

# 演化式計算於共同基金投資組合建構與交易策略挖掘之研究

黃怡婷

輔仁大學資訊管理研究所  
研究生

leenallison821@gmail.com

林文修

輔仁大學資訊管理研究所  
教授

wslin1949@gmail.com

## 摘要

本研究以遺傳演算法與基因表達規劃法為基礎，期望以演化式計算為投資人建構最適基金投資組合及交易策略的推薦模型。本研究採用具有組合最佳化能力的遺傳演算法來建構 TSGA 模型，模型中藉由納入不同的基金類型來建構投資組合，增加投資組合的投資標的，以降低投資組合風險。而交易策略的部份則使用遺傳演算法與基因表達規劃法共演化的方式來建構交易策略，希望將演化式計算於雙軌投資策略上，能使投資策略更負彈性。在研究中發現模型無論在訓練期或是測試期均具有良好的穩定度，且研究結果顯示 TSGA 單筆投資的平均績效表現相較累積績效前十名的基金佳，而由 Co-GAGEP 所建構出來的交易策略的投資績效也比定期定額來得好，故此 TSGA 與 Co-GAGEP 模型在實務上具有其適用性。

**關鍵詞：**共同基金、遺傳演算法、基因表達規劃法、投資組合、交易策略。

## 1. 前言

近年來國內共同基金漸漸成為投資人重要的投資工具之一，國內投信基金的定期定額扣款人數在今年年初也開始屢創新高，根據中華民國證券投資信託暨顧問商業同業公會在 2010 年 9 月公佈的新數據，定期定額扣款人數高達 581,130 人，較上月增加 2,491 人，創下 25 個月來新高；且國內的境內基金的資金規模也來到 1.84 兆元，比 8 月增加了 160.56 億元。

根據 Lipper 統計至 2010 年 9 月底止，過去兩年有 936 檔基金為正報酬。由此可知，從 2008 年的金融海嘯以來，投資人若能選對基金且把握時機逢低買進，在市場低迷時建立部位，便可得到相當可觀的獲利。本研究認為其中有三點值得注意：第一，雖然有高達 936 檔基金為正報酬，但並非全數基金都是正報酬的情形，因此如何挑選出好的基金做投資，是獲利與否的關鍵之一。然而，共同基金的種類與數量眾多，若要建構出一個好的投資組合，也需考慮配合不同的投資標的，才能藉此達到真正分散風險的目的。第二，定期定額的投資缺點在於當情勢屬於上漲波段時，會隨著漲勢不斷而墊高投資成本，倘若當時能將單筆投資納入操作，使定期定額與單筆投資相輔相成，想必獲利應該更可觀。第三，從 2008 年的金融海嘯到現今的景氣回升，更突顯了景氣的興衰交替，也是影響投資報酬率重要的因素之一，故投資人應善加利用景氣循環的特性，將之納入投資策略的考量因素中。

根據以上評比機構的統計數據與景氣循環相關文獻(Brocato & Steed, 1998; 劉文祺等, 2000; 張桂莉, 2000; 陳嘉惠等, 2002; 張婉蘭, 2002)，本研究認為投資人可利用景氣多空循環的特性，建立分散風險的投資組合，再配合一套適當的交易策略，便可獲得更好的報酬。近十年投資領域的研究上，已有不少研究使用各類的柔性計算方式，例如遺傳演算法(Genetic Algorithm, GA)、遺傳規劃法(Genetic Programming, GP)、類神經網路(Neural Network, NN)、粒子群最佳化演算法(Particle Swarm Optimization, PSO)等，近年來更有新的結合技術基因表達規劃法(Genetic Expression

Programming, GEP), 應用於投資組合、資產配置及擇時等研究 (劉貴強, 2004; 王冠弼, 2006; 陳與泳, 2004; 李曉玲, 2008; 林玉萍, 2008; 陳愷瑜, 2008; 劉金宜, 2009), 都有不錯的成效。因此, 本研究將使用柔性計算的相關技術, 為投資人建構包含投資組合與交易策略的共同基金推薦模型。

## 2. 文獻探討

### 2.1 遺傳演算法 (Genetic Algorithms, GA)

#### 2.1.1 遺傳演算法基本原理

遺傳演算法是 John Holland(1975)在“Adaptation in Natural and Artificial Systems”一書中所提出, 其論點主要是運用達爾文的進化論中「物競天擇, 適者生存」的概念所發展而來, 模擬自然界的演化方式, 尋找出問題的最佳解。遺傳演算法有三種機制, 分別為選擇 (Selection)、交配 (Crossover) 與突變 (Mutation), 其目的是為了模擬自然界的演化過程, 最後利用適應函數來評估個體的優劣, 擁有較佳適應值的個體將會有較大的機會存活, 接著繁衍下一代, 即如同自然環境中的淘汰機制。而本研究所使用到的選擇方式為輪盤法, 是在 1989 年由 Goldberg 所提出, 其作法是將適應值加總, 並依個體適應值與全體適應值的比值來做為被選中的機率, 進而進入下一代演化。

#### 2.1.2 GA 相關文獻

過去也有不少研究將遺傳演算法使用到共同基金領域上, 其中也獲得不錯的成效。其中, 王冠弼(2006)利用遺傳演算法結合動態模糊化策略, 建構指數型基金的投資組合, 研究結果發現遺傳演算法確實能建構出追蹤績效不錯的投資組合。

而在建構組合型基金商品的研究中, 劉貴強(2004)利用遺傳演算法在組合型基金商品設計的研究上, 欲在研究中建構一個「好」的組合型基金商品, 其所設計的

系統可在短時間內產生一個組合型基金商品, 其可隨環境變化而調整且獲利性佳的組合型基金, 且系統具有高度的穩定性。但其研究對象只限於股票型、債券型及平衡型三種基金類型。另外, 陳愷瑜(2008)利用共演化遺傳算法於全球股票型基金資產配置, 研究結果發現, 確實能找出優於傳統遺傳演算法與打敗市場的投資組合。但以上兩個研究只針對股票型基金做為研究對象, 故在資金分配中並無使用其他標的。

綜合以上研究, 無論是在投資組合或資產配置上, 遺傳演算法運用於共同基金這項投資工具, 都有相當不錯的成效。然而, 本研究發現在這些研究中, 大部份的研究只針對股票型基金做為研究對象, 但在實務上, 不同的標的才能達到真正分散風險的目的, 故本研將會考量到不同的共同基金投資標的, 來為投資人建構一個分散投資標的的基金投資組合。

### 2.2 基因表達規劃法 (Genetic Expression Programming, GEP)

#### 2.2.1 基因表達規劃法基本原理

基因表達規劃法是近年來由葡萄牙學者 Ferreira(2001)以生物遺傳的基因表達規則為借鏡, 結合 GA 與 GP 的優點, 其作法是以類似 GA 的固定長度的染色體作為基因型, 以類似 GP 的樹狀結構作為表現型, 實現了簡單編碼解決複雜問題的目的。傳統遺傳演算法採用簡單固定長度的編碼方式, 操作簡易, 但在解決較複雜的問題時較具難度, 而遺傳規劃法是採用非線性、不等長的資料結構來編碼, 較能解決複雜問題, 但在操作上也較複雜。

自動定義函數為 Koza(1992)首先在 GP 所提出的概念, 主要將演化過程中較複雜或較常使用的子樹, 利用封裝的概念, 將之轉為一個 ADF, 其用意是在於可重新使用編碼, 節省空間上的浪費, 另一方面也能保護某些區塊使之不會在演化過程中被破壞, 藉此增加 GEP 的演化效率。Ferreira(2006)把 ADF 結構應用於 GEP 上,

將每個基因視為一個 ADF，並取消 ADF 在個數上的限制，使得因搭配 ADF 的結構讓染色體的變化更自由，圖 1 為示意圖 (Ferreira, 2006)，其中 a 部份為染色體的基因型，非粗體的部份構成 ADF，b 部份為每個 ADF 的表現型，而 c 部份為染色體粗體部份的表現型。

a. 012345601012345601012345601012345601234560123  
/+a/abbba+\*-abbabb/b\*+abaab+Q/0\*210212

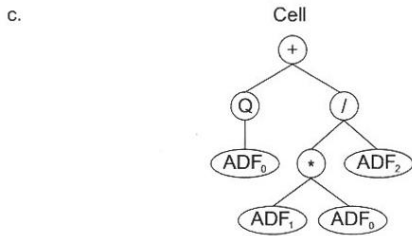
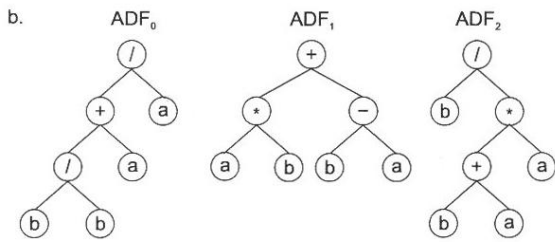


圖 1 ADF 示意圖

### 2.2.2 基因表達規劃法的編碼形式

GEP 的編碼採用等長的線性符號串，每個 GEP 的基因是由頭部和尾部所組成，頭部可包含內部節點與終端節點，但尾部只能包含終端節點。內部節點指的是運算符號或是一個子程式的結構，而終端節點為程式中的輸入、常數或是沒有參數的函數。

GEP 基因的長度是先定義出頭部的長度，再算出尾部的長度，然後再將頭尾加總而得來。而頭部的長度通常取決於要解決的問題，尾部的長度可從下列公式算出：

$$t = h \times (n_{\max} - 1) + 1$$

其中，t 為尾部長度，h 為頭部長度， $n_{\max}$  為所使用的函數集需要最多變量的函數的參數個數。例如終端節點集合為  $\{+, -, \times, /\}$ ，則

$n_{\max} = 2$ 。假設  $h=4$ ，由上式算出  $t$  為 5，則基因的總長度  $g$  為  $4+5=9$ 。所以一個合法 GEP 的基因型如圖 2 所示，其中粗體字表示尾部，而表現型如圖 3 所示。

012345678  
**\*\*a+ababb**

圖 2 基因型

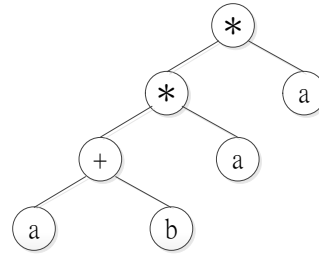


圖 3 表現型

### 2.2.3 GEP 相關文獻

由於 GEP 是一個相當新的演算法，在研究上的數量還不甚豐富，目前 GEP 的相關研究有用於建構分類器 (陳玲蘋, 2008) 以及建構時間序列預測模型 (吳佳玲, 2008) 都有不錯的效果，其中本研究在交易策略模組的染色體設計也有部份參考到陳玲蘋在 2008 年建構分類器的研究。在財務領域上目前也有一些以 GEP 為基礎的研究，也有不錯的研究結果。其中，蔡慧菊 (2010) 以 GEP 為基礎提出 DGEP (Dynamic GEP)，進行台股期貨的價格發現，DGEP 是將 GEP 結合模糊理論動態調整突變率的方式，來改善原本 GEP 模型在收斂時易陷入區域最佳解的問題。研究結果經由成對樣本 T 檢定，發現 DGEP 與 GP、GEP 模型有顯著不同，在執行效率上也明顯優於 GP。

另外，陳奕帆 (2010) 提出 CO-EC 模型，運用 GEP 與 GA 共演化的機制，建置以台灣股票為基礎的股票型共同基金資產配置模型。研究顯示，CO-EC 模型建置出來的投資組合報酬率高於以傳統 GA 建置出來的投資組合報酬率。然而，目前尚未看到 GEP 運用於基金投資策略的研究，故本研究期望能使用 GEP 建構出基金的交易規則。

## 2.3 共同基金與投資組合

### 2.3.1 共同基金簡介

共同基金是在 1822 年起源於荷蘭，台灣的第一家指證券投資信託公司(以下簡稱投信公司)則是在 1983 年成立，到了 1992 年投信公司開始如雨後春筍般陸續成立。共同基金是指由投信公司以發行股份或受益憑證的方式，募集證券投資基金，委託專業經理人進行投資，並由銀行保管該基金資產。

共同基金的特性為投資門檻低，故能讓小額投資人能用小錢進行投資，藉由共同基金來達到規模效益，也就是匯集小額投資人的資金投資多種標的，所以投資人相當於花小錢就能擁有許多標的，另一方面也可以達到分散風險的目的。

自 1960 年代以來，許多學者紛紛針對共同基金的績效提出評估方法(Treynor, 1965; Sharpe, 1966; Jensen, 1968)，至今這些指標還是被廣為使用，本研究也採用這些指標做為挑選基金的依據之一。

### 2.3.2 投資組合

「投資組合」是指投資人以不同資產所組成的財富，如所擁有的資產為有價證券，即為「證券投資組合」(徐俊明, 2007)。Treynor et.al(1973)建議，最適的投資組合應包括積極與保守的組合，且 O'Neal(1997)指出，含有多重基金的投資組合風險遠低於持有單一基金之投資組合。故本研究將會著重在分散基金投資組合的投資標的，以期建構出包含多種資產的投資組合來分散風險。

最早提出投資組合概念的是 Markowitz (1952)，他所提出的「效率前緣(Efficient Frontier)」，其代表意義為：「在總風險相同時，相對可獲得最高之預期報酬率」或「預期報酬相同時，相對上總風險最低」之投資組合所連成的線，即為效率前緣。其假設投資人為理性且具有趨避風險等特性，故在承受同樣風險下，投資人

會尋找報酬較高之投資組合，或是在同樣報酬率的情況下，投資人會偏好風險較低的投資組合。然而本研究會遵循著 Markowitz 的效率前緣理論，在建構投資組合時會同時考慮風險與報酬。

### 2.3.3 遺傳演算法於投資組合相關研究

目前有許多將遺傳演算法使用在共同基金投資組合的研究，都有不錯的效果。如 Lin et. al(2008)使用遺傳演算法來建構具有效率的組合型基金，以最大化投資組合報酬率及最小化投資組合風險為目標，研究中設計四種模型來比較不同限制條件下所構成的投資組合的差異，實驗結果發現使用 Sharpe 指標挑選出來的基金組合，有較好且較穩定的報酬績效。Kin et. al(2006)運用兩階段的遺傳演算法建構股票的投資組合，階段一先對所有股票做排序，階段二再演化投資組合中的股票權重，結果顯示有使用 GA 設定權重的投資組合有較高的報酬。

除了單純的使用遺傳演算法之外，也有結合不同方法的相關研究，Jeurissen 和 van den Berg (2008)利用遺傳演算法結合數學的二次規劃法，試圖找出能與 AEX 指數的追蹤誤差最小的投資組合。遺傳演算法主要用於在 AEX 指數的成份股中，尋找出適合的股票組合；而二次規劃法即為找出最適合之權重值。研究結果發現，此研究所使用 GA 與二次規劃法的方式所建構出來的投資組合表現較隨機選取投資組合來得好四倍。從相關研究可發現到，使用具有最佳化能力的遺傳演算法來建構投資組合均有良好的成效。

## 3. 研究方法

### 3.1 研究架構

本研究的研究目的為利用演化式計算為投資人建構最適投資組合及最適交易策略的基金投資策略推薦模型，如圖 4 所示。故在最適投資組合建置的部份，將會利用

兩階段遺傳演算法 (Two Stage Genetic Algorithm, TSGA)，第一階段先演化最適投資標的組合，第二階段再挑選適合的基金與資金配置，利用此兩階段的方式建構出最適投資組合。在最適交易策略部份，本研究將透過共演化遺傳基因表達規劃法 (Co-Evolution Genetic Algorithm & Genetic Expression Programming, Co-GAGEP) 演化出最適投資策略。

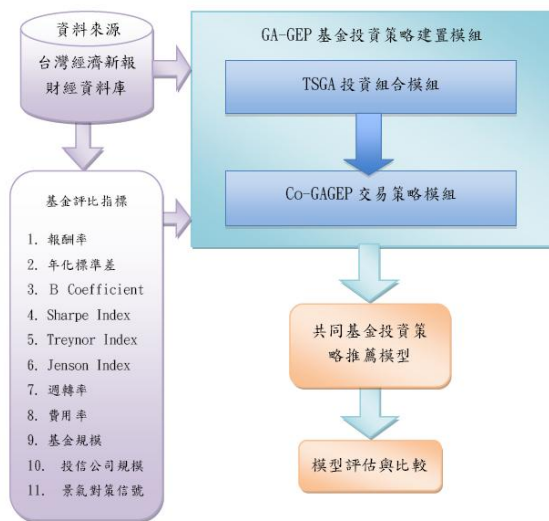


圖 4 系統架構圖

本研究資料來源取自於台灣經濟新報基金資料庫，因模型欲參考穩定的中長期績效，故以國內投信成立發行超過三年開放型的共同基金資料為主要研究對象，其中若資料不齊全的基金，將不列入研究對象。

### 3.2 TSGA 投資組合模組設計

在投資組合建置中，由「資產配置模組」與「基金挑選配置模組」所組成。因投資組合為組合最佳化問題，故此部份使用具有組合最佳化能力的遺傳演算法。一般投資人在挑選基金時，通常會考慮到基金的投資標的與投資區域，但本研究的對象為國內投信發行且投資國內的基金，故本研究只考慮投資標的情況下，替投資人演化出最佳的投資類型配置，再進一步挑選最適基金納入投資組合中，整個過程為一

個兩階段的遺傳演算機制，圖 5 為演化流程示意圖。

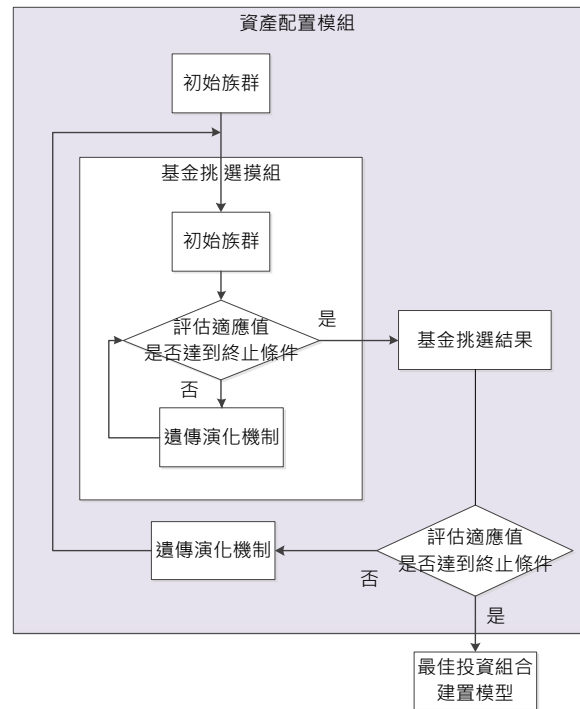


圖 5 投資組合建置流程圖

### 3.2.1 資產配置模組

因考慮到投資標的與投資地區，本研究將參考中華民國證券投資信託暨顧問商業同業公會截至 2010 年 9 月對於境內基金之分類以及台灣經濟新報的分類，在考量同類型基金數量關係下，將其整理過後分為 7 大分類，分別為：國內股票型、國內科技股、國內小型股、指數型、債券型、平衡型與現金。



圖 6 資產配置染色體編碼示意圖

本研究針對一般投資大眾為主，故設計投資組合最多納入 10 支基金，故染色體的長度為 10 個區塊，代表的是投資類型情況，其中包含分類代碼與是否包含該分類，長

度分別為 1、3 與 1，單個基因區塊的長度為 5，染色體總長度為 50，其編碼方式如圖 6 所示。

### 3.2.2 基金挑選配置模組

本研究以期為投資人提供有效分散風險的投資組合，故將投資組合數目設定為 3 到 10 支基金。因此染色體由 10 個基因所組成，每個基因代表一支基金是否被挑選以及配置的情形，每個基因內包含基金代碼、持有比例以及是否持有該基金，長度分別為 4、5 與 1，單個基因區塊的長度為 10，故整個染色體總長度為 100，染色體設計如圖 7 所示：

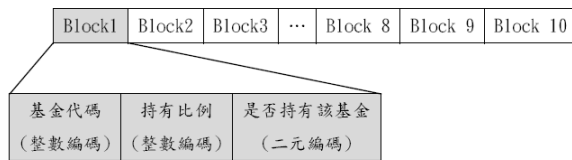


圖 7 基金挑選配置染色體編碼示意圖

### 3.2.3 適應函數

適應函數會同時考量到基金評比指標與景氣循環指標，由於為兩階段的建置方式，所以會有兩個適應函數。根據馬可維茲的現代投資學理論，本研究會在基金適應函數的設計中，同時考量到投資組合的風險與報酬。

在基金挑選模組的適應函數設計將基金每月的平均報酬率扣掉無風險利率後，再除以各支基金的年化標準差，最後加上每個基金各項評比指標分數，乘上各別的基金權重。此處的無風險利率均為台灣銀行定存的年利率。基金挑選配置模組適應函數如以下公式所示：

$$F1 = \sum_{i=1}^n \frac{R_i - R_f}{\sigma_i} + I_a$$

其中， $R_i$ : 第  $i$  支基金報酬率

$\sigma_i$ : 第  $i$  支基金年化標準差

$R_f$ : 無風險利率

$I_a$ : 基金評比指標總分

$$I_a = \sum_{i=1}^m \left( \sum_{j=1}^k I_{aj} \times w_i \right)$$

$i$ : 投資組合中第  $i$  支基金

$j$ : 基金評比指標

$w_i$ : 投資組合中第  $i$  支基金權重值

基金挑選配置模組前 10% 最佳的適應值平均將會以回饋的方式，加入資產配置模組的適應值當中，以達到評估投資組合整體績效的目的，再加上投資組合當中每支基金類型的相關係數，因考慮到相關係數最後的值可能會過小而在適應值中不具影響力，故乘以一個權重值使之變大，最後加上景氣指標分數。資產配置適應函數如以下所示：

$$F2 = F1 + \rho_p \times w_p + I_b$$

其中， $R_p$ : 投資組合報酬率

$\rho_p$ : 投資組合的相關係數

$w_p$ : 相關係數權重值

$I_b$ : 景氣循環指標總分

$$I_b = \sum_{i=1}^m \left( \sum_{j=1}^n I_{bj} \times w_i \right)$$

$i$ : 投資組合中第  $i$  支基金

$j$ : 投資標的對應分數

$w_i$ : 投資組合中第  $i$  支基金權重值

### 3.3 Co-GAGEP 模組設計

本研究設計此交易策略模組以單筆投資策略為目的，與先前投資組合建置部分的定期定額配置相互配合，期望為投資人建構定期定額與單筆投資的綜合策略。本研究將交易投資策略分為買入與賣出規則。本研究擬使用遺傳演算法先演化出多個單筆交易策略，但由於策略應是一套的交易策略，故再將演化出來的單筆投資策略，視為基因表達規劃法中個別的自動定義函數(ADF)，使之建構為基因表達規劃法的終端節點部份，故此模組為遺傳演算法與基

因表達規劃法的共演化(Co-Evolution)機制所運作，其共演化運作方式圖 8 所示。

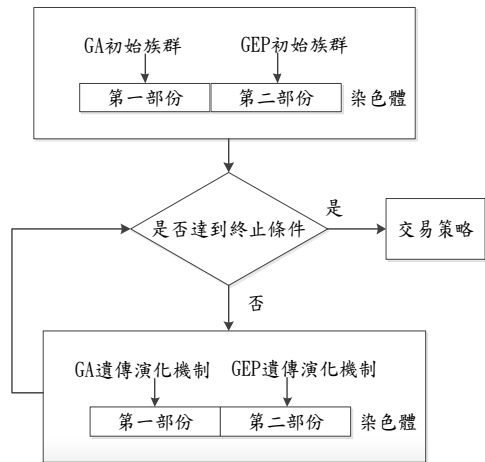


圖 8 GA 與 GEP 共演化機制運作方式

### 3.3.1 編碼方式

染色體編碼方式為二元編碼，分為兩部份，第一部份由遺傳演算法負責此處的自動定義函數的演化，自動定義函數的內容包含了基金影響因子、比較運算元與影響因子之數值，自動定義函數是由相對應的位置所組成。在第一部分染色體區塊的最尾端基因代表訊號強度單位，採分級制，分成強、中、弱三等級。

第二部分由基因表達規劃法負責此處的演化，頭部長度為 4，尾部長度則為  $4 \times (2-1) + 1 = 5$ ，終端節點為 1~4 的常數，其代表第一部分的自動定義函數，而內部節點為 {AND, OR, NOT}，此部份演化的目的是要將自動定義函數組合成一套交易策略。在此處的交易策略共有二棵子樹，分別代表的是買入、贖回策略，以 + 做為連結函數。而所構成買賣訊號的基金影響因子為：基金報酬率變動率、景氣對策信號值、年化標準差、停損點及停利點。染色體設計如圖 9 所示，每一條代表一個子樹，而表現型如圖 10 所示。

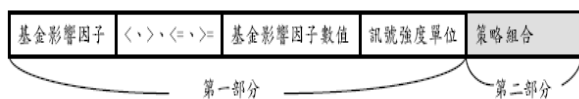


圖 9 交易策略染色體編碼示意圖

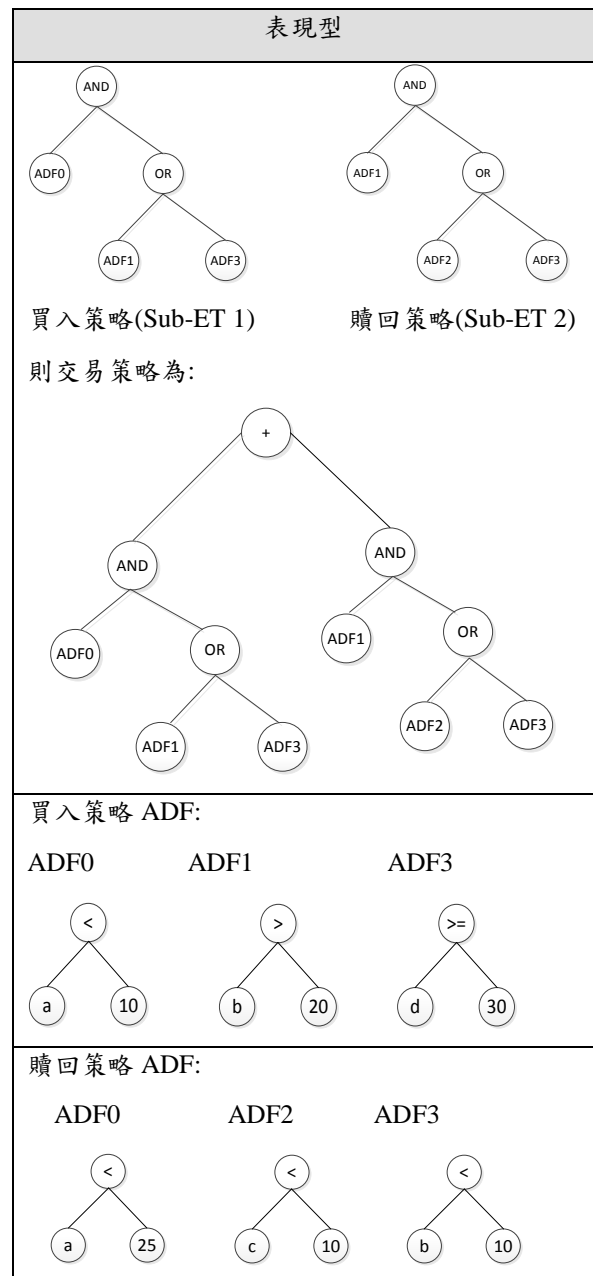


圖 10 表現型

因考量到每次單筆交易的金額不應該相同，故本研究設計訊號強度的來算出此交易策略的投資比例，投資比例的計算方式為訊號強度平均值乘上交易金額單位，因考慮到單筆投資有 1 萬的最低限制，故本研究將交易金額單位設為 1 萬。訊號強度由弱到強為 1、2、3。若訊號強度單位為 10，代表弱訊號比例為 0%~10%，中訊號比例為 10%~20%，強訊號比例為 20% 以上。

訊號強度的判別，會依照基金對於買賣規則的符合超過程度來界定，以各項條

件分開計算，再予以加總平均，即為此次的訊號強度平均值。以上面例子來說，若符合程度小於 10%，則判定為弱訊號，給予其條件 1 分；符合程度界於 10%~20%之間，視為中訊號，給予其條件 2 分；而符合程度超過 30%則視為強訊號，給予其條件 3 分。最後將所有條件得分加總並除以條件個數，所算得即為訊號強度平均值，依此平均值判定此交易訊號的強度值。以下舉一詳細例子說明：

買入策略為「若基金報酬率變動率 <15% 且年化標準差 <10% 則買入」，且訊號強度單位為 20。

故依此例子中計算訊號強度比例：弱訊號比例為 20% 以下，中訊號比例為 20%~40%，強訊號比例為 40% 以上。

若現有某支基金滿足此買入策略，且報酬率變動率為 16%，年化標準差為 0.05，則此基金報酬率變動率超過符合程度為  $(0.16-0.15)/0.15=6\%$ ，視為弱訊號；年化標準差超過符合程度為  $(0.1-0.05)/0.1=50\%$ ，則視為強訊號，故此基金的買入訊號平均強度值為  $(1+3)/2=2$ ，故此單筆投入 2 萬。

### 3.3.2 適應函數

本研究將以執行此投資策略的報酬率與標準差列入適應函數當中。適應函數的設計如以下公式所示，以交易策略的模擬投資做為報酬率，以每支基金分開算的方式再加總。

$$F = \frac{R_p - R_f}{\sigma_p}$$

其中， $R_p$ : 投資組合報酬率  
 $R_f$ : 無風險利率  
 $\sigma_p$ : 標準差

## 3.4 參數設定說明

### 3.4.1 TSGA 參數設定

本研究依照染色體內部的編碼方式不同，所採用的方法也會不同，選擇方式使用菁英法與輪盤法。其中菁英法為前 10%

直接挑選進入下一代。交配則採兩點交配的方式，整數編碼與二元編碼分成兩塊個別運作。突變方式為在染色體整數編碼的部份，隨機選擇兩個基因，將內部的值相加除以二，變成兩個新的值。而二元編碼則為單點突變，即將選到的基因 1 變 0，0 變 1。TSGA 中，兩個模組所使用的參數值相同，如表 1 所示。

表 1 TSGA 參數設定表

參數名稱	模組參數值
族群大小	50
終止條件	50
交配率	0.6
突變率	0.08

### 3.4.2 Co-GAGEP 參數設定

選擇方式使用菁英法與輪盤法。其中菁英法為前 5% 直接挑選進入下一代，族群大小為 200，終止條件為 300 代，其餘參數設定如表 2 及表 3 所示。

表 2 交易策略 GEP 參數設計表

演化參數	參數值
選擇方式	菁英法與輪盤法
基因重組率	0.3
雙點重組率	0.3
基因轉位率	0.1
突變方法	隨機替換基因片段
突變率	0.4
IS 轉位率	0.1
RIS 轉位率	0.1

表 3 交易策略 GA 模組參數設定表

參數名稱	參數值
交配方法	多點交配
交配率	0.6
突變方法	單點突變
突變率	0.05



## 4. 實驗設計

### 4.1 研究對象

以國內投信成立發行超過三年且投資區域為國內的開放型共同基金資料為主要研究對象，期間從 2007 年 1 月 1 日至 2010 年 12 月 31 日止，共計 48 個月。研究資料取自於台灣經濟新報所提供之基金資料庫與總體經濟資料庫，經刪除資料不齊全的基金，共計 213 支基金。

### 4.2 實驗設計與結果分析

#### 4.2.1 實驗一：模型訓練測試長度分析

在模型適用性實驗中分為兩個階段，目的是為了找出適合模型的訓練期長度與測試期長度。首先採用不同訓練期，以相同測試期的方式，找出模型的較合適的訓練期長度。接著以兩種訓練期所演化出來的投資組合，在測試期間的績效來決定實驗第二階段模型的訓練期長度，再以不同的測試期長度來找出模型較合適的測試期長度。訓練方式均使用移動視窗的方式，一次移動 3 個月，各執行 3 次。

##### (1) 訓練期長度分析

因基金投資屬於較中長期的投資，故本研究的訓練期的設定較長，分別設定為 24 個月與 18 個月，而測試期同為 12 個月。表 4 為訓練與測試期區間說明。

第一階段測試訓練期將選出來的投資組合做定期定額與單筆投資的比較，實驗結果呈現於圖 11 與圖 12。顯示訓練期 24 個月的投資組合，在未來的測試期 12 個月內，每次的表現比訓練 18 個月的績效較佳。TSGA 在訓練 24 個月後，測試期間在單筆投資平均報酬率較訓練 18 個月的平均報酬率約高了 4.3%，而在定期定額平均報酬率也高了大概 1%，且在 3 次執行中，訓練 24 個月的表現都較訓練期 18 個月佳，代表

此模型在建構基金最適投資組合上，適合較長的訓練期間。

表 4 訓練與測試期區間

期間	組別	訓練期	測試期
1	1	2007/1/1~2008/12/31	2009/1/1~2009/12/31
	2	2007/7/1~2008/12/31	
2	1	2007/4/1~2009/3/30	2009/4/1~2010/3/1
	2	2007/10/1~2009/3/30	
3	1	2007/7/1~2009/6/30	2009/7/1~2010/6/30
	2	2008/1/1~2009/6/30	
4	1	2007/10/1~2009/9/30	2009/10/1~2010/9/30
	2	2008/4/1~2009/9/30	
5	1	2008/1/1~2009/12/31	2010/1/1~2010/12/31
	2	2008/7/1~2009/12/31	

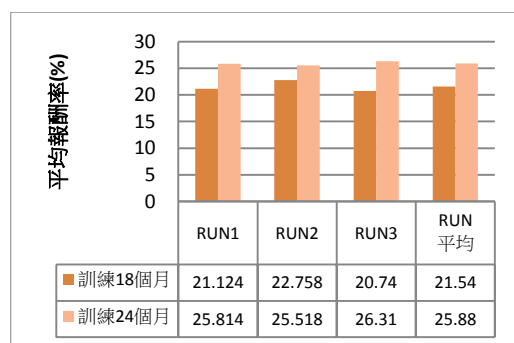


圖 11 單筆平均報酬率

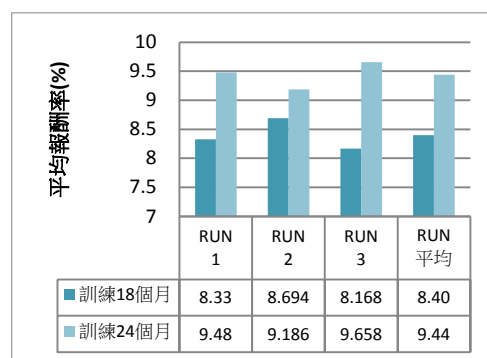


圖 12 定期定額平均報酬率

##### (2) 測試期長度分析

從第一階段訓練期長度測試中發現，以 24 個月為訓練期長度的績效較佳，故使用 24 個月做為第二階段的統一訓練期長度，而測試期分別設為 12 個月與 6 個月，移動區間為 3 個月，試圖找出模型較合適的測

試期長度，表 5 為訓練與測試期區間說明。

表 5 訓練與測試期區間

期間	組別	訓練期	測試期
1	1	2007/1/1~2008/12/31	2009/1/1~2009/12/31
	2		2009/1/1~2009/6/30
2	1	2007/4/1~2009/3/30	2009/4/1~2010/3/1
	2		2009/4/1~2009/9/30
3	1	2007/7/1~2009/6/30	2009/7/1~2010/6/30
	2		2009/7/1~2010/12/31
4	1	2007/10/1~2009/9/30	2009/10/1~2010/9/30
	2		2009/10/1~2010/3/31
5	1	2008/1/1~2009/12/31	2010/1/1~2010/12/31
	2		2010/1/1~2010/6/30

實驗結果顯示於圖 13 與圖 14，從中可看出測試期 6 個月的績效較 12 個月的佳，在單筆投資的平均報酬率與 12 個月測試期的平均報酬率約高了 8.9%，而在定期定額的平均報酬率也高了 8.7%，故此模型所適合之測試期以 6 個月為最佳。

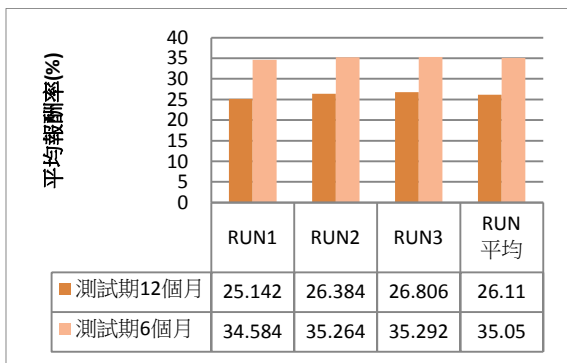


圖 13 單筆平均報酬率

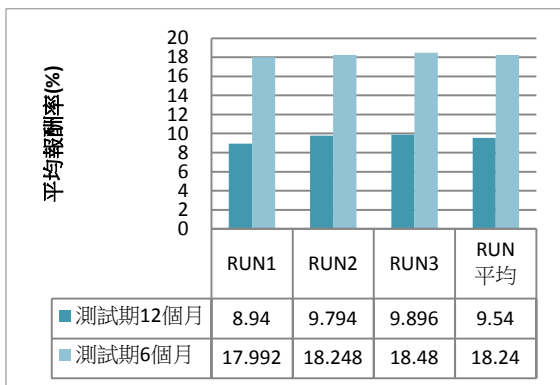


圖 14 定期定額平均報酬率

## 4.2.2 實驗二：模型穩定度分析

此實驗將會把 TSGA 與 Co-GAGEP 的演化適應值結果記錄下來，並分別比較訓練期的適應值與測試期的績效表現，來檢視模型在不同時期的表現是否穩定。

表 6 TSGA 適應值

期間		RUN1	RUN2	RUN3	avg	std
1	avg	52.15	53.42	52.36	52.64	0.68
	best	68.03	67.41	67.46	67.63	0.34
2	avg	47.83	47.66	48.79	48.09	0.61
	best	60.11	60.68	60.40	60.40	0.29
3	avg	43.14	44.58	44.23	43.98	0.75
	best	52.14	53.00	52.67	52.60	0.43
4	avg	39.83	39.20	40.70	39.91	0.75
	best	49.07	47.80	48.60	48.49	0.64
5	avg	37.90	37.40	38.06	37.79	0.34
	best	48.60	48.50	48.55	48.55	0.05

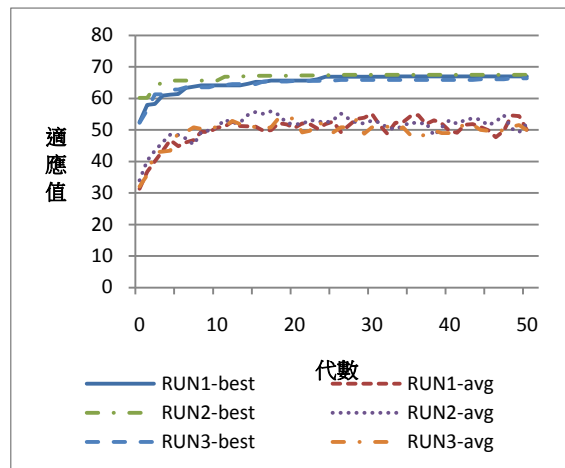


圖 15 TSGA 期間 1 演化收斂圖

TSGA 以 24 個月為訓練期，12 個月為測試期，TSGA 在訓練期的適應值結果如表 6 所示。從表 6 中可發現，同期間的平均適應值及最佳適應值在不同次的執行下都相當接近，且標準差也不大。而期間 1 的演化收斂圖如圖 15 所示，從圖中可看出在 3 次執行的演化過程不會差異太大，也大約在 20 至 30 代之間找到最佳解。從圖

16 與圖 17 中也可看出，在 5 個測試期間，3 次執行的表現無論在定期定額還是單筆投資上的績效表現皆一致，代表模型不僅在訓練期還是測試期的表現都相當穩定。

太多。

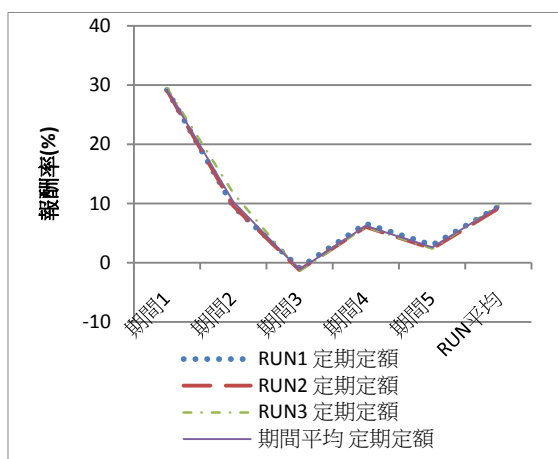


圖 16 TSGA 不同期間定期定額報酬率

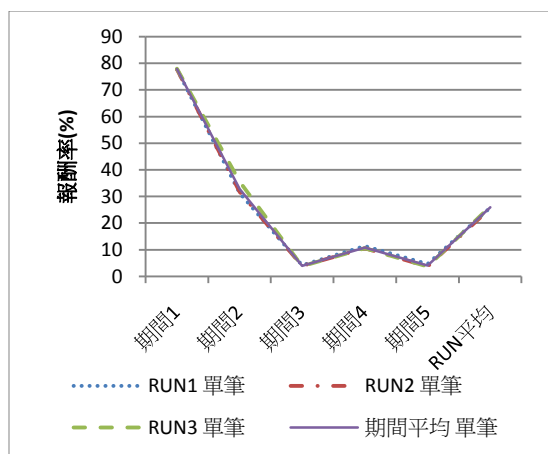


圖 17 TSGA 不同期間單筆報酬率

因避免交易策略的訓練與測試長度太長使得交易策略易產生鈍化，故 Co-GAGEP 以 12 個月為訓練期，6 個月為測試期，從 2008 年 1 月 1 日開始，每次移動 3 個月，共執行 7 次。

期間 1 的演化收斂情形如圖 18 所示，從圖中可看出平均值在演化過程中起伏較大，推測原因是研究中使用的是 GA 與 GEP 的共演化機制，故在變動時會互相影響，而變動的幅度在越後面的代數會越大，但是 3 次的執行下，變動區間都差不多，包含最佳值在 3 次執行下的變動也不會相差

表 7 Co-GAGEP 適應值

期間		RUN1	RUN2	RUN3	avg	std
1	avg	17.24	16.81	16.90	16.98	0.23
	best	21.65	21.44	21.56	21.55	0.11
2	avg	18.98	19.22	19.03	19.08	0.13
	best	22.20	22.20	21.05	21.82	0.66
3	avg	21.93	21.99	21.78	21.90	0.11
	best	25.92	24.79	25.28	25.33	0.57
4	avg	25.21	24.58	25.25	25.01	0.38
	best	28.92	27.30	33.98	30.07	3.48
5	avg	24.45	22.91	24.76	24.04	0.99
	best	40.66	27.36	52.42	40.15	12.54
6	avg	21.00	20.00	21.54	20.85	0.78
	best	23.93	23.37	25.20	24.17	0.94
7	avg	20.32	20.00	19.09	19.80	0.64
	best	23.12	23.13	23.15	23.13	0.02

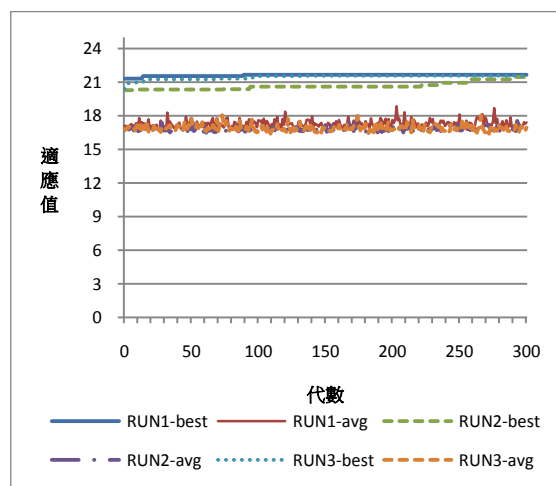


圖 18 Co-GAGEP 期間 1 演化收斂圖

訓練期的適應值結果如表 7 所示，從表中發現同期間的平均適應值，在不同次執行的結果大致接近，只有最佳適應值較有差異，其中以期間 4(2008/10/1~2009/9/30)與期間 5(2009/1/1~2009/12/31)的差異相對較大，本研究推論此情形與當時的情勢波動有關，在較具震盪的情況下，交易策略的解答空間較大，也造成期間 4 與期間 5 的無論在平均適應值或最佳適應值都比其

他期間高。本研究也發現，期間 5 第 2 次執行的最佳適應值比另外 2 次執行的最佳適應值來得低出許多，可能的原因是解答空間變大，使得模組較容易陷入區域最佳解，才會有此情形發生。

測試期的報酬率結果顯示於圖 19，從中也可看出，在不同的期間時，3 次的績效表現還算是穩定，雖然在先前提到期間 4 與期間 5 的訓練期處於大盤情勢較不穩定的期間，但在此期間內訓練的交易策略，在 3 次測試期的表現結果還是相當穩定，縱使在期間 5 的第 2 次執行結果陷入區域最佳解，但在測試期的表現也不會太差，因此 Co-GAGEP 模型的不僅在訓練期還是測試期的表現都屬於穩定。

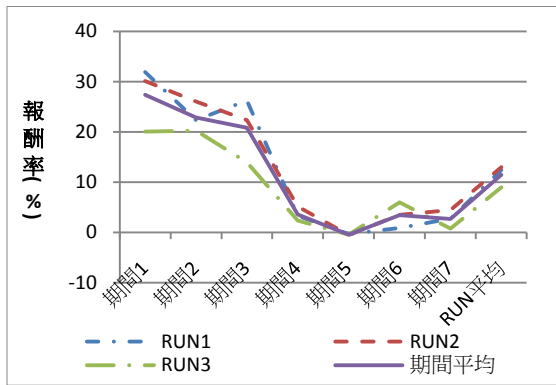


圖 19 Co-GAGEP 不同期間報酬率

#### 4.2.3 實驗三:模型適用性分析

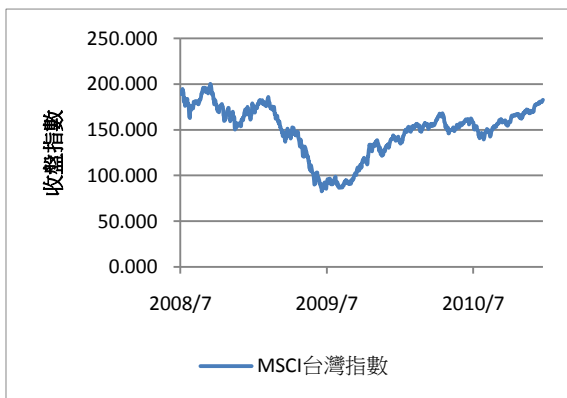


圖 20 MSCI 台灣指數

從 MSCI 台灣指數可看出，從 2008 年年底開始到 2009 年年底結束，是一個波動

較明顯的區段，從 2009 年年初指數呈現下滑的趨勢，直到 2009 年年中才開始向上攀升，這樣的趨勢除了在 2010 年 7 月開始有一點小跌幅外，大部份還是保持著上升的趨勢。

然而，從實驗一中發現單筆的報酬率均比定期定額來得高，且從選出來的基金表現中發現，此段時間的淨值均為上漲的趨勢，在此情況下，定期定額的成本會因此而墊高，相對於單筆投資的固定成本來說，在此期間單筆的表現自然會比定期定額好。

表 8 報酬率比較表(%)

基金	期間一	期間二	期間三	期間四	期間五	avg	std
ML01 保德信高成長	32.1	45.4	28.1	2.5	-7.9	20.0	22.0
CA03 群益馬拉松	34.9	42.5	23.5	-1.3	-13.9	17.1	24.0
0023 群益中小	43.4	51.2	32.2	2.5	-14.5	23.0	27.9
0005 匯豐成功	29.3	36.9	19.9	6.5	-4.1	17.7	16.6
0002 第一金福元	42.4	36.9	11.1	-2.0	-21.0	13.5	26.5
0028 群益店頭	53.0	<b>68.9</b>	39.1	1.9	-9.0	30.8	33.3
ML02 保德信金滿意	37.0	43.2	26.8	-0.4	-13.7	18.6	24.6
0014 統一統信	27.8	39.6	29.2	0.9	-15.6	16.4	22.9
0018 JF 台灣	37.4	41.9	25.5	-2.7	-19.8	16.5	26.7
PS01 統一全天候	38.0	53.4	<b>40.6</b>	1.5	-16.6	23.4	29.5
TSGA	<b>83.3</b>	56.2	26.2	<b>11.3</b>	<b>-1.8</b>	<b>35.1</b>	32.1

另外，從圖 14 可看出定期定額在不同期間的表現，期間 1(2009/1/1~2009/12/31)相對於其他期間的表現明顯較好，主要是這段期間的起伏較大，定期定額自然較能

顯現其功能。至於單筆在這段期間的表現也是優於其他期間，本研究發現此與 TSGA 所選出來的投資組合有相當大的關係，在這個投資組合內的基金，在期間 1 與期間 2 的基金淨值漲幅較大，而在期間 3 到期間 5 的表現雖平穩，但整體還是呈現上漲的趨勢，這就是單筆投資在每次實驗下來都會比定期定額報酬率來得高的原因。

以截至 2010/12/31 時的自成立以來的績效表現前十名來做每個期間的買入持有績效比較，等同以單筆報酬率來做比較，比較期間為 6 個月，結果如表 8 所示。從表 8 中可發現，TSGA 所建構出來的投資組合報酬率比績效前 10 名的基金在五個時期的平均為最高，雖然在期間 2 與期間 3 的績效表現並非最好，但在期間 2 仍高於一半數量的基金，且期間 3 的表現也不會太差。而在期間 5 時，所有基金均屬於負報酬，雖然 TSGA 的報酬率也是負報酬，但在表現上依然是最佳。從標準差方面來看，TSGA 的標準差是以 3 次執行的結果所計算出來，雖然以 TSGA 的標準差來說頗高，原因是期間 1 的績效遠高於其他期間，故標準差被影響到，但是與平均報酬率最高的前 2 名基金的標準差也是相當高，且前 10 名基金的標準差多數都比平均高，只有 TSGA 與匯豐成功基金的標準差比平均來得低，以 Sharpe 值的方式衡量承受一單位風險時可獲得的多少的超額報酬，TSGA 為 0.010572，匯豐成功基金為 0.009926，代表承受一單位風險時 TSGA 可獲得的報酬較多，證明 TSGA 所建構出來的投資組合在實務上還是具有適用性。

#### 4.2.4 實驗四：模型交易策略實用性分析

此實驗是透過 TSGA 在訓練期長度為 24 個月所挑出來的基金投資組合，將之分別使用單筆、定期定額與 Co-GAGEP 所建構的交易策略做模擬投資，比較期間為 12 個月。本研究因考慮到交易策略在長期訓練下較易鈍化的可能性，且基金的交易策略應屬中長期，故 Co-GAGEP 的訓練長度為 12 個月，測試期 6 個月，故等同於在 12

個月的比較期間內，會每 6 個月換一次交易策略，本研究也將會把這 12 個月執行 2 個交易策略的績效平均後，再與定期定額與單筆做比較。

定期定額與 Co-GAGEP 所建構出來的交易策略比較結果如圖 21 所示，從圖中可發現 Co-GAGEP 的績效表現比定期定額來得佳。而期間的比較如圖 22 所示，從圖中可看出除了在期間 5 的表現較差之外，Co-GAGEP 在其他期間的表現平均都比定期定額好，其中以期間 2 與 3 的績效更明顯，尤其在期間 3 時，Co-GAGEP 的表現明顯優於定期定額的負報酬，證明 Co-GAGEP 在市場績效不佳的情況下，還是能具有抗跌的效果，本研究推論原因為在期間 3(2009/7/1~2010/6/30)的大盤趨勢為一開始有下滑的傾向，隨後的 6 個月才拉上來，可見得使用 Co-GAGEP 的交易策略進出場的較合適的時機為基金較有震盪的時後。

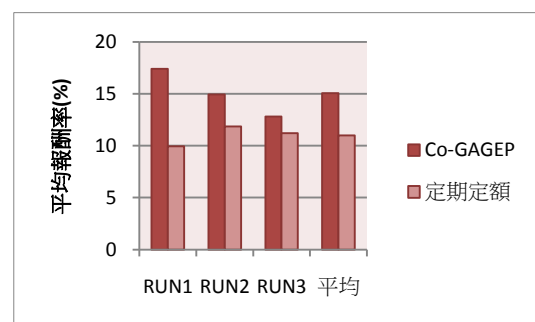


圖 21 交易策略績效比較

然而，從實驗中發現績效會受大盤影響而起伏，尤其以單筆投資的影響較大，主因為單筆投資為單一成本的特性，若是在模擬投資初期買入的價格較低，則單筆的獲利會較同時以定期定額方式買入的投資報酬率還要大，例如實驗中的期間 1 就是如此，才會造成單筆有相當高的超額報酬，反之則所承受的虧損會較大，如期間 3 的單筆獲利只有 3% 左右，與期間 1 相差了將近 80%，故單筆是一個風險較大的投資方式。

至於定期定額的特性為在大盤往上的

情況下會不斷的墊高成本，而影響到投資績效，而以 Co-GAGEP 所建構出來的交易策略來說，是以定期定額為基準，單筆為輔的雙軌投資方式，所以在交叉的策略執行下，若是在測試期當中，單筆沒有買到較低點，或是在低點無法累積較多的部位，其績效就極有可能會比定期定額或單筆投資來得差。

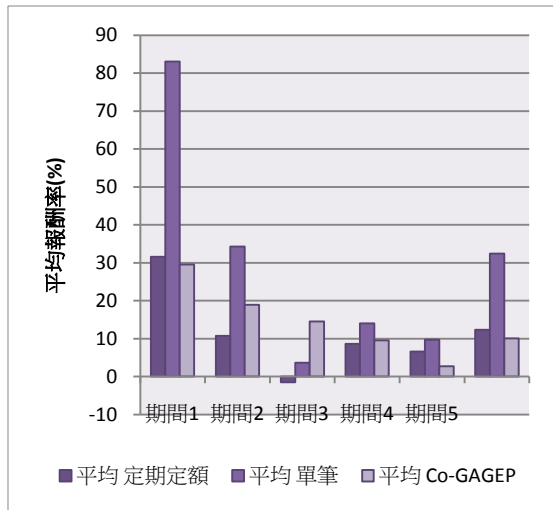


圖 22 不同期間平均報酬率比較

## 5. 結論與建議

本研究以遺傳演算法與基因表達規劃法為基礎，設計出兩階段遺傳演算法 (TSGA) 以及遺傳演算法與基因表達規劃法共演化的 Co-GAGEP 模型來建構出最適投資組合與交易策略模型，輔以不同的實驗設計來探討模型的適用性、穩定性與效用。從實驗一中得知 TSGA 模型的訓練期效果以 24 個月為最佳，測試期效果則以 6 個月為最佳。實驗二中發現，不同期間報酬率在每次施行實驗的結果都相當穩定，除了單筆投資績效的起伏較大之外，大致上的表現還算穩定。而在 TSGA 績效的表現上，不管是在測試期 12 個月還是表現較佳的 6 個月，都呈現不錯的績效。而實驗三顯示，由 TSGA 所建置的投資組合都有不錯的投資報酬率，且以整個期間的平均來說，也高於累積報酬率為前十名的基金。

雖然在情勢較高漲的 2009 年波段使得

單筆投資的平均總報酬率為最高，但是在實驗四顯示以 Co-GAGEP 所建構的雙軌投資法使得整體平均表現還是高於定期定額投資報酬率，且各期間均為正報酬，實驗結果代表 TSGA 與 Co-GAGEP 模型在實務上均有其適用性。

## 參考文獻

- [1] 王冠弼，應用遺傳演算法與動態模糊化調整策略於指數型基金商品設計之研究 - 以台灣 50 指數為例，天主教輔仁大學資訊管理學系碩士論文，2006。
- [2] 李曉玲，演化式粒子群演算法在共同基金組合之設計，天主教輔仁大學資訊管理學系碩士論文，2008。
- [3] 吳佳玲，以適應性權重函數建構時間序列預測模型，元智大學資訊管理學系碩士班碩士論文，2008。
- [4] 林玉萍，使用遺傳演算法建構不同風險承受度的基金投資組合與資產配置，國立中山大學資訊管理研究所碩士論文，2008。
- [5] 張桂莉，資產配置之最適策略，國立政治大學企業管理學研究所碩士論文，2000。
- [6] 張婉蘭，因應台灣景氣循環的最適資產配置投資組合之研究，國立高雄第一科技大學金融營運系碩士論文，2002。
- [7] 陳奕帆，演化式計算在股票型共同基金最適資產配置之研究，天主教輔仁大學資訊管理學系碩士論文，2010。
- [8] 陳玲穎，以 GEP 為基礎的分類規則產生器，元智大學資訊管理學系碩士班碩士論文，2008 年。
- [9] 陳愷瑜，共演化遺傳演算法在共同基金全球資產配置模型建置之研究，天主教輔仁大學資訊管理學系碩士論文，2008。
- [10] 陳與泳，組合式基金資金分配策略-蟻元合作系統與遺傳演算法之應用，國立台灣科技大學資訊管理系碩士學位論文，2004。

- [11] 陳嘉惠、高郁惠、劉玉珍，投資入偏好與資產配置，台灣管理學刊，第 1 卷第 2 期，2002/02，頁 213-232。
- [12] 劉文祺、張淑怡、詹麗錦，不同景氣循環階段最佳投資工具之選擇，產業金融第 108 期，2000，頁 50-52。
- [13] 劉貴強，遺傳演算法於組合型基金商品設計之研究，天主教輔仁大學資訊管理學系碩士論文，2004。
- [14] 蔡慧菊，基因表示規劃法於台股期貨價格發現之研究，天主教輔仁大學資訊管理學系碩士論文，2010。
- [15] 徐俊明，投資學理論與實務，台北:新陸書局，2007。
- [16] Brocato, Joe and Steed, Steven, Optimal Asset Allocation Over the Business Cycle, *Financial Review*, August, 1988, pp.129-148.
- [17] C. Lin, S. Hsu and S. Chen, Apply Genetic Algorithm to Explore Fund of Funds on Efficient Frontier and Product Design, *Proceedings of the IEEE The 3<sup>rd</sup> International Conference on Innovative Computing Information and Control, ICICIC 2008*, pp. 318-322.
- [18] Ferreira, C. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, Springer-Verlag, Germany, 2006.
- [19] Ferreira, C., Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, *Complex Systems*, 13, 2001, pp.87-129.
- [20] Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1989.
- [21] Holland, J. H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 1975.
- [22] Jeason, M., The Performance of Mutual Funds in the Period, *Journal of Finance*, 23(2), 1968, pp.389-416.
- [23] Kin K. L., Lean Y., Shouyang W., and Chengxiong Z., A Double-Stage Genetic Optimization Algorithm for Portfolio Selection, *International Conference on Neural Information Processing, ICONIP 2006*, pp. 928-937.
- [24] Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, Cambridge, MA: MIT Press, 1992.
- [25] Markowitz, Harry M., Portfolio Selection, *The Journal of Finance*, Vol. 7, 1952, pp.71-91.
- [26] O'Neal, E. S., How Many Mutual Funds Constitute a Diversified Mutual Fund Portfolio? , *Financial Analysts Journal*, (53), 1997, pp. 37-46.
- [27] Jeurissen , R. and van den Berg, J., Optimized Index Tracking using a Hybrid Genetic Algorithm, *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2008*, pp. 2327-2334.
- [28] Sharpe, W. F., Mutual Fund Performance, *Journal of Business*, 39(1), 1966, pp.119-138.
- [29] Treynor, J.L., How to Rate Management Investment Fund, *Harvard Business Review*, Vol. 43, Jan/Feb 1965, pp.63-75.
- [30] Treynor, J. and F. Black, How to Use Security Analysis to Improve Portfolio Selection, *Journal of Business*, January, 1973, pp.66-86.