此文中利用主成份分析法(Principal Component Analysis)找出風險報酬因子，再利用因子篩選股票並建立投資組合，檢視因子是否能有效解釋報酬，此研究發現要確保所有因子對報酬的解釋能力，除了市場因子，必須加入產業因子進行解釋，且總共約需要 20 個左右的因子才能較好的解釋報酬，過於複雜且部分的因子解釋報酬能力極小；Avellaneda and Lee (2010)也提供額外的方法，利用產業 ETF 作為

預測股票報酬的因子，發現產業 ETF 同樣具有解釋報酬的能力，接著利用個股及其產業之 ETF 建構配對組合，利用 ETF 取代市場風險因子抵銷市場報酬，藉此賺取市場中立報酬，其方法不利用共整合檢定法先行篩選好的配對組合標的，而是直接選股跟找出其產業之 ETF 進行配對交易，使殘差項為穩定序列，再進行交易策略建構及回測；在建構交易策略訊號方面，則將剩餘殘差假設為 Ornstein-Uhlenbeck process 及 Auto Regression 模型，接著將共整合殘差 (Cointegral Residual) 進行標準化調整，並設定上門檻及下門檻，當標準化後的殘差超過設定的門檻，便認定價格出現異常，交易便會啟動，Avellaneda and Lee (2010) 同樣設定了交易出場的門檻值，當標準化後的殘差落回設定的門檻變平倉出場，本文以此方法建構完整的配對交易策略。

國內論文方面羅君昱 (2005) 以台灣五十成分股為標的進行配對交易，同樣選用 Engle and Granger (1987) 二階段共整合模型，進行兩檔股票共整合關係的檢定，將通過檢定的股票進行配對交易回測，交易策略方面參考 Gatev, Goetzmann and Rouwenhorst (1999)，對共整合殘差計算樣本時間內的標準差，並以殘差穿過兩倍標準差為進場訊號，殘差回歸一倍標準差內則進行平倉，以策略計算樣本時間內之績效，再與大盤報酬率進行比較發現，其結果與 Gatev, Goetzmann and Rouwenhorst (1999) 的論文結果相同；大盤表現佳時，配對交易的投資組合表現較差；當大盤表現差時，配對交易的投資組合表現較好，兩者報酬率呈現負相關。

陳岱佑 (2012) 非使用共整合法，而以技術指標中的布林通道 (Bollinger Bands) 來建構配對交易策略，以合約規模及兩資產價格間的迴歸係數來定義價差，標的資產為台指期貨、電子期貨、金融期貨及台灣 50ETF，啟動進場訊號的方式為當價差穿越布林通道上(下)緣並再次回到上(下)緣時才進場建倉，此策略設計是為了確保價差未來將收斂而非趨勢的開始，並以形成期的績效作為篩選標準，挑選

勝率高於六成，且獲利因子高於損益兩平因子的配對於交易期交易，發現形成期間高勝率以及正報酬率無法保證交易期的績效，顯示此策略的不穩定，仍有改良的空間。



## 第二節 機器學習建構交易策略

在股價預測及交易策略相關的研究中，有許多文章利用機器學習(Machine Learning)、類神經網路(Neural Network)模型，期望藉由更複雜的模型，建構更優良的交易策略及預測能力；配對交易中，除了利用計量方法研究統計套利配對交易策略，也有研究將機器學習、類神經網路模型加入。

Chen, Shih and Wu (2006)利用機器學習方法預測六個亞洲股票市場走勢，比較支援向量機(Support Vector Machine)模型和非線性複雜度更高的倒傳遞類神經網路(Back Propagation Artificial Neural Networks)模型，是否有優良的預測能力，研究中使用簡單的模型，輸入經過數學處理過的歷史價格差分資料，並預測未來五天後市場的價格，研究結果發現不論是 SVM 模型或 NN 模型皆有比傳統計量模型好的預測能力，但作者也表示結果預測並非完美，可以嘗試利用更複雜的模型進行預測。

Kara, Boyacioglu and Baykan (2011)中，作者同樣運用倒傳遞類神經網路(NN)及支援向量機(SVM)預測未來股價走勢，除了使用股票價格時間序列外，本篇使用了更多的技術指標作為模型的輸入變數，並以明天的股票走勢方向作為輸出變數，研究結果發現不論是 SVM 模型或 NN 模型皆比 ARIMA 計量模型有更好的預測能力，且研究中類神經網路模型有 75%左右的預測能力比 SVM 模型的 71% 更優良，但作者認為仍然有改進的空間，可嘗試改變輸入技術指標或加入總體經濟數據變數，或許能增加模型的預測能力。



Madhavaram (2013)中，參照了 Avellaneda, Marco and Lee (2010)中股票報酬率的分析方法，探討了風險因子的統計套利概念，利用主成份分析法篩選出因子，並統計套利理論建構配對交易組合，配對交易策略則利用假設殘差過程為 O-U Process 並進行標準化建立交易進出場訊號，檢視其每支配對交易策略的交易重要績效參數；接著 Madhavaram (2013)利用機器學習模型 SVM，將交易績效評斷參數作為 SVM 的分類標準，進一步找出更優良的配對組合建構投資組合，研究結果發現利用風險因子建構的配對交易組合結合 SVM 進一步篩選後投資組合相較於傳統只利用風險因子建構的投資組合，雖然報酬率沒有優良的改進，但成功的降低了投資組合的波動度，使得投資組合更加穩定。

機器學習在交易策略及預測價格走勢的研究中，有許多的應用方法，如上述介紹：時間序列的預測、配對組合的篩選等，本研究將使用具有較佳預測能力的類神經網路模型，參考 Chen, Shih and Wu (2006)及 Kara, Boyacioglu and Baykan (2011)輸入變數方法及模型架構，進行預測未來的時間序列走勢的實證研究，並利用預測結果結合技術指標進行交易策略的建構。

### 第三章 研究方法

#### 第一節 投資組合理論

Avellaneda and Lee (2010)將報酬劃分為兩種類：指數性(Indexers)和市場中立性。指數性的報酬通常與市場或者是特定產業有關係，是藉由調整市場或是特定產業中的股票權重，來得到指數性的報酬；市場中立性的報酬則是與市場無關聯。

我們可將 N 檔不同股票一個交易日的報酬率以 $\{R_i\}_{i=1}^N$ 表示，且利用 F 代表市場因子同一時間內的報酬率(例如：台灣加權指數的報酬率)， $\beta_i$ 為因子與股票之間的關聯性，我們可以將每檔股票的報酬表示成以下方式：

$$R_i = \beta_i F + \tilde{R}_i \quad (1)$$

利用簡單線性迴歸來分解股票報酬為與市場報酬相關的 $\beta_i F$ 及與市場無相關的 $\tilde{R}_i$ 兩部分；除此之外此式子可以延伸為多元因子模型的形式：

$$R_i = \sum_{j=1}^m \beta_{ij} F_j + \tilde{R}_i \quad (2)$$

上列式子便是利用 m 個市場因子建立而成的系統性投資組合與 $\tilde{R}_i$ 去解釋股票的報酬率。因此若要建構一個市場中立性的投資組合我們可以利用 N 檔不同的股票，給予每檔股票不同的購買權重 $\{Q_i\}_{i=1}^N$ ，進行投資組合的建構，得到下列的結果：

$$\bar{\beta} = \sum_{i=1}^N \beta_{ij} Q_i = 0, j = 1, 2, \dots, m. \quad (3)$$

係數 $\bar{\beta}_j$ 即為投資組合每個對應因子的係數，我們通過找出適當的權重，可以消除掉投資組合裡的每個市場因子報酬，建立一個真正有著市場中立性(Market neutrality)的投資組合，與所有的市場因子無相關性，而此投資組合的報酬可表示為：

$$\begin{aligned} \sum_{i=1}^N Q_i R_i &= \sum_{i=1}^N Q_i \left[ \sum_{j=1}^m \beta_{ij} F_j \right] + \sum_{i=1}^N Q_i \tilde{R}_i \\ &= \sum_{j=1}^m \left[ \sum_{i=1}^N \beta_{ij} Q_i \right] F_j + \sum_{i=1}^N Q_i \tilde{R}_i \\ &= \sum_{i=1}^N Q_i \tilde{R}_i \end{aligned} \quad (4)$$

由此我們可以知道，市場中立性的投資組合只受到每檔股票市場中立性報酬影響，Avellaneda and Lee (2010)研究須利用超過 20 個因子去分析股票的報酬才能找到真正的市場中立報酬，但當使用越多因子時也發現，剩餘的市場中立的報酬很小，獲利能力不足，因此文中試著利用 ETF 取代多因子，與個股建立配對交易，並利用均值回歸的概念，找到能夠獲利的交易方式，而非不斷的增加因子解釋報酬。

因此在建構配對交易策略時我們同樣能利用兩檔股票相關性很高，有多個共同因子的股票進行配對組合，找到一個適當股票持有比例便可以得到只含有市場中立報酬的配對交易組合。

## 第二節 單根檢定

陳旭昇 (2013)時間序列資料對外生衝擊(Impulse)的反應大致上可分為兩類：定態(Stationary)時間序列及非定態(Non-stationary)時間序列。一個外生衝擊對非定態時間序列造成的影響可能不會隨時間消散，使其在時間的變動過程中留下永久性的影響並不斷累積，導致其具有隨機趨勢，即序列的平均數、變異數及各期間的共變異數並不固定且隨時間變動；定態時間序列僅會對外生衝擊產生暫時性的反應，一段時間過後此衝擊造成的影響將逐漸消失，使這樣的時間序列隨時間經過仍能保持穩定。若一時間序列  $R_t$  滿足下列性質，則我們稱  $R_t$  為定態：

1.  $E(R_t) = \mu$ ，即期望值 $\mu$ 為固定常數
2.  $Cov(R_t, R_{t-l}) = \gamma_l$ ，不同期間的共變異數只與落後期數 $l$ 相關

上述定義為弱定態(Weak Stationary)，在大多數時間序列模型分析及迴歸分析中皆需要資料為弱定態，Granger and Newbold (1974)提出若以非定態變數進行迴歸分析將出現假性迴歸(Spurious Regression)的現象，即兩個不相關的序列變數出現極高的相關係數  $R^2$ ，因此在對時間序列資料進行分析前一般需要先檢定該序列是否為定態序列。

單根檢定(Unit Root Test)為檢定時間序列是否為定態的方法，若一時間序列具有單根則該序列有隨機趨勢，表示該序列為非定態序列，換而言之可藉由檢定一時間序列是否具有單根來確認其是否為定態序列。

本研究採用的單根檢定為 ADF 單根檢定，落後項設定為 0，即為 Dickey and Fuller (1979)提出的 DF 檢定(Dickey-Fuller Test)，其模型為：

$$\Delta y_t = \phi y_{t-1} + e_t \quad (5)$$

以上之模型檢定假設為：

$$H_0: \phi = 0 (\text{存在單根}); H_1: \phi < 0 (\text{不存在單根，即序列為定態})$$

其檢定統計量為：

$$DF = \frac{\hat{\phi} - 1}{\sqrt{\text{Var}(\hat{\phi})}} \quad (6)$$

如果拒絕了虛無假設，表示  $y_t$  沒有單根，為一定態時間序列；反之，該序列則為具單根現象之非定態序列。

此 DF 檢定中殘差項為白噪音(White Noise)，而實際上此檢定的殘差項經常具有序列相關性，因此 Said and Dickey(1984)提出修正後的 ADF 檢定(Augmented Dickey-Fuller Test)，於模型中加入差分項以控制殘差的序列相關問題。ADF 檢定考慮三個 AR(p)形式的模型，若  $y_t$  為時間序列資料，分別為：

1. 無截距項及時間趨勢項：

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + e_t \quad (7)$$

2. 含截距項：

$$\Delta y_t = \alpha_0 + \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + e_t \quad (8)$$

3. 含截距項及時間趨勢項：

$$\Delta y_t = \alpha_0 + \beta t + \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + e_t \quad (9)$$

其中 $\alpha_i$ 、 $\beta$ 為常數， $e_t$ 為殘差項， $p$ 為最落後期數，最落後期數須由模型決定。

而 ADF 檢定的虛無假設及對立假設分別為：

$$H_0: \phi = 0 (\text{存在單根}); H_1: \phi < 0 (\text{不存在單根，即序列為定態})$$

其檢定統計量為：

$$ADF = \frac{\hat{\phi} - 1}{\sqrt{\text{Var}(\hat{\phi})}} \quad (10)$$

在具有單根的虛無假設下，ADF 統計量之臨界分配並非常態，因此需要查詢臨界值分配表來得到臨界值，若檢定結果為顯著拒絕虛無假設，則  $y_t$  不具有單根，為一定態時間序列，反之若無法拒絕  $y_t$  具有單根的虛無假設則  $y_t$  為一非定態的時間序列。

許多研究會探討落後其數本與時間趨勢對單根檢定的影響，而研究為簡化配對篩選流程，以利於在大量的台灣上市上櫃股票中選取配對，僅選用無差分項  $\Delta y_{t-i}$  的 DF 檢定，且不考慮截距項  $\alpha_0$  及時間趨勢  $t$ ，以簡化單根檢定對後續配對交易之影響。

### 第三節 共整合檢定

Engle and Granger (1987)提出共整合的理論，是指兩個或多個時間序列存在一個長期穩定均衡的關聯性，若兩個以上非定態的時間序列變數經過線性組合後形成的新序列為定態，則稱他們之間具有共整合關係。具體來說此多個序列分別為非定態，但具有共同的趨勢，使其經線性組合後可將共同隨機趨勢消除，成為一個新的定態時間序列。

在檢定兩個序列是否具有共整合關係前，須先確定兩序列的整合階次(Order of Integration)相同，一時間序列的整合階次為該序列成為定態序列前須經過的最少差分(Difference)次數，舉例來說若時間序列  $r_t$  經過  $d$  次差分後始為定態，則  $r_t$  之整合階次為  $d$ ，一般表示成  $r_t \sim I(d)$ ，整合階次相同的兩序列才有存在共整合關係的可能，大部分的總體經濟變數或價格序列整合階次通常為 1，經過一階差分後即可成為定態序列。

Engle and Granger (1987)提出二階段共整合檢定法，用以檢定兩  $I(1)$  序列是否具有共整合關係，假設  $x_t$ 、 $y_t$  為一組非定態時間序列，兩者經一階差分後皆為定態，且假設此組序列配對僅存在一組共整合關係，首先估計兩者關係：

$$y_t = \beta_0 + \beta_1 x_t + e_t \quad (11)$$

得到係數  $\hat{\beta}_1$ 、 $\hat{\beta}_0$ ，接著對殘差項  $\hat{e}_t$  進行 ADF 單根檢定，Engle and Granger (1987)中建議使用 ADF 檢定模型中不含截距項及時間趨勢的模型估計較佳：

$$\Delta \hat{e}_t = \phi \hat{e}_{t-1} + \sum_{i=1}^p \alpha_i \Delta \hat{e}_{t-1} + \varepsilon_t \quad (12)$$



而此處 ADF 統計量的臨界值 Engle and Granger (1987)中建議使用 Phillips and Ouliaris (1990)推倒之漸進分配臨界值，而非一般 ADF 檢定之臨界值。

本研究選擇使用更簡單的模型，選用不含截距項及時間趨勢，且無落後項之 DF 檢定，同樣的若檢定結果無法拒絕 $H_0: \phi = 0$ ，表示殘差仍為非定態，兩序列無法藉由線性組合消除彼此之隨機趨勢，因此該組合不具共整合關係；若顯著拒絕虛無假設，則殘差為定態，兩序列間存在共整合關係，共整合係數估計值為 $\hat{\beta}_1$ 。

另外，Engle and Granger 二階段檢定法建立於兩個變數序列間只存在一個共整合關係的假設前提下，因此無法處理多個共整合關係的存在，若為三個(或以上)變數的組合則需使用 Jonansen and Juselius (1990)提出的 Johansen 共整合檢定法尋找共整合向量。本研究只單純考慮兩兩一組的配對方式，且期望以較有效率的方式從大量股票中篩選出配對組合，故選用 Engle and Granger 二階段共整合檢定法進行共整合檢定，作為本次台灣股票市場配對交易實證研究之共整合檢定方法。

## 第四節 布林通道技術指標

Bollinger (1992) 中提出了布林通道(Bollinger Bands)的概念，此通道為一個技術分析工具，主要以一條  $n$  日移動平均線為中心線，向上和向下各加減  $c$  倍標準差後，就構成了上界和下界，而由此上下界所圍成的帶狀區間即為布林通道。布林通道的原始設定為：

1. 上界(Upper Bond) = 中心線 + 2 個 20 日標準差
2. 中心線 = 20 日移動平均線
3. 下界(Upper Bond) = 中心線 - 2 個 20 日標準差

布林通道結合了移動平均線和統計學的標準差的概念。其基本的型態，是由 3 條軌道線組成的帶狀通道。中心線移動平均線為平均成本的概念，而上界、下界則為股價壓力及支撐訊號的判斷指標，利用統計的概念在價格偏離平均值過多時，認為價格會回歸均值，利用此訊號進場獲利，一般為逆勢策略，在股票碰到上軌時做空標的，在股票碰到下軌時做多標的，布林通道的策略變化性多，本研究僅使用了最原始的核心原理，選用 20 日為建構布林通道的參數，應用於配對交易之上，建構配對交易策略。

## 第五節 倒傳遞類神經網路

### 一、類神經網路之架構

類神經網路之研究起始於 McCulloch and Pitts (1943)，兩人所提出來結合神經生理學以及邏輯數學的神經網路模型。類神經網路之基本架構主要可分成三個層次：處理單元(Processing Element, PE)、層(Layer)、網路(Network)，其結構如圖 3-1：

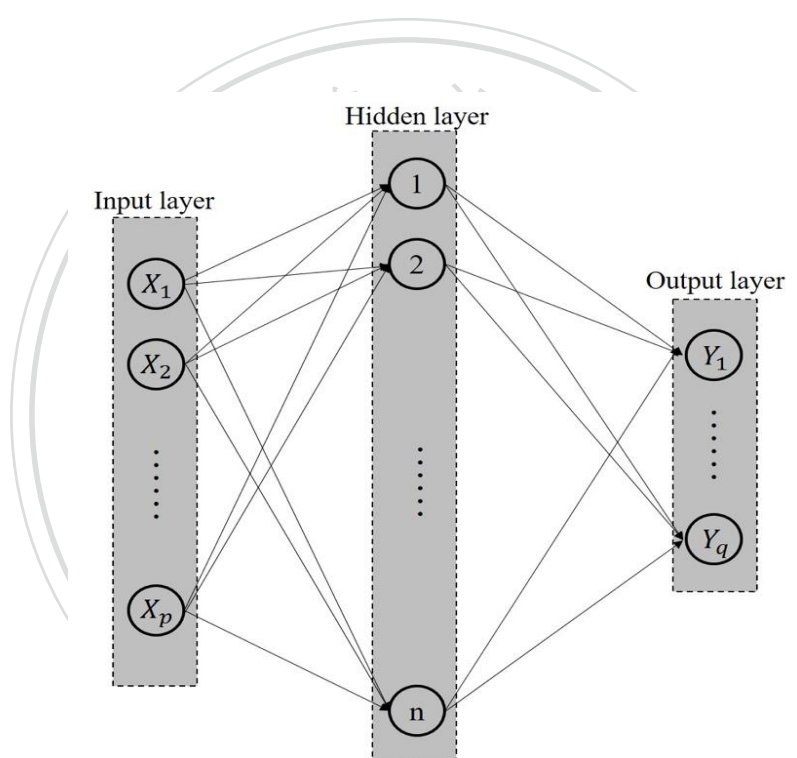


圖 3-1 類神經網路模型基本架構圖

為一個三層包含輸入層(Input Layer)、隱藏層(Hidden Layer)及輸出層(Output Layer)之網路結構，其中每層皆由最基本的處理單元——人工神經元(Artificial Neuron)組成，而人工神經元之結構如圖 3-2：

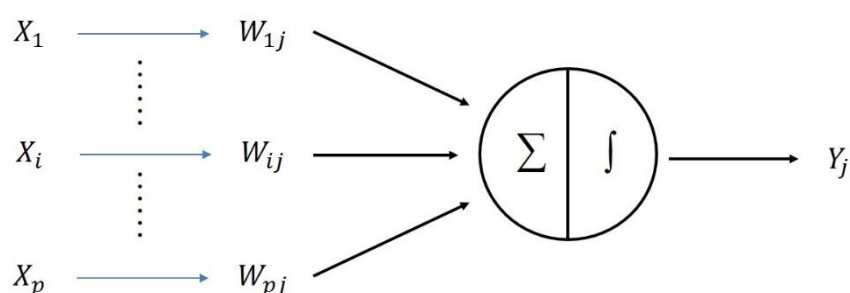


圖 3-2 處理單元示意圖

每一個人工神經元(又稱處理單元)以扇狀輸出，成為其他處理單元的輸入，藉此將資料訊息傳遞，此理單元其輸出值與輸入值的關係式，可用輸入值的加權乘積和之函數表示：

$$Y_j = f \left( \sum_i w_{ij} X_i + \theta_i \right), i = 1, 2, \dots, p; j = 1, 2, \dots, n \quad (13)$$

其中： $X_i$  為輸入之變數，類似生物神經元模型的輸入訊號； $Y_j$  為輸出之變數，類似生物神經元模型的輸出訊號； $f$  為激發函數(Active Function)，是一個非線性函數； $W_{ij}$  為加權值，似生物神經元的突觸強度； $\theta_j$  為門檻值(Threshold)，似生物神經元模型的閾值。

## 二、 倒傳遞類神經網路

倒傳遞類神經網路 (Back-Propagation Neural Network, BPNN) 是由 Rumelhart、McClelland 和 PDP 研究群於 1986 年提出，是一種監督式學習的類神經網路，屬於層狀前饋式網路架構 (Layered Feed-Forward Network)，常用的非線性轉換函數為雙彎曲函數 (Sigmoid Function) 和雙曲線正切函數 (Tanh Function)，適合用來做預測和診斷，是最被廣為應用的類神經網路。其基本原理

乃是利用梯度下降法(Gradient Descent Method)的觀念，採用倒傳遞式學習演算法(Back-Propagation Learning Algorithm)，將錯誤的訊號以回饋方式修正網路上的連結權重，使誤差函數最小化之下，調整網路權重值成為最適合的權重，使演算的輸出值能夠最接近目標輸出值。網路結構包含了一組輸入層、一組或多組的隱藏層以及一組輸出層。倒傳遞類神經網路是在所有類神經網路模型中，應用最廣、最具規模的模型。

本研究以倒傳遞演算法，建立模型，預測共整合殘差序列未來的走勢，並探討布林通道結合到傳遞類神經網路所建構之機器學習交易策略投資組合是否有更好的績效。

本研究激發函數採用雙曲線正切函數(Tanh Function)，其函式表示：

$$\tanh x = \frac{1 - e^{-2x}}{1 + e^{2x}} \quad (14)$$

為一個非線性的轉換函式，如圖 3-3 利用此函示將我們在神經元內計算的值，經過非線性轉換調整數值使其介於 1 到-1 間，其目的是為了使神經元找出之特徵更明顯。

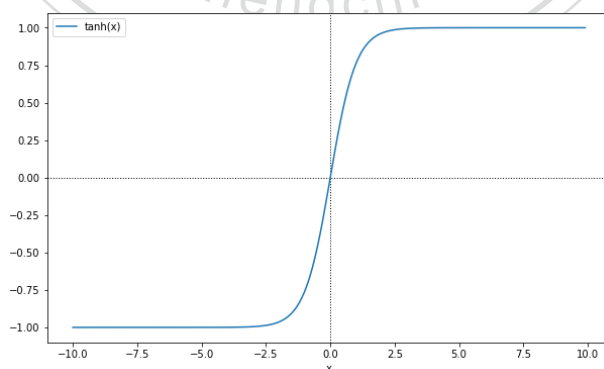


圖 3-3 雙曲線正切函數轉換示意圖

所建構的類神經網路模型為利用 Python3.5 的 Tensorflow1.7 套件包建構；最佳化方法使用 Adam(adaptive moment estimation) Optimier，此演算法是根據損失函數(Loss Function)對每個參數梯度的一階動差及二階動差來動態調整每個參數的學習率(Learning Rate)及梯度方向，簡而言之，Adam 是一個基於梯度下降方法的演算法，但是每次迭代參數的學習率及梯度方向會隨著計算而調整，使學習效率比傳統固定學習率的梯度下降法更佳；在網路輸入層變數則嘗試以 Chen, Shih and Wu (2006)研究中的價格時間序列差分資料作為輸入參數再進一步參考 Kara, Boyacioglu and Baykan (2011)以技術指標作為輸入參數，增加布林通道技術指標，輸入層總共 12 個輸入參數。

表 3-1 為選取之十二個輸入變數及其計算方式，本實驗為了讓模型能夠預測不同的股票標的，故在利用資料訓練模型前，我們會將實驗中每個配對的共整合殘差 $\hat{e}_t$ 先進行標準化再輸入，此動作目的是為了避免某一個配對組合的特徵過於明顯，造成我們的模型訓練產生問題，過度配適於某幾個配對交易組合，而其他輸入變數若須藉由共整合殘差計算，皆使用標準化後的共整合殘差加以計算得出。此處選擇之輸入變數除了利用一般的時間序列資料依序輸入，我們嘗試加入其差分値及時間序列之平方項，此外也利用本次選用之技術指標—布林通道，利用殘差與通道上下緣之差值作為輸入參數。

Variable	Calculation
Residual	$e_t$
Residual Square	$e_t^2$
Difference Residual-1(DR-1)	$e_t - e_{t-1}$
DR-1 Square	$(e_t - e_{t-1})^2$
DR-5	$e_t - e_{t-5}$
DR-10	$e_t - e_{t-10}$
DR-15	$e_t - e_{t-15}$
DR-20	$e_t - e_{t-20}$
MA-20	$\sum_{i=t-20}^t \frac{(e_i - MA_{20})}{20}$
SD-20	$(\sum_{i=t-20}^t \frac{(e_i - MA_{20})^2}{20})^{1/2}$
D-UP	$e_t - (MA_{20} + 2 * SD_{20})$
D-DOWN	$(MA_{20} - 2 * SD_{20}) - e_t$

表 3-1 到傳遞類神經網路輸入變數

網路架構方面 Kara, Boyacioglu and Baykan (2011)的建構方式以單層隱藏層三十個神經元為模型架構，我們將利用此模型進行共整合殘差未來走勢的預測，其輸出層有兩個輸出結點分別代表明天上漲及下跌的機率，若真實資料明天上漲(下跌)，我們將給予上漲之輸出節點機率值 1(0)，下跌之輸出節點為機率值 0(1)，而在訓練之後模型之預測我們將會選擇兩節點中機率較大者，作為我們的預測方向。例如:若明天上漲機率之節點 0.6，下跌機率之節點為 0.4，我們將預測明天為上漲。

為了測試模型的穩健性，我們將會將樣本外時間在進行切割，僅使用部分樣本外時間資料訓練類神經網路模型，並展延時間長度，檢視模型在訓練期間外是否同樣能有良好表現。



## 第六節 配對交易策略

### 一、布林通道配對交易策略

通過篩選之配對，表示在樣本內期間內兩檔股票經 Engle-Granger 二階段檢定具有共整合關係，利用樣本內期間 OLS 估計之參數 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ ，在樣本外期間利用 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ 計算殘差 $\hat{e}_t$ ：

$$\hat{e}_t = Y_t - \hat{\beta}_1 X_t - \hat{\beta}_0 \quad (15)$$

其中  $X_t$ 、 $Y_t$  為前述經過資料處理的樣本外時間區間股票收盤價資料序列。

接著我們利用技術指標布林通道判斷進場做多及放空配對時機的訊號，可分為兩部分策略，其規則為：

1. 當殘差項  $\hat{e}_t$  由下方向上穿越兩倍布林通道上緣  $MA(\hat{e}_t, d) + 2 * SD(\hat{e}_t, d)$  時，於隔日進場放空一股  $Y$  做多  $\hat{\beta}_1$  股  $X$ ，並於  $\hat{e}_t$  由上方向下穿越一倍布林通道上緣  $MA(\hat{e}_t, d) + SD(\hat{e}_t, d)$  時，於隔日平倉出場。
2. 當殘差項  $\hat{e}_t$  由上方向下穿越兩倍布林通道下緣  $MA(\hat{e}_t, d) - 2 * SD(\hat{e}_t, d)$  時，於隔日進場做多一股  $Y$  放空  $\hat{\beta}_1$  股  $X$ ，並於  $\hat{e}_t$  由下方向上穿越一倍布林通道下緣  $MA(\hat{e}_t, d) - SD(\hat{e}_t, d)$  時，於隔日平倉出場。

其中  $MA(\hat{e}_t, d)$  為前  $d$  天的  $\hat{e}_t$  之移動平均， $SD(\hat{e}_t, d)$  為前  $d$  天的  $\hat{e}_t$  之標準差，本文皆使用最基本的 20 天作為輸入參數，不另外進行參數的選擇，避免過度最佳化的問題。策略訊號的產生概念如圖 3-4：

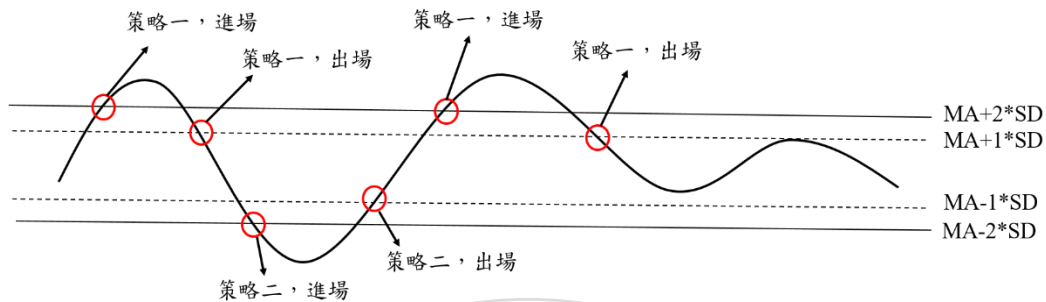


圖 3-4 布林通道配對交易策略進出場方式示意圖

此布林通道策略的基本概念為當殘差項  $\hat{e}_t$  偏離過多超出布林通道時，我們便認為有利可圖，進場建倉，建立多空對沖的投資組合，等未來價差收斂回歸均值時平倉獲利。

本研究為求策略的簡易選用較基本的布林通道門檻值，僅以一倍標準差及兩倍標準差做為判訊號之通道，不再另用最佳化找尋最適合配對交易的通道標準差倍數，減少我們後面探討配對交易投資組合時產生過度最佳化問題。

為了確保能夠確實完成進場部位的建構而不出現實務委託上的無法成交情形，進場訊號出現後隔天便以市價單於開盤前掛單進場，即以「開盤價」買入或放空標的資產，出場則在訊號出現隔天以「開盤價」平倉出場，若部位持有至模擬期間結束則以模擬期間結束日之「收盤價」平倉出場，此外本研究為探討配對交易之市場中立報酬是否真實存在，因此在回測及交易中不考慮手續費及交易稅等交易成本，期望能真實看出配對交易的市場中立報酬效果。

## 二、機器學習配對交易策略

機器學習的配對交易策略應用，我們將嘗試與本文前段利用技術指標建構之配對交易策略進行結合；以機器學習預測共整合殘差 $\hat{e}_t$ 明天之走勢，再結合布林通道技術指標，找出真正的共整合殘差反轉點，其策略如下：

1. 當殘差項 $\hat{e}_t$ 大於兩倍布林通道上緣  $MA(\hat{e}_t, d) + 2 * SD(\hat{e}_t, d)$  且類神經網路模型預測明天殘差項 $\hat{e}_t$ 下跌時，於隔日進場放空一股 Y 做多 $\hat{\beta}_1$ 股 X，並於 $\hat{e}_t$ 由上方向下穿越一倍布林通道上緣  $MA(\hat{e}_t, d) + SD(\hat{e}_t, d)$  時，於隔日平倉出場。
2. 當殘差項 $\hat{e}_t$ 由上方向下穿越兩倍布林通道下緣  $MA(\hat{e}_t, d) - 2 * SD(\hat{e}_t, d)$  且類神經網路模型預測明天殘差項 $\hat{e}_t$ 下跌時，於隔日進場做多一股 Y 放空 $\hat{\beta}_1$ 股 X，並於 $\hat{e}_t$ 由下方向上穿越一倍布林通道下緣  $MA(\hat{e}_t, d) - SD(\hat{e}_t, d)$  時，於隔日平倉出場。

基本概念與技術指標配對交易策略相同，當殘差項 $\hat{e}_t$ 偏離過多超出布林通道上緣(下緣)時，我們便認為有利可圖，但為了避免過早進場無法將策略建構於最高點(最低點)，故以類神經網路模型之預測作為第二層篩選，當類神經網路模型預測明天殘差將下跌(上漲)時，再進場建倉，建立多空對沖的投資組合，等未來價差收斂回歸均值時平倉獲利。

機器學習配對交易策略與布林通道策略相同，為了確保能夠確實完成部位的建構而不出現實務委託上無法成交情形，進場訊號出現隔天以市價單於開盤前掛單進場，即以「開盤價」買入或放空標的資產，出場則在訊號出現隔天以「開盤價」平倉出場，若部位持有至模擬期間結束則以模擬期間結束日之「收盤價」平倉出場，同樣為了能看出市場中立報酬，此處不考慮手續費及交易稅等交易成本。

## 第七節 配對交易投資組合的建構方法

投資組合的報酬率計算為求簡易，我們將投資組合的建構方式設定為給予每對通過二階段共整和檢定的配對組合相同的投資資金，並在每日交易後重新平衡每個配對組合的資金淨值，即減少超過平均資金淨值的配對組合資金，並給予低於平均資金淨值的配對組合資金，進行平衡，使每對配對組合隔日的可投資資金淨值相同。

投資組合的每日總報酬率計算方式以每日每個配對組合的獲利進行加總，再除上前一日投資組合的總資金，交易期間的總報酬率則以複利連乘的方式以日報酬率進行計算。

本研究將以相同的方法為不同的配對交易策略進行投資組合的建構，比較布林通道配對交易策略與類神經網路結合布林通道配對交易策略兩個不同策略在台灣股票市場的表現，以利於瞭解兩策略的優缺點及策略特性。

## 第八節 交易策略績效評估指標

除了報酬率外，本研究使用兩樣指標最為策略績效的評估指標，分別為最大策略虧損報酬率及夏普值，以下將列舉說明。

### 1. 最大策略虧損報酬率

此指標概念為每承受損失 1 元的風險，所產生的預期獲利金額，交易策略的此項指標要能夠為正，並且數值越高表示績效越佳。分母採用最大連續虧損金額 (Maximum Drawdown, MDD)，是一個將交易策略對連續虧損狀況承受能力的衡量指標。最大策略虧損報酬率的計算方法如下：

$$\text{NPMDD Ratio} = \frac{\text{NetProfit}}{\text{MDD}} \quad (16)$$

其中 NetProfit 為策略淨利，MDD 為最大的策略連續虧損金額。

### 2. 夏普值

夏普值(Sharpe Ratio)為一個考慮報酬相對於風險的指標，其概念為每單位的風險能賺的報酬，數值越高表示績效越佳，最早於 Sharpe (1966)中被使於評估共同基金(Mutual Funds)之風險，在被廣泛應用後於 Sharpe (1994)正式發表取名。夏普值計算方法如下：

$$\text{Sharpe Ratio} = \frac{E(R_p) - R_f}{\sigma_p} \quad (17)$$

其中  $E(R_p)$  為投資組合預期報酬率，一般使用過去的平均報酬， $R_f$  為無風險利率，本研究使用美國 3 個月公債殖利率 1.8%； $\sigma_p$  為投資組合的風險，一般使用投資組合報酬的標準差。

## 第四章 實證分析

本研究之實證分為三個階段，首先利用研究方法中的二階段檢定法找出台灣股票市場可行的配對組合，再利用布林通道進行樣本外交易回測，觀察配對組合績效。第二部分為利用配對組合及布林通道策略建構投資組合，觀察投資組合其績效。第三部分利用類神經網路學習布林通道策略，觀察模型對交易策略的學習效果及其建構出的投資組合績效。

### 第一節 實證資料與研究期間

本研究所使用之資料來源為台灣經濟新報(TEJ)資料庫，資料包含台灣所有上市及上櫃公司普通股總共 1683 檔股票(截至 2017 年 12 月 31 日，公開交易之股票)，每日調整後之開盤價格及收盤價格，此外使用上市及上櫃公司的基本資料，包含由台灣經濟新報(TEJ)所給予公司的子產業分類。資料期間為 2007 年 1 月 1 日到 2017 年 12 月 31 日，在資料期間內我們將分為樣本內形成期間由 2007 年 1 月 1 日到 2013 年 12 月 31 日，及樣本外的測試期間由 2014 年 1 月 1 日到 2017 年 12 月 31 日，表 4-1 為台灣股票市場交易資訊簡介。

本文使用之台灣上市上櫃公司產業分類方法為根據台灣經濟新報(TEJ)依公司主要營業內容進行分類之 TEJ 子產業類別(目前分為 167 類)，作為我們進行產業分類的依據。

表 4-1 台灣證券交易市場簡介

交易時間	盤中	9:00~13:30
	盤後	14:00 ~ 14:30
	零股	13:40 ~ 14:30
交易單位	盤中	張
	盤後	張
	零股	股
交易方式		T+2 現金交割方式
股票單位		一張=1000 股
股價		10 塊=每股 10 元(每張=10000 元)
股本		10*股數
漲跌停		昨日收盤價上下 7%

資料來源:台灣證券交易所





## 第二節 台灣股票配對組合篩選結果

### 一、 單根檢定

台灣上市及上櫃股票 1683 檔公開發行時間不同，且收盤價格時間序列存在缺漏值，本研究統一使用 2007 年 1 月 1 日至 2013 年 12 月 31 日期間，進行 DF 單根檢定，為確保資料時間長度足夠，讓模型具有一致的最小樣本數方便比較，若股票上市日期在 2012 年 1 月 1 日後，我們不對其進行檢定，也不將其納入後面配對交易研究討論之成分股票；暫停交易等因素造成的資料缺漏值部分，不論是開盤貨收盤價格，接以前一日收價格進行補齊，避免在後續研究中缺漏值時間內有過大的股價波動，造成交易訊號的不準確，也可以避免買不到股票造成策略及投資組合損益的計算有問題。

經資料整理過後，將每檔股票進行 DF 檢定，若其檢定之 p-values 大於顯著水準 1%，表示其股票價格不為定態時間序列，便將此股票納入後續研究樣本，經過此步驟我們將 1668 檔股票篩選減少成 1306 檔，後續的配對交易策略及因子研究，皆以此篩選後的 1306 檔股票為資料樣本進行研究。

### 二、 共整合檢定

經過單根檢定後，將通過檢定的 1306 檔皆為  $I(1)$  序列的股票標的，進行兩兩配對，在樣本內期間 2007 年 1 月 1 日至 2013 年 12 月 31 日間進行 Engle-Granger 二階段檢定，檢測共整合關係，此處若遇到兩檔股票上市時間不同，造成資料有缺漏，將選取較晚上市的股票的上市日期做為共整合檢定的樣本時間之起始點；在 Engle-Granger 二階段檢定中，先以兩股票  $X_t$ 、 $Y_t$  估計迴歸式：

$$Y_t = \beta_0 + \beta_1 X_t + e_t \quad (18)$$

接著對其殘差項 $\hat{e}_t$ 進行 DF 單根檢定，若檢定之 p-values 小於 1%則通過共整合檢定，代表此配對具有共整合性質，我們之後將通過樣本內共整合檢定的配對在進行樣本期間外(2014 年 1 月 1 日至 2017 年 12 月 31 日)的交易策略回測，為了模擬實際交易的清況，樣本外測試期間不需要再重複進行線性迴歸估計樣本外其間的 $\beta_0$ 、 $\beta_1$ ，而是直接使用樣本內期間估計得到之 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ 來計算樣本外期間之殘差項 $\hat{e}_t$ ，我們後續將會利用樣本外時間之 $\hat{e}_t$ 作為我們建構策略之資料。

經過共整合檢定篩選後，得到 98187 個 p-values 小於 1%在統計上具有共整合性質的配對組合，為了減少配對的數量及增加配對組合的穩健性，確保選出的配對具有長期市場中立的特性，如同 Avellaneda and Marco and Lee (2010)考慮產業因子來解釋股票報酬，我們需要進一步對配對進行產業的篩選。

### 三、 產業篩選

經過上述篩選交易配對的共整合檢定，接著我們進一步對股票配對進行同產業的篩選，此步驟是為了減少共整合檢定後過多的配對組合，此外也是為了避免篩選出的股票配對組合，兩公司在經濟意涵上無相關連性，而兩間公司對彼此的影響力僅在樣本期間內有統計上有顯著結果，若有上述問題將有可能於樣本外期間失去獲利能力，甚至造成虧損。經過共整合檢定及產業篩選後，我們將所有可能的 $C_2^{1306}$ 種股票配對組合，篩選至剩下 1628 組。

經過共整合檢定及產業篩選後，我們發現利用迴歸估計的參數 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ 在樣本外 $\hat{\beta}_0$ 容易有偏離的現象，其原因可能是因為 $\hat{\beta}_0$ 中包含了未被配對交易組合消除的市場大盤的情勢，若樣本內、外市場大盤的情況差異很大會使我們的 $\hat{\beta}_0$ 不準確造成殘差 $\hat{e}_t$ 有向上或向下偏移(Shift)之現象，以加高(8182)、希華(2484)兩檔股票之配對為例，由其兩檔股票之收盤價及殘差之圖 4-1，可以發現其殘差項 $\hat{e}_t$ 在樣本外期間(紅色去格線右側)有些微偏移，但整體的關係還是維持定態。因此我們嘗試在選擇配對交易策略時，選擇使用會隨著時間變動調整的技術指標－布林通道，一個對時間適應性較佳，且能夠判斷配對之發散收斂趨勢的策略，取代傳統配對交易策略給予一個固定上界及下界作為進出場訊號之方法。



圖 4-1 加高與希華兩股票之收盤價與共整合殘差圖

最後，在建構交易策略前，我們會先判斷其 $\hat{\beta}_1$ 係數之正負號，若配對組合之 $\hat{\beta}_1$ 係數為負，代表兩股票間之關係為負相關，在我們利用殘差 $\hat{e}_t$ 進和交易訊號判斷時，我們無法利用殘差值之正負判斷股票對組合中價格過高的股票為兩檔股票中何檔，在此狀況下變無法進行交易策略的建構，因此本研究會進一步將 1628 檔股票中， $\hat{\beta}_1$ 係數為負的配對組合排除，利用剩下 $\hat{\beta}_1$ 係數為正的配對組合共計 1494 組進行交易策略的回測及建構投資組合。

圖 4-2 為各產業配對交組合數目統計圖，由於產業數目過多本圖僅列出配對組合數目大於 10 組的產業，可發現其中 IC 設計產業有 491 個配對組合與第二多的產業相差約四倍，由此圖可發現具共整合性質的配對交易組合產業分布不均。

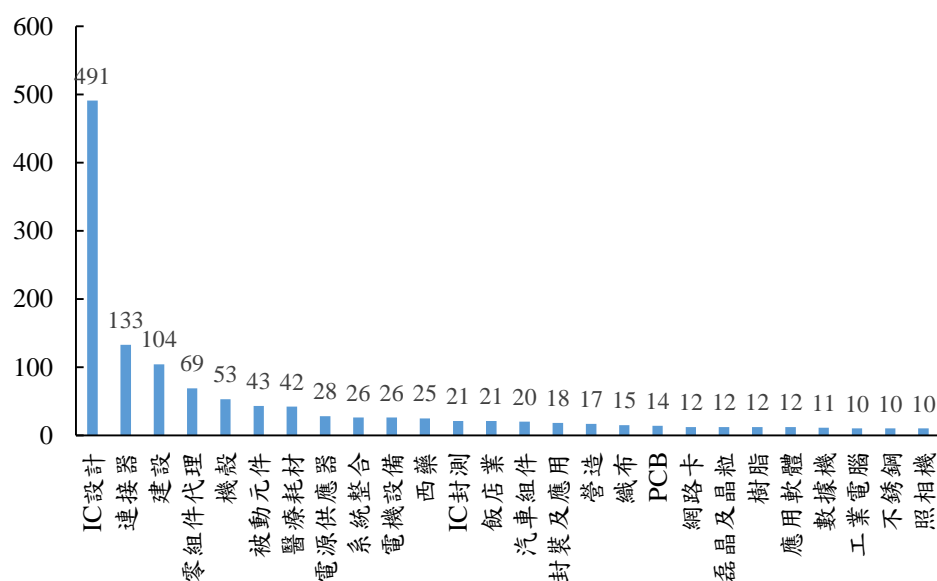


圖 4-2 配對交易組合數目之產業分配統計圖

### 第三節 布林通道配對交易策略實證結果

#### 一、 布林通道配對交易策略進出場訊號及績效

篩選出 $\hat{\beta}_1$ 為正、同產業且具有共整合性質的 1494 組配對交易組合後，我們利用布林通道建構交易策略進行回測，而策略績效的計算及回測我們只利用樣本外期間 2014 年 1 月 1 日至 2017 年 12 月 31 日測試，計算過程皆使用報酬率當計算依據，即每一塊錢的資金在策略每次的進出場後可以取得的獲利，本研究計算的是樣本外期間內之總報酬率，以利於我們去計算評估績效的最大策略虧損報酬率。

以宏亞(1236)、天仁(1233)兩檔股票之配對為例，圖 4-3 為兩檔股票樣本內及樣本外期間之收盤價及共整合殘差，接著我們將利用樣本外時間進行交易回測，利用布林通道策略產生策略訊號，如圖 4-4 為宏亞(1236)、天仁(1233)兩檔股票配對交易績效圖及共整合殘差訊號圖，績效圖中之績效是以買進(放空)一股宏亞(1236)並放空(買進) $\hat{\beta}_1$ 股天仁(1233)計算其交易的賺取的點數；而下方的訊號圖中紅色訊號代表放空宏亞(1236)、買進天仁(1233)之策略一進場期間，綠色訊號代表買進宏亞(1236)、放空天仁(1233)之策略二進場期間，可發現此配對交易策略部位持有時間都較短，主要是於配對組合的成分股票價格出現明顯的異常時才進場獲利，並且認為配對組合殘差具有均值回歸的特性，未來在價格回歸正常時反向平倉出場，為一個短進短出的交易策略。

表 4-2 為宏亞(1236)、天仁(1233)兩檔股票之配對交易策略績效參數，由參考的績效評估指標可發現此布林通道配對交易策略為一個交易次數較少、勝率較高、回檔(Drawdown)較小之交易策略，但因此其報酬率也較低此配對四年僅有 22%

之獲利能力。此處額外考慮單一組配對較易策略回測之勝率(Win Rate)是因為配對交易策略的單次金出場獲利不高，但虧損時卻相對容易是大金額的虧損，因此勝率對於衡量配對交易績效相當重要，回測績效未超過 50%的配對交易策略應難以實際應用於市場並獲取利益。勝率的計算公式如下：

$$\text{Win Rate} = \frac{\text{Win Trades}}{\text{Total Trades}} \quad (19)$$

其中 Win Trades 為所有交易中獲利為正的交易次數，Total Trades 為總交易的次數，進場建倉至平倉出場視為一次交易。



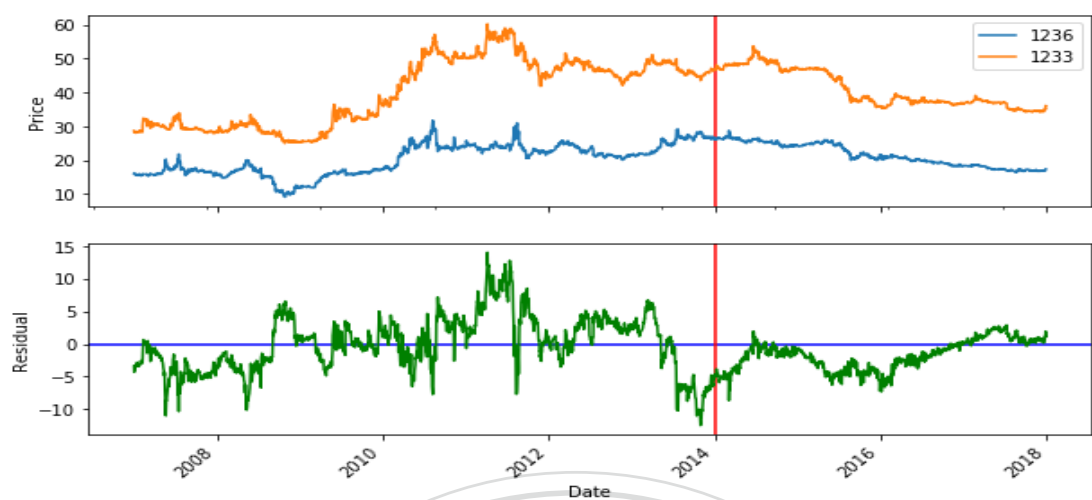


圖 4-3 宏亞與天仁收盤價與共整合殘差圖

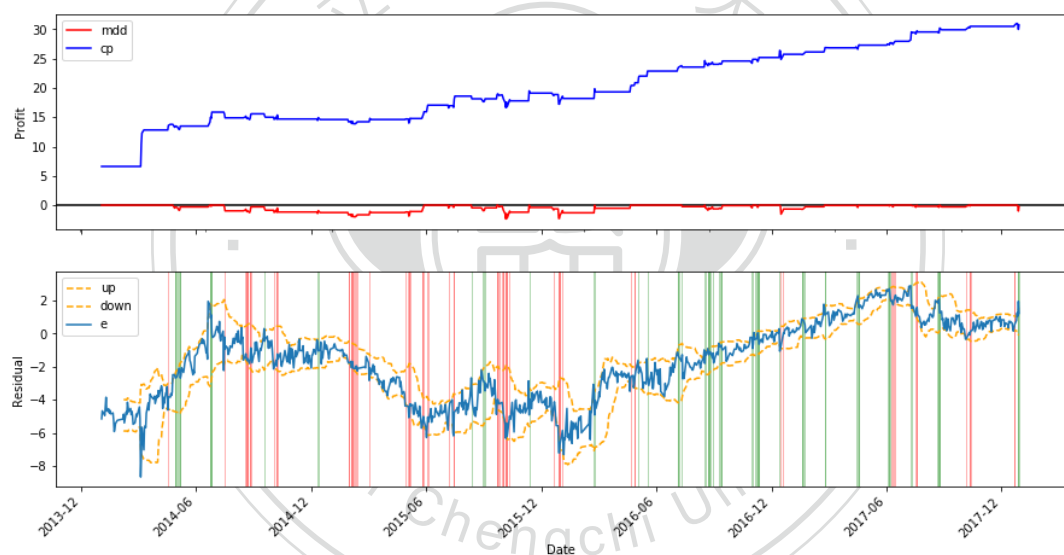


圖 4-4 宏亞與天仁布林通道交易策略績效及進場訊號圖

表 4-2 宏亞與天仁布林通道交易策略績效參數

最大報酬率	最大虧損(點)	最大策略虧損報酬	交易次數(次)	勝率	產業
0.2244	-2.3510	7.8302	45	0.7556	其他食品



篩選出的配對 1494 組配對交易組合，我們皆以與上述宏亞(1236)、天仁(1233)之配對交易組合相同之策略及方法進行回測，並得到各個配對交易組合的績效參數，其敘述統計如表 4-3，可發現確實我們選用的布林通道策略，屬於交易次數相對少，勝率高的交易策略，但由於我們使用的是最基本的布林通道策略，參數也沒有進行最佳化，故在 1494 組配對交易組合中，有些組合在回測結果上其實並不理想，甚至有策略的最大虧損達到 2210 點，最大策略虧損報酬率的平均值僅有 1.4596，不過我們也有此可知，市場上確實存在著市場中立的報酬，我們也確實能利用簡單的布林通道策略去發現價格的異常並進場賺取之，但其中之利潤並不大；因此若單以一個配對交易組合，作為交易的策略或許不是理想的作法，因此本研究嘗試接著探討以多個配對交易組合建立投資組合，觀察其穩定性。除了單獨探討配對交易策略的績效參數，我們嘗試將產業加入，觀察各產業之配對交易組合，是否存在著績效的差異，如附錄，可發現不同的產業確實最大策略虧損報酬率有差異。

表 4-3 布林通道配對交易策略績效參數之敘述統計表

	最大報酬率	最大虧損	最大策略 虧損報酬	交易次數	勝率
平均數	1.3812	-21.0562	1.4596	42.5027	0.6380
標準誤	0.0098	1.6554	0.0498	0.0962	0.0020
中間值	1.2828	-10.4413	1.0377	42.0000	0.6429
眾數	1.0000	NA	NA	43.0000	0.6667
標準差	0.3784	63.9832	1.9236	3.7190	0.0755
變異數	0.1432	4093.8486	3.7001	13.8309	0.0057
峰度	12.3479	920.6525	4.4082	-0.3193	0.1604
偏態	2.7323	-27.2873	1.6938	0.0599	-0.2664
範圍	3.6814	2210.3361	15.1666	24.0000	0.5665
最小值	1.0000	-2210.7705	-0.9497	31.0000	0.3030
最大值	4.6814	-0.4343	14.2170	55.0000	0.8696

## 二、 布林通道配對交易投資組合實證結果

圖 4-5 為布林通道配對交易投資組合在樣本外期間的績效，可發現布林通道投資組合為一個波動較小的策略，但由於其賺取的是市場中立的報酬故四年僅賺取 19% 的報酬，但此策略也不容易造成巨大虧損，穩定性高，僅在 2015 年下半年有較大的回檔(Drawdown)。

由圖 4-5，布林通道配對交易投資組合之績效數據投資組合之最大策略虧損報酬有 11.4244，夏普值為 2.4514，由此可知此策略有不錯的樣本外回測績效，為一個穩定且具有獲利能力的交易策略。

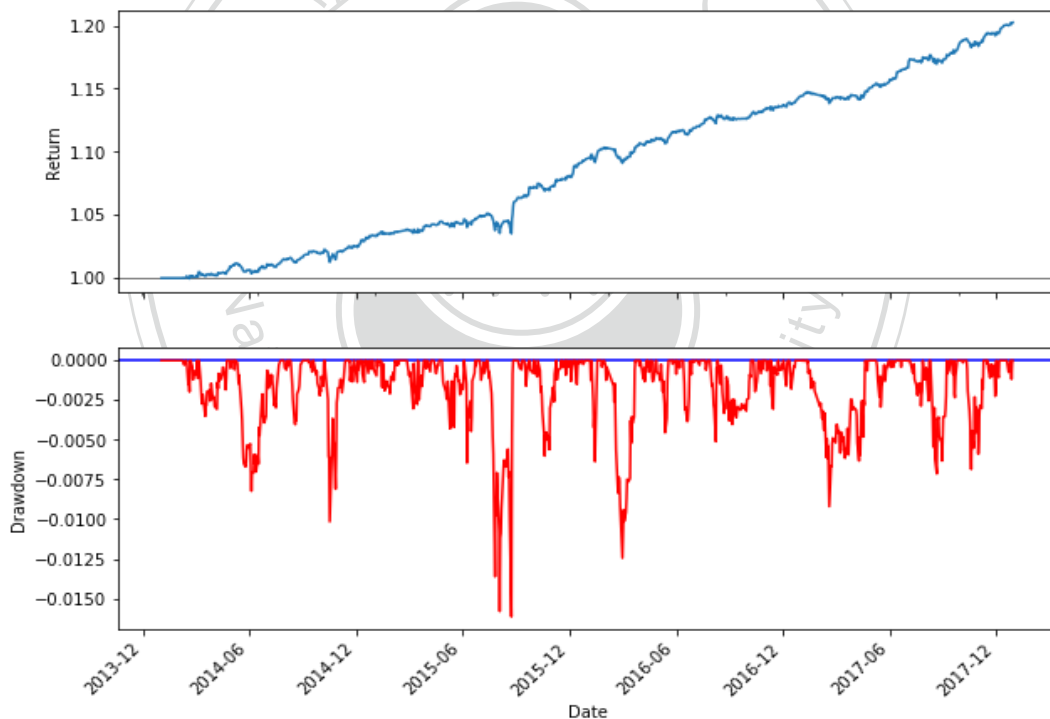


圖 4-5 布林通道配對交易投資組合績效圖與虧損圖

圖 4-6 為布林通道配對交易投資組合與台灣股價加權指數報酬率之比較，由圖可知在樣本外期間 2014 年 1 月 1 日至 2017 年 12 月 31 日間，台灣加權股價指數(TAIEX)，有很大的波動度，但我們的布林通道配對交易策略在此時間，

以穩定的報酬率獲利，風險相對小。表 4-4 中也能從績效參數觀察到，布林通道配對交易策略的績效項對於買進持有大盤有著良好的獲利能力。

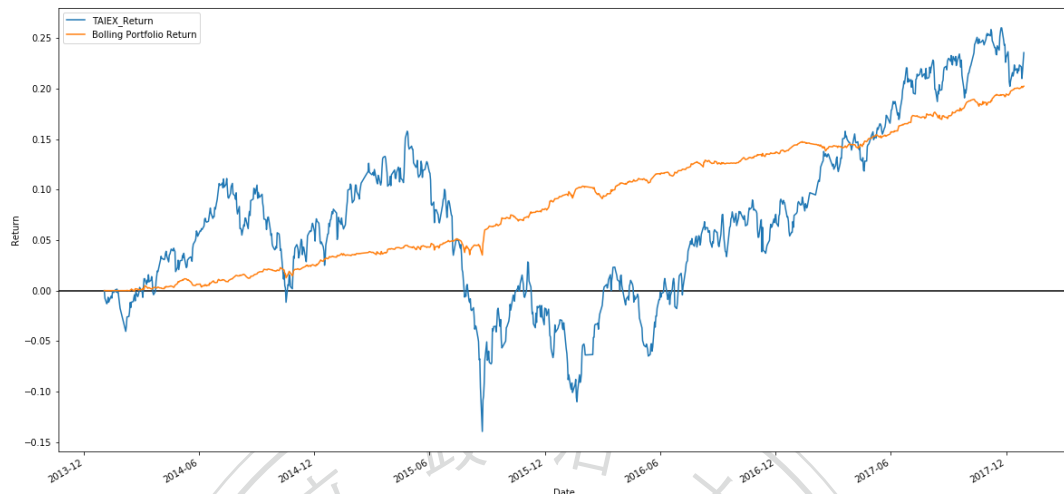


圖 4-6 布林通道配對交易投資組合與大盤走勢圖

表 4-4 布林通道配對交易投資組合績效參數

	最大報酬	最大虧損	最大策略 虧損報酬	平均日報酬	日標準差	夏普值
布林通道	19.0208%	-1.6649%	11.4244	0.0178%	0.1148%	2.4513
加權指數	26.0322%	-29.7564%	0.8748	0.0246%	0.7814%	0.4980

圖 4-7、圖 4-8、圖 4-9 為布林通道配對交易投資組合日損益分析、月損益分析及年損益分析。由圖 4-7 可知，布林通道配對交易投資組合的日報酬雖然小，但以大於零的正日報酬多於負的日報酬；由圖 4-8 可知，其每個月報酬以獲利的正報酬率月份居多，四年的時間僅有九個月是虧損，且也沒有出現連續虧損多個月的问题；由圖 4-9 可知，此投資組合在樣本外四年內其每年的獲利狀況皆為正，且每年的獲利差異並不大，僅 2015 年有明顯優於其他年份的獲利。

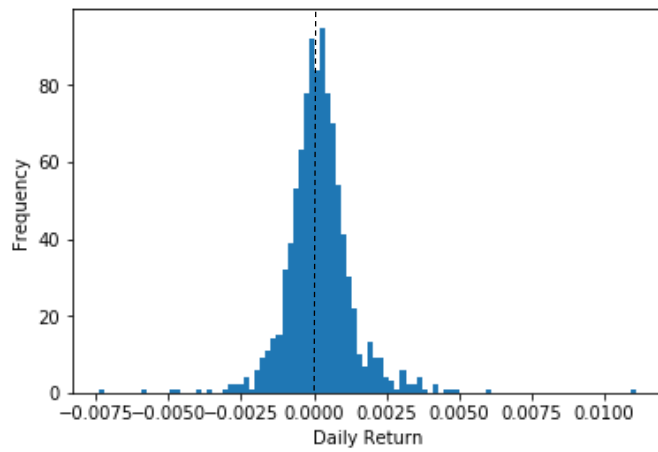


圖 4-7 布林通道配對交易投資組合日損益分析圖

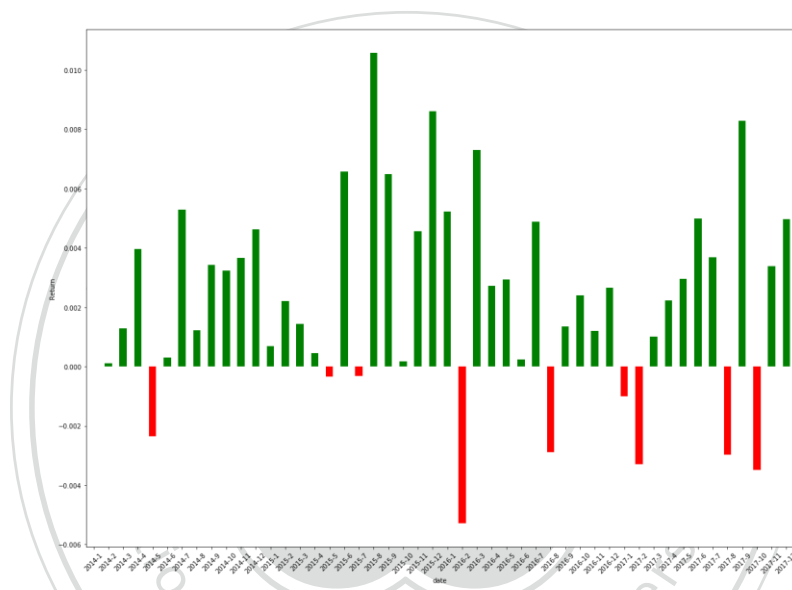


圖 4-8 布林通道配對交易投資組合月損益分析圖

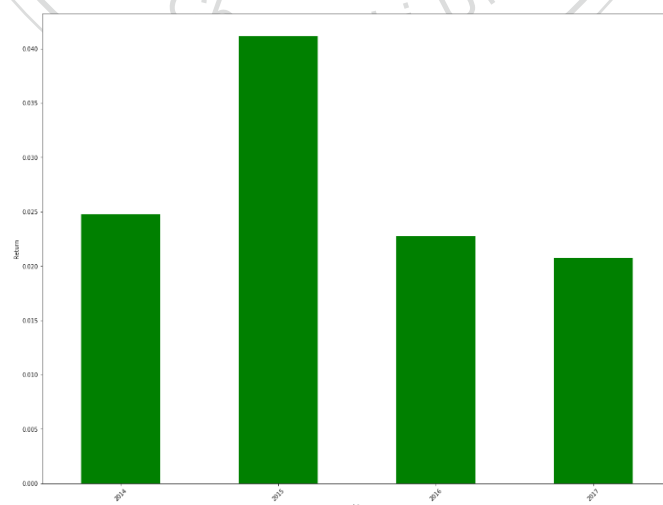


圖 4-9 布林通道配對交易投資組合年損益分析圖

## 第四節 機器學習配對交易策略實證結果

### 一、類神經網路預測共整合殘差走勢能力分析

類神經網路模型的訓練我們將使用 Tensorflow 中以梯度下降法極小化損失函數的 Adamoptimizer 演算法，極小化的損失函數為誤差(MSE)，演算法學習率初始值設定為 0.01。我們將會將交易回測的樣本外期間 2014 年 1 月 1 日到 2017 年 12 月 31 日在進行分割，僅利用 2014 年 1 月 1 日到 2016 年 12 月 31 日前間最為模型訓練期，而最後的一年做為測試期，期望能藉此觀察模型在訓練樣本外的預測能力。

類神經網路模型在經過訓練一萬次後，觀察模型在訓練過程中誤差及預測能力(Accuracy)之情形，如圖 4-10 為一層隱藏層 30 個神經元節點之類神經網路模型之訓練過程中誤差和預測能力變化圖，可發現隨著訓練次數的增加誤差在初期快速的下降且速度快速減慢，而預測準確率則隨著訓練的次數慢慢地上升，最後達到訓練期間約 54%的準確率。

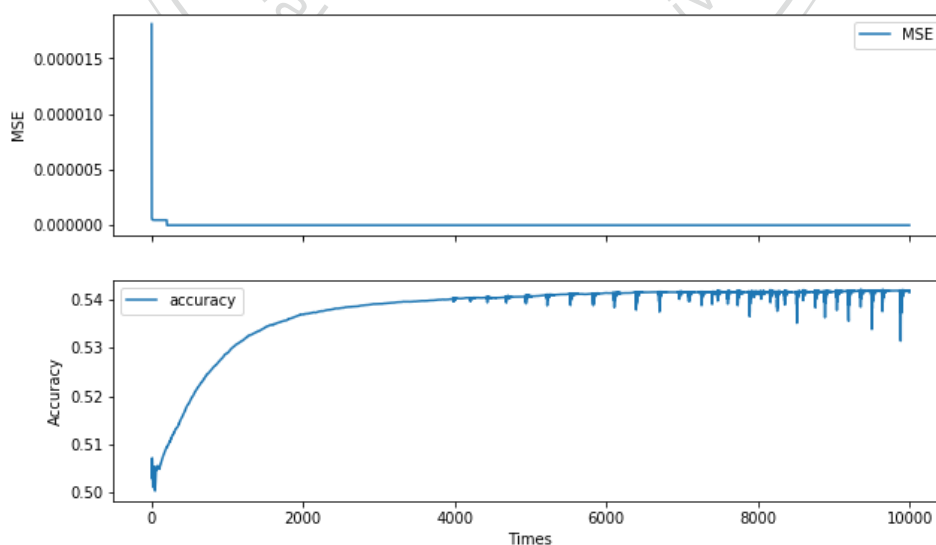


圖 4-10 類神經網路模型訓練過程之誤差及預測能力圖

為了確保我們的模型的穩健性，避免過度最佳化的問題，我們嘗試將隱藏層的結點數作改變，我們嘗試將不同節點數的類神經網路模型進行相同的訓練方法，經過一萬次的訓練後結果如表 4-5，可發現我們不同節點數的模型，預測能力差異並不大，各模型在樣本內對於全部資料的預測能力約 54%，而訓練期間外對全部資料的預測能力也都大約 53%，但在預測我們訊號需要判斷殘差走勢的區間(共整合殘差穿越布林通道上下界的時間點)樣本內預測率約 55%，樣本外預測率約 54%，相較於全樣本有較好的預測能力，或許是與我們輸入模型的變數有關係；雖然預測能力不到完美，我們還是嘗試將其結合進我們的配對交易策略中，期望能有更多的發現，接下來之策略建構及交易回測皆使用隱藏層節點數 30 個進行實證研究。

表 4-5 隱藏層節點改變之穩健性分析表

隱藏層 節點數	10	15	30	60	90
誤差(MSE)	4.62E-07	4.62E-07	4.63E-07	4.63E-07	4.62E-07
全樣本準確率 (訓練期間內)	0.5413	0.5410	0.5411	0.5415	0.5422
全樣本準確率 (訓練期間外)	0.5336	0.5334	0.5317	0.5334	0.5332
訊號區準確率 (訓練期間內)	0.5567	0.5571	0.5579	0.5565	0.5580
訊號區準確率 (訓練期間外)	0.5479	0.5470	0.5462	0.5466	0.5462

## 二、機器學習配對交易策略績效

本文利用類神經網路模型對共整合殘差進行預測，嘗試在殘差穿越布林通道上下界時，不立即進場建倉，而是等到類神經網路模型預測殘差反轉時再進場，期望利用此方式能抓到真正的反轉點，使我們的交易策略有更好的獲利能力。

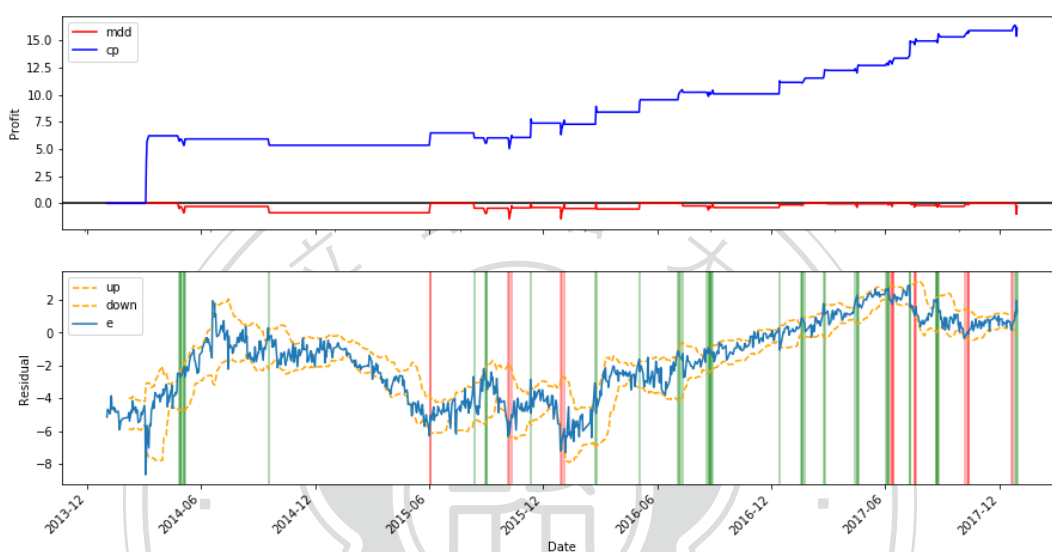


圖 4-11 宏亞與天仁機器學習交易策略績效及進場訊號圖

以宏亞(1236)、天仁(1233)兩檔股票之配對為例，圖 4-11 為宏亞(1236)、天仁(1233)兩檔股票配對交易之績效圖及共整合殘差訊號圖，績效圖中之績效是以買進(放空)一股宏亞(1236)並放空(買進) $\hat{\beta}_1$ 股的天仁(1233)計算其交易的賺取的點數；而下方的訊號圖中紅色訊號代表放空宏亞(1236)、買進天仁(1233)之策略一進場期間，綠色訊號代表買進宏亞(1236)、放空天仁(1233)之策略二進場期間，同樣可發現此配對交易策略部位持有時間都較短，此相較於僅使用布林通道的配對交易策略，進場的次數更少，表示其進場的條件更加嚴苛，雖然因此賺取的金額減少，但其進場時機更為準確，賠錢的交易更少，且在模型訓練期間之外之 2017 年，此策略也有不錯的表現。表 4-5 為宏亞(1236)、天仁(1233)兩檔股票之



配對交易策略績效參數，由表 4-6 中參考的績效評估指標可發現機器學習配對交易策略相較於僅使用布林通道，有更小的最大虧損、更好的報酬率，使得最大策略虧損報酬達到 11.1031，且其交易次數減少至 24 次，勝率也提高至 0.7917，表示利用類神經網路改良策略，確實有改進布林通道交易策略，使策略有更好的績效。

表 4-6 宏亞與天仁機器學習交易策略績效參數

最大報酬率	最大虧損(點)	最大策略虧損報酬	交易次數(次)	勝率	產業
0.2515	-1.4505	11.1031	24	0.7917	其他食品

此處我們同樣利用此改進之策略回測 1494 組配對交易組合，並得到各個配對交易組合的績效參數，其敘述統計如表 4-7，可發現經過類神經網路模型對殘差的預測作為近一步的篩選，確實使 1494 組配對組合回測的最大虧損下降至 15.7477 點，最大虧損之配對也從約 2210 點變為約 1368，確實利用類神經網路做進一步的篩選，能夠幫助我們改善布林通道策略，更準確地找出真正的殘差反轉時間點，而不會在殘差因市場狀況的偏移或在市場波動較小時，因為只依靠單一技術指標而誤判，進場後有嚴重的虧損；但由於進場次數少及勝率提高，使我們的平均最大報酬從可賺取 1.3812 倍本金，下降至平均 1.2504 倍，也因此造成平均最大策略虧損報酬率下降至 1.3047。同樣由於一組配對交易組合，其獲利較少，進場時間短，故單以一組標的進行投資穩健性不足，我們將嘗試利用與前述相同的投資組合建構方法，建構機器學習配對交易投資組合，檢視在樣本外期間，此投資組合之穩健性與獲利能力。

表 4-7 機器學習配對交易策略績效參數之敘述統計表

	最大報酬	最大虧損	最大策略 虧損報酬	交易次數	勝率
平均數	1.2504	-15.7477	1.3047	25.0214	0.6471
標準誤	0.0064	1.0934	0.0514	0.1393	0.0026
中間值	1.1854	-8.0596	0.7685	25.0000	0.6500
眾數	1.0000	NA	NA	25.0000	0.6667
標準差	0.2456	42.2630	1.9853	5.3828	0.0988
變異數	0.0603	1786.1633	3.9413	28.9748	0.0098
峰度	14.0303	707.1651	7.7139	0.2338	0.2061
偏態	2.7260	-22.9555	2.1910	0.3669	-0.3184
範圍	2.7034	1367.8178	18.3399	38.0000	0.6833
最小值	1.0000	-1368.1387	-0.9438	9.0000	0.2667
最大值	3.7034	-0.3209	17.3961	47.0000	0.9500

### 三、 機器學習配對交易投資組合實證結果

圖 4-12 為機器學習配對交易投資組合在樣本外期間的績效，可發利用類神經網路進一步的篩選進場時間點後，相較於單純使用布林通道策略之投資組合，最大虧損下降，原本策略 2015 年下半年較大的回檔也變得更小，但也因為進場條件更加嚴苛、進場時間更少，使加入類神經網路之策略的總報酬也隨之減少。

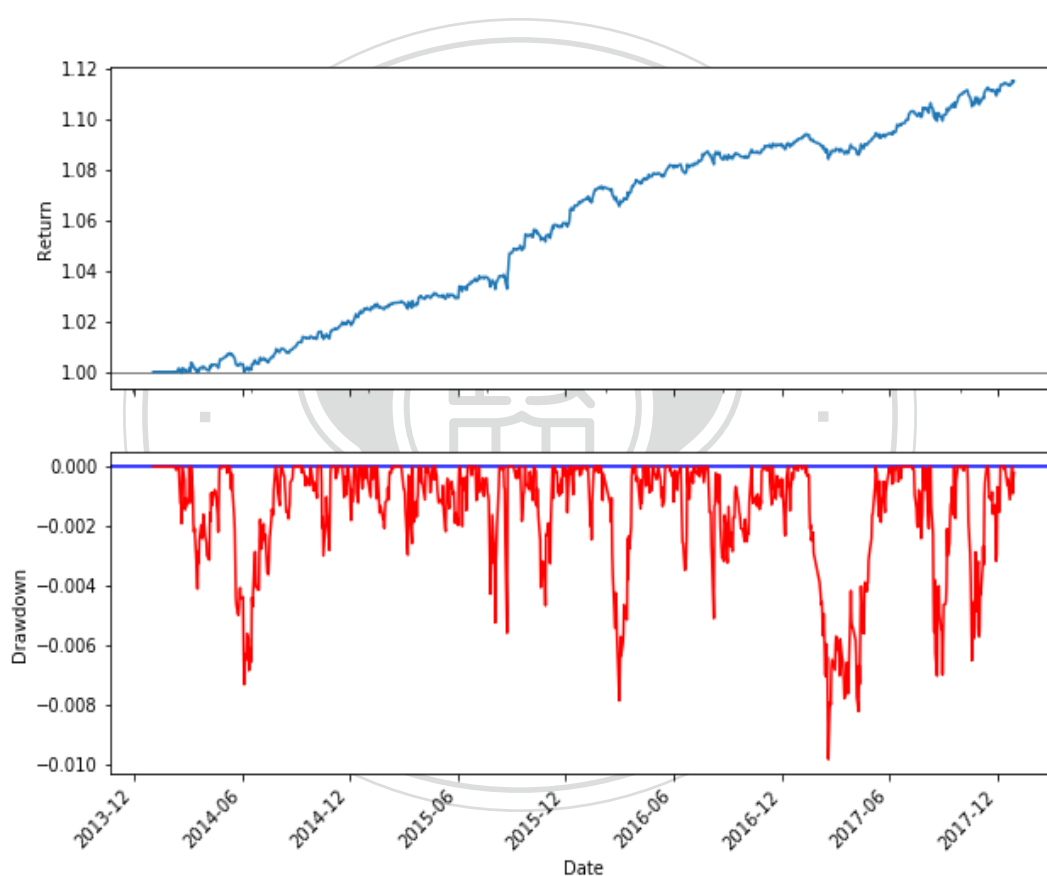


圖 4-12 機器學習配對交易投資組合績效圖與虧損圖

圖 4-13 為機器學習配對交易投資組合與台灣股價加權指數，可發現機器學習配對交易投資組合相對於加權指數，非常的穩定，波動極小，也不易產生巨額虧損；但也因此，從表 4-8 中可知其四年獲利僅有 11%，相對於加權指數的 26% 非常的少，但其最大策略虧損報酬及夏普值皆高於加權指數許多。

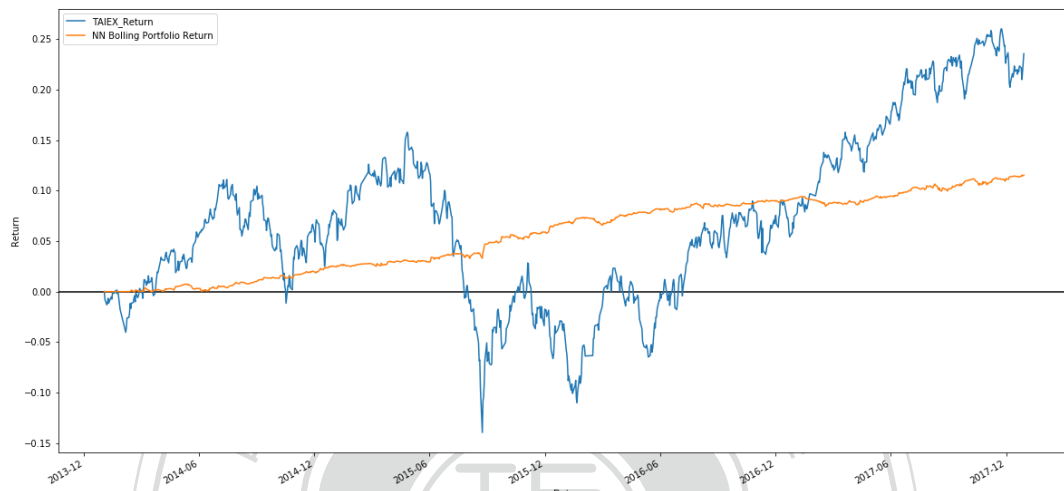


圖 4-13 機器學習配對交易投資組合與大盤走勢圖

表 4-8 機器學習配對交易投資組合績效參數

	最大報酬	最大虧損	最大策略 虧損報酬	平均日報 酬	日標準差	夏普值
機器學習	11.5479%	-0.9816%	11.7640	0.0111%	0.0803%	2.1951
布林通道	19.0208%	-1.6649%	11.4244	0.0178%	0.1148%	2.4513
加權指數	26.0322%	-29.7564%	0.8748	0.0246%	0.7814%	0.4980

表 4-8 為機器學習配對交易投資組合之績效與其他投資組合之比較，投資組合之最大策略虧損報酬有 11.7640，與布林通道配對交易投資組合相比有更好的績效表現，但增加並不明顯；夏普值為 2.1951，與布林通道配對交易投資組合相比稍為的下降，其原因可能是由於策略總報酬率相對減小過多而造成，可能是此策略主要的問題。

圖 4-14、圖 4-15、圖 4-16 為機器學習配對交易投資組合日損益分析、月損益分析及年損益分析。由圖 4-14 可知機器學習配對交易投資組合與僅使用布林通道投資組合一樣，日報酬雖然小，但以大於零的正日報酬多於負的日報酬；在圖 4-15 中，為機器學習配對交易投資組合每個月報酬以獲利的正報酬率月份居多，且幾乎沒有出現連續虧損多個月的問題，最多僅連虧損三個月，相較於僅使用布林通道之投資組合，同樣沒有連續虧損的問題，且虧損相對更小，但也出現了相對較多個月虧損的問題；在圖 4-16 中，機器學習配對交易投資組合與僅使用布林通道之投資組合同年獲利情況相似，此投資組合在樣本外四年內其每年的獲利狀況皆為正，且每年的獲利能力差異並不大，雖然同樣 2015 年有較明顯高於其他年份的獲利，但 2015 年和其他年份的獲利差異變小了。

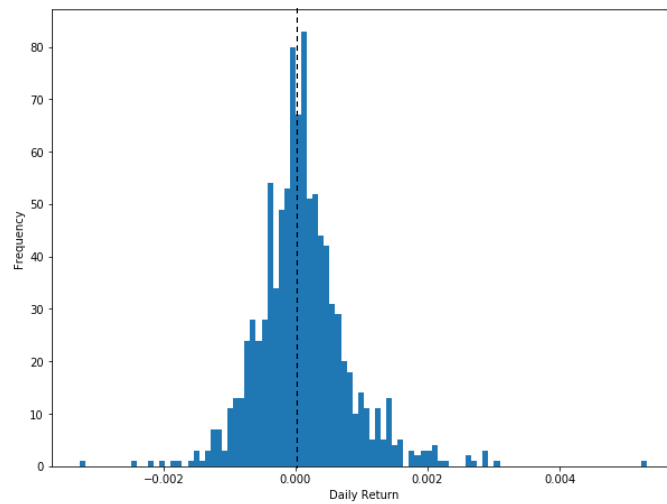


圖 4-14 機器學習配對交易投資組合日損益分析圖

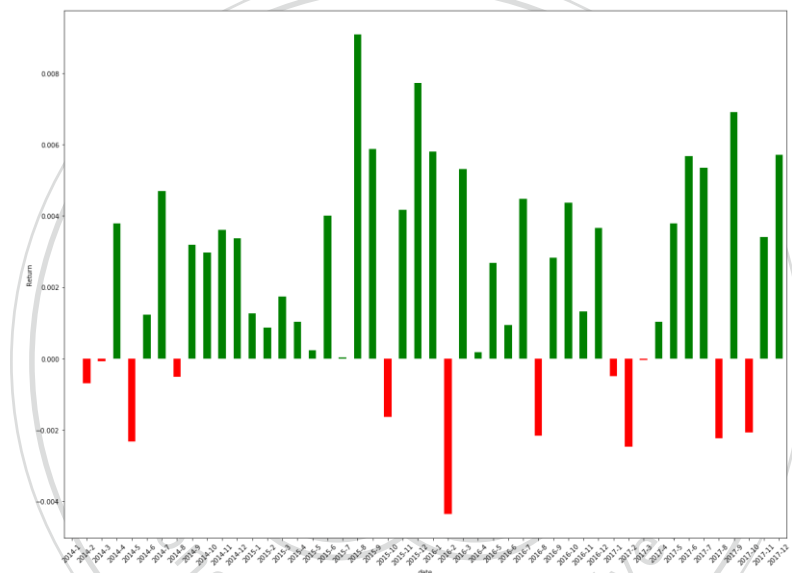


圖 4-15 機器學習配對交易投資組合月損益分析圖

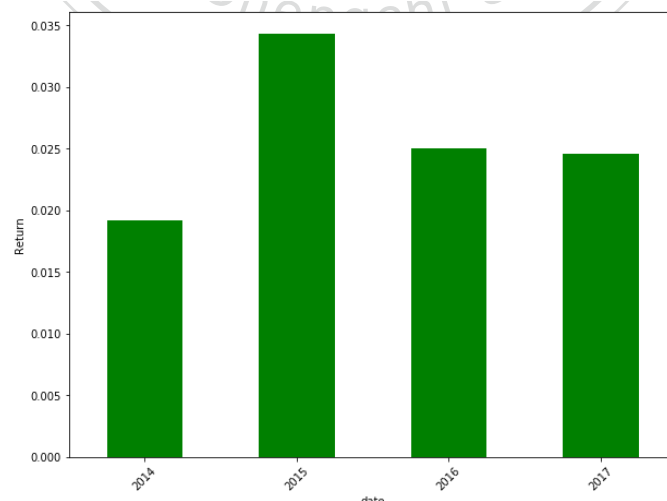


圖 4-16 機器學習配對交易投資組合年損益分析圖

## 第五章 結論

本研究根據 Vidyamurthy (2004)中的統計套利方法，依據 Engle and Granger (1987)二階段共整合檢定法，篩選配對交易組合，並利用技術指標及機器學習模型進行交易策略的建構，藉由交易策略績效，探討市場中立性報酬是否存在，並嘗試建構投資組合觀察此策略在市場中的穩定性。我們使用上述理論進行實證，使用台灣上市、上櫃股票作為實證標的，觀察配對交易策略投資組合在市場上之獲利情形。

根據實證結果，我們發現在樣本內期間 2007 年 1 月 1 日到 2013 年 12 月 31 日間，可以從台灣上市、上櫃股票中可以找到 1494 組通過檢定、相同產業且係數為正之配對交易組合，接著利用布林通道建構策略，發現布林通道確實能夠找出股票配對價格異常的時間點，並進場賺取異常報酬；在回測 1494 組配對交易組合後，嘗試建構投資組合，我們發現投資組合之績效優秀，最大策略虧損報酬為 11.4244，夏普值為 2.4513，相對於台灣加權股價指數其報酬波動度小，獲利穩定，為一個不錯的策略，但由於賺取的是市場中立報酬，投資組合整體之獲利較少，故此策略還有改良的空間。

我們嘗試利用利用類神經網路預測共整合殘差未來走勢，期望能改良布林通道配對交易策略，我們發現在類神經網路模型在訓練期內有約 54%的準確率，訓練期外的測試也有 53%，有足夠的穩定性；接著利用類神經網路結合布林通道建構機器學習配對交易策略，回測 1494 組配對交易組合，並嘗試建構投資組合，我們發現此測略相較於單純使用布林通道確實有較好的績效，最大策略虧損報酬為 11.7640 高於布林通道配對交易策略之 11.4244，但夏普值為 2.1951 小於布林



通道配對交易策略之 2.4513，確實利用類神經網路模型的多一層條件篩選降低了最大虧損，但也因此使投資組合總報酬下降，這也是夏普值下降的主要原因，故我們雖然利用類神經網路模型使最大虧損下降，但我們的獲利也隨之減少，這是一個可以進一步研究探討的問題。

本研究對於未來的研究提出了下面幾點建議與研究方向：

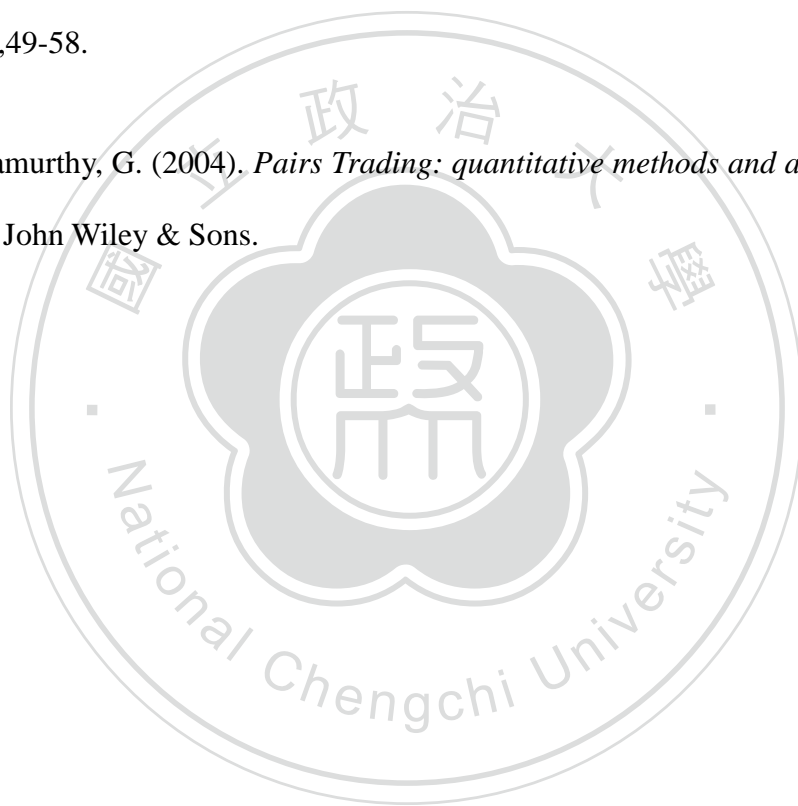
1. 嘗試使用不同的技術指標作為判斷股票價格異常的訊號，更可以進一步利用多個技術指標建構複雜度更高的配對交易策略。
2. 在篩選配對組合時嘗試利用更多的方法減少配對組合數目，如利用基本面資料等做進一步篩選，找出其中較好的投資標的，再進行投資組合建構。
3. 對共整合檢定方法進行更多測試，改變單根檢定使用的模型，除了改用 ADF 檢定並考慮落後期數等方法亦可考慮其他不同的檢定方式，找出共整合關係更穩定的股票配對。
4. 類神經網路模型可使用更多隱藏層的方式，利用深度學習(Deep Learning)嘗試提高預測能力，亦可嘗試使用 RNN 等網路結構更複雜的模型進行研究。
5. 投資組合的建構可以利用一些篩選機制，嘗試給每個配對交易組合不同的投資組合權重，使每組配對組合有不同的投資資金，讓績效好的配對組合更多的資金籌碼，賺取更多的報酬。
6. 嘗試將實驗期間進行改變，比較在金融市場波動度大及波動度小的時候，配對交易的獲利能力是否不同。

## 參考文獻

- [1] 沈宣佑(2015)。三檔股票交易設計並與傳統配對交易之績效表現比較。交通大學財務金融研究所學位論文，1-92。
- [2] 陳旭昇，2013。時間序列分析：總體經濟與財務金融之應用。臺灣東華。
- [3] 陳岱佑, & 王克陸. (2012)。台灣指數期貨與 ETF 價差交易之研究-以台股期貨，電子期貨，金融期貨與台灣 50ETF 為例。未出版之碩士論文，國立交通大學，財務金融研究所。
- [4] 羅君昱(2005)。台灣股票市場執行統計套利之可行性分析。未出版之碩士論文，國立政治大學，經營管理研究所。
- [5] Chen, W. H., Shih, J. Y., & Wu, S. (2006). Comparison of support-vector machines and back propagation neural networks in forecasting the six major Asian stock markets. *International Journal of Electronic Finance*, 1(1), 49-67.
- [6] Guenster, N., Kole, E., and Jacobsen, B. (2009). Riding bubbles, *Working paper*.
- [7] Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.
- [8] Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251-276.

- [9] Gatev, E., GOETZMANN, W., & ROUWENHORST, K. (1999). Pairs trading: performance of a relative value Arbitrage rule; *Working Paper 7032*, National Bureau of Economic Research, Cambridge.
- [10] Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2), 111-120.
- [11] Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and statistics*, 52(2), 169-210.
- [12] Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Applications*, 38(5), 5311-5319.
- [13] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [14] Madhavaram G. R. (2013) Statistical arbitrage using pairs trading with support vector machine learning. *Working paper*. Saint Mary's University.
- [15] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.
- [16] Phillips, P. C., & Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica: Journal of the Econometric Society*, 165-193.

- [17] Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. (1986). Sequential thought processes in PDP models. *Parallel distributed processing: explorations in the microstructures of cognition*, 2, 3-57.
- [18] Said, S. E., and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3), 599-607.
- [19] Sharpe, W. F. (1994). The sharpe ratio. *Journal of portfolio management*, 21(1), 49-58.
- [20] Vidyamurthy, G. (2004). *Pairs Trading: quantitative methods and analysis* (Vol. 217). John Wiley & Sons.



## 附錄

### 布林通道配對交易策略之產業分析

	配對 組合數	勝率		最大策略虧損報 酬		交易次數	
		平均	標準差	平均	標準差	平均	標準差
IC 設計	491	0.6497	0.0734	1.7977	1.8113	43.0550	3.5935
連接器	133	0.6392	0.0696	0.7044	1.5956	41.6541	3.5228
建設	104	0.6388	0.0720	1.4475	1.5402	41.6442	3.8058
零組件代理	69	0.6281	0.0692	1.5836	1.5845	43.4203	4.0198
機殼	53	0.6123	0.0776	0.4619	1.1365	41.6226	3.6702
被動元件	43	0.6136	0.0834	0.4865	1.3859	41.5349	3.0265
醫療耗材	42	0.6220	0.0719	0.9637	1.7470	41.2143	3.0567
電源供應器	28	0.6240	0.0610	1.3646	1.4518	43.6429	4.4490
系統整合	26	0.6398	0.0788	2.3711	3.0119	43.0385	4.5737
電機設備	26	0.6522	0.0709	1.5286	1.4018	43.4615	3.6795
西藥	25	0.6656	0.0807	1.9132	2.1341	41.6000	3.4521
IC 封測	21	0.6233	0.0872	1.0842	1.4414	42.4762	3.2805
飯店業	21	0.6714	0.0877	2.6627	2.9788	45.9048	3.3601
汽車組件	20	0.6662	0.0587	1.9477	3.2053	44.3500	2.8335
封裝及應用	18	0.6028	0.0668	1.0020	1.6227	41.8889	3.5956
營造	17	0.5627	0.0580	-0.1109	0.4427	40.9412	2.5853
織布	15	0.6441	0.0735	1.9071	2.7243	41.8000	3.6489
PCB	14	0.6240	0.0821	0.9464	1.3830	42.3571	3.3879
網路卡	12	0.6874	0.0505	3.7515	2.9201	43.5833	3.6546
磊晶及晶粒	12	0.6556	0.0536	1.4489	1.1803	41.3333	2.9644
樹脂	12	0.6571	0.0350	0.2142	0.5454	39.5000	2.9695
應用軟體	12	0.5996	0.0935	0.0325	1.3451	39.2500	2.3012
數據機	11	0.6480	0.0921	1.8526	2.1289	41.9091	5.2049
工業電腦	10	0.6130	0.0806	1.7074	3.0340	41.1000	3.1073
不銹鋼	10	0.6630	0.0666	1.2688	1.4890	42.7000	2.1628
照相機	10	0.6601	0.0839	2.0673	2.3880	45.4000	4.2479

	配對 組合數	勝率		最大策略虧損報 酬		交易次數	
		平均	標準差	平均	標準差	平均	標準差
天然氣	9	0.6628	0.0663	4.6221	2.8613	46.0000	2.2913
金融控股	9	0.6359	0.0984	2.7444	2.1497	39.8889	4.2850
工具機	8	0.6505	0.0717	1.4212	2.2814	40.3750	4.0686
3C 通路	7	0.6676	0.0738	2.9250	1.8582	44.1429	2.7343
太陽能電池	7	0.6195	0.0486	1.1929	1.1023	42.1429	2.9114
光纖通訊	7	0.6002	0.0836	0.2463	1.0036	43.8571	4.8107
其他化學	7	0.6685	0.0602	1.8723	2.9296	45.0000	5.3852
廠辦機電	7	0.6266	0.0552	0.9540	1.5492	45.0000	3.7859
證券	7	0.6320	0.0525	2.3113	2.6179	43.8571	3.5790
加工絲	6	0.6184	0.0801	1.4609	1.5359	39.3333	3.2042
線材	6	0.6004	0.0824	0.7619	0.6828	42.8333	3.8166
觸控面板	6	0.5467	0.0995	-0.5391	0.2067	38.1667	3.3714
PC 週邊	5	0.6367	0.0449	0.9493	1.8829	41.8000	3.8987
POS	5	0.6582	0.0818	1.7408	2.8672	40.0000	2.9155
油脂飼料	5	0.5452	0.0336	-0.4738	0.3516	41.2000	3.1145
散熱零組件	5	0.6164	0.0619	0.4165	0.7276	41.8000	1.9235
電線	5	0.5554	0.0858	-0.4920	0.2128	42.6000	4.8785
二極體整流	4	0.5727	0.0508	-0.1773	0.4784	38.0000	3.5590
水泥製造	4	0.6025	0.0862	2.2212	1.5210	44.0000	1.4142
石英元件	4	0.6686	0.0474	1.4292	0.3752	45.0000	2.5820
沖壓零組件	4	0.6437	0.0691	3.0097	2.3586	42.0000	5.4160
其他食品	4	0.6021	0.1041	2.7957	3.4000	40.2500	3.6856
其他電子	4	0.6540	0.0472	1.5173	1.7351	41.7500	2.6300
掃瞄器	4	0.5651	0.0435	1.1111	0.4139	45.0000	2.5820
貨運倉儲業	4	0.6147	0.0629	1.1722	1.0917	44.2500	4.8563
散裝輪	4	0.6893	0.0465	0.8397	0.7217	44.5000	3.6968
電機	4	0.6384	0.0497	1.5826	0.8207	45.0000	1.4142
鋼板	4	0.6595	0.0433	1.0000	1.0606	43.2500	2.0616
LCD 原料	3	0.5702	0.0788	-0.0110	0.6007	40.6667	3.2146
PCB 上其他	3	0.6490	0.0181	1.3404	0.9961	42.6667	3.0551
本國銀行	3	0.6746	0.1249	6.0124	2.3710	45.6667	2.8868
印刷	3	0.6587	0.1311	3.8354	4.0054	41.0000	1.7321

	配對 組合數	勝率		最大策略虧損報 酬		交易次數	
		平均	標準差	平均	標準差	平均	標準差
泛用塑膠	3	0.6368	0.1161	3.8960	2.1273	47.0000	1.0000
背光模組	3	0.6456	0.0557	2.3123	1.8472	44.3333	2.3094
音響設備	3	0.6679	0.0470	1.9935	0.9947	46.3333	4.9329
貨櫃輪	3	0.6305	0.1121	2.0813	0.5311	40.0000	4.0000
軟板	3	0.6354	0.0892	1.5970	1.4935	42.6667	2.0817
通訊設備	3	0.6376	0.0197	2.3826	2.1912	40.6667	5.5076
造紙	3	0.5709	0.0718	0.2790	0.5980	41.0000	2.6458
3C 二次電池	2	0.6555	0.1337	1.4365	1.9562	42.5000	2.1213
FCCL	2	0.6512	0.0658	-0.3512	0.3331	43.0000	0.0000
LCD-其他	2	0.4737	0.0000	-0.8165	0.0260	38.0000	0.0000
NB	2	0.6699	0.1133	1.8352	2.9300	41.5000	3.5355
光碟	2	0.6322	0.1498	0.3301	0.7213	40.0000	2.8284
百貨購物中	2	0.7049	0.0227	1.3532	1.3854	44.0000	1.4142
其他零組件	2	0.5220	0.1321	-0.6205	0.0246	37.0000	2.8284
玻纖紗. 布	2	0.6018	0.1132	-0.9627	0.0128	45.0000	1.4142
記憶體模組	2	0.6053	0.0744	-0.0140	0.2210	38.0000	0.0000
晶圓材料	2	0.5212	0.0607	-0.3424	0.0924	42.5000	4.9497
貴金屬回收	2	0.7374	0.0425	4.8815	1.8425	42.0000	1.4142
監控系統	2	0.6533	0.1213	0.1841	0.5847	41.5000	6.3640
監視器	2	0.6872	0.0181	2.8760	0.8923	41.5000	2.1213
餐飲業	2	0.5647	0.1106	0.8698	1.4259	39.5000	3.5355
主機板	1	0.5581	NA	-0.6673	NA	43.0000	NA
加油站	1	0.6458	NA	0.3239	NA	48.0000	NA
生物科技	1	0.5854	NA	2.2684	NA	41.0000	NA
自行車組件	1	0.6905	NA	-0.4111	NA	42.0000	NA
其他服務	1	0.6000	NA	-0.6844	NA	45.0000	NA
矽晶圓	1	0.5250	NA	-0.4440	NA	40.0000	NA
金屬製品	1	0.5581	NA	0.3594	NA	43.0000	NA
便利商店	1	0.5476	NA	1.3605	NA	42.0000	NA
保全業	1	0.6667	NA	0.2963	NA	42.0000	NA
玻璃陶瓷	1	0.4750	NA	1.5014	NA	40.0000	NA
家電	1	0.6279	NA	-0.3751	NA	43.0000	NA



	配對 組合數	勝率		最大策略虧損報 酬		交易次數	
		平均	標準差	平均	標準差	平均	標準差
特用 IC	1	0.8043	NA	2.4893	NA	46.0000	NA
產險業	1	0.7442	NA	9.0158	NA	43.0000	NA
貿易業	1	0.6596	NA	2.0022	NA	47.0000	NA
塑膠加工	1	0.5000	NA	-0.0333	NA	38.0000	NA
電視廣播	1	0.5128	NA	-0.5135	NA	39.0000	NA
網路服務	1	0.6905	NA	2.7725	NA	42.0000	NA
聚酯絲棉	1	0.7568	NA	3.8083	NA	37.0000	NA
模具零件	1	0.6905	NA	2.9131	NA	42.0000	NA
鞋及旅行箱	1	0.7143	NA	0.5244	NA	42.0000	NA
橡膠原料	1	0.6410	NA	0.2353	NA	39.0000	NA
鋼筋	1	0.6500	NA	-0.4261	NA	40.0000	NA
鋼管	1	0.6829	NA	-0.3778	NA	41.0000	NA