

國立政治大學應用數學系

數學教學碩士在職專班

碩士學位論文

多變量模糊時間數列分析與轉折區間檢測

**Multivariate Fuzzy Time Series Analysis with  
Change Periods Detection**

碩專班學生：廖俊銘 撰

指導教授：吳柏林 博士

中華民國 九十九 年 六 月 二十八 日

# 多變量模糊時間數列分析與轉折區間檢測

## 摘要

近年來，隨著科技的進步與工商業的發展，預測技術的創新與改進愈來愈受到重視，同樣地，對於預測準確度的要求也愈來愈高。尤其在經濟建設、人口政策、經營規畫、管理控制等問題上，預測更是決策過程中不可或缺的重要資訊。有鑑於此，本論文嘗試應用模糊關係方程式，提出多變量模糊時間數列建構過程及轉折區間檢測模式理論架構。另一方面，多變量模糊時間數列模式建構過程，研究者曾提出很多轉折點之偵測與檢定方法，然而在實際的例子中，時間數列之結構改變所呈現出來的是一種緩慢的改變過程，即轉折點本身就是模糊不確定。這個概念在建構不同模式分析各國經濟活動數據時更顯重要。本論文針對轉折區間之檢測提出一個完整的認定程序。多變量時間數列系統中的隸屬度函數等於在計算成果指標群時的群集中心。應用本論文提出的方法，我們以德國、法國及希臘之總體經濟指標 GDP 為例，考慮通貨膨脹率、GDP 成長率及投資率來進行轉折區間的檢測。

關鍵詞：模糊關係、模糊馬可夫關係矩陣、多變量模糊時間數列、模糊規則庫、平均預測秩階準確度。

## Abstract

In recent years, along with the technological advancement and commercial development, the creation and improvement of forecasting techniques have more and more attention. Especially at the economic developments, population policy, management planning and control, forecasting gives necessary and important information in the decision-making process. Regarding stock market as the example, these numerals of closing price are uncertain and indistinct. Again, the factors of influence on quantity are numerous, such as turnover, exchange rate etc. Therefore, if we consider merely the closing price of front day to build and forecast, we will not only misestimate the future trend, but also will cause unnecessary damage.

Owing to this reason, we propose the procedure of multivariate fuzzy time series model constructed and theory structure by fuzzy relation equation. Combining closing price with turnover, we apply our methods to build up multivariate fuzzy time series model on Taiwan Weighted Index and predict future trend while examine the predictive results with average forecasting accuracy.

A fuzzy time series  $C_t^*$  is defined on averages of cumulative fuzzy entropies of the tree time series. Finally, an empirical study about change periods identification for Germany, France and Greece major macroeconomic indicators are demonstrated.

**Keyword: Fuzzy relation, fuzzy markov relative matrix, multivariate fuzzy time series, fuzzy rule base, average forecasting accuracy Change periods, Fuzzy statistics analysis.**

# 目錄

## 內容

多變量模糊時間數列分析與轉折區間檢測.....	1
第一章 前言.....	4
第二章 多變量模糊時間數列模式建構與預測.....	7
2.1 多變量模糊時間數列模式建構.....	7
2.2 如何由模糊規則庫進行屬性判別.....	8
2.3 多變量模糊時間數列之預測.....	11
2.4 平均預測秩階準確度.....	12
第三章 模糊決策過程.....	14
3.1 模糊轉折.....	14
3.2 群落中心.....	15
第四章 實證分析.....	19
第五章 結論.....	23
參考文獻.....	24

## 第一章 前言

近年來，隨著科技的進步與工商業的發展，預測技術的創新與改進愈來愈受到重視，相對地，對於預測準確度的要求也愈來愈高。尤其在經濟建設、人口政策、經營規畫、管理控制等問題上，預測更是決策過程中不可或缺的重要資訊。由於企業運作的規模及複雜性日益增加，若錯估未來市場供需，將造成不必要的損失甚至危及企業生存。在此危機意識下，唯有透過合理、正確的預測方法，進而觀察市場走勢及供需結構變化，配合系統模式運作，才能充分發揮經營決策的效率。這對於提高經營目標水準和獲取最佳經濟效益亦有重大貢獻。

在時間數列分析中，資料的走勢型態可以作為判斷事件發生的基礎，如：遞增或遞減、季節性循環或突發暴漲等。是故，根據所觀察的特性，可藉由先驗的模式族中，如：ARIMA 模式族、ARCH 模式族或門檻模式族等，挑選出最佳的配適模式。但由於資料收集的誤差、時間的延遲(lag)或變數之間的交互影響，使得單一度量的數值，形式上看似一精確值，而實際上所隱含的卻是某一區間範圍的可能值。在此情況下，我們若以傳統的模式建構與分析方法，來配適出一數學模式，以解釋時間數列資料與走勢，可能會產生模式過度配適的危險。

模糊集合(邏輯)的概念由 Zadeh (1965)首先提出，其乃參考人腦思維方式對動態環境所使用模糊測度與分類原理，對多元複雜的不確定現象，給予較為穩健描述的處理方法。由於，模糊理論(fuzzy theory)本身具有語言變數(linguistic variables)蘊含特性，此種特性可以減少在處理不確定性問題時可能造成的困擾。因此，模糊理論目前已被廣泛地應用於各種領域，如：航空、機械、醫學、電力、地質等。其中在模糊控制系統方面的應用更是不勝枚舉，較具代表性的專書可參考 Nguyen 與 Sugeno (1998)。

此外，近年來也逐漸應用在社會科學方面，例如 Lowen (1990)、Ruspini (1991)、Dubois 與 Prade (1991)分別提出理念概似(approximate reasoning)的計量方法。Clymer、Corey 與 Gardner (1992)提出離散事件的機場起落模糊控制方法，Cutsem 與 Gath (1993)曾提出利用模糊分類的程序來偵測離群值及穩健地估計參數，Hathaway 與 Bezdek (1993)則對模糊迴歸模式進行參數估計與分類。Yoshinari、Pedrycz 與 Hirota (1993)經由模糊分類法建立模糊模式，Romer、Kandel 與 Backer (1995)在統計推論上採用模糊分割理論及可能性理論，Wu 與 Sun (1996)將模糊統計應用於社會調查分析上。在模糊決策分析方面，Werners (1987)、Zimmermann (1991)、Tseng 與 Klein (1992)等，Yang (1993)對於模糊分類(fuzzy clustering)領域亦有一較完整的探討。

近來，模糊邏輯應用於動態資料分析上亦有逐漸增加的趨勢。如 Song 與

Chissom (1993a, b, 1994)、Park 等(1995)、Chen (1996)、Song 等(1997)利用模糊理論建立趨勢型模糊時間數列。Wu 與 Hung (1999)提出模糊認定法則，以作為 ARCH 模式族與 Bilinear 模式族的決策判定標準。Wu 與 Chen (1999)利用模糊分類法來檢定時間數列資料結構轉變的轉折區間。Tseng 與 Tzeng 等(2001)考慮以傳統時間數列 ARIMA 模式與模糊迴歸模式結合，提出了模糊 ARIMA 模式來預測新台幣對美元的匯率。

然而，將模糊邏輯應用於時間數列分析過程時，首要步驟就是考慮要如何結合語言變數分析方法，以解決資料的不確定性問題。針對這點，Tong (1978)提出邏輯檢查方法(logical examination method)並利用決策表來描述模糊模式，但此方法很難推廣至多變數的系統。所以，為了獲得更精確的模糊模式，Graham 與 Newell (1989)、Xu 與 Lee (1987)在語言方面提出了具有學習能力的方法去修正模糊模式，Chiang 等(2000)提出模糊語言概念系統(fuzzy linguistic summary system)來收集時間數列資料以發現有用的資訊。另外，像是以嘗試錯誤過程(trial-and-error procedure)的方法去選擇適當的加權因子，亦是相當麻煩。事實上，若是由模糊關係方程式著手，是較決策表或決策法則容易理解與應用的。有鑑於此，大部分的學者常採用模糊關係方程式來求解。如 Song 與 Chissom (1993a, b)就利用模糊關係方程式，提出詳細的模糊時間數列建構過程及模式理論架構。Song 與 Chissom (1993a, 1994)並將此法應用在阿拉巴馬大學新生註冊人數之預測。而 Lee 等(1994)提出了兩階段的認定過程，並結合了語言方法及模糊關係方程式數值解來認定模糊模式。

在人文社會科學的測度理論裡，模糊統計與模糊相關性日漸受到重視，這應是複雜的人文社會現象，無法以傳統數值模型充分合理解釋的一種自然發展結果。以股票市場為例，收盤價這個數字的本身具有不確定性與模糊性，再加上影響此數值的因素眾多，如成交量、匯率等。因此，若僅考慮前一日收盤價為其決定因素來建構模式以進行預測，將會錯估未來市場走勢，且造成不必要的損失。在以往的文獻中，大部分僅探討單一變量的模糊時間數列，而對於多變量的動態資料皆未加以進一步的研究。有鑑於此，本論文乃嘗試提出多變量模糊時間數列建構過程及模式理論架構，並提出轉折區間的檢測法。最後，利用此方法，結合通貨膨脹率(inflation rate)、GDP 成長率(percentage growth)及投資率(investment rate)三項因素，對德國、法國及希臘多變量模糊時間數列模式，並進行轉折區間的檢測。

以建構模式來分析時間數列的方式存在一個缺點，就是無法在整個動態過程中產生顯著的結構轉變。因此，常常有兩個問題被提起：(1)是否存在一個良好的模式可以引導這個動態過程？(2)單一模式便足夠配適時間數列的動態過程嗎？有需要使用多個模式來配適嗎？



概括來說，時間數列的模式認定對於從模式庫系統中挑選最適配模式有著相當大的影響，見 Wu (1994)。只有在選到合適先驗模式族的情況下，整個模式建構的步驟如參數估計、分析、預測等才顯得有意義。另一方面，如果時間數列本身展現出某種程度的結構轉變，當然在建構整個模式過程前，我們要試著去檢測出其結構轉變的點與區間。若是不進一步找到轉折點與轉折區間(change periods)落在何處，而直接使用傳統的模式建構技術，這樣我們是很難建構出一個能適切解釋時間數列的良好模式。

時間數列轉折點的檢定問題已被很多專家學者研究過。比方說，Tsay (1991)曾提出針對單變數時間數列的外離(outliers)、級別變換(level shift)及變異數改變(variance changes)的檢定程序，他所提出的方法不僅有用而且非常容易實行。然而 Balke (1993)指出 Tsay 所提出的程序對於級別變換的檢定結果並不總是令人滿意。Inclan 與 Tiao (1994)則提出利用中心化的累加平方和(a centered version of the cumulative sums of squares)來偵測變異數改變的疊代法。

關於轉折點的檢定統計量(testing statistics)還包括 Page (1955)提出的 MPAGE (modified PAGE)、Hinkley (1971)的 CUSUM (cumulative sum)。Hsu (1982)提出的序列變異數變換檢定法避免了沉重的計算負荷，並讓研究者能聚焦在轉折點的檢測上。

本論文的架構規劃如下：在第二章我們定義時間數列的模糊表示法及引入景氣循環的趨勢測量方法，在第三章則採用模糊分類及模糊決策來檢測轉折區間的發生時機，而第四章則以歐盟三個重要經濟體（德國、法國及希臘）之總體經濟指標(GDP)為例說明，第五章則提出結論。

## 第二章 多變量模糊時間數列模式建構與預測

### 2.1 多變量模糊時間數列模式建構

在多變量模糊時間數列模式的建構過程中，有許多的情況是需要注意的。我們分別敘述如下：

由於所收集的資料，其型態可能為數值的或屬性的資料、也有可能是語言值 (linguistic values) 的資料（例如：透過品嚐所獲得的資料等等），針對這些型態的資料，很難只利用傳統的時間數列方法來加以分析。因此，倘若能使用模糊集合的方法，則不會受資料型態的影響，且更能建立合適的模式。

在架構模糊論域集合時，其分割的集合數，應該要多少個才適當，目前尚無定論。一般來說，分割的集合數越多，其準確度越高，然而運算的複雜度也相對地增加。所以，準確度與複雜度的取舍就應視個人的需要而定。

在資料的模糊化方面，我們是使用標準的隸屬度函數將其模糊化。在此過程中，其典型值的取得尚未有固定的方法，我們認為可以用每個集合中其元素的中位數、平均數或分類的群落中心為其典型值。但是，若以群落中心為其典型值，則易造成一集合中有多個典型值，此時該集合為梯形隸屬度函數，將會使得轉換過程較為複雜。反之，若使用中位數或平均數為其典型值，則每一個集合中其典型值只有一個，此時該集合為三角形隸屬度函數，轉換過程將會較為簡單。

資料結構是否穩定，是模式建構過程中研究的重點。在我們探討的問題中，其資料大致趨於穩定，故可直接進行模式建構步驟。然而，在一般情況中，若收集的資料是由數值資料轉換而成的模糊資料，且此時走勢具有遞增或遞減的傾向，其處理的方法是將原始資料差分，使之達到穩定，然後再繼續前述的步驟，來進行模式的建構。反之，若所收集的資料已具備模糊結構，且檢定出確有模糊趨勢，則目前尚未有方法可加以解決。

階次的認定在多變量模糊時間數列分析上，佔著相當重要的地位。若我們能找到正確的階次，才能掌握影響資料走勢的因素，進而建構符合實際狀況的數學模式。在實務應用上，尤其是股票、匯率、期貨等資料，其時間數列走勢常呈現高者愈高，低者愈低的非線性型態，且大部分金融市場上的運作情形皆符合馬可夫性質。因此，可藉由實際的狀況，來決定其階次的選取。

當選定多變量模糊時間數列的階次後，利用所定義的計算方法，並結合影響資料的多項因素進而得到模糊馬可夫關係矩陣  $\mathfrak{R}$ 。雖然，Lee 等(1994)提出利用實際的模糊輸出值與估計的模糊輸出值，可得較小誤差，但是這種計算方式牽涉到太多的數值運算，過於複雜。因此，本論文仍採用 Wu (1986)及 Sugeno 與 Tanaka (1991)提出的方法來計算模糊馬可夫關係矩陣  $\mathfrak{R}$ 。



有了以上的說明，我們便可以分析多變量模糊時間數列模型。在此，我們只探討多變量一階自迴歸模糊時間數列，其定義如下：

### 定義 2.1 多變量一階自迴歸模糊時間數列模式

一多變量模糊時間數列  $\{(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t})\}$ ，若對任意  $t$  可寫成

$$(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t}) = (FX_{1,t-1}, FX_{2,t-1}, \dots, FX_{k,t-1}) \begin{pmatrix} \mathfrak{R}_{11} & \dots & \dots & \mathfrak{R}_{1k} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathfrak{R}_{k1} & \dots & \dots & \mathfrak{R}_{kk} \end{pmatrix},$$

其中  $\mathfrak{R}_{ij}$  是第  $i$  個變量對第  $j$  個變量的模糊馬可夫關係矩陣， $i, j = 1, 2, \dots, k$ 。我們稱此多變量模糊時間數列  $\{(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t})\}$  是階次為 1 的多變量模糊自迴歸(auto regressive)模式，記做 VFAR(1)。在此模式中，因為  $(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t})$  僅與  $(FX_{1,t-1}, FX_{2,t-1}, \dots, FX_{k,t-1})$  有關，故此模式又可稱為馬可夫過程(Markov process)。

## 2.2 如何由模糊規則庫進行屬性判別

在多變量模糊時間數列中，如何將模糊數值（隸屬度函數）轉換成所屬的語言變數（屬性）亦是一研究重點。通常，是以其最大隸屬度函數所在的位置來加以判定。但如果其最大隸屬度函數不只一個時，該如何做選擇以決定其屬性？至今，還沒有一定的規則可循。為此，本論文定義一語言向量指標函數來處理最大隸屬度函數不只一個的情形。

再者，本論文例 2.2 所使用的時間數列資料是台灣地區加權股票指數，而要探討的是其漲跌之所屬的語言變數。然而，我們都知道影響股票走勢的因素甚多，如：成交量、匯率、利率及政治因素等。故在一即時、開放的市場，股票漲跌幅的決定，往往以前一日的收盤價為最大考量因素，其次才是成交量的多寡。這是由於成交量是股價的先行指標且成交量的持續擴大亦是股價上漲的必要條件，而對於長期趨勢的部份就當作論域之等級參考。另外，在傳統的時間數列分析中，由於前一期的資訊與當期的資訊有高度的自相關。故針對此兩項特點，可建立起多變量模糊時間數列。因此，經由模糊馬可夫關係矩陣，我們可由前一筆的模糊數值，而得到這一筆的模糊數值。但如何將其模糊數值轉換成語言變數呢？首先，我們必須給定以下的定義：

## 定義 2.2 語言向量指標函數

令  $L = \{(L_{11}, \dots, L_{1r}), \dots, (L_{k1}, \dots, L_{kr})\}; L_{ij}$  為語言變數,  $F\hat{X}_t$  為一多變量模糊時間數列相對於  $L$  的隸屬度函數。設  $F\tilde{X}_t$  為語言向量指標函數, 令  $F\tilde{X}_t = \{(I_{11}, \dots, I_{1r}), \dots, (I_{k1}, \dots, I_{kr})\}; I_{ij} = 1 \text{ or } 0\}$ , 其中

$$I_{ij} = \begin{cases} 1, & \text{if } \mu_{L_{ij}}(F\hat{X}_t) \geq k \\ 0, & \text{if } \mu_{L_{ij}}(F\hat{X}_t) < k \end{cases},$$

其中  $\mu_{L_{ij}}(F\hat{X}_t)$  為  $F\hat{X}_t$  在語言變數  $L_{ij}$  中所屬的隸屬度函數。

例 2.2. 令  $L = \{(L_{11}, L_{12}, L_{13}, L_{14}, L_{15})(L_{21}, L_{22}, L_{23}, L_{24}, L_{25})\}; L_{11} = \text{大跌}, L_{12} = \text{跌}, L_{13} = \text{平盤}, L_{14} = \text{漲}, L_{15} = \text{大漲}; L_{21} = \text{低}, L_{22} = \text{偏低}, L_{23} = \text{中等}, L_{24} = \text{偏高}, L_{25} = \text{高}\}$ , 若一筆二變量時間數列資料經由模糊馬可夫關係矩陣  $\mathfrak{R}$  運算後且相對於  $L$  的隸屬度函數為  $F\hat{X}_t = (1, 1.5, 2, 2, 1.5)(1, 1.5, 1.5, 2, 1.5)$ , 即其屬於  $L_{11}$  的隸屬度函數為 1, 屬於  $L_{12}$  的隸屬度函數為 1.5, 屬於  $L_{13}$  的隸屬度函數為 2, 屬於  $L_{14}$  的隸屬度函數為 2, 屬於  $L_{15}$  的隸屬度函數為 1.5; 屬於  $L_{21}$  的隸屬度函數為 1, 屬於  $L_{22}$  的隸屬度函數為 1.5, 屬於  $L_{23}$  的隸屬度函數為 1.5, 屬於  $L_{24}$  的隸屬度函數為 2, 屬於  $L_{25}$  的隸屬度函數為 1.5。由定義 2.2, 我們可得到  $F\tilde{X}_t = (0, 0, 1, 1, 0)(0, 0, 0, 1, 0)$ 。

根據定義 2.2, 我們可將經由多變量模糊時間數列模式預測後所得的模糊數值轉換成語言向量指標函數。但是, 我們該如何經由語言向量指標函數來判定其所代表的語言變數? 我們可根據上述的定義, 並利用模糊推理的方式來建立一門檻函數, 進而得到一模糊規則庫以分析其輸出的語言變數。

至於, 模糊規則庫該如何得到? 顧名思義, 規則庫中所放置的就是一些推理的法則, 即是以一條條的語意(linguistic)規則來描述其推理策略。而模糊規則庫就是針對日常生活中一些模糊的現象或知識, 所建立的一套專家系統(法則)。但規則該如何產生呢? 對於模糊時間數列而言, 至今還無一明確的方法可依循。所以, 本論文中的規則是基於傳統的時間數列本身乃統計相關, 利用自相關函數(autocorrelation function, ACF)與偏自相關函數(partial autocorrelation function, PACF), 來找出時間數列模式的係數。但針對非線性時間數列而言, 其自相關函數與偏自相關函數並非明確清晰, 故可依照傳統的自迴歸整合移動平均(auto regressive integrated moving average, ARIMA)模式, 利用三階段建構步驟, 即(1)階次認定(order identification)(2)參數估計(parameter estimation)(3)診斷檢定(diagnostic checking), 來協助我們得到最佳的配適模式。因此, 本論文所建立的模糊規則, 基本上都是得自以上的觀念與經驗的累積, 且根據模糊推理而得, 故

建立的方式是直覺而主觀的。

根據本論文例 2.2 所使用的時間數列資料中，針對漲跌幅與成交量差兩項因素而言，我們架構其論域分別為{大跌，跌，平盤，漲，大漲}與{低，偏低，中等，偏高，高}，故將  $n$  設為 5。而我們考慮以  $(I_{k1}, \dots, I_{k5})$  為模糊推理指標，又其  $I_{kj} = 0$  或 1， $j = 1, 2, \dots, 5$ ，因此我們可建立起 32 個語言向量，但需扣除  $(0, 0, 0, 0, 0)$  這個向量，因為其無法代表任何一個語言變數。但是，要如何將 31 個語言向量分類為其所屬的語言變數，並非是一件容易的事。因為，若語言向量中只有一個 1 其餘為 0，則輸出的便是 1 其所在位置代表的語言變數。例如： $(0, 0, 0, 1, 0)$  表示屬於漲之隸屬度函數為 1，故輸出「漲」這個語言變數。

然而，當語言向量中 1 不只一個時，我們又該如何認定？假使我們將每一筆模糊時間數列的語言向量中的  $I_{kj}$ ， $j = 1, 2, \dots, 5$ ，從其  $I_{k1} = 0$  或 1 開始，依序判別至  $I_{k5} = 0$  或 1 結束，再依據我們所累積的知識與經驗，決定其所代表的語言變數，這將是件耗時的工作。倘若，我們改以整個語言向量來作判定，藉由經驗法則，我們可以輕易的判定出其所代表的語言變數。例如： $(0, 0, 0, 1, 1)$  表示屬於漲與大漲之隸屬度函數各為 1，依據經驗法則，則我們判定其輸出的語言變數為「大漲」。同理， $(1, 1, 0, 0, 0)$  亦是指其輸出的語言變數為「大跌」。

因此，根據上述的推導方式，我們可分別針對不同的論域分割集合，來考慮一門檻函數  $H_t$ 。

$$H_t = \begin{cases} \text{大跌(低),} & \text{當 } K_t \leq -2 \\ \text{跌(偏低),} & \text{當 } -2 < K_t \leq -1, \text{ 或, 當 } K_t = -2 \text{ 與 } \sum_{j=1}^5 I_{kj} \geq 3 \\ \text{平盤(中等),} & \text{當 } K_t = 0 \\ \text{漲(偏高),} & \text{當 } 1 \leq K_t < 2, \text{ 或, 當 } K_t = 2 \text{ 與 } \sum_{j=1}^5 I_{kj} \geq 3 \\ \text{大漲(高),} & \text{當 } 2 \leq K_t \end{cases},$$

其中， $K_t = \sum_{j=1}^5 (j-3)I_{kj}$ 。

最後，我們再依據此門檻函數  $H_t$  來建立以下的模糊規則庫(fuzzy rule base)。

### 模糊規則庫

(1) 假若  $F\tilde{X}_t \in \{(1, 0, 0, 0, 0), (1, 1, 0, 0, 0), (1, 0, 1, 0, 0), (1, 1, 1, 0, 0)\}$ ，則輸出的

語言變數為“大跌(低)”

(2) 假若  $F\tilde{X}_t \in \{(0, 1, 0, 0, 0), (1, 1, 0, 1, 0), (1, 1, 1, 0, 1), (1, 1, 0, 0, 1), (1, 0, 0, 1, 0),$

$(1, 1, 1, 1, 0), (0, 1, 1, 0, 0), (1, 0, 1, 1, 0)\}$ ，則輸出的語言變數為“跌(偏低)”

(3) 假若  $F\tilde{X}_t \in \{(0, 0, 1, 0, 0), (1, 0, 1, 0, 1), (1, 0, 0, 0, 1), (1, 1, 1, 1, 1), (0, 1, 0, 1, 0),$

$(1, 1, 0, 1, 1), (0, 1, 1, 1, 0)\}$ ，則輸出的語言變數為“平盤(中等)”

(4) 假若  $F\tilde{X}_t \in \{(0, 0, 0, 1, 0), (0, 1, 0, 1, 1), (1, 0, 1, 1, 1), (1, 0, 0, 1, 1), (0, 1, 0, 0, 1),$

$(0, 1, 1, 1, 1), (0, 0, 1, 1, 0), (0, 1, 1, 0, 1)\}$ ，則輸出的語言變數為“漲(偏高)”

(5) 假若  $F\tilde{X}_t \in \{(0, 0, 0, 0, 1), (0, 0, 0, 1, 1), (0, 0, 1, 0, 1), (0, 0, 1, 1, 1)\}$ ，則輸出的

語言變數為“大漲(高)”

在例 2.2 中，我們得到  $F\tilde{X}_t = (0, 0, 1, 1, 0)(0, 0, 0, 1, 0)$ ，故經由上述的模糊規則庫可知，對於漲跌幅與成交量差而言，其所輸出的語言變數分別為「漲」與「偏高」。

### 2.3 多變量模糊時間數列之預測

預測是決策過程中不可或缺的重要資訊，且精確的預測結果可以提供決策者更多的訊息，有利於做出正確的決策與適當的反應。所以，我們利用所得的多變量模糊時間數列模式來進行預測，以得知其預測之效果。在此，多變量模糊時間數列模式之預測的定義給定如下：

定義 2.3 多變量一階自迴歸模糊時間數列模式的預測

設一階多變量模糊自迴歸模式為

$$(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t}) = (FX_{1,t-1}, FX_{2,t-1}, \dots, FX_{k,t-1}) \begin{pmatrix} \mathfrak{R}_{11} & \cdots & \cdots & \mathfrak{R}_{1k} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathfrak{R}_{k1} & \cdots & \cdots & \mathfrak{R}_{kk} \end{pmatrix}$$

且已知觀測值  $(FX_{1,t}, FX_{2,t}, \dots, FX_{k,t}), t = 1, 2, \dots, n$ ，則往前  $l$  期的預測值為

(1) 當  $l = 1$  時

$$(FX_{1,n}(1), FX_{2,n}(1), \dots, FX_{k,n}(1)) = (FX_{1,n}, FX_{2,n}, \dots, FX_{k,n}) \begin{pmatrix} \mathfrak{R}_{11} & \cdots & \cdots & \mathfrak{R}_{1k} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathfrak{R}_{k1} & \cdots & \cdots & \mathfrak{R}_{kk} \end{pmatrix}$$

(2) 當  $l = 2$  時

$$(FX_{1,n}(2), FX_{2,n}(2), \dots, FX_{k,n}(2)) = (FX_{1,n}, FX_{2,n}, \dots, FX_{k,n}) \begin{pmatrix} \mathfrak{R}_{11} & \dots & \dots & \mathfrak{R}_{1k} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathfrak{R}_{k1} & \dots & \dots & \mathfrak{R}_{kk} \end{pmatrix}^2$$

(3) 往前  $l$  期的預測值為

$$(FX_{1,n}(l), FX_{2,n}(l), \dots, FX_{k,n}(l)) = (FX_{1,n}, FX_{2,n}, \dots, FX_{k,n}) \begin{pmatrix} \mathfrak{R}_{11} & \dots & \dots & \mathfrak{R}_{1k} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathfrak{R}_{k1} & \dots & \dots & \mathfrak{R}_{kk} \end{pmatrix}^l$$

## 2.4 平均預測秩階準確度

建構了多變量模糊時間數列模式與模糊規則庫之後，為了比較經由模式預測後所得的語言變數與實際的語言變數之間的誤差，本論文利用秩階(rank)的觀念，針對每一個語言變數分別賦予其代表值，如：大跌為-2、跌為-1、平盤為0、漲為1及大漲為2，因而定義其所謂的平均預測秩階準確度。

定義 2.4 平均預測秩階準確度

令  $\{RL_t, t = 1, \dots, n\}$  為一時間數列所表示的語言變數，且  $\{FL_t, t = 1, \dots, n\}$  為經由模糊時間數列模式預測所得的語言變數。設  $L = \{(L_1, L_2, \dots, L_r) = (-(r-1)/2, -(r-3)/2, \dots, (r-1)/2)\}$ ； $L_j$  為語言變數； $j = 1, \dots, r$ ，則

$$P = 1 - \frac{\sum_{t=1}^n \frac{|FL_t - RL_t|}{r-1}}{n}$$

稱為平均預測秩階準確度，其中  $r$  表示語言變數的個數。

例 2.4. 設一時間數列所表示的語言變數為{跌，平盤，跌，大漲，平盤，跌，大漲，跌，平盤，大跌}，則其語言變數的代表值為{-1, 0, -1, 2, 0, -1, 2, -1, 0, -2}；而經由模糊時間數列模式預測所得的語言變數為{跌，平盤，大跌，大漲，平盤，平盤，大漲，跌，大漲，平盤}，則其語言變數的代表值為{-1, 0, -2, 2, 0, 0, 2, -1, 2, 0}。根據定義 2.4，我們可得其平均預測秩階準確度為



$$P = 1 - \frac{\sum_{t=1}^{10} \frac{|FL_t - RL_t|}{4}}{10} = 1 - \frac{6}{40} = 0.85.$$

我們給定多變量模糊時間數列模式建構之整合過程如下：

### 多變量模糊時間數列模式建構之整合過程

- 步驟 1：觀察時間數列  $\{X_{1,t}\}, \dots, \{X_{k,t}\}$ 。分別針對  $\{X_{1,t}\}, \dots, \{X_{k,t}\}$ ，決定其論域與語言變數  $\{L_1, L_2, \dots, L_r\}$  的範圍。
- 步驟 2：分別計算  $\{X_{1,t}\}, \dots, \{X_{k,t}\}$  的模糊時間數列  $\{FX_{1,t}\}, \dots, \{FX_{k,t}\}$ ，並根據其最大隸屬度函數所在位置來判定其所屬的語言變數。
- 步驟 3：找出  $(\{FX_{i,t}\}, \{FX_{j,t}\})$ ， $i, j = 1, 2, \dots, k$  之間的模糊關係，並計算其模糊馬可夫關係矩陣。
- 步驟 4：根據步驟 3，依照  $(\{FX_{i,t}\}, \{FX_{j,t}\})$ ， $i, j = 1, 2, \dots, k$  之間的關係，可得最大的模糊馬可夫關係矩陣  $\mathfrak{R}$ ，進而建構一多變量模糊時間數列模式。
- 步驟 5：運算後的多變量模糊時間數列，針對各單一模糊時間數列，檢定其最大隸屬度函數是否唯一。若為唯一，則直接輸出其所屬的語言變數；若不唯一，則依照模糊規則庫輸出其所屬的語言變數。
- 步驟 6：最後，經由此模式進行預測。

### 第三章 模糊決策過程

傳統上時間數列的模式建構是基於貝氏經驗(Bayesian experience)，即從模式庫中挑選一個能良好解釋數列動態行為的好模式。然而如果一個時間數列表現出明顯型態的結構轉變，我們需要一個有效率的轉折點或轉折區間檢定法。以往時間數列的檢測法都假設時間數列本身會呈現出相當明顯的轉折點特徵。不過，在處理來自不同領域的時間數列時，不能僅考慮到轉折點的偵測，同時也要一併考量轉折區間的特徵。

#### 3.1 模糊轉折

時間數列在發生結構變化時的型態並不一定是立即且顯著的，反而常常呈現出一種持續的變化過程。因此，對於此種現象不能僅當作是在單一時刻的突發性轉向來看待，舉例來說：(1)匯率可能會受到新的財經政策的實施而逐步地上升或是下降；(2) $M_1$  或  $M_2$  的貨幣供應量會隨著當下國家經濟狀況而增減。事實上，「轉折點」的語義本來就不是非常清楚地被理解，見 Wu 與 Chen (1999)。

為了符合實際日常生活中的情形，最好是採用「轉折區間」的概念來取代「轉折點」，並且在腦海中抱持著「時間數列從某一型態轉換到另一型態的過程並不是一下子就切換過去，而是經歷一段調整適應期才轉換過去」這種想法。本論文嘗試應用轉折區間的想法來取代轉折點在分析結構變化過程中的角色，為採用與偵測轉折點不同的觀點，本論文將先研究轉折區間的性質，為此我們先定義模糊時間數列。

##### 定義 3.1 模糊時間數列

令  $\{X_t \in R, t = 1, 2, \dots, n\}$  為一個時間數列，且  $U$  為其論域。令  $\{P_i; i = 1, 2, 3, \dots, m, \bigcup_{i=1}^m P_i = U\}$  為  $U$  的一個次序分割集合(ordered partition set)，且其相對於語言變數為  $\{L_i, i = 1, 2, \dots, m\}$ 。若在  $\{L_1, L_2, \dots, L_m\}$  上相對於  $X_t$  的模糊集合  $F_t$  有隸屬度函數為  $\{\mu_{t1}, \mu_{t2}, \dots, \mu_{tm}\}$ ， $0 \leq \mu_{ti} \leq 1$ ， $i = 1, 2, \dots, m$ ，則我們稱  $\{F_t\}$  為佈於  $\{X_t\}$  上的一個模糊時間數列。並且記為

$$F_t = \frac{\mu_{t1}(X_t)}{L_1} + \frac{\mu_{t2}(X_t)}{L_2} + \dots + \frac{\mu_{tm}(X_t)}{L_m} \quad (3.1)$$

其中，「+」表示連結符號， $\mu_{ti}: R \rightarrow [0,1]$ ，且  $\sum_{i=1}^m \mu_{ti} = 1$

在本論文中我們以  $\{F_t\}$  表示為佈於  $\{X_t\}$  上的一個模糊時間數列，並且為了方

便起見，我們把  $F_t$  簡寫為  $F_t = \langle \mu_{t1}, \mu_{t2}, \dots, \mu_{tm} \rangle$ 。而集合  $\{L_i, i = 1, 2, \dots, m\}$  被視為一個語言變數的序列，而模糊時間數列  $\{F_t, t = 1, 2, \dots, m\}$  的元素則是由相對於語言變數的隸屬度所組成，意即對任意一個  $F_t (t = 1, 2, \dots, m)$ ， $F_t$  皆含有對應於每一個  $L_i (i = 1, 2, \dots, m)$  的隸屬度。

舉例來說，若有一個記錄每日溫度的時間數列，選擇論域  $U = \{L_1 = \text{非常冷}, L_2 = \text{冷}, L_3 = \text{熱}, L_4 = \text{非常熱}\} = \{\text{非常冷}, \text{冷}, \text{熱}, \text{非常熱}\}$ ，則我們可以藉由使用語言變數替代時間數列  $\{X_t, t = 1, 2, \dots, n\}$  來得到模糊時間數列，下面的例子將更清楚地說明此一想法。

例3.1 令  $\{X_t, t = 1, 2, \dots, n\}$  是一個記錄台灣每日降雨量的時間數列，若我們選擇語言變數集  $U = \{L_i, i = 1, 2, \dots, 5\} = \{\text{非常少}, \text{少}, \text{普通}, \text{多}, \text{非常多}\}$  來取代時間數列  $\{X_t, t = 1, 2, \dots, n\}$ 。

### 3.2 群落中心

應用模糊理論檢定是否發生轉折區間時，首先應將時間數列模糊化，然後再對模糊時間數列進行分類。在分類的過程中，我們會應用到模糊群落中心、隸屬度及模糊熵等觀念。因此將其定義如下：

定義 3.1：模糊時間數列群落中心

設  $\{F_t\}$  為佈於  $\{X_t\}$  上的一個模糊時間數列， $t = 1, \dots, N$ ， $k$  為一正整數。若存在一集合  $C = \{C_1, C_2, \dots, C_m\}$ ，其中  $C_i = \langle c_{i1}, c_{i2}, \dots, c_{ip} \rangle$  使得歐基里得距離平方和最小，即

$$\min \sum_{i=1}^m \sum_{t=1}^N \|F_t - C_i\|^2 \quad (3.2)$$

其中  $\|F_t - C_i\|^2 = \sum_{j=1}^p (\mu_{tj} - c_{ij})^2$ ， $c_{ij}$  是第  $i$  個群落中心對第  $j$  個語言變數的隸屬度。

則我們稱集合  $C = \{C_1, C_2, \dots, C_m\}$  是模糊時間數列  $\{F_t\}$  的群落中心集合。

定義 3.2 時間數列群落隸屬度

設  $\{y_t\}$  為一模糊時間數列， $t = 1, \dots, N$ ，且  $C = \{C_1, C_2, \dots, C_m\}$  是模糊時間數列  $\{y_t\}$  的群落中心集合。令  $\mu_{it}$ ， $i = 1, 2, \dots, m$  表示時間數列中的元素  $y_t$  對每個群落中心  $C_i$  隸屬度，則定義隸屬度為

$$\mu_{it} = 1 - \frac{(y_t - C_i)^2}{\sum_{j=1}^m (y_t - C_j)^2}, \quad i = 1, 2, \dots, m; \quad t = 1, \dots, N \quad (3.3)$$

### 定義 3.3. 模糊熵

設  $\{y_t\}$  為一模糊時間數列， $\mu_{it}$ ， $i = 1, 2, \dots, m$  表示時間數列中的元素  $y_t$  對每個群落中心  $C_i$  隸屬度。 $y_t$  的模糊熵定義為：

$$\delta(y_t) = -\left(\frac{1}{m}\right) \sum_{i=1}^m \mu_{it} \ln(1 - \mu_{it}) \quad (3.4)$$

在模糊識別過程中，模糊熵是類似於熱力學中能量的含蘊標準單位，但它和傳統之熵的意義不同。在定義模糊熵時並非使用機率(probability)論的觀念，而是使用可能性(possibility)理論的觀念。模糊熵表示模糊集合的平均內部訊息量，此訊息量是作為對模糊集合所描述的對象進行分類時，分類的判定標準。因此，將一時間數列化為起因於內部可能訊息量的度量（以模糊熵表示），可用以有效地辨認此時間數列模型是否有結構性改變的發生。此外，並利用  $t$  個時間的平均累加模糊熵來觀察模糊熵的訊息變化情形，並據以做為模型轉折分類的標準。平均累加模糊熵定義如下：

$$MS\delta(y_t) = \frac{1}{t} \sum_{i=1}^t \delta(y_i) \quad (3.5)$$

經由觀察平均累加模糊熵的走勢，便可以判定此時間數列是否有結構性改變。換言之，若平均累加模糊熵發生顯著的群聚（連串）變化時，即意謂模型發生結構性變化；反之，若模型無結構性的變化，則平均累加模糊熵的走勢是相當穩定，而不會有顯著的群聚變化。此外，無論是自然或人文社會科學的研究，時間數列群落中心的認定相當主觀與分歧，若事先無特定的評量尺度，必會增加實証分析的困難度。因此，模糊分類通常會設定一  $\lambda$  門檻水準(threshold level)， $0 < \lambda < 1$ ，來決定分類的標準。根據實証經驗， $\lambda$  值不應選取太大，否則會造成無法分類；但也不應太小，否則會有分類過多的情形發生，通常我們視情況選取  $\lambda$  在 0.1 至 0.001 之間。

利用  $\lambda$  門檻水準將欲檢定之模糊時間數列予以分類後，若相對於每個群落中心的分類結果皆一致，則不用再對此分類結果做調整。但若在某些時點發生分類不相同的情形時，則以相同時點之隸屬度最高者的分類結果為標準，其原因在於隸屬度最高者表示在某一時點對某個群落中心有最高的歸屬度，則其分類結果將

會比隸屬度較低者更為準確。

在完成統一分類結果的步驟之後，接著要應用模糊統計來檢定分類的轉型是否具有統計顯著性。模型結構改變的檢定法則為： $H_0$ :模型無結構改變， $H_1$ :模型有結構改變。決策標準為：設定判定水準 $\alpha$ （通常取 $\alpha$ 在0.1至0.05之間），且 $\alpha$ 值的選取應考慮時間數列的樣本數 $N$ 。因此，我們以 $[\alpha N]$ （ $[\alpha N]$ 為小於或等於 $\alpha N$ 的最大整數）作為分類是否顯著的判定標準。若連串的樣本數大於 $[\alpha N]$ ，則此一連串的樣本歸屬於同一類組。當分類的組數超過一組時，則拒絕 $H_0$ :模型無結構改變，表示此一時間數列之模型在 $\lambda$ 顯著分類水準下發生結構性改變，進而找出其轉折區間。

### 轉折區間偵測演算法

1. 給定一個多變量時間數列 $\{y_t^k\}$ ，其中 $k=1, 2, \dots, s$ 。對每一個 $k=1, 2, \dots, s$ ，尋找時間數列 $\{y_t^k\}$ 的群落中心 $C_i^k$ （ $i=1, 2, \dots, m$ ），將時間數列 $\{y_t^k\}$ 分成 $m$ 群。
2. 令 $\mu_{it}^k$ 表示時間數列中的元素 $y_t^k$ 對每個群落中心 $C_i^k$ 的隸屬度。對 $k=1, 2, \dots, s$ ， $\mu_{it}^k$ 的隸屬度計算方式為

$$\mu_{it}^k = 1 - \frac{(y_t^k - C_i^k)^2}{\sum_{j=1}^m (y_t^k - C_j^k)^2}, \quad i=1, 2, \dots, m; t=1, 2, \dots, N \quad (3.6)$$

3. 使用 $y_t^k$ 的隸屬度來計算 $y_t^k$ 的模糊熵，計算公式為

$$\delta(y_t^k) = -\left(\frac{1}{m}\right) \sum_{i=1}^m \mu_{it}^k \ln(1 - \mu_{it}^k) \quad (3.7)$$

4. 計算對每個 $k$ 及 $t$ 值的平均累加模糊熵，計算公式為

$$MS(\delta(y_t^k)) = \frac{1}{t} \sum_{j=1}^t \delta(y_j^k), \quad t=1, 2, \dots, N \quad (3.8)$$

並計算 $s$ 個平均累加模糊熵的乘積，計算公式為

$$MS_t = MS(\delta(y_t^1))MS(\delta(y_t^2))\dots MS(\delta(y_t^s)), \quad t=1, 2, \dots, N \quad (3.9)$$

尋找 $MS_t$ 的眾數，並表示為 $Med(MS_t)$ 。

5. 選取一個適當的門檻水準 $\lambda$ ，將模糊時間數列分類，分類方式為

$$C_t^* = \begin{cases} 0, & \text{若 } MS_t \in [-\infty, Med(MS_t) - \lambda] \\ 1, & \text{若 } MS_t \in [Med(MS_t) - \lambda, Med(MS_t) + \lambda] \\ 2, & \text{若 } MS_t \in [Med(MS_t) + \lambda, +\infty] \end{cases} \quad (3.10)$$

6. 設定檢定法則為： $H_0$ :模型無結構改變， $H_1$ :模型有結構改變。設定轉折區間



的判定水準  $\alpha$ ，即對每一個類組，若是類組 1 含有連串的樣本數大於  $[\alpha N]$ ，則拒絕  $H_0$ ：模型無結構改變。



## 第四章 實證分析

為驗證本論文所建議的轉折區間演算法之成效，我們取歐盟 3 個重要經濟體（德國、法國及希臘）之總體經濟指標 GDP 為例來進行檢測。在變量的部份我們考慮通貨膨脹率(inflation rate)、GDP 成長率(percentage growth)及投資率(investment rate)等 3 個變量，資料來源為 1972-2006 年 OECD 的資料庫，研究目的為偵測德國、法國及希臘之總體經濟指標 GDP 轉折區間，並將轉折區間發生的時間點與歷史上歐盟重大事件來比較其間的相關性。從歷史上來看，在 1971-2006 年間較顯著之事件有：

1985, 1986 – 申根和約(Schengen Agreement)及歐洲單一法案(Single European Act)簽署

1989, 1990 – 柏林圍牆倒塌及歐盟提出農村發展政策(NRD)

1990 – 德國、法國及比荷盧三國關稅同盟簽署申根和約(Schengen Agreement)旨在消除國境關卡的管制

1992 – 希臘認可歐洲聯盟的協議

1999 – 歐元開始流通

2001 – 希臘加入歐元區

### 德國、法國及希臘之轉折區間偵測演算

1. 取得德國、法國及希臘在 1972-2006 年間之通貨膨脹率、GDP 成長率及投資率等 3 個變量的數據（即  $k = 3$ ， $N = 35$ ）。並利用 K-means 演算法計算 3 個群落中心（即  $m = 3$ ），並將時間數列分成 3 群，此乃根據一般的實証分析經驗及時間數列的走勢研判的 (Wu 與 Chen, 1999)。

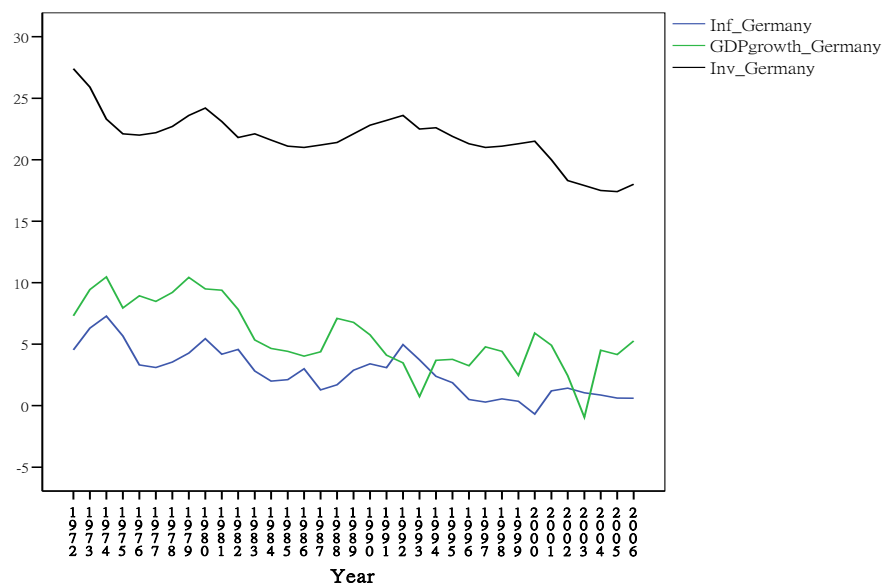


圖 4.1 德國的通貨膨脹率、GDP 成長率及投資率

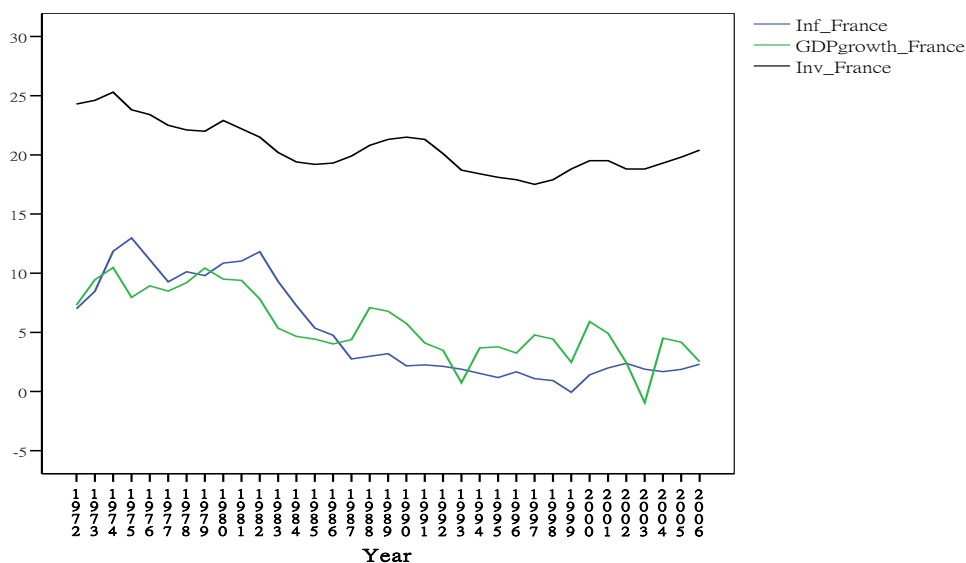


圖 4.2 法國的通貨膨脹率、GDP 成長率及投資率

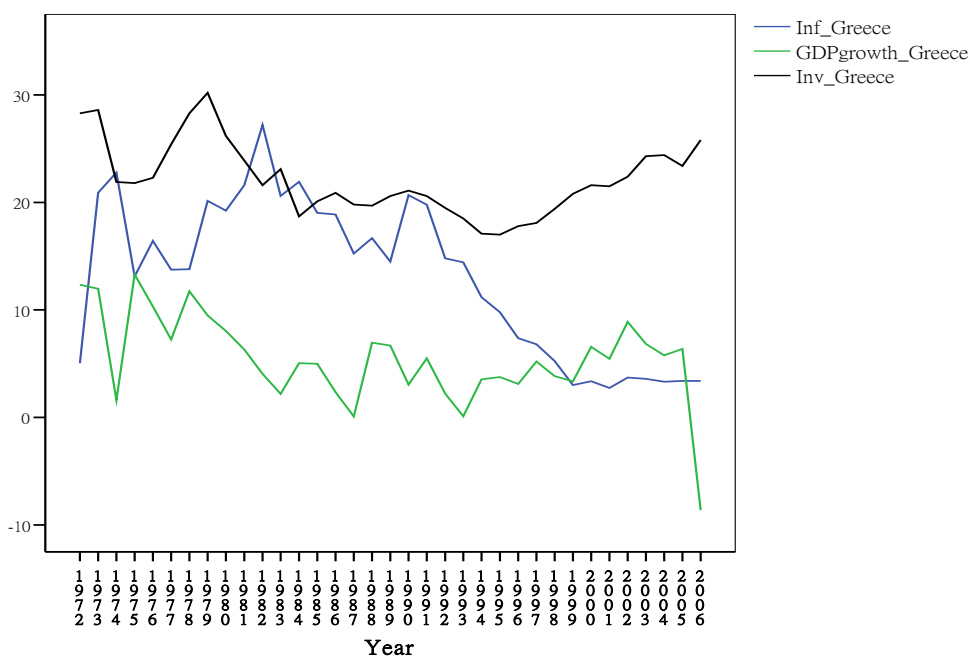


圖 4.3 希臘的通貨膨脹率、GDP 成長率及投資率

2. 利用公式 3.6 計算時間數列中的元素對 3 個群落中心的隸屬度。
3. 利用公式 3.7 計算模糊熵。
4. 利用公式 3.8 計算平均累加模糊熵，並利用公式 3.9 計算 3 個平均累加模糊熵的乘積，並從乘積中尋找眾數。
5. 選取一適當的門檻水準  $\lambda$ ，並依公式 3.10 將模糊時間數列分 3 類，並將  $C_t^*$  圖

描繪出來。

6. 設定檢定法則為： $H_0$ : 模型無結構改變， $H_1$ : 模型有結構改變。設定轉折區間判定水準  $\alpha = 0.01$ ，即對每一類組，若是類組 1 含有連串的樣本數大於  $[\alpha N]$  (即  $[0.01 \times 35] = 0$ )，則拒絕  $H_0$ 。

德國的  $C_t^*$  圖如圖 4.4，觀察此圖以  $\alpha = 0.01$  的顯著水準來說，因為 1988 至 1989 的連串樣本數大於  $[\alpha N]$ ，所以我們可以拒絕「德國在 1988 至 1989 年沒有結構改變」的假說。若參考歷史事件來看，當年柏林圍牆的倒塌及歐盟的農村發展政策對德國經濟有著不小的衝擊。

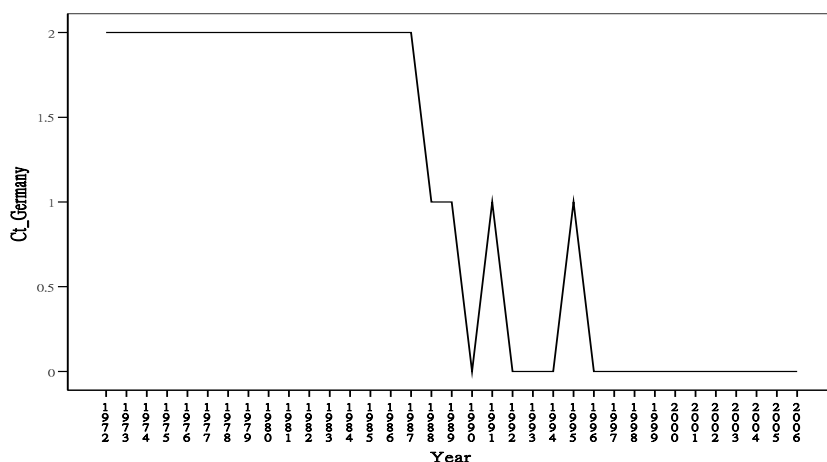


圖 4.4 德國的  $C_t^*$  圖

法國的  $C_t^*$  圖如圖 4.5，觀察此圖以  $\alpha = 0.01$  的顯著水準來說，因為 1991 至 1992 的連串樣本數大於  $[\alpha N]$ ，所以拒絕「法國在 1991 至 1992 年沒有結構改變」的假說。對此，我們認為此乃法國簽署旨在漸次消除國境關卡管制的申根和約 (Schengen Agreement) 所帶來的改變。

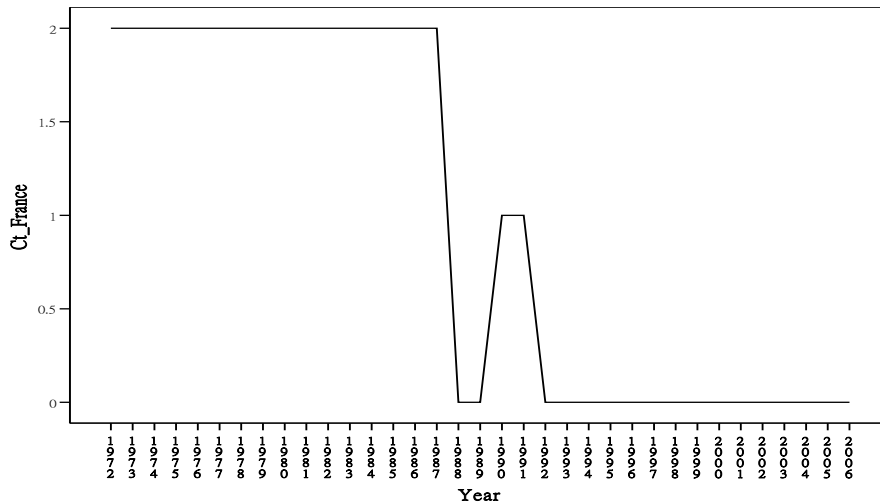


圖 4.5 法國的  $C_t^*$  圖

希臘的  $C_t^*$  圖如圖 4.6，觀察此圖以  $\alpha = 0.01$  的顯著水準來說，因為 1988 至 1989 及 1993 至 1994 的連串樣本數大於  $[aM]$ ，所以拒絕「希臘在 1988 至 1989 及 1993 至 1994 年沒有結構改變」的假說。若與歷史事件對比，前者 1988 與 1989 的改變可能源自於希臘於 1988 年所實施的間接稅改革措施，後者則可能是導因於 1992 年希臘認可歐洲聯盟的協議。

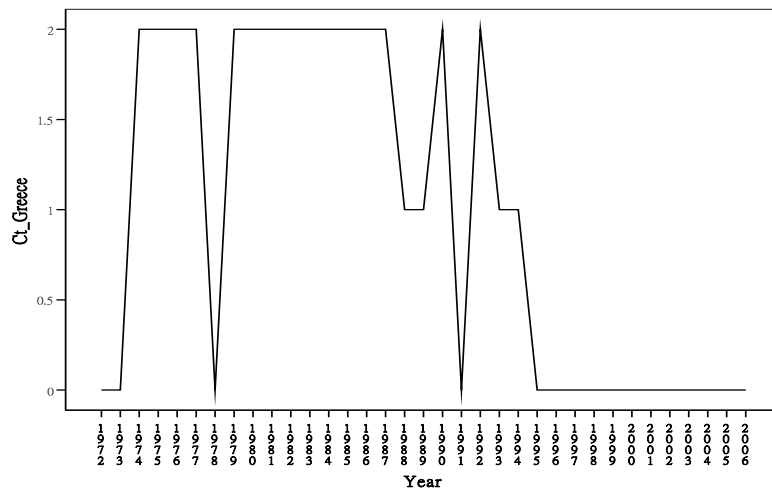


圖 4.6 希臘的  $C_t^*$  圖



## 第五章 結論

在時間數列的分析過程中，統計數值資料本身的不確定性與模糊性，可說是傳統時間數列模型不易建構的困難所在，Manski (1990)曾指出，數值資料有過度解釋的危險。倘若我們利用此假性的精確值，來作因果分析或計量度量，可能造成因果判定偏差、決策模式誤導或擴大預測結果和實際狀態之間的差異。人類的思維模式中，計量過程本身即充滿著模糊性，這似乎是單純化的結果，但過分單純化又會阻礙計量方法與標準的客觀性。若能結合較穩健性與符合實際狀況的模糊統計分析，應能避免這樣的情況發生。

自從 Zadeh 教授在 1965 年提出了模糊集合的概念，做為測試不明確隸屬度函數的工具以來，許多的模糊研究嘗試以此方法設定理論架構，並試圖以此一理論作為建構更為先進模糊理論的基礎。雖然，過去模糊理論的應用範疇，大部份侷限在計算機科學領域裡的專家系統研究中。但近來，由於人文社會科學的測度理論發展過程裡，複雜的人文社會科學現象，無法以傳統數值模型充分合理解釋，進而導致模糊統計與模糊相關性日漸受到重視，這應是一種自然的發展結果。有鑑於此，目前將模糊理論應用在人文社會科學研究中的情形已相當普遍。

模糊分類(fuzzy clustering)是一種用來處理資料分類及型態識別的新方法，在本論文中我們提出以轉折區間的概念來取代傳統時間數列發生結構改變的轉折點，並提供一套程序來檢測在顯著水準  $\alpha$  下時間數列是否存在轉折區間。轉折區間的存在與否幫助我們可以辨別時間數列是否有結構上的轉變，以本論文中的實例來看，我們所建議的程序可以有效的檢測出轉折區間。

若是跟傳統的方法相比，本論文所提出的方法具有幾個優點：

1. 時間數列的結構訊息對於本論文的轉折區間檢測程序來說並不需要。
2. 此法亦適用於歷史資料的語意處理，處理過程與本論文所建議十分相似，只需在步驟 2 加上學者經驗判斷即可。

儘管具有上述的優點，但在本論文中尚有待解決的問題，亦是未來值得研究的方向，分述如下：

1. 本論文雖提及多變量模糊時間數列模式建構與預測，但由於時間的關係，並未於文中舉出實例的建構過程。
2. 字詞的語意應該細心地定義以求穩定，像是「轉折點」所代表的究竟是平均數的改變、變異數的改變、特徵參數的改變或是模式的改變，這點必須在檢測程序實施前先行做好。

## 參考文獻

- Balke, N. S. (1993), Detecting level shifts in time series, *Journal of Business Economic Statistics*, 11(1), 81-92.
- Chen, S. M. (1996), Forecasting enrollments based on fuzzy time series, *Fuzzy Sets and Systems*, 81, 311-319.
- Chiang, D., L. Chow, and Y. Wang (2000), Mining time series data by a fuzzy linguistic summary system, *Fuzzy Sets and Systems*, 112, 419-432.
- Clymer, J., P. Corey, and J. Gardner (1992), Discrete event fuzzy airport control, *IEEE Transactions on Systems, Man, and Cybernetics*, 22(2), 343-351.
- Cutsem, B. V., and I. Gath (1993), Detection of outliers and robust estimation using fuzzy clustering, *Computational Statistics and Data Analysis*, 15, 47-61.
- Dubois, D. and H. Prade (1991), Fuzzy sets in approximate reasoning, Part I : Inference with possibility distributions, *Fuzzy Sets and Systems*, 40, 143-202.
- Graham, B. P. and R. B. Newell (1989), Fuzzy adaptive control of a first-order process, *Fuzzy Sets and Systems*, 31, 47-65.
- Hathaway, R. J., and J. C. Bezdek (1993), Switching regression models and fuzzy clustering, *IEEE Transactions of Fuzzy Systems*, 1, 195-204.
- Hinkley, D. V. (1971), Inference about the change point from cumulative sum test, *Biometria*, 26, 279-284.
- Hsu, D. A. (1982), A Bayesian robust detection of shift in the risk structure of stock market returns, *Journal of the American Statistical Association*, 77, 29-39.
- Inclan, C. & Tiao, G. C. (1994), Use of cumulative sums of squares for retrospective detection of changes of variance, *Journal of the American Statistical Association*, 89(427), 913-924.
- Lee, Y. C., C. Hwnag, and Y. P. Shih (1994), A combined approach to fuzzy model identification. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(5), 736-743.
- Lowen, R. (1990), A fuzzy language interpolation theorem, *Fuzzy Sets and Systems*, 34, 33-38.
- Manski, C. (1990), The use of intention data to predict behavior: a best case analysis, *Journal of the American Statistical Association*, 85, 934-940.
- Nguyen, H. and M. Sugeno (1998), *Fuzzy Modeling and Control*, CRC Press.

- Page, E. S. (1955), A test for change in a parameter occurring at an unknown point, *Biometrika*, 42, 523-527.
- Park, Y. M., U. C. Moon, and K. Y. Lee (1995), A self-organizing fuzzy logic controller for dynamic systems using fuzzy auto-regressive moving average (FARMA) model, *IEEE Transactions on Fuzzy Systems*, 3(1), 75-82.
- Romer, C., A. Kandel, and E. Backer (1995), Fuzzy partitions of the sample space and fuzzy parameter hypotheses, *IEEE Transactions on Systems, Man and Cybernetics*, 25(9), 1314-1321.
- Ruspini, E. (1991), Approximate reasoning: past, present, future, *Information Sciences*, 57, 297-317.
- Song, Q. and B. S. Chissom (1993a), Forecasting enrollments with fuzzy time series—Part I, *Fuzzy Sets and Systems*, 54, 1-9.
- Song, Q. and B. S. Chissom (1993b), Fuzzy time series and its models, *Fuzzy Sets and Systems*, 54, 269-277.
- Song, Q. and B. S. Chissom (1994), Forecasting enrollments with fuzzy time series—Part II, *Fuzzy Sets and Systems*, 62, 1-8.
- Song, Q., R. P. Leland, and B. S. Chissom (1997), Fuzzy stochastic fuzzy time series and its models, *Fuzzy Sets and Systems*, 88, 333-341.
- Sugeno, M. and K. Tanaka (1991), Successive identification of a fuzzy model and its applications to prediction of a complex system, *Fuzzy Sets and Systems*, 42, 315-334.
- Tong, R. M. (1978), Synthesis of fuzzy models for industrial processes, *International Journal of General Systems*, Vol.4, 143-162.
- Tsay, R. S. (1991), Detecting and modeling non-linearity in univariate time series analysis, *Statistica Sinica*, 1(2), 431-451.
- Tseng, F., G. Tzeng, H. Yu, and B. Yuan (2001), Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems*, 118, 9-19.
- Tseng, T. and C. Klein (1992), A new algorithm for fuzzy multicriteria decision making, *International Journal of Approximate Reasoning*, 6, 45-66.
- Werners, B. (1987), An interactive fuzzy programming system, *Fuzzy Sets and Systems*, 23, 131-147.
- Wu, B. (1994), Identification environment and robust forecasting for nonlinear time series, *Computational Economics*, 7, 37-53.
- Wu, W. (1986), Fuzzy reasoning and fuzzy relational equations, *Fuzzy Sets and Systems*,

20, 67-78.

- Wu, B. and C. Sun (1996), Fuzzy statistics and computation on the lexical semantics, *Language, Information and Computation (PACLIC 11)*, 337-346, Seoul, Korea.
- Wu, B. and M. Chen (1999), Use of fuzzy statistical technique in change periods detection of nonlinear time series, *Applied Mathematics and Computation*, 99, 241-254.
- Wu, B. and S. Hung (1999) , A fuzzy identification procedure for nonlinear time series: with example on ARCH and bilinear models, *Fuzzy Sets and Systems*, 108, 275-287.
- Xu, C. W. and Y. Z. Lee (1987) , Fuzzy model identification and self-learning for dynamic systems, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SCM-17, 683-689.
- Yang, M. (1993), A Survey of Fuzzy Clustering, *Mathematical and Computer Modelling*, 18(11), 1-16.
- Yoshinari, Y., W. Pedrycz, and K. Hirota (1993), Construction of fuzzy models through clustering techniques, *Fuzzy Sets and Systems*, 54, 157-165.
- Zadeh, L. A. (1965), Fuzzy Sets, *Information and Control*, 8, 338-353.
- Zimmermann, H. J. (1991), *Fuzzy Set Theory and Its Applications*, Boston: Kluwer Academic.