

馬氏-田口系統之特性探討及其於信用評分之應用

研究生：王慧君

指導教授：蘇朝墩博士

國立交通大學

工業工程與管理系

摘要

田口方法已廣泛地被應用在許多工程方面以改善產品及製程的品質。馬氏-田口系統(Mahalanobis-Taguchi System, MTS)是田口博士提出的新方法，它結合品質工程原理及馬氏距離，透過建構多變量量表來進行診斷及預測以提供決策之參考。

本研究首先探討 MTS 之特性，藉由改變各種實驗條件，包括馬氏空間、抽樣策略及訊號雜音比等，以評估 MTS 在分類之績效，並且利用紫羅蘭資料進行 MTS 特性之說明，以擴展此方法未來之發展與應用。其次，本研究將 MTS 應用在信用評分模式上，信用評分模式已廣泛被使用在信用決策上，主要可減少信用分析的時間與成本，雖然目前已存在不少信用評分模式，但傳統的模式並沒有考慮雜音對模式之影響，本研究以 MTS 建構一信用評分模式，並利用國內某銀行之實例驗證模式之可行性。此外，針對 MTS 處理類別資料之缺點，本研究亦提出一整合羅吉斯迴歸與田口方法之修正模式，以改善其績效；並將 MTS 與修正模式之實證結果與決策樹以及鑑別分析進行比較，發現 MTS 及其修正模式在信用評分的領域上可獲得很好的應用。

最後，本研究也針對 MTS 的幾個議題及未來的研究方向進行說明。

關鍵詞：馬氏-田口系統、馬氏距離、直交表、訊號雜音比、羅吉斯迴歸、信用評分

Analysis of MTS: Application of Credit Scoring

Student: Huei-Chun Wang

Advisor: Dr. Chao-Ton Su

Department of Industrial Engineering and Management

National Chiao Tung University

Abstract

Taguchi Methods have been successfully applied to many engineering application to improve the performance of product and process. The Mahalanobis-Taguchi System (MTS) developed by Dr. Taguchi is a relatively new technology for diagnosis and forecasting using multivariate data. MTS is based on the principles of quality engineering and Mahalanobis distance (MD). This technology is aimed at providing a better prediction for multivariate data through the construction of a multivariate measurement scale. The results of MTS can be a reference for making decisions.

For development and applications of MTS, to understand the features is very important. In this dissertation, first, we study the features of MTS, and evaluate the performance of MTS. Through changing the conditions of experiments including Mahalanobis Space, Sampling and Signal to Noise ratios, the performance of MTS was evaluated. Iris data was used to illustrate the performance of MTS. Secondly, the credit scoring system based on MTS is proposed. Credit scoring is widely used to make credit decisions, to reduce the cost of credit analysis and to make decisions fast. There are many credit scoring techniques, however, traditional credit scoring models do not consider the influence of noises. The effectiveness of the proposed MTS approach is demonstrated by real case data from a large Taiwanese bank.

In addition, we propose a modified approach that is combined logistic regression (LR) with Taguchi's approach for improving the shortcoming of categorical data of MTS. Results are benchmarked against two traditional methods including decision tree and linear

discriminant analysis. Due to the results of the above, both MTS and modified approach (LRTM) can be applied to credit scoring system.

Finally, several important issues regarding to the MTS and the suggestions of future research are summarized in the conclusion.

Key words: Mahalanobis-Taguchi System (MTS), Mahalanobis distance, orthogonal array, Signal-to-Noise ratio, Logistic regression, Credit scoring.



誌 謝

畢業是一個階段的結束，也是另一個階段的開始。回首這五年博士班的生涯，由於有許多貴人的相助，才能使我在家庭、工作與學業的壓力下順利完成學業，得以圓人生的一個夢。

本論文得以順利完成，首先要感謝指導教授 蘇朝墩 博士的悉心指導與訓練，從老師的身上看到了治學的精神、嚴謹的態度，為我開啟學術研究的一扇窗，師恩之浩瀚沒齒難忘，自此將秉持這樣的精神，在學術的領域中繼續努力。

在論文的口試期間，感謝 李榮貴 教授、沙永傑 教授、陳飛龍 教授、邱志洲 教授、駱景堯 教授及 陳穆臻 教授在百忙中撥冗審查並給予許多寶貴的建議，使我受益良多。並感謝 邱志洲 教授提供本次研究之相關資料，以協助此論文的完成。

同時要感謝 戴久永 教授及 唐麗英 教授在五年前為我所寫的推薦信，使得我能夠順利進入交大攻讀博士學位。在修業的過程中，感謝大華技術學院工管系主任 蔡志弘 博士及 姚銀河 博士的關心，感謝同仁 嘉麗、玫玲、永愛、幼珍、金帶、于婷、伶恩、麗妃、芬如、佩芬、惠珍 及 何東洋 博士的鼓勵及 陳烈 博士的協助，更感謝大華技術學院校長 陳文洲 教授的提攜。

在學過程中，感謝有 世洲、心怡、春美 同學的相互鼓勵與好友 璧玉 的支持及學弟 志華 的協助，使得我更堅定地走完這條漫長的路。

感謝父母養育之劬勞，感謝母親、公婆替我分擔家務及照顧孩子的辛勞，感謝先生全心全力的支持與體諒，感謝兩個孩子的體貼，有了你們的付出及陪伴才能讓我完成這個夢想。

最後，謹以此論文獻給我最敬愛的父親，以慰父親在天之靈。

目錄

中文摘要.....	I
英文摘要.....	II
謝誌.....	IV
目錄.....	V
圖目錄.....	VII
表目錄.....	IX
第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究目的.....	2
1.3 研究架構.....	3
第二章 相關研究.....	4
2.1 資料探勘.....	4
2.2 資料探勘中分類技術之探討.....	6
2.2.1 分類決策.....	7
2.2.2 分類技術的探討.....	8
第三章 MT 系統.....	15
3.1 田口的多變量系統.....	15
3.2 MT 系統之介紹.....	18
3.2.1 馬氏距離.....	18
3.2.2 MT 系統.....	19
3.3 MT 系統之執行步驟.....	22
3.3.1 MTS 之執行步驟.....	23
3.3.2 MTGS 之執行步驟.....	24
3.3.3 MTA 之執行步驟.....	25
3.4 MT 系統與其他多變量方法的比較.....	26
3.5 MT 系統之相關議題.....	28
3.5.1 直交表及訊號雜音比的角色.....	28
3.5.2 雜音的處理.....	32
3.5.3 閾值的決定.....	33
3.5.4 其他議題.....	34
3.6 結論.....	35
第四章 MTS 特性之探討與比較.....	37
4.1 資料說明與實驗規劃.....	37
4.2 馬氏空間之選擇.....	39
4.2.1 實驗說明及過程.....	39
4.2.2 結論.....	52

4.3 抽樣策略之探討.....	54
4.3.1 實驗說明與過程.....	54
4.3.2 結論.....	61
4.4 訊號雜音比之影響.....	62
4.4.1 實驗說明與過程.....	62
4.4.2 結論.....	64
4.5 閾值的討論.....	64
4.6 MTS 與其他方法之比較.....	66
第五章 MTS 在信用評分之應用.....	68
5.1 信用評分及現況.....	68
5.1.1 信用評分及相關研究.....	68
5.1.2 評估信用評分模式的準則.....	70
5.1.3 訓練及測試樣本的抽樣策略.....	72
5.1.4 國內目前常用的信用評分方式.....	74
5.2 MTS 信用評分模式之建構.....	75
5.3 MTS 信用評分模式之實證研究.....	80
5.3.1 資料之前處理與檢查.....	80
5.3.2 各種 MTS 信用評分模式之建構.....	83
5.3.3 以 MTS 建構信用評分之優缺點探討.....	92
5.4 MTS 之修正模式.....	93
5.5 各種信用評分模式比較.....	101
第六章 結論.....	103
6.1 研究結果與結論.....	103
6.2 未來研究建議.....	104
參考文獻.....	105

圖目錄

圖 2.1 知識發掘過程.....	5
圖 2.2 資料探勘的三階段.....	5
圖 2.3 資料探勘應用架構.....	6
圖 2.4 分類決策的兩種方法.....	7
圖 3.1 多變量系統圖.....	15
圖 3.2 田口式品質工程的整體觀.....	16
圖 3.3 GSP 程序.....	21
圖 3.4 MT 系統之執行步驟.....	22
圖 3.5 MT 系統之選用流程.....	36
圖 4.1 紫羅蘭之盒鬚圖.....	38
圖 4.2 紫羅蘭之散佈圖.....	39
圖 4.3 馬氏空間之實驗規劃.....	40
圖 4.4 實驗 4-1 全因子模式訓練樣本之馬氏距離分配圖.....	42
圖 4.5 實驗 4-1 全因子模式測試樣本之馬氏距離分配圖.....	42
圖 4.6 實驗 4-1 全因子模式之 ROC 曲線.....	42
圖 4.7 實驗 4-1 因子效果圖.....	44
圖 4.8 實驗 4-1 之盒鬚圖.....	44
圖 4.9 實驗 4-1 重要因子模式訓練樣本之馬氏距離分配圖.....	45
圖 4.10 實驗 4-1 重要因子模式測試樣本之馬氏距離分配圖.....	45
圖 4.11 實驗 4-1 重要因子模式之 ROC 曲線.....	45
圖 4.12 實驗 4-2 全因子模式測試樣本馬氏距離分配圖.....	48
圖 4.13 實驗 4-2 全因子模式之 ROC 曲線.....	48
圖 4.14 實驗 4-2 因子效果圖.....	48
圖 4.15 實驗 4-2 重要因子模式測試樣本馬氏距離分配圖.....	49
圖 4.16 實驗 4-3 全因子模式測試樣本馬氏距離分配圖.....	50
圖 4.17 實驗 4-3 全因子之 ROC 曲線.....	50
圖 4.18 實驗 4-3 因子效果圖.....	51
圖 4.19 實驗 4-3 重要因子模式訓練樣本馬氏距離分配圖.....	51
圖 4.20 實驗 4-3 重要因子模式測試樣本馬氏距離分配圖.....	51
圖 4.21 三組實驗 ROC 曲線之比較.....	53
圖 4.22 抽樣策略之實驗規劃.....	54
圖 4.23 實驗 4-4 全因子模式訓練樣本馬氏距離分配圖.....	56
圖 4.24 實驗 4-4 全因子模式測試樣本馬氏距離分配圖.....	56
圖 4.25 實驗 4-4 全因子模式之 ROC 曲線.....	56
圖 4.26 實驗 4-4 因子效果圖.....	58

圖 4.27 實驗 4-5 全因子模式訓練樣本馬氏距離分配圖.....	59
圖 4.28 實驗 4-5 全因子模式測試樣本馬氏距離分配圖.....	60
圖 4.29 實驗 4-5 因子效果圖.....	60
圖 4.30 實驗 4-6 之因子效果圖.....	63
圖 5.1 信用評分模式之階段.....	69
圖 5.2 傳統信用評分與 MTS 信用評分之比較.....	75
圖 5.3 以資料探勘流程建構信用評分的程序.....	78
圖 5.4 MTS 信用評分模式流程圖.....	79
圖 5.5 實驗 5-1 三組訓練及測試樣本分類正確率之 ROC 曲線.....	84
圖 5.6 實驗 5-1 重要因子效果圖.....	86
圖 5.7 實驗 5-2 之重要因子效果圖.....	87
圖 5.8 實驗 5-3 全因子模式訓練及測試樣本分類正確率之 ROC 曲線.....	88
圖 5.9 實驗 5-3 第一組資料之重要因子效果圖.....	89
圖 5.10 實驗 5-3 第二組資料之重要因子效果圖.....	89
圖 5.11 實驗 5-3 第三組資料之重要因子效果圖.....	89
圖 5.12 實驗 5-3 重要因子模式訓練及測試樣本分類正確率之 ROC 曲線.....	90
圖 5.13 LRTM 信用模式.....	94



表目錄

表 2.1 各種技術的分類一覽表.....	9
表 2.2 資料驅動與模式驅動方法之比較.....	9
表 3.1 MT 系統方法一覽表.....	22
表 3.2 MT 系統三種方法的比較表.....	26
表 3.3 鑑別分析與馬氏田口系統比較表.....	27
表 3.4 動態訊號雜音比之變異數分析.....	31
表 3.5 MT 系列中各種訊號雜音比.....	32
表 4.1 紫羅蘭資料預檢表.....	38
表 4.2 紫羅蘭之相關矩陣表.....	38
表 4.3 各種實驗規劃一覽表.....	39
表 4.4 實驗 4-1 以 SETASO 為馬氏空間之相關資料.....	41
表 4.5 實驗 4-1 全因子模式訓練及測試樣本之正確率.....	43
表 4.6 實驗 4-1 直交表配置及訊號雜音比.....	44
表 4.7 實驗 4-1 重要因子模式訓練及測試樣本之分類正確率.....	46
表 4.8 實驗 4-1 全因子及重要因子模式之訊號雜音比及正確率之比較表.....	46
表 4.9 實驗 4-2 以 VERSICOLOR 為馬氏空間之相關資料.....	47
表 4.10 實驗 4-2 全因子及重要因子模式之訊號雜音比及正確率之比較表.....	49
表 4.11 實驗 4-3 以 VIRGINICA 為馬氏空間之相關資料.....	50
表 4.12 實驗 4-3 全因子及重要因子模式之訊號雜音比及正確率之比較表.....	52
表 4.13 三組實驗之全因子模式測試樣本正確率.....	54
表 4.14 三組實驗之重要因子模式測試樣本正確率.....	54
表 4.15 實驗 4-4 SETASO 為馬氏空間之相關資料.....	55
表 4.16 實驗 4-4 全因子模式訓練及測試樣本之分類正確率.....	56
表 4.17 實驗 4-4 重要因子訓練及測試樣本之分類正確率.....	57
表 4.18 實驗 4-4 全因子及重要因子模式之訊號雜音比及正確率比較表.....	58
表 4.19 實驗 4-5 以 SETASO 為馬氏空間之相關資料.....	59
表 4.20 實驗 4-5 全因子模式訓練及測試樣本之分類正確率.....	60
表 4.21 實驗 4-5 重要因子模式訓練及測試樣本之分類正確率.....	61
表 4.22 實驗 4-5 全因子及重要因子模式之訊號雜音比及正確率比較表.....	61
表 4.23 抽樣策略之全因子測試樣本正確率比較表.....	62
表 4.24 實驗 4-6 之 ANOVA 分析表.....	63
表 4.25 實驗 4-6 全因子與重要因子模式之訊號雜音比及正確率之比較表.....	63
表 4.26 兩組實驗之訊號雜音比的增益及正確率比較.....	64
表 4.27 HOTELLING T^2 與 MTS 閾值之轉換表.....	65
表 4.28 以 HOTELLING T^2 所建立之閾值與正確率.....	65
表 4.29 三種分類技術之分類結果比較表.....	66
表 5.1 混淆矩陣.....	70

表 5.2 國內目前常用的信用評分方法一覽表.....	74
表 5.3 信用資料一覽表.....	81
表 5.4 客戶基本資料之次數分配表.....	81
表 5.5 客戶基本資料之敘述統計表.....	82
表 5.6 基本資料與信用類別交叉分析結果一覽表.....	82
表 5.7 客戶基本資料之 T 檢定一覽表.....	82
表 5.8 相關分析一覽表.....	83
表 5.9 各種信用評分模式一覽表.....	83
表 5.10 實驗 5-1 之馬氏距離一覽表.....	84
表 5.11 各組的訓練樣本在不同閾值下分類正確率一覽表.....	85
表 5.12 實驗 5-1 之直交表配置.....	86
表 5.13 實驗 5-2 之直交表配置.....	87
表 5.14 實驗 5-3 之馬氏距離一覽表.....	88
表 5.15 實驗 5-3 之直交表配置.....	89
表 5.16 實驗 5-3 之三組資料重要因子一覽表.....	90
表 5.17 實驗 5-3 之三組資料重要因子模式之正確率.....	90
表 5.18 LRTM 實驗規劃一覽表.....	95
表 5.19 實驗 5-5 訓練及測試樣本分類正確率.....	95
表 5.20 實驗 5-5 羅吉斯迴歸模式迴歸係數一覽表.....	97
表 5.21 LRTM 模式與 MTS 模式分類正確率之比較表--全因子模式.....	99
表 5.22 LRTM 模式與 MTS 模式分類正確率之比較表--重要因子模式.....	99
表 5.23 實驗 5-6 訓練及測試樣本分類正確率.....	100
表 5.24 實驗 5-6 與實驗 5-3 正確率之比較表--全因子模式.....	100
表 5.25 實驗 5-6 與實驗 5-3 正確率之比較表--重要因子模式.....	100
表 5.26 四種分類技術之分類結果比較表.....	101

第一章 緒論

1.1 研究背景與動機

不論企業或個人都會面臨到需要進行決策的狀況，而一個決策的進行必須考量許多的因素，往往這些因素間又是彼此相關的，若忽略因素間的關係很可能會做出錯誤的判斷。這種綜合的考量，就形成了多變量的問題。

田口博士以其穩健設計而聞名，其靜態與動態參數設計方法已成功地應用在產品及製程品質的改善上，而 MT(Mahalanobis-Taguchi)系統則是最近提出的新方法，整合數學、統計及穩健設計之原理，主要用來進行多變量的診斷及預測。MT 系統結合馬氏距離(Mahalanobis Distance, MD)及田口方法(Taguchi Methods)，其中馬氏距離是用來量測事物間的相似性，而田口方法則是以有效率方式找出變數之最佳組合。MT 系統是一種藉由建構多變量量表(Measurement Scale)來協助進行決策的方式，主要目的是透過量表來量測觀察值異常的程度以提供決策之參考。

MT 系統包括馬氏-田口系統(Mahalanobis-Taguchi System, MTS)、MTGS (Mahalanobis-Taguchi Gram Schmidt)、MTA (Mahalanobis-Taguchi Adjoint)等方法，其中 MTGS 及 MTA 是由 MTS 延伸而來，發展 MT 系統的主要目的是為克服傳統多變量技術的缺點，田口博士以一個不同的角度來看待有關多變量決策的問題。根據田口博士的說法有關 MT 系統的應用，包括疾病診斷、火警偵測、地震預測、氣象預報、聲音辨識及信用評分等(Taguchi, 2001)。

雖然田口博士提出了 MT 系統進行多變量的診斷，其利用馬氏距離當成量表以量測觀察值異常程度及田口方法最佳化量表的概念是相當新穎的，但有關 MT 系統的相關問題仍有頗多值得研究之處。因此，本論文乃針對 MTS 之特性及以 MTS 為分類方法之應用作一探討。

1.2 研究目的

本研究主要目的有以下幾點：

一、研究 MTS 之特性。由於 MTS 是一個新的方法，有關此方法的特性並沒有相關的研究，瞭解方法之特性與優缺點，對於此方法的發展及應用有其必要性。本研究探討馬氏空間的選擇、抽樣策略及訊號雜音比(Signal-to-Noise Ratio, SN 比)是否對 MTS 的分類績效有所影響，及有關閾值訂定的問題。並利用費氏紫羅蘭(IRIS)的資料檔來進行相關的說明。

二、MTS 在信用評分應用之探討。由於信用評分是一種綜合的評價，信用的辨識因子間彼此都有關連，這符合 MTS 方法的特質，可將 MTS 應用於信用評分中。儘管目前已有許多研究探討信用評分的問題(Thomas, 2000)，但大多數研究均以提昇分類之正確率為主，較少研究考慮雜音對信用評分系統的影響。本研究利用 MTS 處理信用評分的議題主要有以下的考量：(1)由於 MTS 有衡量異常嚴重程度的能力，所以除了提供分類結果之外，還可以衡量客戶信用異常的程度，決策者可根據此程度彈性調整閾值以進行決策；(2)利用 MTS 概念來設計具穩健性的信用評分系統。

三、MTS 缺點之改善。由於 MTS 仍存在一些問題，本研究嘗試針對 MTS 在類別資料處理上的缺點提出修正之建議，並針對 MTS 應用在信用評分上時之缺點提出一個修正模式。利用羅吉斯迴歸(Logistic Regression, LR)來修正馬氏距離的想法，因為羅吉斯迴歸具有不受變數的分配限制外，還具有預測分類之組別及計算事件發生機率的特性，將羅吉斯迴歸結合田口方法，以 LRTM 代表所提之修正模式，其中 LR 代表羅吉斯迴歸，而 TM 則代表田口方法，以改善 MTS 處理類別變數之缺點。

1.3 研究架構

本研究首先整理 MT 系統。有關 MT 系統之文獻及研究相當少，本論文的第一部份是整理 MT 系統中 MTS、MTGS、MTA 之間的關係及差異，包括執行步驟、應用時機及方法間彼此的關係。

其次，針對發展時間相對較長的 MTS 之特性加以研究，此部份分成兩個議題討論：(一)MTS 特性的探討；(二)MTS 與其他分類方法之比較。有關特性探討的部份，包括(1)馬氏空間的選擇對於 MTS 分類績效之影響。由於 MTS 的使用過程中必須選擇適當的馬氏空間以建構量表，面對分類的問題應該選擇哪一類別當成馬氏空間，是值得探討的問題。(2)抽樣策略的選擇對 MTS 分類績效之影響。在此的抽樣策略是指樣本的比率，也就是觀察訓練樣本與測試樣本的比率變化對 MTS 分類績效之影響。傳統分類方法之績效會受到訓練樣本與測試樣本比例的影響，但 MTS 是否也會受抽樣策略的影響亦是本研究探討的議題之一。(3)訊號雜音比的使用對 MTS 分類績效之影響。田口建議使用望大或動態特性之訊號雜音比；然而應該使用何種訊號雜音比較為恰當？本研究利用望大特性、動態特性來探討田口所提出之建議。(4)閾值的決定，本研究嘗試利用 Hotelling T^2 管制界限及最小化總錯誤率的方式來探討有關閾值的決定。

由於 MTS 是一種可進行綜合診斷的方法，其應用的領域包括疾病診斷、火警偵測、地震預測、氣象預報、聲音辨識及信用評分等，本研究將 MTS 應用在信用評分上。除探討 MTS 在信用評分應用之績效外，並討論 MTS 在信用評分應用之優缺點，針對其缺點提出改善的方法，最後將 MTS 與修正的方法和傳統分類方法進行比較。

本論文共分成六章，第一章為緒論，說明研究動機、目的與研究架構，第二章相關研究整理回顧，第三章探討並整理 MT 系統之概念及內容，第四章為 MTS 特性之探討與比較，第五章探討 MTS 於信用評分之應用，藉由國內某一銀行信用卡部門所提供之實際資料為分析的對象來驗證 MTS 在信用評分應用之可行性，並針對 MTS 應用在信用評分上之缺點提出改善，最後將結果與其他方法進行比較，第六章則為結論與未來研究之建議。

第二章 相關研究

2.1 資料探勘

1980 年代末期出現知識發掘過程(Knowledge Discovery in Databases, KDD)這個名詞，很快地被人工智慧及機器學習的專家所使用。知識發掘過程的定義：「為發現『新的、有意義的』關係、樣式、趨勢的程序」，它的步驟包括資料選取(Selection)、資料前置處理(Pre-processing)、資料轉換(Transformation)、資料探勘(Data Mining) 與結果的詮釋及評估(Interpretation /Evaluation)等。完整的知識發掘過程步驟，如圖 2.1 所示(Fayyad, *et al.*, 1996)，從上述的定義，可知資料探勘是知識發掘過程的步驟之一。

由於資訊科技及資料庫的快速成長，導致必須發展新的方法及工具來處理資料以成為有用的資訊，對於現代的企業而言，資料的獲取不再是個問題，無法從大量資料中產生有用的資訊才是問題，因此資料探勘成為近年來最重要的研究領域之一。資料探勘是一種可在大量資料中尋找可用資訊的技術，它用來萃取隱藏、未知、潛在的有用資訊(Weiss, *et al.*, 1998)，透過有效地使用資料倉儲(Data Warehouse)，資料探勘可以增加企業資訊資產之價值。由於資料探勘能有效地幫助企業處理龐大的資料，並從中萃取出關鍵的資訊以作為管理者決策過程之參考，所以資料探勘的方法已被大量使用於企業的經營活動中。

值得注意的，資料探勘所要強調的不只是一些探勘的演算法而是整個資料粹取的過程，這個過程的目的在解決企業所定義的問題，利用各種技術來分析儲存在大型資料庫中的資料以找尋適當的解決方案。雖然資料探勘所用的技術很多是來自統計或人工智慧等領域，這些都並非是新的技術，但對於現代的企業而言，從資料庫中粹取知識的想法卻是新穎的。一個理想的資料探勘技術是能自動從大型資料庫中發現知識以協助決策，Moxon(1996)對資料探勘的定義是『資料探勘是一組技術，用於自動的程序中，可徹底探索及帶出大型資料庫中複雜的關係』。不過 Moxon 這個定義容易讓人誤解，以為只要使用探勘技術就可神奇地獲得答案，然而，整個資料探勘的過程離「自動」還很遙遠，還必須有人為的協助，實務的解決方案也不只是利用幾個演算法就可獲得的(Liu, 2002)。

資料探勘的本質可利用三個階段來說明，如圖 2.2 所示。

階段 1：定義企業的問題以決定目標

根據企業所設定的目標，蒐集相關資料或直接或由資料倉儲中擷取可用資料或從不同的來源獲得資料，不論如何這些資料必須轉換成資料探勘技術可接受的標準格式，在此階段資料探勘的專家必須協同領域專家來定義問題並蒐集相關的資料。

階段 2：利用各種不同的資料探勘技術來分析資料與建立模式

在階段 2 中，不斷地執行一個反覆的程序以尋找最佳解，這個反覆的程序包括以下四個子程序：(1)與輸入資料有關的子程序；(2)核心的演算法；(3)與模式有關的子程序；(4)評估模式的子程序。在階段 2 中必須進行許多的判斷，例如應該採用哪一種技術、訓練樣本需要多少及如何調整參數等。

階段 3：將所找出的方案應用在實務中並且驗證其績效

隨著時間及使用環境的改變，模式會逐漸喪失其正確性，一段時間就必須要重新建模。

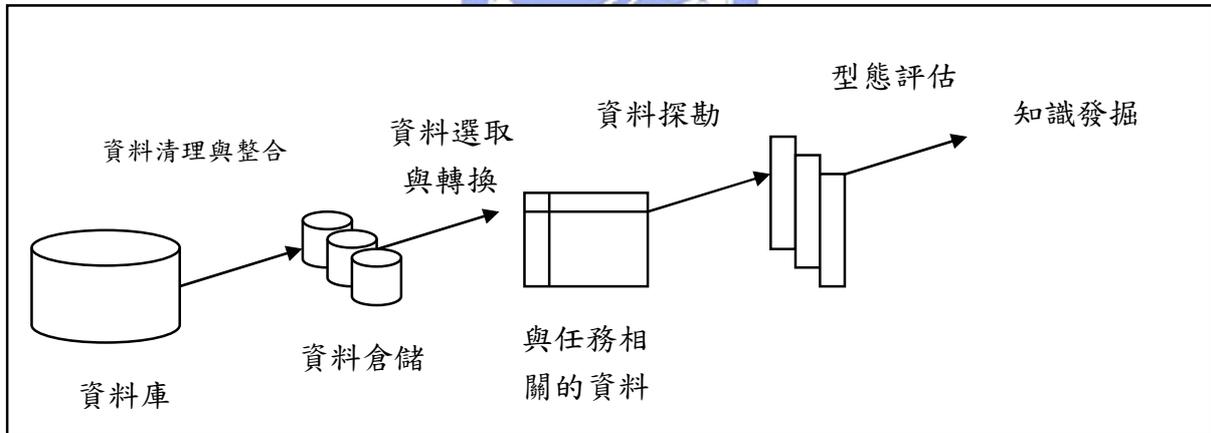


圖 2.1 知識發掘過程(Fayyad, et al., 1996)

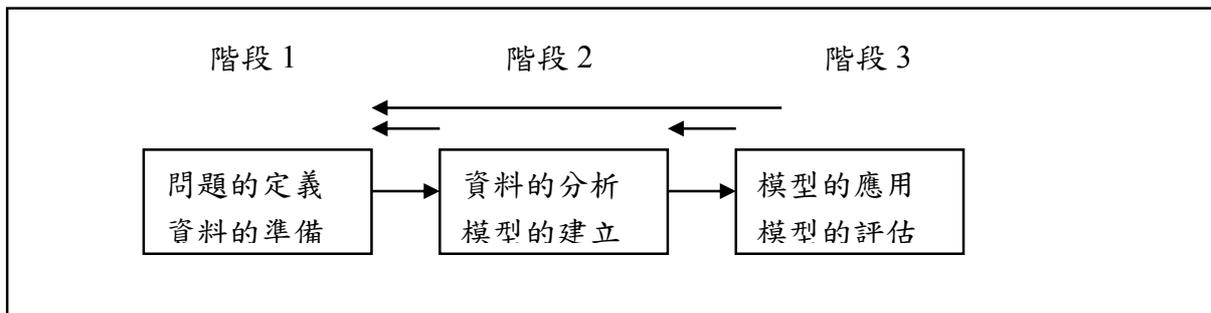


圖 2.2 資料探勘的三階段(Liu, 2002)

實務上，資料探勘可能是一個反覆的程序。例如資料過時，必須蒐集新的資料(回到階段 1)，或隨著時間、環境的改變以致資料有所變動時，模型必須重建(回到階段 2)。尤其是應用在財務問題時，這種反覆的程序是更頻繁。所以，實務上階段 1 及階段 3 可能無法自動，「自動」--資料探勘的理想目標，只能在階段 2 實現。階段 2 的目的是為所定義的問題產生一個最佳解，在此階段的技術，可分成核心演算法(Core Algorithms)及週邊技術(Peripheral Techniques)。核心演算法用來產生模式，週邊技術則用以處理輸入的資料及評估所獲得的模式。利用不同的週邊技術，包括前處理或後處理的技術以協助確認資料間的關係及評估所獲得的結果。通常，「自動」會在核心演算法中產生，而所謂「自動」的意義是核心演算法可以不需要對資料進行任何的假設，它透過一個資料驅動(Data-Driven)的方法來發現資料的型態及關係(Liu, 2002)。基本上，資料探勘與傳統資料分析不同之處，在於資料探勘不需對資料提出假設，反而是在其探勘的過程中發掘資料存在的各項條件，因此更能反映出資料隱藏的特性(Ciso, *et al.*, 1998)。資料探勘的應用涵蓋許多方面，一個三層的應用結構可以描述資料探勘的多樣性，如圖 2.3 所示，其包括應用面(Applications)、方法及問題型態(Approaches)、演算法及模式(Algorithms and Models)。

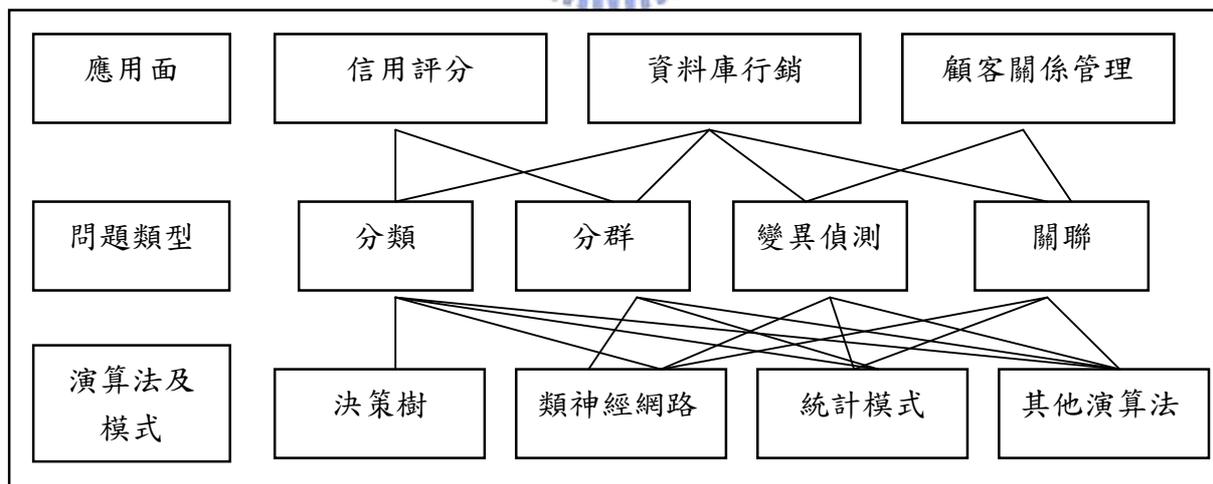


圖 2.3 資料探勘應用架構

2.2 資料探勘中分類技術之探討

隨著資料探勘的技術愈受重視，其應用範圍愈來愈廣。在資料探勘的應用模式方面，較常見的模式分別為：(1)分類模式；(2)分群模式；(3)變異偵測；(4)關聯分析；(5)

統計摘要模式；(6)順序分析。分類模式是資料探勘中最主要的問題之一，本章主要討論分類的決策及常用的分類工具，包括線性鑑別分析、羅吉斯迴歸、決策樹、最鄰近法及類神經網路等。

2.2.1 分類決策

分類的任務是根據事物之特性將事物指派到某一類別的過程，一般而言，我們會依賴以前的經驗來進行分類決策，這種經驗通常以兩種方式獲得，一為歸納法(Inductive)，另一為演繹法(Deductive)，如圖 2.4 所示。演繹法通常由知識為基礎之系統來執行，又稱為專家系統(Expert System)，以相關專家所獲得的知識為基礎之電腦系統來自動進行決策。在歸納法中，分類模式的建構則是由無數的資料所歸納出來的，模式用來決定未知個案之類別，以歸納法建構分類的模式是資料探勘技術中的一個主要任務。以上兩種方法各有其限制，以演繹法發展出的系統通常被質疑的是此系統的能力會超過專家嗎？另一個質疑是要建立並維持一個知識基礎系統，需要與專家晤談並捕捉知識及維護系統，這是相當費時且不容易的。以歸納法為基礎的分類，常因為訓練樣本的不足而使得預測的正確率受到限制，並且樣本的選擇及結果的解釋都需要專家的協助，以此方法所發現的規則有時不見得合理，通常須經該領域的專家來修改及認同(Liu, 2002)。

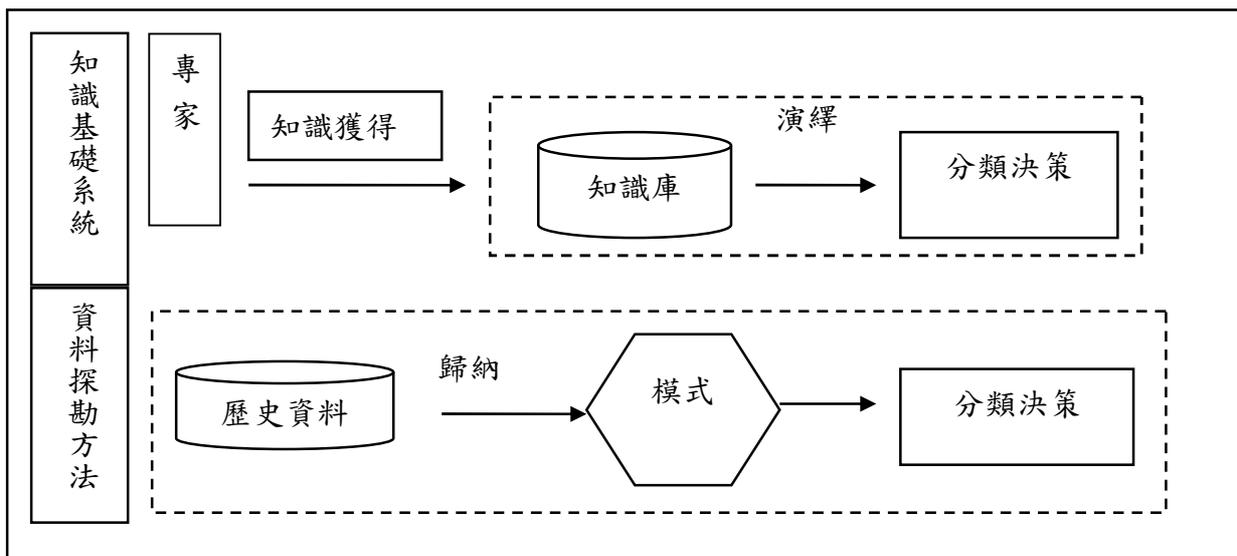


圖 2.4 分類決策的兩種方法(Liu, 2002)

本研究所討論的分類程序是以歸納法為主，分類程序大致可分成三個步驟，第一是取得訓練樣本，這組訓練樣本中包括了已知的類別以及用來區分類別的預測變數；第二步驟是建立分類模式；第三步驟則為分類。樣本大小、抽樣策略、變數特性及分類的定義都會影響分類的績效。因為樣本大小及代表性的限制，使得透過歸納法所學習出來的分類規則之績效可能不好。其次，若所選用之變數辨識能力很差時，則無法有好的績效。最後，類別定義的不確定性增加了分類學習上的複雜度，實務上，有時不見得可以對類別做出正確的定義。例如在信用評分問題中正確的風險類別可能要經過一段時間後才會知道，這段時間就稱為結果週期(Outcome Period)，我們無法了解結果週期需要多長，通常需要透過專家的協助來選擇。簡言之，以歸納法為基礎的學習可能因為上述理由而無法獲得精確的結果，事實上，如果忽略了要解決的問題的前後關係，只從資料中直接學習其結果是有限的，成功的分類不僅與分類技術有關，並且與所選的資料有關。

2.2.2 分類技術的探討



Desai(1997)將分類技術大致分成兩類，如表 2.1 所示，某些分類技術以驗證為基礎(Verification-based)，使用這些技術必須假設資料間存在某種關係並驗證此假設，這種以驗證為基礎的技術，其缺點為費時且其對於非線性模式的處理能力較為不足。例如線性鑑別分析法必須假設母體為常態分配，並在此假設下估計母體的參數。另一種方法則以探索為基礎(Discovery-based)，利用這些演算法來決定資料中複雜的關係，這些方法是運用電腦強大的計算能力來反覆尋找，直到找到最佳的解，但這種以資料驅動為主的方法，其缺點為對雜音敏感且有過度配適(Over-fitting)的情況。因為它們是根據所找出的型態自動產生模式，所以又稱為無母數方法或資料驅動方法。資料驅動方法不像傳統的理论驅動方法需費時以檢查資料的假設，也就是使用者不需要有太多統計知識就可解決問題。

透過一個資料驅動的方法來發現資料的型態及關係，對推論複雜的非線性關係而言，資料驅動是一個較好的方法，但是它所產生的模式通常是相對較大、較特殊也較困難解釋。對於資料驅動方法，模式只是建立在所抽取的樣本上，因此模式可能會因為資

料不具代表性或是有雜音而容易有過度配適的情形，使得對未來的推論不佳。所以，重要的任務是找出模式複雜程度的適當性，不要過度的配適訓練樣本。例如：修剪決策樹以獲得一簡單的結構，或是在資料有雜音時調整參數。因此對於資料驅動的方法有必要調整模式的複雜度來配適問題，否則模式過度配適訓練樣本時，將會造成偏差。表 2.2 顯示資料驅動與模式驅動方法的比較。

表 2.1 各種技術的分類一覽表(Desai, 1997)

類別	技術	演算法
以驗證為基礎之技術 (Verification-based, Parametric, Theory-driven)	迴歸(Regression)	線性迴歸(Linear Regression)等
	鑑別分析 (Discriminant Analysis)	線性鑑別分析 (Linear Discriminant Analysis)等
以探索為基礎的技術 (Discovery-based, Non-parametric, Data-driven)	決策樹(Decision Tree)	C5.0 等
	類神經網路 (Neural Network)	多層感知機 (Multi-layer Perceptron)等
	惰性學習(Lazy Learning)	最鄰近法(KNN) 局部權重迴歸 (Locally Weighted Regression)等
	演化論計算法 (Evolutionary Computing)	基因演算法(Genetic Algorithm)等

表 2.2 資料驅動與模式驅動方法之比較(Liu, 2002)

	資料驅動	模式驅動
變數的關係	複雜的及非線性關係	直線關係
模式產生方式	自資料中自動產生	假設檢定
主要任務	調整模式的複雜性	參數估計
缺點	1)對雜音資料敏感 2)過度配適	1)必須檢查假設條件 2)不容易處理非線性關係

以下將針對幾個常見的分類技術進行說明。

(1) 鑑別分析(Discriminant Analysis, DA)

鑑別分析由Fisher(1936)提出，基本概念是根據預測變數 X_1, \dots, X_k 的線性組合做為觀察值分組的基礎，這個線性組合即為鑑別函數(Discriminant Function)。當處理兩個以上群體的分類問題時，若各群體的預測變數是多元常態分配，則可以導出待分類樣本與該群

體的中心點距離平方的鑑別函數，距離越小則屬於這個群體的可能性越高，如果待分類的樣本與某一群體的鑑別函數值最小，則將待分樣本歸到該群體。進行鑑別分析時必須符合某些假設：(1)多變量的常態性；(2)各母體的共變數矩陣均相等；(3)預測變數之間沒有複共線性(Johnson & Wichern, 1998)。若資料不符合常態性的假設，將影響鑑別函數估計的正確性，也會影響統計顯著性的檢定；若各群體的共變數矩陣不相等，對分類的過程也會有不利的影響；變數間複共線性的問題也會影響分類的結果，尤其在利用逐步估計法來估計模式時，其影響更為嚴重。鑑別分析可以透過鑑別函數或是馬氏距離來完成，鑑別函數的重點在於使組間變異對組內變異之比值最大，而馬氏距離法則是在預測變數所形成之多維空間中計算距離，利用距離的大小當成分類的依據。估計鑑別函數的方法有兩種，一種是同步估計法(Simultaneous Estimation)和逐步估計法(Stepwise Estimation)。所謂的同步估計法就是在計算鑑別函數時同時考慮所有的變數，不論每一個預測變數的鑑別力如何，都納入考量。而逐步估計法則根據每一個預測變數鑑別力的大小，每次納一個變數到鑑別函數中，逐步估計法可剔除那些不具預測力的變數，但逐步迴歸容易受到樣本數與變數個數比例之影響。鑑別分析可以將分類成本納入模型中，當成本納入模型時，其分類的準則就是使得錯誤分類的期望成本(Expected Cost of Misclassification)最小(Sharma, 1996)。如果各母體的共變數矩陣均相等的假設不成立時，可用二次鑑別函數(Quadratic Discriminant Function)，但二次的鑑別函數的表現相對線性鑑別函數而言較不穩健(Martell, *et al.*, 1981)。

(2) 羅吉斯迴歸(Logistic Regression, LR)

鑑別分析必須滿足預測變數為多元常態且兩組樣本之共變數矩陣相等的假設下，才有最佳的預測結果。羅吉斯迴歸不限制變數的分配，且當預測變數同時存在離散型、連續型及虛擬變數(Dummy Variable)時也可採用，即使滿足鑑別分析所要求之假設時，羅吉斯迴歸模式仍有良好的表現(Hosmer, 1989)。此外，羅吉斯迴歸模式除了提供預測分類之組別外，還可計算出事件發生的機率。羅吉斯迴歸模式的基本假設是(1)資料必須來自隨機樣本；(2)因變數為多個預測變數之函數；(3)羅吉斯迴歸對複共線性

(Multicollinearity)也敏感，所以必須假設預測變數之間沒有複共線性(王濟川等, 2003)。羅吉斯迴歸模式是為修正線性機率模式之缺點所發展出來的，以線性模式預測事件發生之機率會有殘差項變異數異質性的問題，且因變數的估計值可能會落在(0,1)之外，而與條件機率期望值之條件衝突。

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = w_0 + w_1x_1 + \dots + w_kx_k \quad (2.1)$$

公式(2.1)為羅吉斯迴歸方程式，其中 logit 是將勝算比(Odds)取對數，這個函數稱為連結

函數(Link Function)，將(2.1)轉換成 $p = \frac{\exp(w_0 + \sum_{i=1}^k w_i x_i)}{1 + \exp(w_0 + \sum_{i=1}^k w_i x_i)}$ ，代表具有 (x_1, x_2, \dots, x_k) 的觀察

值會屬於某一組別之機率，預測變數與機率 p 的關係是非線性的，所以羅吉斯迴歸在解釋時要特別注意。公式(2.1)中的參數則是利用最大概似法(Maximum-Likelihood Method)所獲得，最大概似估計法通常是比較複雜的，因為它利用反覆(Iterative)的程序來獲得最佳的參數，比線性迴歸需要更多的計算時間，但電腦的強大計算能力也使得這一點不再是個問題(Thomas, 2000)。



(3) 決策樹(Decision Tree)

決策樹可以說是歸納式學習法中最簡單的一支，它是一種樹結構(Tree Structure)可以產生易於了解的規則，主要功能是藉由分類已知的觀察值來建立一樹狀結構，並從中歸納出觀察值間的某些規則。決策樹以樹狀的方式由上而下表現出規則，將特定的物件集合，隨著樹的成長，逐步地分割成更小的子集合，且方便將規則轉換為 SQL 語法進行運用，使用決策樹進行分析資料時，主要的優點是在於能透過樹狀圖之方式，將資料的結果完整呈現出來，因此較具視覺性與容易解釋的特性。

一般而言，使用決策樹來進行資料探勘的工作，通常過程為：

1. 選擇使用的決策樹演算法：不同的決策樹演算法有不同的特性。
2. 決策樹的建構：運用訓練資料加以建構決策樹。
3. 決策樹的修剪：加以評估修剪決策樹。

目前最被廣泛使用的決策樹演算法包括分類迴歸樹(Classification and Regression Trees, CART)、卡方自動互動檢視法(Chi-square Automatic Interaction Detector, CHAID)、C4.5 等三種。

CART 演算法是由 Breiman(1984)等人發展出一種演算法，透過二元分割的過程來分析龐大的資料集，並透過遞迴(Recursive)的程序，CART 會依照預測變數與其相對的各項指標將既有的訓練樣本劃分成數個已知的類別，並產生一組分類的規則，再透過此規則對新樣本進行分類。CART 的最大優點之一，就是演算法會自動檢驗模式，找出最佳的一般模式。針對任何節點，CART 會先檢查節點中的資料是否屬於同一類別，若節點中的資料只屬於同一類別，則此節點不需再分割。若節點中仍有二個以上的類別，則 CART 會測試所有資料屬性。通常使用 CART 進行資料分析時，可同時處理連續型與間斷型資料，若為間斷型資料會檢查所有合併方式後，選擇最佳方法將屬性值合併為兩組；連續型資料則會選擇一適當之分割點將資料一分為二，然後計算每個屬性分割條件下的吉尼指數(Gini Index)，能降低吉尼指數最多的屬性就成為分割條件。CART 的運算過程中，不論是否有任何限制的假設，都會產生包括許多節點的最大樹狀結構，然後再依照所挑選的準則進行決策樹的修剪(Pruning)，直到滿足修剪的準則為止。這樣的流程雖然花費較多時間但卻能避免資料被隱藏。

CHAID演算法是由Kass(1980)發展出來，主要是利用卡方檢定(Chi-square Test)來檢驗分割變數是否對於資料的分割有顯著差異，以此將資料進行分割。CHAID演算法只適用於類別型的資料，連續變數必須被區隔成幾個區域範圍，CHAID 在建立決策樹方面和CART 很類似，不過選擇分割的做法不同。CART 使用吉尼係數來決定最佳的分割，而CHAID則用卡方檢定測試來決定。

C4.5 決策樹演算法，是澳洲學者 Quinlan 發展出來，它的基本理論是從 ID3 學習系統改良而來的，C4.5 具有處理遺失值的能力，C4.5 在每一個節點產生不同數目的分支，這是與 CART 的一個差異，CART 在每個節點都呈現二分法。C4.5 對於連續性變數的處理方式和 CART 類似，但對類別變數的處理方式就相當不同，當 C4.5 評估一個類別變數的值作為分割變數，其方式是變項的每一個值都有一個分支。C4.5 和 CART 另一個

不同的地方是修剪決策樹的方式，CART 使用測試樣本來修剪用訓練樣本所得出的分支樹，而 C4.5 則直接使用訓練樣本來修剪決策樹。

(4) KNN 法(K Nearest Neighbor)

K個最鄰近法的分類，假設幾何空間中有許多已知屬於那些群體的參考樣本；分類時，先選取與待分類樣本點 X 在幾何空間上最近的 K 個點，這 K 個點中那個群體的樣本數最多，即判定 X 屬於這個群體。對有 p 個預測變數的任何樣本 $X(x_1, x_2, \dots, x_p)$ ，可以將樣本視作 p 維空間中的一個點。若有已知屬於二個以上群體的 N 個樣本作為模式建立樣本，則分類的步驟：

1. 在 N 個已知屬於那些群體的樣本中，找出 K 個樣本點與待分類點 X 的最小馬氏距離。
2. 統計 K 個點中，屬於群體 i 的樣本數。
3. 計算樣本 X 屬於每個群體的機率。
4. 將待分類樣本 X 歸到機率最高的群體。



(5) 類神經網路

類神經網路是處理分類問題主要的技術之一，由於類神經網路具有嚴謹的數學推論、容錯能力、大量平行處理的能力、高聯想力以及能過濾雜訊等特性外，並可應用於建構非線性之模式，能彌補傳統統計模式建構時須滿足某些假設條件之缺點(Rumelhart, *et al.*, 1986)。截至目前為止，類神經網路已發展出許多不同的模式，依網路模式架構大致可分成三種，分別是聯想式、監督式及非監督式學習網路，其中監督式神經網路的建模資料包括輸入及輸出變數，其目的在尋找輸入與輸出變數之間的對應關係，較為常見的模式為倒傳遞類神經網路(Backpropagation Neural Network, BPN)及 Radial Basis Function (RBF)。在眾多監督式學習網路的模式中，BPN 是應用最為廣泛的(Fish, *et al.*, 1995)，但由於其建模的時間常因問題的複雜度的增加而所有延長，因此也造成網路模式發展的主要限制，除此之外，其網路的訓練階段過程有如黑箱作業(Black Box)以致不易解釋，這也是此技術的缺點。類神經網路雖可自動進行型態辨識、學習、分類及推論，

且不需要對資料進行任何的假設，但常會因為要修正模式內的大量神經元連接鍵係數而比其他技術花費更多的學習時間(Craven, *et al.*, 1997)。



第三章 Mahalanobis-Taguchi 系統

馬氏-田口(Mahalanobis-Taguchi, MT)系統主要是將田口方法應用在多變量資料的診斷及預測上。MT 系統是利用馬氏距離來建構多變量量表以衡量事物異常的程度，並利用直交表與訊號雜音比來最佳化量表。

本章於 3.1 節中說明田口多變量系統及討論 MT 系統在田口式品質工程的定位，3.2 節與 3.3 節中分別解釋 MT 系統的概念及其執行之步驟。在 3.4 節中將 MTS 與其他多變量技術進行比較，在 3.5 節中討論其他與 MT 系統相關之議題，最後 3.6 節則為結論。

3.1 田口的多變量系統

建立多變量系統的目的通常是為協助決策的進行，因此，了解系統的本質是很重要的。首先針對田口博士所提的多變量系統加以說明，圖 3.1 代表一個典型的多變量系統圖，圖中 X_1, X_2, \dots, X_k 代表 k 個控制變數，這些變數是提供判斷及決策時的重要資訊。這些變數及所蒐集到的正常觀察值組成一個資料矩陣(Data Matrix)，這個資料矩陣是進行多變量分析的基礎，在此稱這個資料矩陣為參考空間(Reference Space)。雜音因子 N 則代表會影響系統績效的因子，通常是指不可控或是控制的成本很高的因子。一個好的決策支援系統應該是無論雜音的狀況如何，都可以做出穩健且正確的決策，研究者有必要了解在執行判斷的過程中可能存在的雜音。訊號因子 M 是產品或系統使用者可自行調整的因子，它可能已知也可能未知，若已知則代表系統的真值，輸出 Y 則應該可以反映輸入 M 的情況，也就是 Y 必須與 M 成比例，未來則根據 Y 值進行預測(Taguchi, 2002)。

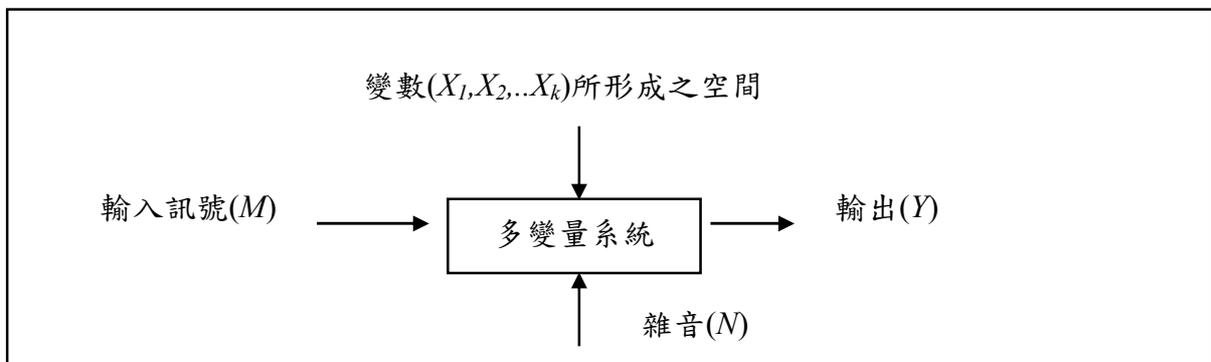


圖 3.1 多變量系統圖(Taguchi, 2002)

許多的決策機制都可視為多變量系統，不論是產品檢查、疾病診斷、聲音或臉部辨識、信用評分、或是甄試入學等均可視為多變量系統。對於多變量系統而言，必須有一個量表以衡量事物間異常的程度。例如：疾病診斷系統中用來診斷罹患某種疾病的嚴重程度；在產品檢查系統中用來了解產品異常的情形；在信用評分系統中用來衡量客戶信用異常的狀況，利用多變量系統中的特徵變數來建立量表可以協助決策者進行判斷。

MT 系統是田口式品質工程在多變量系統的應用，所以在介紹 MT 系統前，先審視田口式品質工程之整體觀，如圖 3.2 所示。圖 3.2 上半段代表在製程及產品上的應用，這分成線外品質工程(Off-line QC)及線上品質工程(On-line QC)，所謂線外品質工程是指研究開發與設計等上游的階段，而線上品質工程是從製造到市場的中下游階段。圖 3.2 下半段則代表田口方法在多變量的應用，左邊代表 MT 系統，右邊則代表多變量系統製作的評價問題。

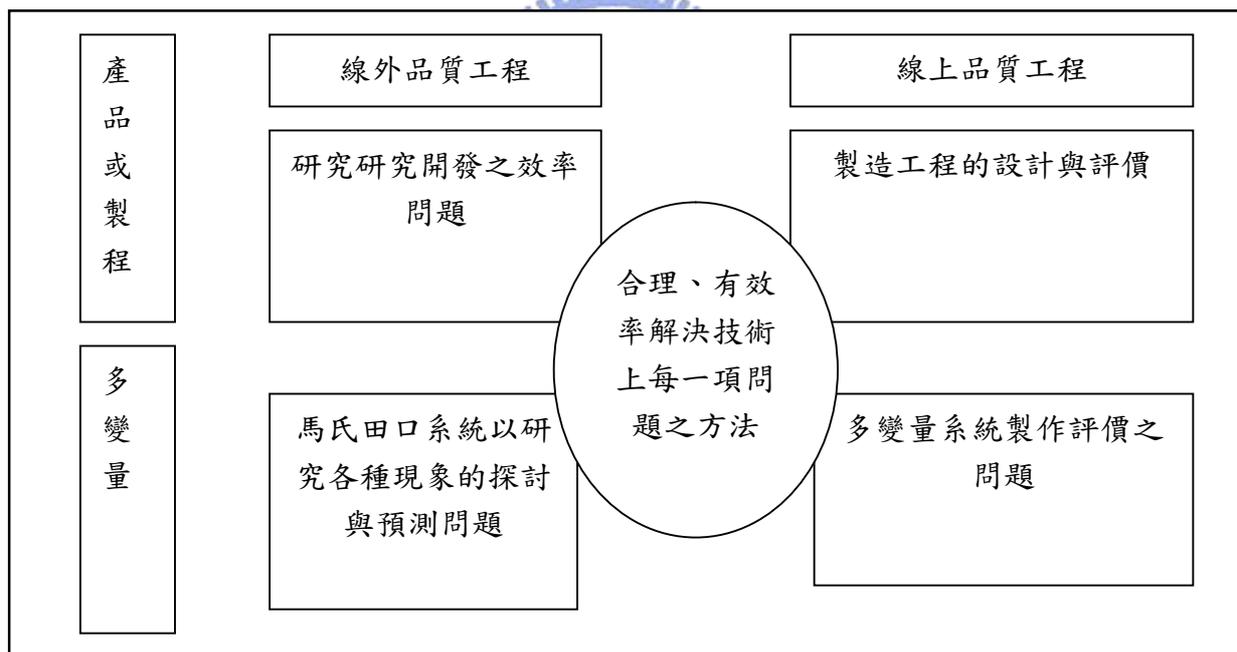


圖 3.2 田口式品質工程的整體觀

穩健設計是田口博士的一種思考哲學，田口方法以工程的角度去事先瞭解品質問題，利用社會損失成本作為衡量產品品質的依據，穩健設計提供一個具成本效益的方法以提升市佔率。在田口的概念中，品質分成兩種：(1)顧客導向的品質；(2)工程品質。當顧客導向的品質提升時，市佔率也會逐漸提升，顧客導向的品質定位在產品規劃階

段，對於創造新的市場而言是相當重要的，而工程品質則包括不良、失效、雜音、震動、污染等，工程品質的好壞可利用以與理想績效之偏差來衡量，顧客導向的品質定義了市場的區隔，而工程的品質則協助贏得市佔率，穩健設計則以改善工程品質為目標。在多變量的應用中，馬氏距離與工程品質很類似，因為它可測量某事物偏離已知的參考空間或正常群體的程度。

為更清楚的瞭解 MT 系統的想法，茲將田口方法的基本原理扼要說明如下：

(1)利用能量轉換(Energy Transformation)以衡量機能

田口方法中最重要部分就是要找出理想的機能，能衡量出系統的機能以才能夠改善產品品質。

(2)利用控制因子與雜音因子間的交互作用

在田口方法中，對控制因子之間的交互作用不感興趣，而是對於控制因子與雜音因子間的交互作用較感興趣，因為這樣可以設計出較為穩健的系統。

(3)使用直交表與訊號雜音比

直交表用來最小化實驗次數，儘管許多人對於直交表在田口方法的應用有些意見，但直交表的應用只是田口方法的一部份。在品質工程中，訊號雜音比是用來進行評價的工具。它以有用的情報(訊號)與有害的情報(雜音)的比值來評價機能，訊號雜音比在對雜音因子進行某些調整後可以捕捉訊號影響的大小，它是一種以判斷錯誤的程度做為量測的誤差大小來進行評價的方法。

(4)兩階段最佳化

在建構實驗後，透過兩階段的協助可以找出因子水準的組合。第一階段最小化變異，第二階段調整平均值至目標值，因為在最小化變異之後較容易調整平均值。

(5)利用品質損失函數進行允差設計及線上品質工程

前面四個原理與參數設計有關，產品及製程的參數設計與其說是技術改善，倒不如說是透過確認實驗的增益之再現性，來檢視實驗者所選擇的基本機能是否適切的方法。第五個原理與允差設計及線上品質工程有關，利用參數設計決定最佳設定後，利用品質損失函數進行允差設計。如果績效偏離目標值，則會造成社會損失，此損失與離差的平

方成比例，利用品質損失函數的方法來設定安全因子，線上品質工程則用來監控系統績效及偵測製程的變化。

3.2 MT 系統之介紹

在工程問題中，產品或製程的設計者擁有選擇設計概念的權利。然而，在選擇設計概念後，必須在各種不同的使用狀況下對此設計之績效加以評估。品質工程的原理可以協助評估的程序，如果此設計的績效不受雜音的影響，則稱為穩健的設計。相同的概念，應用在 MT 系統中，建立 MT 量表所用的變數完全由決策者決定，一旦量表建構完成後，就必須在各種狀況下加以評估。MT 系統是田口博士結合馬氏距離及田口方法所發展的新方法，用來診斷多變量系統。田口認為若只根據個別的變數來判定是無法做出正確的決策，因為系統中的變數是彼此相關的，若忽略變數間的相關，則可能做出錯誤的判斷。MT 系統是一種藉由建構多變量量表來協助進行決策的方法，主要目的是藉由量表來量測事物異常的程度。為達診斷或預測的目的，田口認為發展出適合的量表是非常重要的，而建構量表所需之工具就是所謂的「馬氏距離」。根據資料特性的不同，田口將馬氏距離加以修正，並建構出三種不同的方法，分別為 MTS、MTGS 及 MTA 等，在此統稱為 MT 系統。本節將針對馬氏距離及 MT 系統的內容加以說明。

3.2.1 馬氏距離

符號說明：

MD: 馬氏距離

X_i : $(x_{i1}, x_{i2}, \dots, x_{ik})$ 代表第 i 個觀察值之 k 個變數所形成之向量

X_i^T : 代表轉置矩陣

C_w : 聯合組內共變數矩陣

C_w^{-1} : 聯合組內共變數矩陣之反矩陣

馬氏距離是 1936 年由 Mahalanobis 博士所提出，主要是用來衡量多維空間中兩個樣

本點之間的距離，用以辨識不同的群體。馬氏距離是將傳統的歐氏距離用 C_w 來加以調整，公式如(3.1)所示。由於歐氏距離有兩個缺點：(1)歐氏距離無法提供有關未知點相對於參考群體有多近的統計測量，(2)它只測量離重心的相對距離並不考慮群體的分配狀況，相對於馬氏距離來說，歐氏距離之用處較少。馬氏距離比較適合處理多變量的問題，主要是有考慮變數間相關的問題，且對變數間關係的改變相當敏感。在多變量的分析中，有時會出現複共線性的問題，也可利用馬氏距離來處理。

$$MD_{ij} = (X_i - X_j)^T C_w^{-1} (X_i - X_j) = [x_{i1} - x_{j1}, x_{i2} - x_{j2}, \dots, x_{ik} - x_{jk}] C_w^{-1} \begin{bmatrix} x_{i1} - x_{j1} \\ x_{i2} - x_{j2} \\ \dots \\ x_{ik} - x_{jk} \end{bmatrix} \quad (3.1)$$

在傳統的分類方法中，馬氏距離被用來找出某個未知組別的觀察值與幾個群體平均位置的距離，這個未知的觀察值將被分到離它最近的哪一群。

3.2.2 MT 系統



MT 系統是一種資料分析的方法，藉由建構多變量量表以協助決策的進行，MT 整合兩個方法，一是藉由馬氏距離建立量表，以進行型態的辨識；另一是利用訊號雜音比及直交表進行最佳化及評價量表之優劣。MT 系統分成 3 種方法，分別為 MTS 及 MTGS 及 MTA。首先進行相關符號的說明，其次說明三種方法的內容。

符號說明：

SMD: 調整後之馬氏距離

k : 代表變數的個數

Z_{ij}^T : 轉置矩陣

$Z_{ij} = (X_{ij} - \bar{X}_i) / S_i$ 代表標準化矩陣

$\bar{X}_i = \sum_{j=1}^n X_{ij} / n$ 代表第 i 個變數之平均數

$$S_i = \sqrt{\sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 / (n-1)}$$
 代表第 i 個變數之標準差

C : 相關矩陣

C_w : 聯合組合共變數矩陣

C^{-1} : 相關矩陣之反矩陣

C_w^{-1} : 聯合組合共變數矩陣之反矩陣

C_{adj} : 相關矩陣之伴隨矩陣

$$U_i = Z_i - \sum_{j=1}^{i-1} \frac{U_j^T Z_i}{U_j^T U_j} U_j$$
 代表經由 Gram-Schmidt 直交化後之變數

$i = 1, 2, \dots, k$ 代表變數之個數

$j = 1, 2, \dots, n$ 代表觀察值之個數



在此，首先說明 MTS，MTS 主要是利用公式(3.2)來計算觀察值之馬氏距離。公式(3.2)與原來由 Mahalanobis 所提出的馬氏距離有些不同，主要的差別是公式(3.2)將原來的馬氏距離除以 k 值，並將 C_w 用 C 取代，主要理由是聯合組內共變數矩陣容易受到量測單位的影響，而相關矩陣不易受到量測單位的影響。

$$SMD_j = (1/k) Z_{ij}^T C^{-1} Z_{ij} \quad (3.2)$$

在此為區別這兩個公式，將公式(3.2)的馬氏距離稱之為調整後之馬氏距離(Scaled MD)，以 SMD 表示。此 SMD 用來建構量表並定義一個參考點，以衡量並比較觀察值異常的程度。當量表建構完成後，必須驗證其正確性，並利用直交表及訊號雜音比篩選出重要的變數，最後使用重要的變數所形成之量表進行預測或診斷，此方法詳細之執行步驟於 3.3 節介紹。

有關調整後之馬氏距離的計算方法有三種：(1)相關矩陣；(2) GSP(Gram-Schmidt Orthogonalization Process)；(3)伴隨矩陣。表 3.1 顯示這三種不同的方法。相關矩陣法如前所述，而 GSP 主要的目的是將已知線性獨立向量 Z_1, Z_2, \dots, Z_k ，透過 GSP 程序，以獲得

相互垂直的向量 U_1, U_2, \dots, U_k ，其中 $U_i = Z_i - \sum_{j=1}^{i-1} \frac{U_j^T Z_i}{U_j^T U_j} U_j$ 。由 GSP 計算所出來的調整後之

馬氏距離 SMD，如公式(3.3)所示，圖 3.3 為 GSP 的程序，而利用 GSP 程序所得的調整後之馬氏距離 SMD 來建構量表的方法稱為 MTGS。

$$SMD_j = \frac{1}{k} \left(\frac{u_{1j}^2}{s_1^2} + \frac{u_{2j}^2}{s_2^2} + \dots + \frac{u_{kj}^2}{s_k^2} \right) \quad (3.3)$$

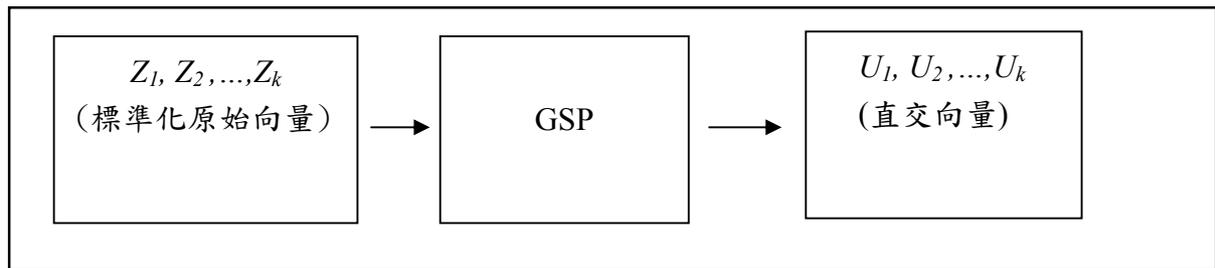


圖 3.3 GSP 程序

發展 MTGS 的原因是為處理異常「方向性」(Direction of Abnormals)的問題，所謂「方向性」所指的是會有「好」及「壞」的異常。例如，入學甄試時會有成績極好或極差的學生，若以 MTS 來處理只能分辨出正常與否，無法分辨異常的方向性，也就是無法分辨極好或極壞的觀察值，因此產生了 MTGS 法。MTS 除了無法解決「方向性」的問題外，另一個無法處理的是「共線性」(Multicollinearity)問題。所謂的「共線性」是由變數間高度相關所引起的，在傳統多變量技術中處理共線性的方法之一，就是去除其中一個變數而保留另一個變數。但田口建議在未進行訊號雜音比分析前不要將變數刪除，因為多變量系統中的每一個變數及其相關矩陣都很重要。當兩個變數間存在高度共線性時，若使用 MTS 來計算 SMD，其相關矩陣的行列式值會很趨近於 0，這使得矩陣幾乎是奇異(Singular)，此時可能無法計算反矩陣或是所得之反矩陣並不正確。故田口提出了 MTGS 及 MTA 兩種方法來解決共線性的問題(Taguchi, 2002)，將原先的公式(3.2)中的 C^{-1} 換成 C_{adj} ，這種利用伴隨矩陣所計算的方法稱為 MTA，公式如(3.4)所示。

$$SMD_j = \frac{1}{k} Z_{ij}^T C_{adj} Z_{ij} \quad (3.4)$$

以伴隨矩陣所形成的馬氏距離有以下三個優點：(1)資料矩陣儘管有共線性存在，仍

可求出距離。(2)比例係數可廣泛使用。(3)容易製作多重馬氏距離。

表 3.1 MT 系統方法一覽表

方法	SMD 之計算
MTS	利用相關矩陣之反矩陣
MTGS	透過 Gram-Schmidt Orthogonalization Process 計算
MTA	利用伴隨矩陣

3.3 MT 系統之執行步驟

本節將 MT 系統執行之主要步驟，以圖 3.4 表示之。首先，第一階段必須定義馬氏空間(Mahalanobis Space, MS)，空間中包括了正常的觀察值及辨識變數，根據資料的屬性或問題的性質，決定利用 MTS、MTGS 或 MTA，當完成 SMD 時，量表也隨之完成。第二階段就是驗證量表，第三階段則利用直交表與訊號雜音比來篩選重要的變數，最後利用重要的變數所形成的新量表進行未來的預測。以下分別對 MTS、MTGS 及 MTA 進行詳細的說明。

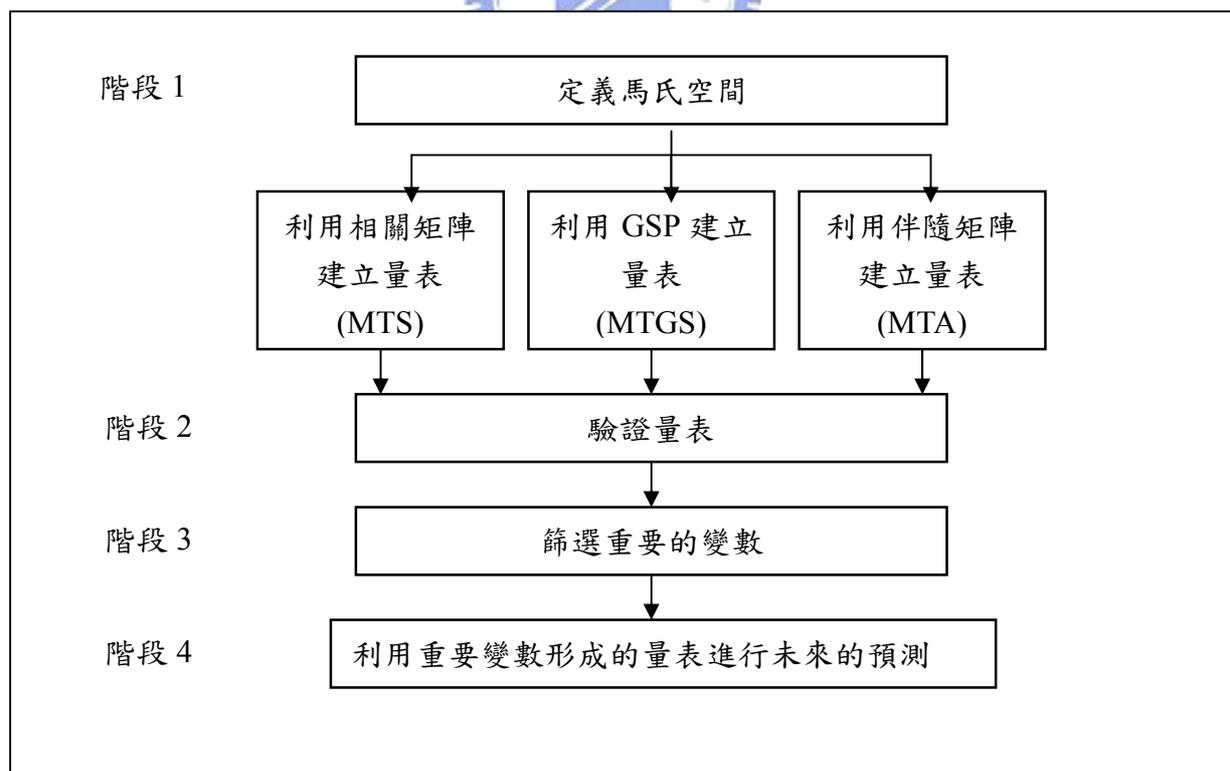


圖 3.4 MT 系統之執行步驟

3.3.1 MTS 之執行步驟

MTS 是 MT 系統中最早被發展出來的方法，我們將其執行步驟分成四個階段來討論。

階段 1：定義馬氏空間，以建構量表

在階段 1 中，首先要先定義何謂正常群體及其相關之變數，根據這些正常的群體，建構出一個基礎空間，我們稱之為「馬氏空間」。馬氏空間的選擇是使用 MTS 方法中最困難的一件事，它是決定 MTS 是否有用的主要原因。馬氏空間的主要作用是建立量表的基準點及單位，它首先將這些變數進行標準化，之後計算其平均數、標準差及相關矩陣，根據公式(3.2)來計算所有觀察值的 SMD。正常群體中的觀察值要儘可能同質性，在馬氏空間中 SMD 的平均值會收斂在 1，故將馬氏空間稱為單位空間。

階段 2：量表驗證

在階段 2 中，找出「異常」的觀察值。計算這些異常的觀察值的 SMD，以確保量表的正確性。這些異常觀察值在進行標準化時，所需要的平均數、標準差及相關係數矩陣都來自於馬氏空間中。如果所發展出來的是一個好的量表，則異常觀察值所對應的 SMD 會較大。

階段 3：篩選重要的變數

在進行決策及判斷的過程中，有時候不見得需要所有的變數，對資料的蒐集及決策者而言，減少變數是很重要的，因為可節省處理資料時間與成本。MTS 利用訊號雜音比及直交表篩選出重要的變數，利用直交表配置出各種變數的組合，並計算此組合下各組「異常」觀察值的訊號雜音比，利用訊號雜音比的「增益」(Gain)來找出重要變數。一個直交表中，每一組實驗都包括了各種變數的水準組合，以研究每個變數對反應值(Response)的影響。在實驗中，每一個變數都被指派到直交表的「行」中，而直交表中的每一列則代表了每一組實驗組合。在 MTS 中有 k 個變數，每個變數有兩個水準(Level)，這兩個水準分別代表「包含」與「不包含」該變數， k 個變數被指派到直交表中。對於每一組實驗(Run)，我們將利用所指派的變數來計算每一個觀察值之 SMD，再

利用這些 SMD 求出訊號雜音比。訊號雜音比被定義為一種測量量表正確性的工具，訊號雜音比的選擇與觀察值異常程度的先驗知識有關。較高的訊號雜音比意味著較低的錯誤，在田口方法中，有許多訊號雜音比，詳細的說明將於 3.5.1 節討論。

階段 4：利用重要的變數所形成的量表進行未來的預測

利用這些重要的變數進行預測或診斷，並根據預測結果採取一些適當的行動。所採取的行動取決於閾值的大小，有關閾值的討論將於 3.5.3 節討論。

3.3.2 MTGS 之執行步驟

MTGS 主要用來處理異常觀察值具方向性及共線性的問題。在多變量的空間中進行直交化，田口博士建議可根據工程、成本、重要度的順序將標準化變數 Z_1, Z_2, \dots, Z_k 排序，而順序的決定則是最重要的步驟，接下來利用直交化程序可獲得新的直交變數 U_1, U_2, \dots, U_k ，如公式(3.3)所示，但經由直交化的 U_1, U_2, \dots, U_k 的變異數不是 1，所以再經過標準化獲得 Y_1, Y_2, \dots, Y_k ，則 Y_1, Y_2, \dots, Y_k 為直交且變異數為 1，再根據 Y_1, Y_2, \dots, Y_k 計算距離。這個方法在樣本小且具共線性的情況時，能可獲得精度較佳的 SMD。MTGS 之執行步驟如下所述：

階段 1：利用馬氏空間，以建構量表

在階段 1 中，定義正常的狀況，接下來根據正常的定義蒐集相關的變數，利用公式(3.3)來計算所有觀察值的 SMD，利用這些 SMD 可以定義原點及單位距離，並將此原點及單位距離當成量尺的基準點。

階段 2：量表驗證

在階段 2 中，確認「異常」的情況。計算這些異常觀察值的 SMD，以驗證量表的能力。如果所發展出來的是一個好的量表，則異常的觀察值所對應的 SMD 會較大。

階段 3：確認重要的變數

利用訊號雜音比及直交表來找出重要的變數。在 MTGS 中，如果淨相關(Partial Correlations)的影響不顯著，且如果能對這些變數定義特別的順序，則可以直接由直交(Orthogonal)向量來計算所有變數之訊號雜音比。然而，如果淨相關的影響顯著，就需要

利用直交表找出有用的變數。通常會建議使用直交表，因為這樣就可以不需要注意變數的順序及淨相關的問題(Taguchi, 2002)。

階段 4：利用重要的變數所形成的量表進行未來的預測

利用這些重要的變數進行預測或診斷，並根據預測結果採取一些適當的行動，所採取的行動取決於閾值的大小。

3.3.3 MTA 之執行步驟

MTA 與 MTS 很類似，當(1)行列式不為 0 時， $MTS=MTA$ ；(2)行列式為 0 時，MTA 優於 MTS(Taguchi, 2002)，MTA 主要解決共線性的問題，利用餘因子矩陣來計算距離。以下為 MTA 之步驟：

階段 1：定義馬氏空間，以建構量尺

在階段 1 中，須定義正常的群體，接下來蒐集相關的變數，並將這些變數進行標準化。利用伴隨矩陣，根據公式(3.4)來計算所有觀察值的 SMD。

階段 2：量表驗證

在階段 2 中，找出「異常」的觀察值。計算這些異常的觀察值的 SMD，以確保量表的正確性。

階段 3：確認重要的變數

利用訊號雜音比及直交表篩選出重要的變數。利用直交表配置出各種變數的組合，並計算此組合下各組異常觀察值的訊號雜音比，利用訊號雜音比的「增益」來找出重要變數。

階段 4：利用重要的變數所形成的量表進行未來的預測

利用這些有用的變數進行預測或診斷，並根據預測結果採取一些適當的行動，而所採取的行動取決於閾值的大小。

表 3.2 整理出 MTS、MTGS 及 MTA 的比較，包括 SMD 之計算方式、異常之方向性及共線性問題的處理等。

表 3.2 MT 系統三種方法的比較表

方法	MTS	MTGS	MTA
計算馬氏距離的方法	相關矩陣	GSP 直交化	伴隨矩陣(餘因素矩陣)
可處理異常的方向性問題	×	○	×
可處理共線性問題	×	○	○
可處理相關程度很低問題	○	×	○
是否為單位空間	○	○	×

3.4 MT 系統與其他多變量方法的比較

Taguchi 及 Rajesh(2000)指出 MTS 與傳統的多變量方法有許多的不同，MT 系統最主要的目的是為發展一量表以執行有效之診斷。本節主要說明幾種常用的多變量技術與 MT 系統在分類及變數篩選上之異同。

(1)主成份分析法(Principal Component Analysis, PCA)

主成份分析是透過原始變數之線性組合以獲得新變數的方法來解釋資料變異-共變異 (Variance-Covariance) 之結構。主成份的目的為：(1)資料縮減；(2)資料解釋，雖然系統總變異產生來自 p 個變數，但大多數的變異可由 k 個主成份來解釋。主成份是由 p 個變數線性組合而成，這些線性組合代表利用最大變異透過對原系統之轉軸的方式來獲得新的座標軸以提供對共變異結構的簡單解釋。因為由原始變數所組合而成的新變數彼此之間是無關的，這有些類似直交 Gram-Schmidt 向量。雖然 MTS/MTGS 與主成份的目的不同，但也可以利用主成份進行 SMD 之計算。

而主成份分析與 MT 系統在縮減變數方面的差異，在於主成份分析以原始變數之線性組合來縮減維度，主成份分析的計算仍需要利用到原始變數，因此，主成份分析對於原變數之縮減沒有幫助。MT 系統主要是提供量表，維度的縮減則是根據量表測量的能力、訊號雜音比及直交表。

(2)鑑別分析(Discriminant Analysis, DA)

有關鑑別分析的說明，在 2.2.2 節中略有說明。鑑別分析是最常被使用於分類問題

的統計技術之一，主要目的是找出預測變數的線性組合，使其組間差異平方和相對於組內差異平方和之比值為最大，其程序為(1)檢定各組重心是否有差異，(2)找出哪些預測變數具有最大的鑑別能力，(3)最後根據新觀察值的預測變數，將該觀察值進行指派。利用鑑別函數或是馬氏距離可執行鑑別分析，實際上也就是找出能最佳分類的預測變數組合。在結果的解釋方面，鑑別分析主要了解哪幾個變數可以區分不同之組別，而這些組別又有何不同；在預測方面，則根據建構出的鑑別函數進行分類，亦即各觀察值帶入函數求得鑑別分數，並根據所得之鑑別分數進行分類。分派的原則是根據最小化錯誤分類之期望成本，而此成本是根據錯誤分類的成本及先驗機率來決定。因此，線性鑑別函數的本質是一種機率的方法，鑑別分析的使用有某些限制。鑑別分析強調分類的結果，對於測量不正常的嚴重程度沒有幫助。表 3.3 說明鑑別分析與馬氏田口系統的差異。

表 3.3 鑑別分析與馬氏田口系統比較表

	鑑別分析	馬氏田口系統
馬氏距離	使用馬氏距離的目的是為分類	馬氏距離用以發展量表
機率屬性	鑑別函數的本質是一種機率的方法	本質上不屬於機率的觀念
閾值的決定	利用先驗機率及錯誤分類的成本	利用品質損失函數，最小化總成本

(3) 逐步迴歸(Stepwise Regression)

逐步迴歸廣泛地應用在篩選重要的變數上。這種反覆的程序，透過在每一個步驟中增加或減少變數的方法，來建構一系列的迴歸模式。根據 F 值來決定增加或減少變數的個數。當變數很多時，這個方法需要許多的反覆。逐步迴歸的方法被質疑，因為無法保證所獲得的為最佳變數。

(4) 複迴歸(Multiple Regression)

複迴歸的主要目的是估計依變數 y ，將 y 視為 k 個自變數的函數。利用最小平方法來發展複迴歸需要有一些假設，例如：誤差項的平均數為 0；誤差項的變異數為常數；兩組誤差項之共變異為 0。複迴歸分析比較大的問題是當變數很多的時候會變得很複雜，並且利用複迴歸並無法輕易測量不正常的嚴重程度。MT 系統與複迴歸相似之處是都可用來判斷出觀察值是否異常，但 MT 系統更可以進一步衡量異常的嚴重程度。

(5)多變量製程管制圖(Multivariate Control Charts)

多變量製程管制圖是單變量製程管制圖的延伸，多變量管制圖的目的為同時監控多個變數的狀況，這些管制圖的使用類似單變量的管制圖，當製程失控時可以進行修正。MTS 中所用的 SMD 是 Hotelling T^2 多變量管制圖乘以一個常數，MT 系統與 Hotelling T^2 管制圖相同之處在於兩者皆可進行製程監控，惟管制圖是根據機率的管制界限來進行判斷，但 MTS/ MTGS 則是根據品質損失函數所獲得的閾值來決定是否異常。

(6)類神經網路(Neural Networks)

有關類神經網路已於 2.2.2 節略有說明。類神經網路與傳統統計方法的最大不同在於它不需指定特定的函數形式，此外在進行計算時有平行處理能力，且更能根據樣本本身進行學習，因此在進行資料分析時不需任何的假設。MT 系統與類神經網路相似之處是不需要對分配進行假設。但類神經網路在使用上有些限制：(1)對於訓練網路而言，資料必須隨機化，因為當我們加入新的資料型態時，網路必須忘記舊的資料型態；(2)降低維度不容易，類神經網路並沒有提供一套簡單的方法來縮減變數；(3)無法利用量表衡量出不正常的程度(Taguchi, 2002)。



3.5 MT 系統之相關議題

由於 MT 系統是個新方法，仍有許多問題必須面對。本節討論以下的議題：(1)直交表及訊號雜音比角色；(2)雜音的處理；(3)閾值的決定；(4)其他議題。

3.5.1 直交表及訊號雜音比的角色

在多變量系統中，減少維度(Dimension Reducing)是一重要的挑戰，高維度的系統不但不容易處理也不容易解釋。通常在沒有先驗知識的情況下，會儘可能將許多的特性納入模式中，之後根據這些變數對分類的總貢獻率來進行修剪。本節主要討論直交表及訊號雜音比的角色如何修剪變數。

(1)直交表的角色

在穩健設計中，使用直交表的目的是利用最少的實驗次數來評估因子與因子間之交互作用效果，直交表讓工程師在考慮雜音及成本因素下設計出一穩健之系統。直交表是一種檢查的裝置，可以防止不良的設計流到下游。在 MT 系統中，直交表的目的是透過利用最少的實驗次數來確認重要的變數。將變數指派到不同的行，其中水準 1 代表「使用」此變數，水準 2 代表「不用」此變數，在 MT 系列中只應用了兩水準的直交表，對應每一組實驗利用水準 1 與水準 2 來決定是否使用該變數計算 SMD。

為了說明直交表在 MT 系統的角色與作法，以 $L_8(2^7)$ 加以說明。假設有四個變數 A,B,C,D，將之配置於直交表的前 4 行，其中第一組組合為 1-1-1-1，代表馬氏空間是由這四個變數所組成的 4*4 相關矩陣所建構，而對應於這一組實驗的異常觀察值的 SMD 也是由這個矩陣計算出來的。第二組組合為 1-1-1-2，則代表馬氏空間是由變數 A,B,C 所組成的 3*3 相關矩陣所建構。對於每一組實驗，只計算對應於異常狀況的 SMD，並不需要考慮對應於正常狀況下的馬氏距離，因為我們知道這一個群體是正常的。而變數的重要性則根據他們測量「異常程度」的能力來決定。然而，在 MTGS 中，如果偏相關的效果不顯著時，則不需要直交表，因為可以直接從直交變數直接估計。

(2) 訊號雜音比的角色

對於多變量的應用而言，確認一組可充分偵測異常程度的變數是非常重要的，同時，對於說明原始系統的績效評估及系統的改善程度而言，也非常重要。在 MT 系統的方法中，訊號雜音比有以下幾個目的：(1) 確認足以偵測異常觀察值的重要變數；(2) 評估系統的績效；(3) 改善系統的績效 (Taguchi, 2002)。訊號雜音比對於改善量表正確性及降低診斷成本而言是相當重要的，訊號雜音比嘗試在對雜音進行某些調整後來捕捉真正資訊的大小。訊號雜音比只利用異常的觀察值來計算，因為它需要以馬氏空間外的觀察值來測試量表的正確性。

在進行直交配置並計算各種組合下異常值之 SMD 後，我們根據訊號雜音比來決定重要的變數。在 MT 系統中，有兩種訊號雜音比：(1) 望大特性；(2) 動態特性。若異常的觀察值的異常程度未知時，通常建議使用望大特性。望大特性的公式如下：

$SN = -10 * \log[\frac{1}{t} \sum_{i=1}^t (\frac{1}{D_i^2})]$ ，其中 SN 代表訊號雜音比， $D_1^2, D_2^2, \dots, D_t^2$ 代表對應 t 個異常觀察值的 SMD。使用望大特性的理由是相對於正常觀察值而言，異常觀察值的 SMD 應該較大。

田口等人(2003)也建議可用動態訊號雜音比，因為動態訊號雜音比可以提供更精確之預測。若採用動態訊號雜音比，則必須在事前瞭解每一個異常觀察值的嚴重程度。用 M_j 來代表 t 個異常觀察值的嚴重程度，此值越大代表異常的程度越嚴重。如果 M_j 未知，則可將異常觀察值依嚴重程度分群，可將每一群 $\sqrt{SMD_j}$ 的平均值當成 M_j ，這些平均值通常稱為“Working Average”。討論以下兩種狀況(1)所有異常狀況均已知時；(2)所有異常狀況均未知，但可利用“Working Average”時。

(1) 所有異常狀況均已知時：

假設 M_1, M_2, \dots, M_t 代表異常觀察值的嚴重程度的真值，並假設存在 $y_i = \beta M_i$ 的關係時，其中 $y_i = \sqrt{SMD_i}$ ， $i=1, 2, \dots, t$ ， β 代表斜率。

Lunain 等人(1997) 及 Tsui (1999) 認為利用變異數分析來解釋動態訊號雜音比是很恰當的，如表 3.4 所示。表中 $SSR = (\sum_{j=1}^t M_j y_j)^2 / r$ 代表由斜率引起的平方和，其期望值

$E(SSR) = \sigma^2 + \beta^2 r$ ， $r = \sum_{j=1}^t M_j^2$ 則代表由訊號因子所引起的平方和，而斜率的估計式為

$\hat{\beta}^2 = \frac{1}{r}(SSR - MSE)$ ，其中 $MSE = (SST - SSR) / t - 1$ 則代表誤差均方和，而動態的訊號雜

音比則為 $10 \log[(\frac{1}{r})(SSR - MSE) / MSE] = 10 \log[\hat{\beta}^2 / MSE]$ 。

(2) 所有異常狀況均未知，但可利用 Working Average：

假設可將異常觀察值根據嚴重度分成 k 群，並假設存在 $y_i = \beta M_i$ 的關係時，其中 $y_i = \sqrt{SMD_i}$ ， $i=1, 2, \dots, k$ ， β 代表斜率。 M_i 代表第 i 群的平均值，假設每一群中有 m 個觀察值，則 $M_i = (\sum_{j=1}^m y_{ij}) / m$ ，其中 $y_{ij} = \sqrt{SMD_{ij}}$ ， $i=1, 2, \dots, k$ ， $j=1, 2, \dots, m$ 。其動態的

訊號雜音比為 $10 \log\left[\left(\frac{1}{r}\right)(SSR - MSE) / MSE\right]$ ，其中 $S_T = \sum_{i=1}^k \sum_{j=1}^m y_{ij}$ ， $r = m\left(\sum_{i=1}^k M_i^2\right)$ ， $SSR = \left(\sum_{j=1}^t M_j y_j\right)^2 / r$ ， y_j 代表第 i 類所有觀察值之總和， $SSE = S_T - SSR$ ， $MSE = SSE / (km - 1)$ 。表 3.5 整理各種型態的訊號雜音比。此表中 D_i^2 代表每一個觀察值的 SMD， $r = \sum_{i=1}^t M_i^2$ ，其中 M_i 代表輸入訊號， S_β 代表因斜率引起之平方和， V_e 代表誤差均方和。

對於已知的變數 X_i ，以符號 $\overline{S/N^+}$ 代表「包含」該變數時所有實驗結果之平均訊號雜音比，而符號 $\overline{S/N^-}$ 代表「不包含」該變數時所有實驗結果之平均訊號雜音比。而「增益」則代表這兩個值的差異，即 $\overline{S/N^+} - \overline{S/N^-}$ 。在 MT 分析中，如果「增益」為正值則保留此變數，否則刪除，「增益」可以說是測量改善程度的一種方式。總之，訊號雜音比的優點：(1) 提供一個簡單的方法確認重要的變數；(2) 可以量測系統的功能；(3) 可以利用各種不同的異常狀況評估績效並針對各種狀況找出重要的變數。

然而，Bovas 等人(2003)提出直交表並非是選擇重要變數之最佳策略，還有其他的方法包括 Forward Selection, Backward Elimination 等方法也都可提供最佳化，田口也回應認為如果有比直交表好的方法也可納入 MT 系統中(Taguchi, 2003)。

表 3.4 動態訊號雜音比之變異數分析

變異來源	平方和	自由度	均方和
β	$SSR = \left(\sum_{j=1}^t M_j y_j\right)^2 / r$	1	$MSR = SSR / 1$
誤差	SST-SSR	$t - 1$	$MSE = (SST - SSR) / (t - 1)$
總變異	$SST = \sum_{j=1}^t y_j^2$	t	

表 3.5 MT 系列中各種訊號雜音比

MTS			
使用條件	望大特性	動態特性	
	真值未知	真值未知但可分類	真值已知
	$SN = -10 * \log[\frac{1}{t} \sum_{i=1}^t (\frac{1}{D_i^2})]$	$SN = 10 * \log(\frac{1}{r} \frac{S_\beta - V_e}{V_e})$	$SN = 10 * \log(\frac{1}{r} \frac{S_\beta - V_e}{V_e})$
MTGS			
使用條件	望大特性	動態特性	
	真值未知	真值未知但可分類	真值已知
	$SN = -10 * \log[\frac{1}{t} \sum_{i=1}^t \frac{1}{(\frac{u_{ji}}{s_j})^2}]$	$SN = 10 * \log(\frac{1}{r} \frac{S_\beta - V_e}{V_e})$	$SN = 10 * \log(\frac{1}{r} \frac{S_\beta - V_e}{V_e})$

3.5.2 雜音的處理



本節討論有關雜音因子的處理，所謂的雜音因子通常指那些會影響系統績效，但不能控制、很難控制或控制成本很高的因子。我們無法知道所有的雜音，但我們必須辨認一些重要的雜音因子，而這些重要雜音產生了系統大部份的變異。對於雜音，若任其發生或採事後的補救的方式來處理，將徒增成本，若系統對雜音不敏感則可增加系統的穩定性。當有雜音因子存在時，很重要的是要了解多變量系統的行為。田口博士在有關 MT 系統相關的研究中對雜音的處理著墨不多，同時也認為這是個未來值得討論的議題，田口認為在各種不同的雜音狀況下最佳化系統可增加決策過程的穩健性，他提供以下幾個方法可以處理雜音：(1)分開處理雜音因子的每一個水準並在這些水準下應用 MT 系統，(2)將雜音視為一個控制變數中並決定它是否有用，(3)將雜音因子進行所有水準的組合，並在所有組合下應用 MT 系統，(4)如果無法量測雜音，則不考慮雜音因子 (Taguchi, 2002)。

3.5.3 閾值的決定

在多變量的系統中，要有效率的進行診斷，閾值的決定是很重要的。在 MT 系統中，閾值是由品質損失函數(Quality Loss Function, QLF)來決定的。在穩健設計中，品質損失函數被用來決定產品的規格界限，而在 MT 系統中，田口也引用品質損失函數的概念來決定閾值。傳統的統計方法是利用最小化錯誤分類的期望成本決定閾值，因此會涉及到機率問題。但田口先生認為在 MT 系統中，有關機率的問題是無意義的(Taguchi, 2002)，因為他認為每一個「異常」的觀察值都是獨立的個體，而不會組成另一個群體。在 MT 系統中的閾值，應該是最底的標準(安全因子)，以建立一個具經濟規模的模式。閾值對診斷過程的正確性有很大的影響，不正確的閾值會造成許多的錯誤，會引起巨大的損失。在 MT 系統中，提到兩種閾值：一般閾值(General Threshold)及特殊閾值(Specific Threshold)。閾值的決定的方式是使得社會損失最小，田口先生建議利用品質損失函數來決定閾值。在品質工程中，品質損失函數有三種：(1)望目特性；(2)望大特性；(3)望小特性。為了進行好的診斷，決策者必須明智地決定閾值，這個閾值類似 Lethal Dosage (LD)₅₀，也就是具有這個閾值的觀察值中有 50% 會被判成「好」，有 50% 會被判成「壞」，可以利用品質損失函數來求出此值。損失函數的曲線可利用 SMD 的平方根與其對應的損失來繪出。在靜態特性中會採用望小特性來計算品質損失函數，是因為 SMD 的目標值為 0。假設 Λ_0 為最低的限度，當某個觀察值 SMD 大於 Λ_0 時，就被認為是異常，如果 A_0 為對應 Λ_0 的損失，則損失函數為 $(\frac{A_0}{\Lambda_0^2}) * SMD$ 。

以下有兩種方法決定閾值：

方法 1：假設以專家來判定某一群觀察值是「異常」，則對應這群觀察值的 SMD 的平均值 Λ^2 就可被視為一個閾值，即 $T = \Lambda^2$ ，也就是完全由專家主觀的判斷。

方法 2：假設 Λ_0 為最低的限度，這是一極重要的距離。如果 A_0 為對應 Λ_0 的損失，則損失函數 $(\frac{A_0}{\Lambda_0^2}) * SMD$ 。故當 T 為一個閾值時，則 $T = \Lambda^2$ 。所以當某一個觀察值的 SMD 為 T 時，則這個觀察值可能是「正常」也可能「異常」，若其 SMD 小於 T 值則為正常；

若 SMD 大於 T 值則為不正常。

3.5.4 其他議題

除上述探討的議題之外，還有一些議題須加注意，包括 MTS 對類別資料的處理、MTS 中如何抽樣及 MTS 之假設條件等議題。

一般而言，資料型態可分成名目、順序、等距及等比尺度。在 MT 系統中，針對等距及等比尺度是可直接使用不需另行處理，但對於名目尺度則採用虛擬(Dummy)變數表示。若變數值的類別多於二個以上時，可以用 1-of-(c-1)編碼，即以 c-1 個虛擬變數來代表 c 個類別，以避免線性相依(Linear Dependency)的問題。針對類別變數的處理，仍遵循前述的四個步驟，包括(1)利用馬氏空間建立量表；(2)驗證量表；(3)確認重要變數；(4)利用重要變數進行未來預測。在 MTGS 中若有使用類別變數，在確認重要變數的階段中，重要變數的決定是靠變數的水準來決定，如果任何一個變數的其中一個水準的訊號雜音比較高，則該變數就是重要的變數。有關計算 SMD 時，類別變數的處理方式是否恰當；及類別過多時應如何處理的部份，田口博士於此著墨甚少，這是未來可討論的議題之一。

其次是抽樣的問題。在 MT 系統中，建構馬氏空間是非常重要的，因為型態的辨識完全依賴此群體。田口認為在驗證階段中，抽樣並不是一個重要的議題，在 MT 系統中，只有在建構馬氏空間時才有必要抽樣。但田口先生並沒有對抽樣的議題多所著墨，例如：如何抽樣及樣本的比例等議題，這是探討的主題之一。

最後，討論假設條件的議題。田口博士一直強調 MT 系統是一套資料分析的技術，就如同類神經網路的方法，使用這些技術都不需要任何分配的假設。這個說法使得 MT 的使用限制大幅的減少，但是 MT 系統所用的 SMD 卻必須建立在常態的假設下，這是 MT 系統受爭議之處(Woodall, *et al.*, 2003; Hawkins, 2003)，這將是未來可探討的議題。

3.6 結論

以上各節將 MT 系列作一整理及說明，本研究將 MT 系統的使用條件進行匯整。圖 3.5 說明 MT 系列之選用流程。首先，必須判斷變數的個數是否過多，何謂多寡尚未有一恰當之定義，此步驟的目的與直交表的使用有關，因為變數過多時，直交表的使用就變得繁複，所以當變數過多時則可採用多重馬氏距離法(Multiple Mahalanobis Distance Method, MMD)來選取子集合。其次是「正常」的定義容易嗎？一般而言，如果有領域專家的協助對於問題的定義會較為容易且恰當，但有些狀況要定義「正常」並不容易，所以利用同質性的概念，計算出將所有觀察值的 SMD 並刪除極端值，利用同質性較高的觀察值建構馬氏空間。一但有適合的馬氏空間後，接下來判斷欲處理的問題是否具有「方向性」，若有則採用 MTGS 來處理，若沒有「方向性」時，則可觀察相關矩陣是否具有「共線性」，若有則可採用 MTGS/MTA 來處理，若沒有「共線性」則可採用 MTS。



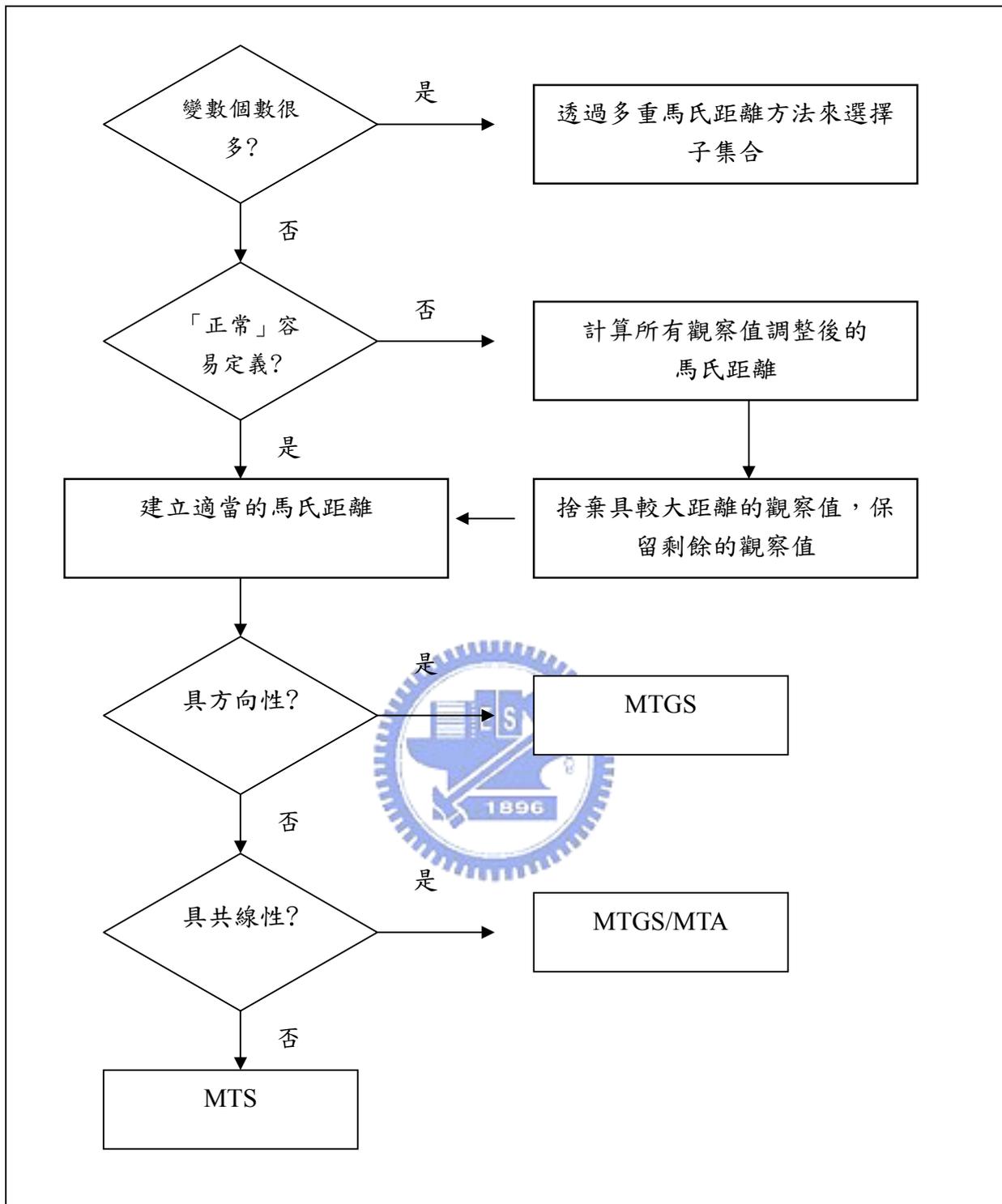


圖 3.5 MT 系統之選用流程

第四章 MTS 特性之探討與比較

本章主要探討 MTS 之特性，透過各種實驗對 MTS 進行瞭解。本章共分成 6 節，4.1 節為資料說明及實驗規劃，4.2 節為 MS 的選擇，4.3 節為抽樣的策略，4.4 節則針對訊號雜音比進行討論，4.5 節提出對閾值之看法，4.6 節將 MTS 與其他方法進行比較。

4.1 資料說明與實驗規劃

為探討 MTS 之特性，本研究欲觀察(1)馬氏空間的選擇；(2)抽樣策略的變動；(3)訊號雜音比的選擇對 MTS 分類績效的影響，並討論閾值的決定。利用一筆紫羅蘭花的資料來驗證，此資料共有 150 筆觀察值，包括三種品種分別是 Setosa、Versicolor, Virginia 以及 4 個辨識變數包括花萼長度、花萼寬度、花瓣長度、花瓣寬度(資料來源 www.ics.uci.edu/mlearn/MLSummary)。分類的主要目的是利用 4 個辨識變數來決定花的品種，此筆資料經常被當成分類之典型範例。

首先進行基本資料之分析，包括敘述統計、常態性的檢查、相關分析及同質性檢定。表 4.1 為此筆紫羅蘭資料的預檢表，可觀察出於花萼長度、花瓣長度及花瓣寬度上，其平均值的大小依序為 Setosa 最小、Versicolor 次之、Virginia 最大，但在花萼寬度上則是以 Setosa 最大、Virginia 次之、Versicolor 最小。且 Setosa 在花瓣長度與寬度上都有比較小的現象，透過無母數 Kolmogorov-Smirnov 進行常態性的檢定，可知並非所有的變數的分配均為常態，再透過同質性的檢定，發現除花萼寬度外，其他三個變數均不具有同質性。

圖 4.1 為此筆資料之盒鬚圖(Boxplot)，圖中發現 Versicolor 及 Virginia 的 4 個變數的排列，由小至大分別為花瓣寬度、花萼寬度、花瓣長度、花萼長度；但 Setosa 的變數排列則為花瓣寬度、花瓣長度、花萼寬度及花萼長度。表 4.2 與圖 4.2 分別為紫羅蘭資料之相關矩陣表及散佈圖，觀察出這 4 個辨識變數中除花萼長度及花萼寬度間沒有顯著的相關性之外，其他變數之間均有顯著的相關性。為瞭解 MTS 之特性，將整筆資料劃分成訓練樣本與測試樣本，以訓練樣本來建立分類模式，以測試樣本來驗證模式。利用

表 4.3 所示之各種實驗來觀察 MTS 的特性。

表 4.1 紫羅蘭資料預檢表

品種		個數	平均數	標準差	常態性檢查(KS)		同質性檢查	
					統計量	顯著性	統計量	顯著性
花萼 長度	Setosa	50	50.06	3.5249	0.115	0.097	7.381	0.001*
	Versicolor	50	59.36	5.1617	0.096	0.2*		
	Virginica	50	65.88	6.3588	0.115	0.096		
花萼 寬度	Setosa	50	34.28	3.7906	0.105	0.2*	0.595	0.553
	Versicolor	50	27.68	3.1392	0.123	0.057		
	Virginica	50	29.74	3.225	0.128	0.04		
花瓣 長度	Setosa	50	14.62	1.7366	0.153	0.005	20.684	0.000*
	Versicolor	50	42.6	4.6991	0.117	0.084		
	Virginica	50	55.52	5.5189	0.114	0.125		
花瓣 寬度	Setosa	50	2.46	1.0539	0.349	0.000	19.652	0.000*
	Versicolor	50	13.26	3.911	0.148	0.008		
	Virginica	50	20.26	2.7465	0.121	0.066		

*在顯著水準為 0.05 時，具統計顯著性

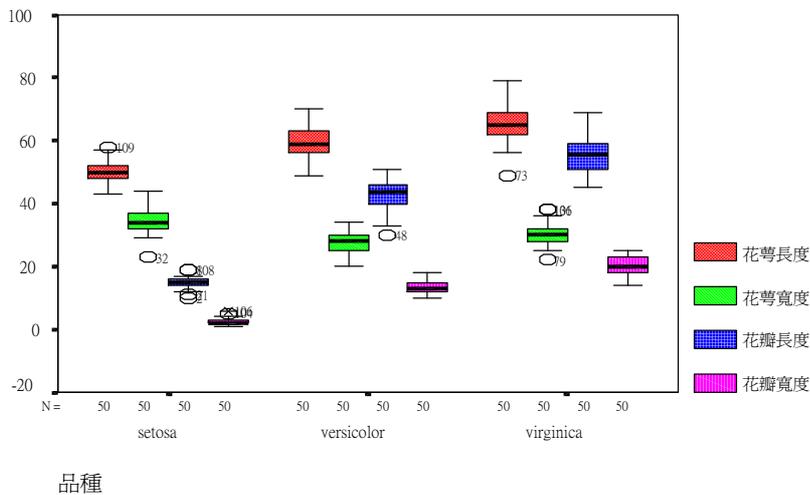


圖 4.1 紫羅蘭之盒鬚圖

表 4.2 紫羅蘭之相關矩陣表

	花萼長度	花萼寬度	花瓣長度	花瓣寬度
花萼長度	1	-0.117	0.872**	0.818*
花萼寬度	-0.117	1	-0.429**	-0.366*
花瓣長度	0.872**	-0.429**	1	0.963*
花瓣寬度	0.818**	-0.366**	0.963*	1

**在顯著水準為 0.01 (雙尾) 顯著相關性

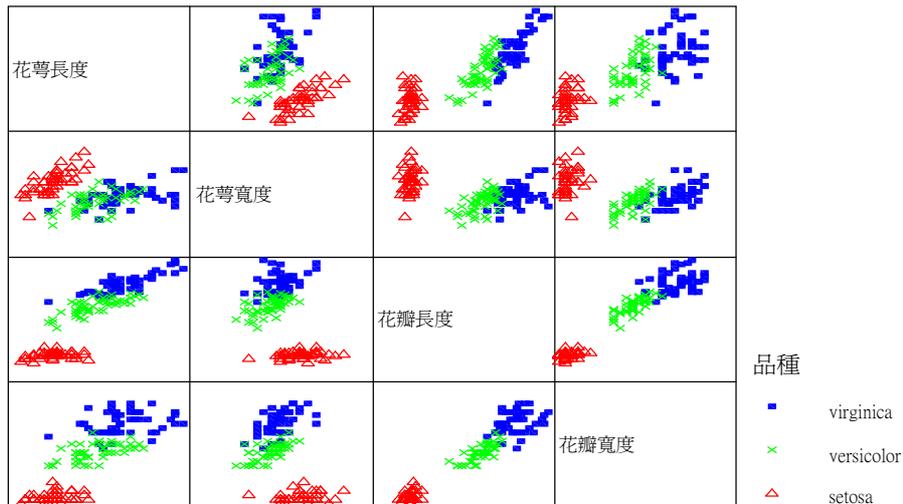


圖 4.2 紫羅蘭之散佈圖

表 4.3 各種實驗規劃一覽表

名稱	馬氏空間	訓練樣本與測試樣本比例	直交表	訊號雜音比
實驗 4-1	Setosa	訓練:測試 2:1(34 /16)	L ₈ 標準配置	望大
實驗 4-2	Versicolor	訓練:測試 2:1(34 /16)	L ₈ 標準配置	望大
實驗 4-3	Virginica	訓練:測試 2:1(34 /16)	L ₈ 標準配置	望大
實驗 4-4	Setosa	訓練:測試 1:1(25 /25)	L ₈ 標準配置	望大
實驗 4-5	Setosa	訓練:測試 4:1(40/10)	L ₈ 標準配置	望大
實驗 4-6	Setosa	訓練:測試 2:1(34 /16)	L ₈ 標準配置	動態

4.2 馬氏空間之選擇

本節主要探討馬氏空間的選擇，對 MTS 在分類問題績效之影響。

4.2.1 實驗說明及過程

為了解馬氏空間的選擇對分類績效之影響，本研究將資料劃分成訓練與測試樣本，其比率約為 2:1，其中訓練樣本各有 34 筆資料，測試樣本各有 16 筆資料。分別進行三次實驗：以 Setosa 為馬氏空間，Versicolor、Virginica 為「異常值」之實驗，稱為實驗

4-1；實驗 4-2 則以 Versicolor 為馬氏空間，Setosa、Virginica 為「異常值」；實驗 4-3 以 Virginica 為馬氏空間，Versicolor、Setosa 為「異常值」，實驗的過程如圖 4.3 所示。

圖 4.3 中利用 Receiver Operating Characteristic(ROC)曲線來描述在不同的閾值下，分類之績效。在分類的問題中，ROC 曲線是一種評估分類模式的方法，它是由一組(FP,TP)所組成軌跡，其中 FP(False Positive)代表 β 值，也就是犯型 II 錯誤之機率；TP (True Positive)代表 $1-\alpha$ 值，也就是將目標樣本正確分類的機率。利用 ROC 曲線評估分類的績效時，若曲線愈高，則代表該分類模式之績效愈好，若曲線愈接近對角線時，則表示分類績效愈差。

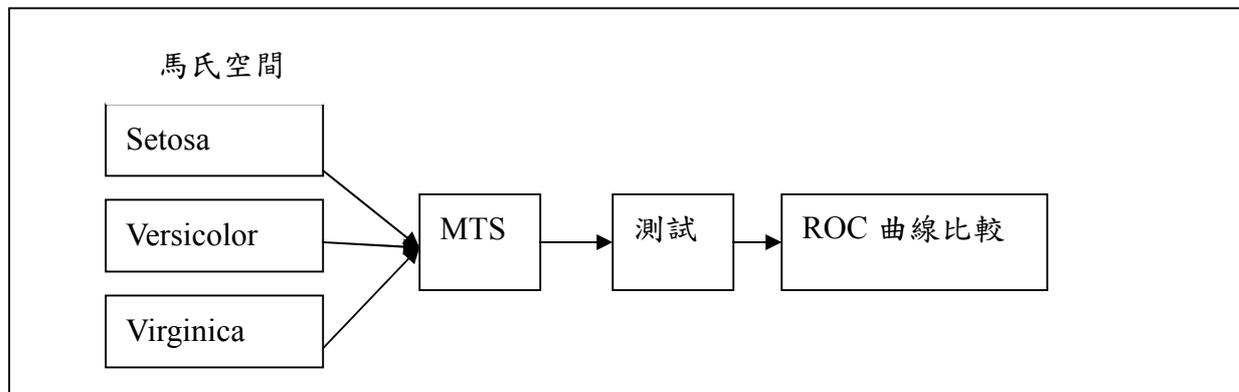


圖 4.3 馬氏空間之實驗規劃

<實驗 4-1>以 Setosa 建構馬氏空間

實驗 4-1 以 Setosa 的 34 個觀察值及 4 個辨識變數所構成之矩陣為馬氏空間，獲得表 4.4 的相關矩陣、反矩陣、平均數及標準差，透過公式(3.2)計算得此空間中最小之 SMD 為 0.096，最大為 2.747，其平均數為 0.9999，符合 MTS 的理論。在此以「全因子模式」代表利用所有的變數所建構出的分類模式，「重要因子模式」則代表經由直交表與訊號雜音比篩選出的重要變數所形成之模式。

接著以 Versicolor、Virginica 為「異常」觀察值，並以馬氏空間中的平均數、標準差、相關矩陣之反矩陣來計算這兩群紫羅蘭之 SMD，可得 Versicolor 之最小值為 42.06，最大值為 145.687；Virginica 之最小值為 140.608，最大值為 288.724。表示這是一個很好的量表，可以很容易區別出 Setosa 與 Versicolor、Virginica 之差別，其 SMD 之分配圖，

如圖 4.4 所示。利用 Setosa、Versicolor、Virginica 的測試樣本所獲得的 SMD 之範圍分別為 0.087~4.713；60.371~153.615；145.879~329.744，如圖 4.5 所示。

由圖 4.4 及圖 4.5 可以瞭解，以 Setosa 為馬氏空間的意義是希望找出與 Setosa 不同的品種，我們很容易區分 Setosa 與 Versicolor 及 Setosa 與 Virginica，但圖形顯示 Versicolor 與 Virginica 有重疊的現象。若以 Setosa 所形成之量表來測量及區分 Versicolor、Virginica 之差異時，必須訂定一個次閾值(Sub-threshold)以協助判別，假設在最小化總錯誤率的方法下，在本例中將次閾值訂定在 141 時，則在訓練樣本中將有 2.9%的誤判，在測試樣本中則有 3%的誤判。

雖然，圖 4.4 及圖 4.5 很容易可以區隔 Setosa 與其他兩品種的不同，但有時不見得能如此清楚地區分，需要閾值來協助，而閾值的決定會影響型 I 及型 II 錯誤。本研究觀察不同的閾值對分類正確率之影響，其 ROC 曲線如圖 4.6 所示。

利用最小化總錯誤率的方式，觀察當閾值設在 3~40 時，訓練樣本中均有 100%之分類正確率。而當閾值為 3、3.5、4 及 4.5 時，其測試樣本之正確率分別為 93.75%、95.83%、97.91%及 97.91%，當閾值訂為 5~40 時，其測試樣本之正確率為 100%，當閾值超過 40 時，訓練樣本之正確率就為 99.01%，資料如表 4.5 所示。

表 4.4 實驗 4-1 以 Setosa 為馬氏空間之相關資料

Correlation Matrix	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)	Inverse matrices	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)
A(花萼長)	1	0.718*	0.463	0.383	A(花萼長)	2.481	-1.528	-0.559	-0.326
B(花萼寬)	0.718*	1	0.3	0.261	B(花萼寬)	-1.528	2.07	0.082	0.011
C(花瓣長)	0.463	0.3	1	0.4	C(花瓣長)	-0.559	0.082	1.377	-0.358
D(花瓣寬)	0.383	0.261	0.4	1	D(花瓣寬)	-0.326	0.011	-0.358	1.265
(表中*者代表高度相關)									
平均值	49.853	34.059	14.441	2.294	標準差	3.448	3.948	1.718	0.924

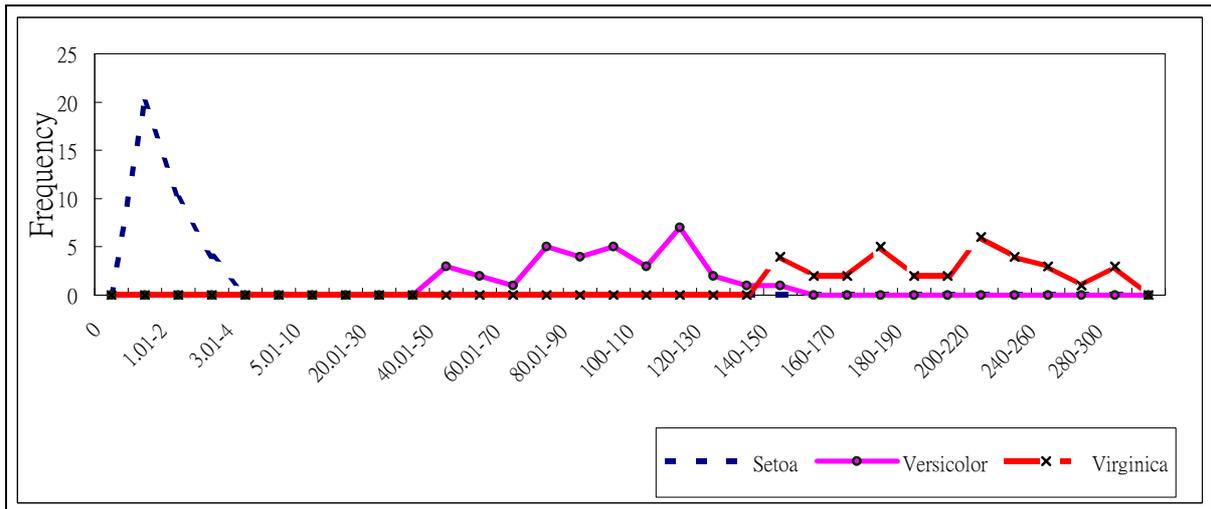


圖 4.4 實驗 4-1 全因子模式訓練樣本之馬氏距離分配圖

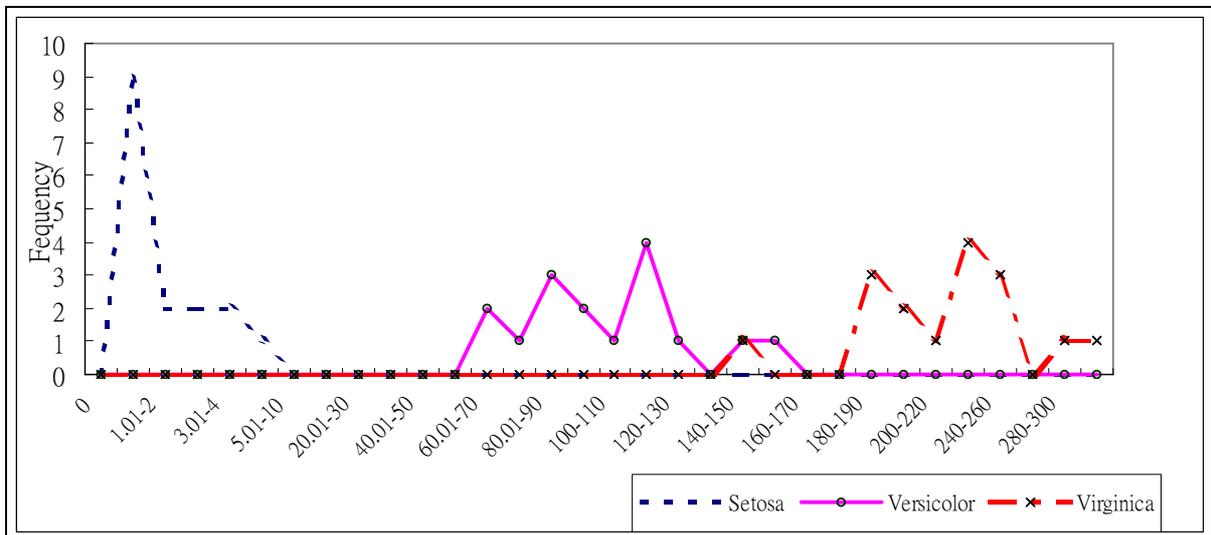


圖 4.5 實驗 4-1 全因子模式測試樣本之馬氏距離分配圖

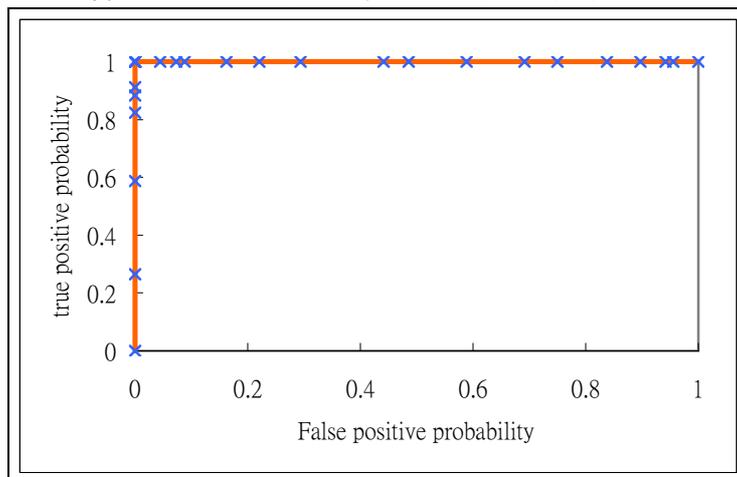


圖 4.6 實驗 4-1 全因子模式之 ROC 曲線

表 4.5 實驗 4-1 全因子模式訓練及測試樣本之正確率

閾值	實驗 4-1 全因子分類正確率			
	4~4.5	訓練	預測 \ 實際	Setosa (34)
Setosa			100%(34)	0%(0)
其他品種			0%(0)	100%(68)
測試		預測 \ 實際	Setosa (16)	其他品種(32)
		Setosa	93.75%(15)	0%(0)
		其他品種	6.35%(1)	100%(32)
閾值 5~40	訓練	預測 \ 實際	Setosa (34)	其他品種(68)
		Setosa	100%(34)	0%(0)
		其他品種	0%(0)	100%(68)
	測試	預測 \ 實際	Setosa (16)	其他品種(32)
		Setosa	100%(16)	0%(0)
		其他品種	0%(0)	100%(32)

其次為篩選重要變數。利用 L8 直交表與望大特性之訊號雜音比處理異常觀察值，隨機抽取共 20 個異常值，其中 Versicolor、Virginica 各 10 個，詳細配置如表 4.6 所示，所獲得的重要因子為 B(花萼寬)、C(花瓣長)及 D(花瓣寬)，這三個因子為區分 Setosa 與其他兩種品種的重要因子，其因子效果如圖 4.7 所示。

觀察此筆資料之盒鬚圖，如圖 4.8 所示，Setosa 的花萼長度與 Versicolor、Virginica 有部分重疊，花萼寬度也是一樣有重疊的現象，但 Setosa 的花萼寬度的範圍較大，且分佈位置與 Versicolor、Virginica 相反，也就是 Versicolor、Virginica 的花瓣長度比花萼寬度大，但 Setosa 卻是花萼寬度大於花瓣長度，Setosa 的花瓣寬度明顯比其他兩種窄。

最後，利用所找出之重要變數 B、C 及 D，形成新的量表以觀察其績效，圖 4.9 為重要因子模式訓練樣本之馬氏距離分配圖，圖 4.10 為重要因子模式測試樣本之馬氏距離分配圖。採用 ROC 曲線判定其重要因子模式之分類能力的好壞，如圖 4.11 所示，利用最小化兩種誤判機率的方式來選取閾值，當閾值為 4~50 時，訓練樣本之正確率為 100%，當閾值為 4~5.8 時，其測試樣本之正確率為 97.92%。當閾值為 5.9~50 時，則測試樣本之正確率則為 100%，資料如表 4.7 所示。

表 4.6 實驗 4-1 直交表配置及訊號雜音比

run	1(A)	2(B)	3(C)	4(D)	5	6	7	MD1	MD2	MD3	MD4	MD5	MD6	MD7	MD8	...	MD18	MD19	MD20	SN比
1	1	1	1	1	1	1	1	106	123	112	77.1	48.6	74.87	111.7	48.77	...	72.03	179.9	185.2	20.3571
2	1	1	1	2	2	2	2	126	134	135	89.3	53.2	95.13	122.4	53.09	...	81.71	180.9	210.9	20.7017
3	1	2	2	1	1	2	2	70.3	112	80.7	56.7	40.5	35.14	97.59	41.66	...	56.71	188.5	147.7	19.5051
4	1	2	2	2	2	1	1	4.3	14.5	34.1	5.58	0	5.582	8.66	0.061	...	5.582	19.3	16.83	-14.29
5	2	1	2	1	2	1	2	79.5	119	88	67.4	48	43.18	107.2	46.72	...	66.13	202.9	158.9	20.0222
6	2	1	2	2	1	2	1	3.2	0.07	0.27	4.17	7.85	3.197	1.642	6.492	...	3.197	1.057	0.6	-1.4044
7	2	2	1	1	2	2	1	170	211	195	123	67.9	122.2	184.3	67.94	...	115.8	303.4	312.5	22.3065
8	2	2	1	2	1	1	2	316	359	359	221	117	238.9	316.3	116.7	...	204.3	477.8	557.2	24.5945

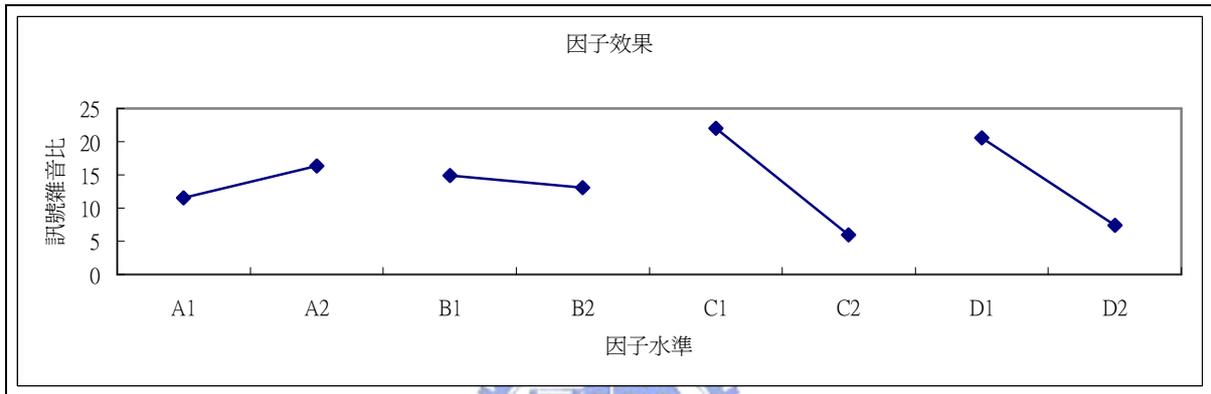


圖 4.7 實驗 4-1 因子效果圖

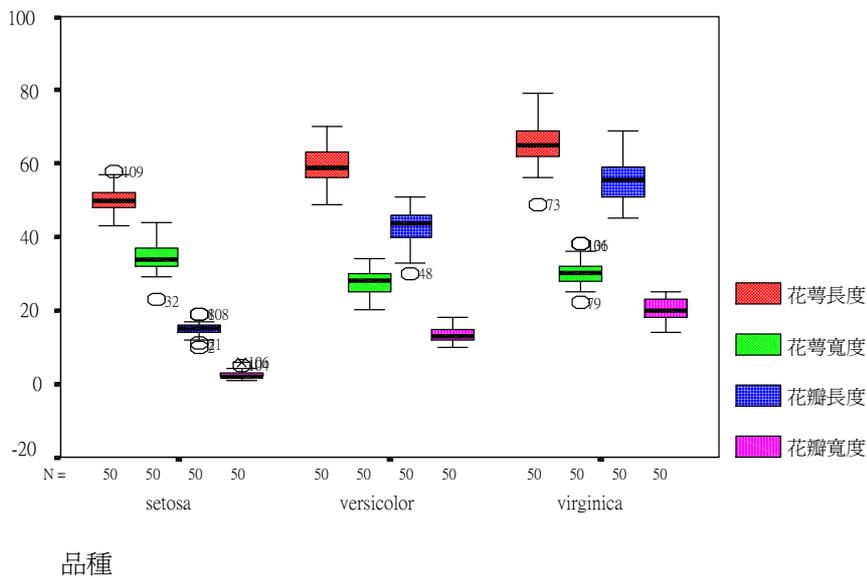


圖 4.8 實驗 4-1 之盒鬚圖

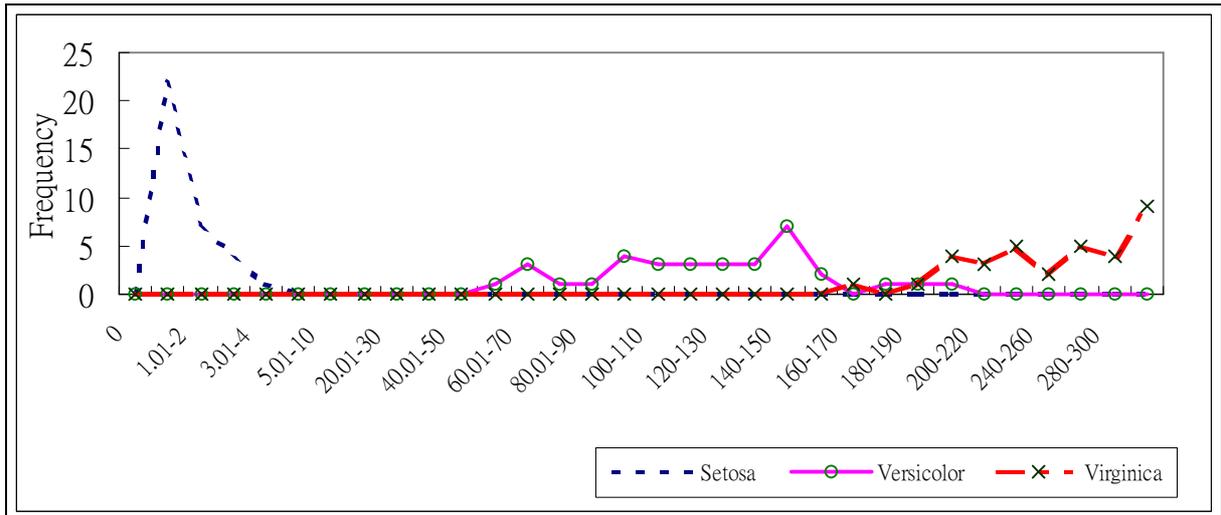


圖 4.9 實驗 4-1 重要因子模式訓練樣本之馬氏距離分配圖

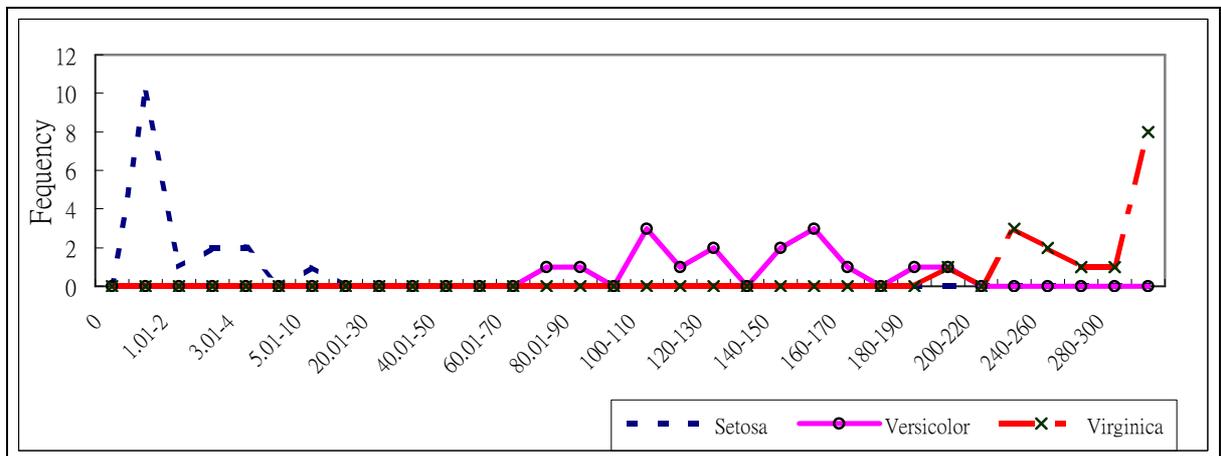


圖 4.10 實驗 4-1 重要因子模式測試樣本之馬氏距離分配圖

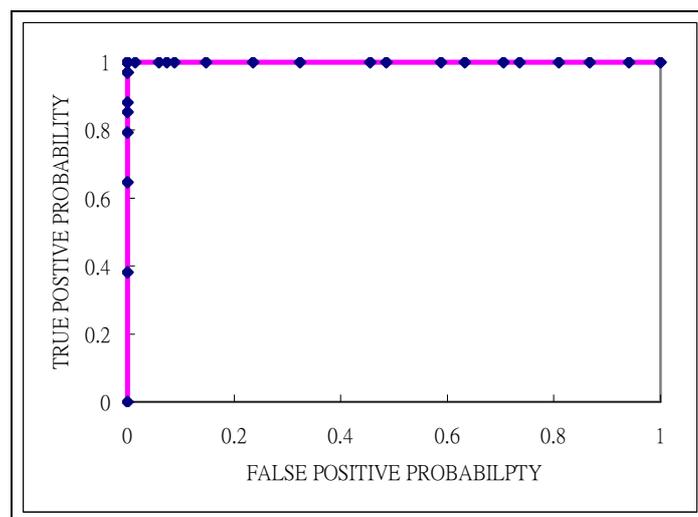


圖 4.11 實驗 4-1 重要因子模式之 ROC 曲線

表 4.7 實驗 4-1 重要因子模式訓練及測試樣本之分類正確率

閾值	實驗 4-1 重要因子分類正確率			
	4~5.8	訓練	預測 \ 實際	Setosa (34)
Setosa			100%(34)	0%(0)
其他品種			0%(0)	100%(68)
測試		預測 \ 實際	Setosa (16)	其他品種(32)
		Setosa	93.75%(15)	0%(0)
		其他品種	6.35%(1)	100%(32)
閾值 5.9~50	訓練	預測 \ 實際	Setosa (34)	其他品種(68)
		Setosa	100%(34)	0%(0)
		其他品種	0%(0)	100%(68)
	測試	預測 \ 實際	Setosa (16)	其他品種(32)
		Setosa	100%(16)	0%(0)
		其他品種	0%(0)	100%(32)

表 4.8 實驗 4-1 全因子及重要因子模式之訊號雜音比及正確率之比較表

	訊號雜音比	總正確率 (Threshold=4~4.5)	總正確率 (Threshold =6~40)
全因子模式 (4 個變數)	20.4	97.91%	100%
重要因子模式 (3 個變數)	22.3	97.91%	100%
增益	1.9		

完成實驗 4-1 之結果，我們比較原系統(全因子模式)與新系統(重要因子模式)的訊號雜音比及正確率，如表 4.8 所示。當閾值為 4~4.5 時，全因子模式與重要因子模式的測試樣本分類正確率為 97.1%，當閾值為 6~40 時，則全因子模式與重要因子模式的正確率為 100%。以訊號雜音比當成一個評估系統的績效之量測工具，表 4.8 提供實驗 4-1 整個系統最佳化前後的訊號雜音比，增益 1.9，此改善相當於減少約 20%的變異範圍。

<實驗 4-2>以 Versicolor 構建馬氏空間

重複實驗 4-1 的程序，將 Versicolor 視為馬氏空間進行實驗 4-2，獲得表 4.9 的相關資料，其最小之 SMD 為 0.087，最大值為 2.984，其平均值為 1.000029，符合 MTS 理論。

以 Setosa, Virginica 為異常觀察值，可得 Setosa 之最小 SMD 為 13.65，最大為 44.32；Virginica 之最小值為 1.534，最大值為 14.328，此量表並不理想。雖然可以很容易區別

出 Versicolor 與 Setosa 但並不容易區分 Versicolor 與 Virginica。利用此訓練樣本獲得的馬氏空間來進行測試，各測試 16 個樣本所獲得的 SMD 分別為 18.52~45.04(Setosa)，0.342~2.466(Versicolor)，2.904~17.467(Virginica)，如圖 4.12 所示。由圖 4.12 中可以瞭解，我們很容易區分 Versicolor 與 Setosa；但圖形顯示 Versicolor 與 Virginica 有重疊，Setosa 與 Virginica 也有重疊，故要以 Versicolor 所形成之量表來測量及區分 Setosa 與 Virginica 之差異時，在閾值為 15 的情況下其訓練樣本會有 3%的誤判，而測試樣本也有 3%的誤判。由圖 4.12 可知要區隔 Versicolor 與其他兩種品種的不同，需要閾值的協助。在總錯誤率最小的狀況下，當閾值為 2 時，其訓練樣本最高之正確率為 97.05%，當閾值為 2.5、3、3.5 及 4 時，其正確率分別為 96.07%、96.07%、93.13%、90.19%。在閾值為 2 的條件下，其測試樣本之正確率為 95.83%。圖 4.13 為實驗 4-2 之 ROC 曲線，比較圖 4.6 與圖 4.13 可以發現，實驗 4-2 比實驗 4-1 的績效差。

利用直交表 L8 與望大特性之訊號雜音比所獲得的重要因子為 A(花萼長)、C(花瓣長)及 D(花瓣寬)，這三個因子為區分 Versicolor 與其他兩種品種的重要因子，如圖 4.14 所示。最後利用所找出來的重要變數 A、C 及 D 來建構新量表以觀察其績效，圖 4.15 為重要因子模式測試樣本馬氏距離分配圖，發現 Virginica 與 Setosa 仍有重疊之現象。

當閾值為 2~2.9 時，其訓練樣本之分類正確率為 96.07%，而測試樣本之正確率為 93.75%。比較原系統與新系統的訊號雜音比及正確率，如表 4.10 所示。資料顯示，實驗 4-2 的新系統其增益及正確率並無改善。

表 4.9 實驗 4-2 以 Versicolor 為馬氏空間之相關資料

Correlation Matrix	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)	Inverse matrices	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)
A(花萼長)	1	0.553	0.81*	0.607*	A(花萼長)	3.165	-0.7	-2.676	0.643
B(花萼寬)	0.553	1	0.557	0.681*	B(花萼寬)	-0.7	2.024	0.468	-1.319
C(花瓣長)	0.81*	0.557	1	0.78*	C(花瓣長)	-2.676	0.468	4.823	-2.456
D(花瓣寬)	0.607*	0.681*	0.78*	1	D(花瓣寬)	0.643	-1.319	-2.456	3.423
(表中*者代表高度相關)									
平均值	59.059	27.529	41.971	13.118	標準差	5.319	3.336	4.854	1.967

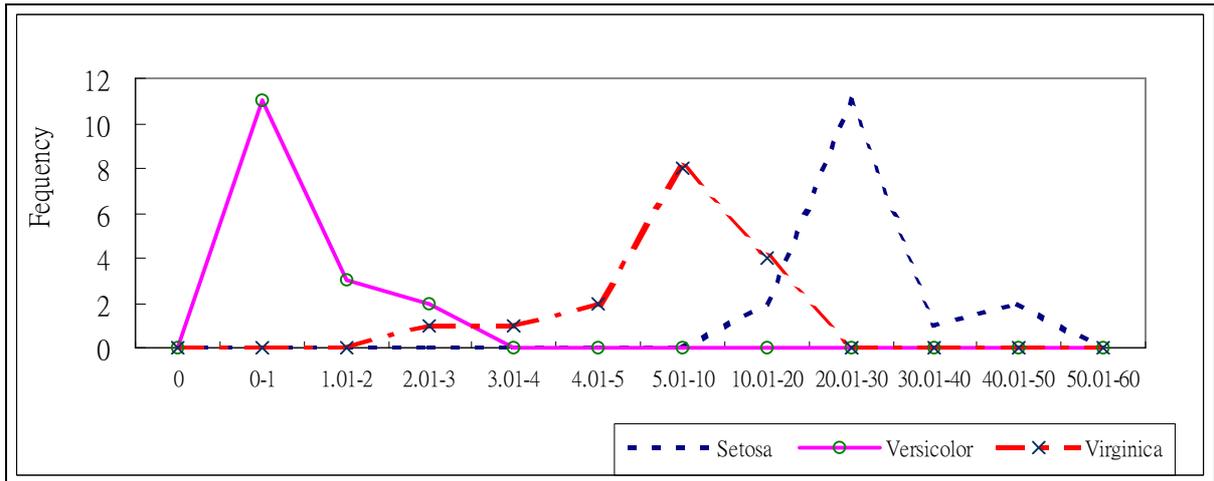


圖 4.12 實驗 4-2 全因子模式測試樣本馬氏距離分配圖

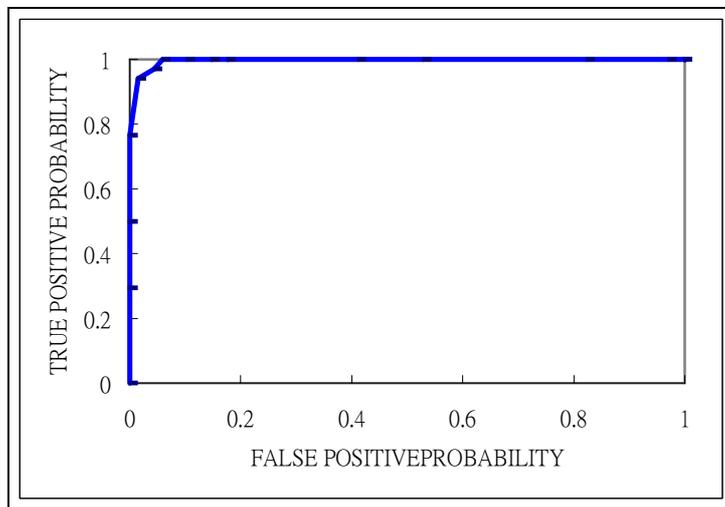


圖 4.13 實驗 4-2 全因子模式之 ROC 曲線

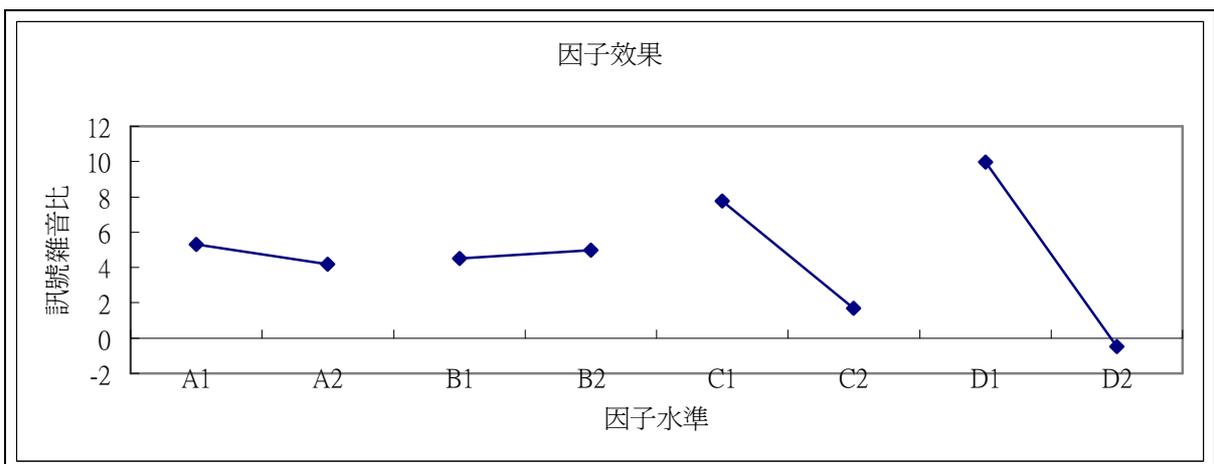


圖 4.14 實驗 4-2 因子效果圖

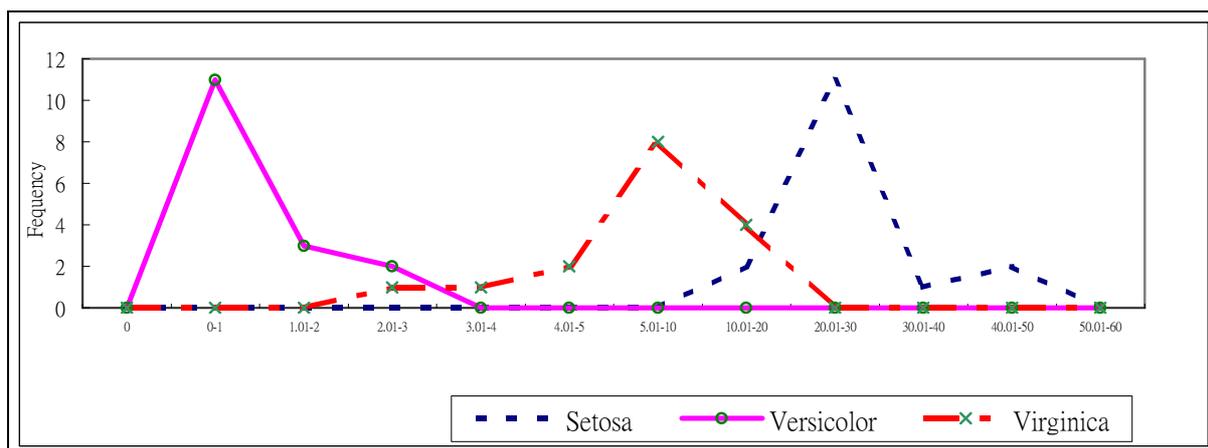


圖 4.15 實驗 4-2 重要因子模式測試樣本馬氏距離分配圖

表 4.10 實驗 4-2 全因子及重要因子模式之訊號雜音比及正確率之比較表

	訊號雜音比	總正確率
全因子模式(4 個變數)	9.53	95.83%
重要因子模式(3 個變數)	9.01	93.75%
增益	-0.52	

<實驗 4-3>以 Virginica 建構馬氏空間

實驗 4-3 以 Virginica 中 34 個觀察值當成馬氏空間，獲得之結果如表 4.11 所示。以 Virginica 為正常群體，所獲得之最小 SMD 為 0.203，最大為 2.833，其平均值為 1.000029，符合 MTS 理論。再以 Setosa, Versicolor 為異常觀察值，來計算這兩組之 SMD，可得 Setosa 最小之 SMD 為 35.46，最大為 83.38，Versicolor 之最小值為 1.438，最大值為 10.21，所以表示這量表也不盡理想，雖然可以很容易區別出 Virginica 與 Setosa 但並不容易區分 Virginica 與 Versicolor。利用此訓練樣本獲得的馬氏空間來進行測試，測試三種品種各 16 個樣本所獲得的範圍分別為 44.99~87.18(Setosa)，0.734~6.458(Versicolor)，0.274~5.976(Virginica)，如圖 4.16 所示。在閾值設定為 2.5 時，其訓練樣本之正確率為 96.07%，測試樣本之正確率為 87.5%，圖 4.17 則為實驗 4-3 之 ROC 曲線。

實驗 4-3 所篩選出的重要因子為 A(花萼長)、C(花瓣長)及 D(花瓣寬)，這三個因子為區分 Virginica 與其他兩品種的重要因子，如圖 4.18 所示。而由重要因子所形成之量表，所測量出的訓練及測試樣本的 SMD 分別如圖 4.19 及圖 4.20 所示。在重要因子的模式中，當閾值為 2 時，其訓練樣本之正確率為 93.13%，測試樣本之正確率為 79.16%。

最後，以重要變數建構新量表以觀察其績效並且比較原系統與新系統的訊號雜音比及總正確率，如表 4.12 所示，我們發現其訊號雜音比及正確率都沒有改善。

表 4.11 實驗 4-3 以 *Virginica* 為馬氏空間之相關資料

Correlation Matrix	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)	Inverse matrices	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)
A(花萼長)	1	0.569	0.845*	0.483	A(花萼長)	3.594	-0.436	-2.832	0.097
B(花萼寬)	0.569	1	0.591	0.671*	B(花萼寬)	-0.436	2.146	-0.327	-1.05
C(花瓣長)	0.845*	0.591	1	0.544	C(花瓣長)	-2.832	-0.327	3.867	-0.516
D(花瓣寬)	0.483	0.671*	0.544	1	D(花瓣寬)	0.097	-1.05	-0.516	1.938
(表中*者代表高度相關)									
平均值	65.029	30.029	54.559	20.412	標準差	5.828	3.204	4.966	2.691

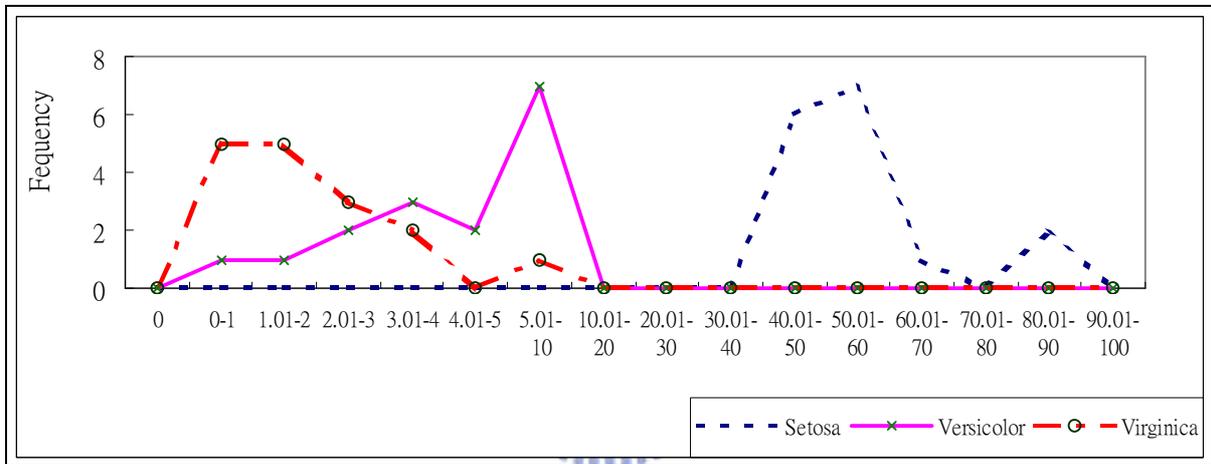


圖 4.16 實驗 4-3 全因子模式測試樣本馬氏距離分配圖

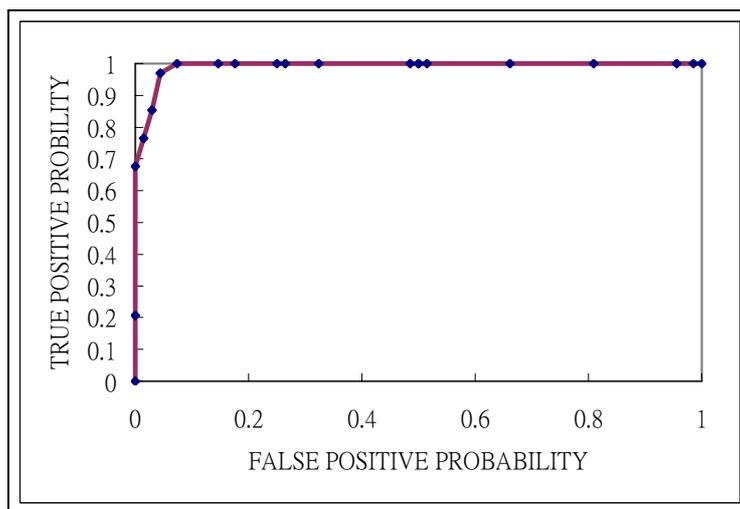


圖 4.17 實驗 4-3 全因子之 ROC 曲線

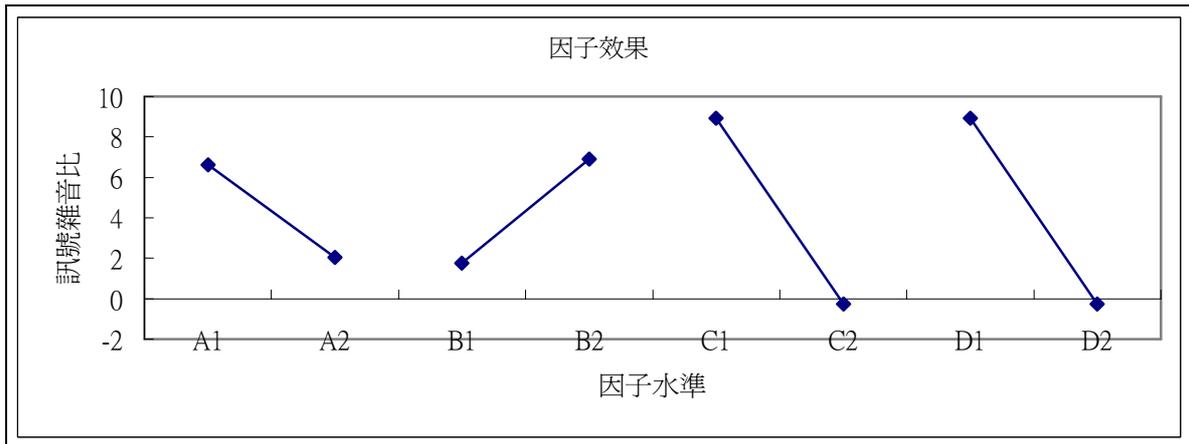


圖 4.18 實驗 4-3 因子效果圖

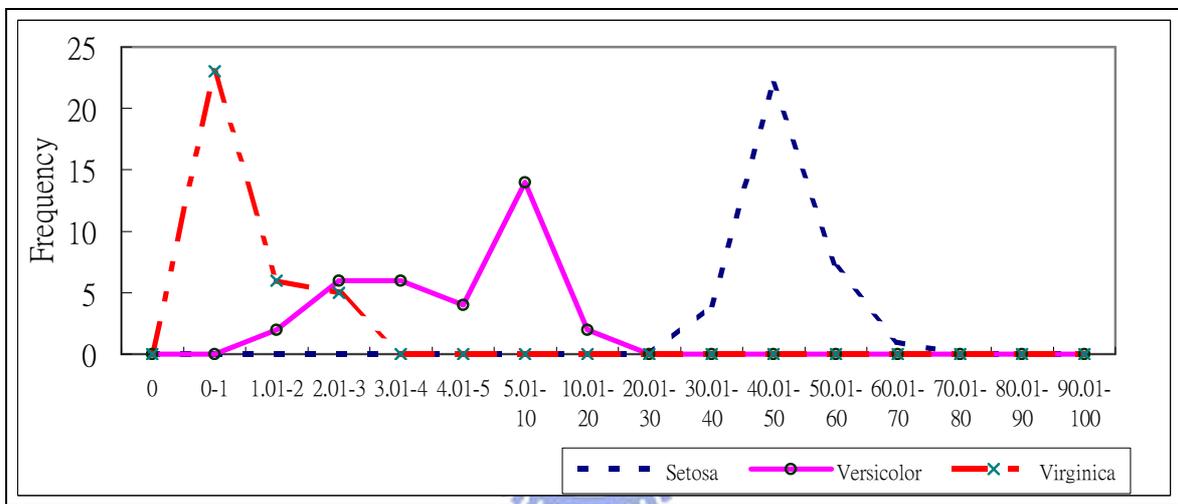


圖 4.19 實驗 4-3 重要因子模式訓練樣本馬氏距離分配圖

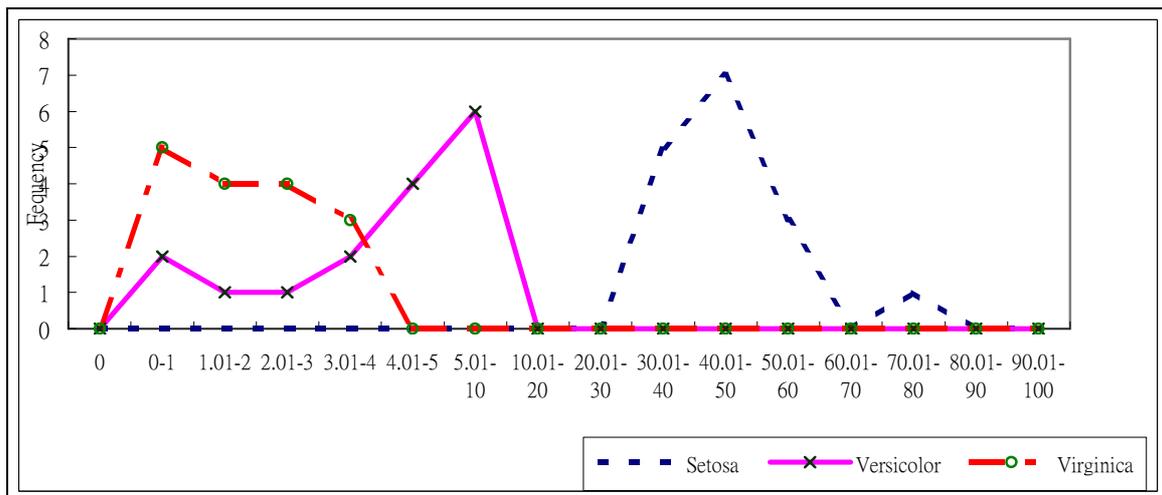


圖 4.20 實驗 4-3 重要因子模式測試樣本馬氏距離分配圖

表 4.12 實驗 4-3 全因子及重要因子模式之訊號雜音比及正確率之比較表

	訊號雜音比	總正確率
全因子模式(4 個變數)	9.11	87.5%
重要因子模式(3 個變數)	8.92	79.16%
增益	-0.19	-8.34%

4.2.2 結論

觀察上述三組實驗可知，MTS 之分類績效會因所選擇的馬氏空間而有所不同，因此馬氏空間的選擇是一重要議題。然而，應該如何選擇適當的馬氏空間，在 MTS 的理論中並沒有提及，田口博士建議由決策者決定。田口等人(2003)提到如果馬氏空間無法滿足平均數為 1 的條件，則需重新選擇不同的觀察值來建構馬氏空間；而測試的結果若與決策者的研判不符合時，則應該選擇不同的變數來建構馬氏空間。

觀察上述三組實驗，發現這三組 SMD 之平均值均為 1，故判定所選擇的樣本觀察值是適當的，但實驗 4-2 及 4-3 卻無法明顯區分欲辨識之目標群體，所以判定這 4 個變數對辨識 Versicolor 及 Virginica 而言，可能不恰當。因為這兩個品種在這 4 個變數上的相似性很高，所以可能需要尋找另外的辨識變數。

藉由比較三組實驗之 ROC 曲線，如圖 4.21 所示，發現不論閾值設定在多少，實驗 4-1 的 ROC 曲線均比實驗 4-2 與實驗 4-3 來得高，這意味著實驗 4-1 所建構之模式比實驗 4-2 與實驗 4-3 佳。此現象與觀察訊號雜音比的增益效果是一致的，研究發現實驗 4-1 的結果優於實驗 4-2 及實驗 4-3。繼續我們觀察三組實驗的相關矩陣，發現實驗 4-2 及 4-3 的相關矩陣中，變數間有高度相關的情況比較多，根據田口的說法，當馬氏空間中變數間具有高度共線性時，MTS 之表現會較差(Taguchi, 2002)，在此我們也觀察出這種現象。此外，從圖 4.8 中觀察 Versicolor 與 Virginica 的相似度較 Setosa 高，就分類問題而言，是否要以相似度較低的品種當馬氏空間其所形成量表會較具辨識的能力，將是可討論的議題。

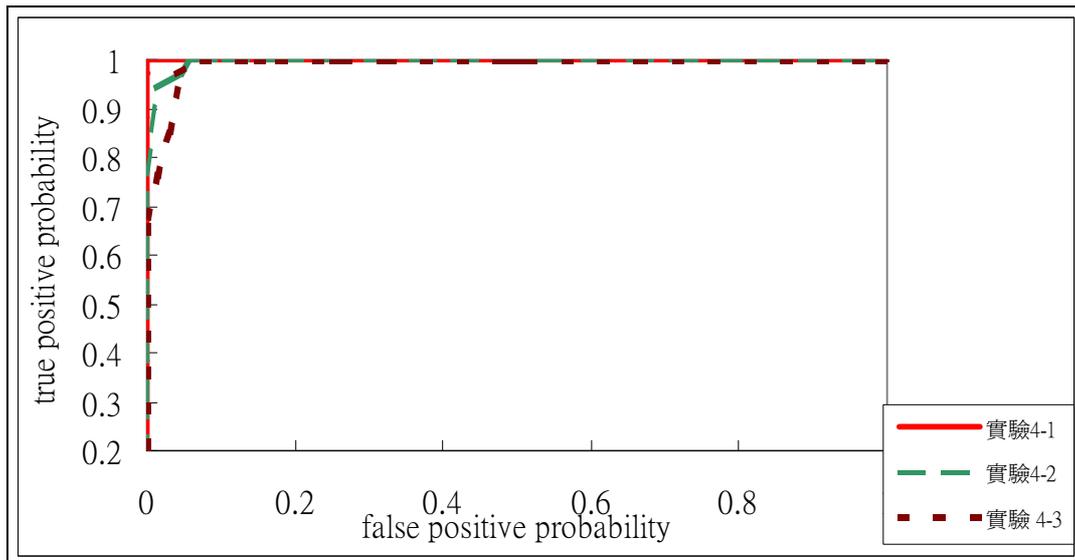


圖 4.21 三組實驗 ROC 曲線之比較

為瞭解馬氏空間的選擇對分類績效之影響是否有統計上的顯著差異，將每一種紫羅蘭均分割成 5 等分，每一等分都有 10 個樣本，利用其中 4 組樣本組合成訓練樣本，留下 1 組樣本成為測試樣本，共有 $c_4^5 = 5$ 種狀況，每一種狀況均重複上述之實驗，其所獲得測試樣本的分類正確率，如表 4.13 所示，利用無母數之變異數分析 Kruskal-Wallis 檢定進行以下之檢定。

首先，建立以下之虛無假設：

H_0 ：不同馬氏空間所建構之量表對分類績效沒有影響。(H0： $\mu_{Setosa} = \mu_{Versicolor} = \mu_{Virginica}$)

H_1 ：不同馬氏空間所建構之量表對分類的績效有影響。(H1： μ_i 不完全相等)

將所獲得的資料進行無母數之變異數分析，結果發現卡方值為 9.177，其 p 值為 0.010 達顯著水準，故判定在全因子模式中馬氏空間的選擇會影響分類的績效，在本例中以 Setosa 來建立馬氏空間最為適當。

除觀察全因子模式外，也觀察建構重要因子模式時馬氏空間之影響，共進行 120 次實驗，結果如表 4.14 所示。表中的英文代碼代表重要因子，括號的數值則代表正確率。經無母數變異數分析結果發現，卡方值為 11.152，其 p 值為 0.04 達顯著水準，故判定在重要因子模式中馬氏空間的選擇也會影響分類的績效。

表 4.13 三組實驗之全因子模式測試樣本正確率(Threshold=4)

MS	1	2	3	4	5
Setosa	100%	93%	100%	100%	100%
Versicolor	90%	93%	83.3%	93.3%	90%
Virginica	83.3%	73%	93.3%	80%	83%

表 4.14 三組實驗之重要因子模式測試樣本正確率(Threshold=4)

MS	1	2	3	4	5
Setosa	C、D(100%)	A、C、D(100%)	A、C、D(100%)	A、C、D(100%)	B、C、D(100%)
Versicolor	A、C、D(86.7%)	A、C、D(93.3%)	B、C、D(90%)	A、C、D(90%)	A、C、D(96.7%)
Virginica	A、C、D(90%)	A、C、D(90%)	A、C、D(86.7%)	A、C、D(86.7%)	A、C、D(86.7%)

4.3 抽樣策略之探討

4.3.1 實驗說明與過程

本節觀察不同的抽樣策略對 MTS 分類績效之影響。本節所謂的「抽樣策略」所指的是「抽樣比率」，許多分類技術之績效會受到訓練與測試樣本比率之影響，故本節觀察訓練樣本量的變化對 MTS 分類績效之影響。由於 4.2 節中觀察以 Setosa 為馬氏空間的分類績效較好，故本節之實驗均以 Setosa 為馬氏空間，分別進行三次實驗，其訓練與測試樣本之比率分別為 1：1、2：1、4：1。透過調整訓練樣本與測試樣本之比例以觀察其結果，過程如圖 4.22 所示。

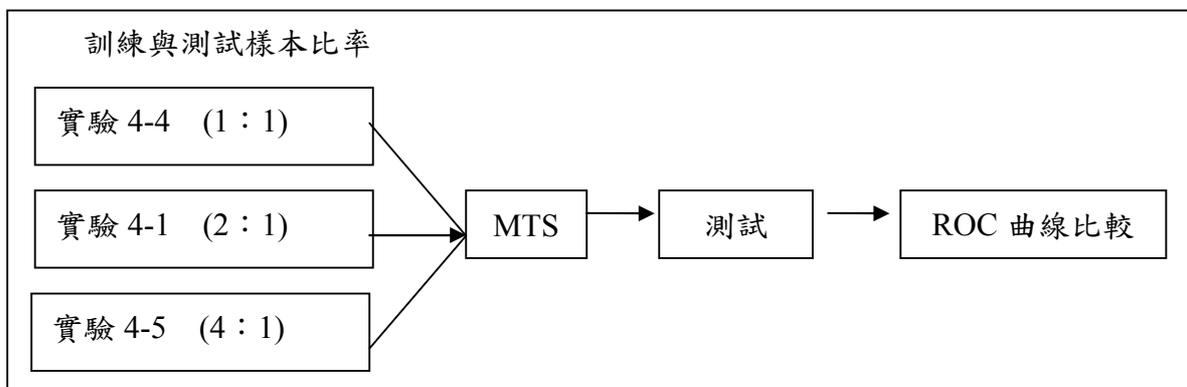


圖 4.22 抽樣策略之實驗規劃

<實驗 4-4>訓練與測試比為 1 : 1

本實驗以 Setosa 的 25 個觀察值為馬氏空間，獲得表 4.15 之相關資料。所獲得 SMD 之最小值為 0.106，最大值為 2.395，其平均值為 1.0004，符合 MTS 理論之要求。以 Versicolor, Virginica 為異常觀察值，得 Versicolor 之最小值為 36.698，最大值為 120.66，Virginica 之最小值為 118.321，最大值為 232.194，顯示這是一個很好的量表，可以很容易區別出 Setosa 與 Versicolor, Virginica 之差別，如圖 4.23 所示。利用 Setosa、Versicolor、Virginica 的測試樣本所獲得之 SMD 之範圍分別為 0.082~5.446；48.533~126.173；122.614~270.974，如圖 4.24 所示。這個結果與實驗 4-1 很類似，圖 4.24 顯示 Versicolor 與 Virginica 有重疊現象，故當次閾值訂於 121 時，在訓練樣本中有 6%之誤判，在測試樣本中則有 4%之誤判。當閾值為 3~30 時，訓練樣本之正確率為 100%，其中當閾值為 3~5.5 時，其測試樣本之正確率為 96%，當閾值為 5.6~30 時，其測試樣本之正確率為 100%，資料如表 4.16 所示。圖 4.25 為實驗 4-4 之 ROC 曲線。

接著篩選重要變數。所獲得的重要因子為 B(花萼寬)、C(花瓣長)及 D(花瓣寬)，表示這三個因子為區分 Setosa 與其他兩種品種的重要因子，如因子效果圖 4.26 所示。

最後，利用因子 B、C、D 所形成之新量表進行分類，結果發現當閾值為 3~45 時，訓練樣本之正確率為 100%。在閾值為 3、4、5、6 及 6.9 時，其測試樣本之正確率分別為 92%、97.3%、97.3%、97.3%及 98.67%，閾值在 7~45 時其測試樣本之正確率則為 100%，資料如表 4.17 所示。比較原系統(全因子模式)與新系統(重要因子模式)的訊號雜音比及正確率，如表 4.18 所示。

表 4.15 實驗 4-4 Setaso 為馬氏空間之相關資料

Correlation Matrix	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)	Inverse matrices	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)
A(花萼長)	1	0.666*	0.446	0.373	A(花萼長)	2.17	-1.258	-0.503	-0.288
B(花萼寬)	0.666*	1	0.253	0.212	B(花萼寬)	-1.258	1.808	0.08	0.045
C(花瓣長)	0.446	0.253	1	0.506	C(花瓣長)	-0.503	0.08	1.503	-0.59
D(花瓣寬)	0.373	0.212	0.506	1	D(花瓣寬)	-0.288	0.045	-0.59	1.397
(表中*者代表高度相關)									
平均數	49.52	34	14.4	2.2	標準差	3.523	3.578	1.855	0.894

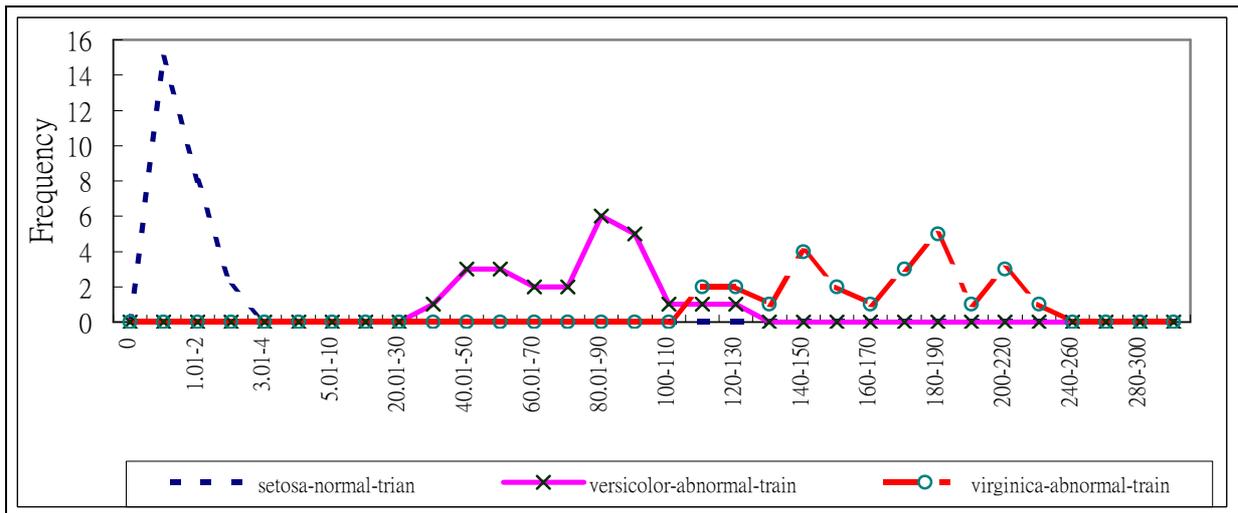


圖 4.23 實驗 4-4 全因子模式訓練樣本馬氏距離分配圖

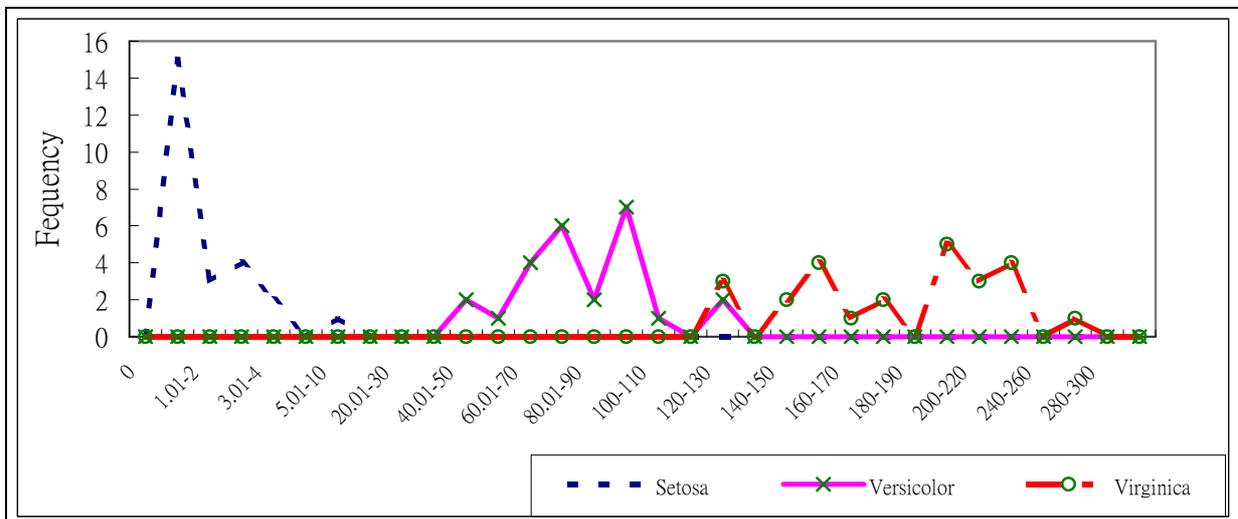


圖 4.24 實驗 4-4 全因子模式測試樣本馬氏距離分配圖

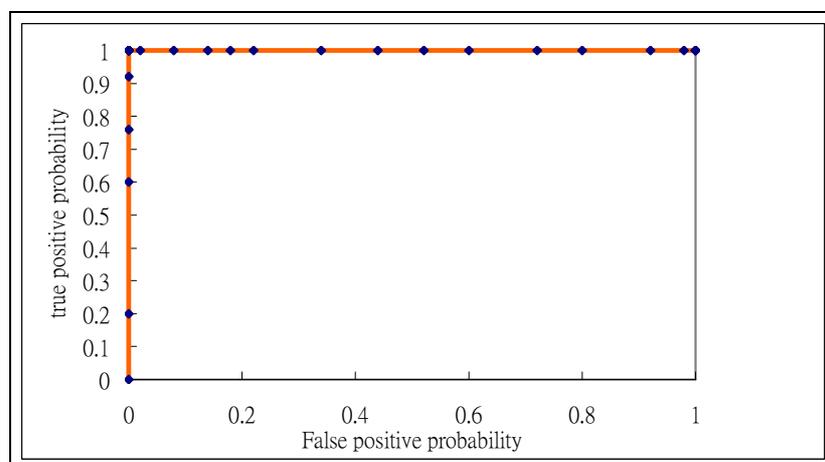


圖 4.25 實驗 4-4 全因子模式之 ROC 曲線

表 4.16 實驗 4-4 全因子模式訓練及測試樣本之分類正確率

閾值	實驗 4-4 全因子分類正確率			
	3~5.5	訓練	預測 \ 實際	Setosa(25)
Setosa			100%(25)	0%(0)
其他品種			0%(0)	100%(50)
測試		預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	96%(24)	0%(0)
		其他品種	4%(1)	100%(50)
閾值 5.6~30	訓練	預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	100%(25)	0%(0)
		其他品種	0%(0)	100%(50)
	測試	預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	100%(25)	0%(0)
		其他品種	0%(0)	100%(50)

表 4.17 實驗 4-4 重要因子訓練及測試樣本之分類正確率

閾值	實驗 4-4 重要因子分類正確率			
	4~6	訓練	預測 \ 實際	Setosa (25)
Setosa			100%(25)	0%(0)
其他品種			0%(0)	100%(50)
測試		預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	96%(23)	0%(0)
		其他品種	4%(2)	100%(50)
閾值 7~45	訓練	預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	100%(25)	0%(0)
		其他品種	0%(0)	100%(50)
	測試	預測 \ 實際	Setosa (25)	其他品種(50)
		Setosa	100%(25)	0%(0)
		其他品種	0%(0)	100%(50)

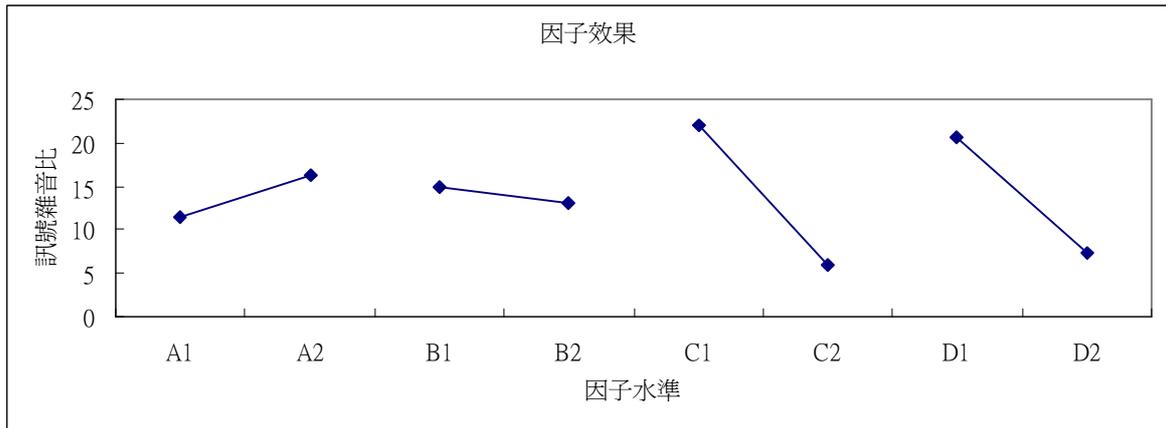


圖 4.26 實驗 4-4 因子效果圖

表 4.18 實驗 4-4 全因子及重要因子模式之訊號雜音比及正確率比較表

	訊號雜音比	總正確率 (Threshold=3)	總正確率 (Threshold=3~5.5)	總正確率 (Threshold=5.5~7)	總正確率 (Threshold=7~30)
原系統 (4 個變數)	20.4	96%	96%	100%	100%
新系統 (3 個變數)	20.7	92%	97.3%	98.67%	100%
增益	0.3				

< 實驗 4-1>訓練與測試比為 2 : 1，已於 4.2 節中執行，不再重述。

< 實驗 4-5>訓練與測試比為 4 : 1

本實驗以 Setosa 的 40 個觀察值為馬氏空間，獲得表 4.19 之相關資料。所獲得之最小值 SMD 為 0.123，最大值為 2.719，平均值為 0.9999，符合 MTS 理論所要求。圖 4.27 為實驗 4-5 之訓練樣本之馬氏距離分配圖。以 Versicolor, Virginica 為「異常」觀察值，可得 Versicolor 之最小值為 30.89，最大值為 123.587，Virginica 之最小值為 103.654，最大值為 247.074，並利用此量表進行測試，利用 Setosa 與 Versicolor, Virginica 的測試樣本所獲得之 SMD 之範圍分別為 0.089~3.335；50.147~99.112；135.946 ~276.478，如圖 4.28 所示。當閾值為 3~30 時，其訓練樣本之正確率為 100%。當閾值為 3 時，測試樣本之正確率為 96.66%，當閾值為 4~30 時，測試樣本之正確率為 100%，資料如表 4.20 所示。利用直交表與訊號雜音比篩選重要變數，所獲得的重要因子為 B(花萼寬)、C(花瓣

長)及 D(花瓣寬)，如因子效果圖 4.29 所示。最後，利用因子 B、C、D 所形成之新量表進行分類，發現當閾值為 3.5~40 時，訓練樣本之正確率為 100%，測試樣本之正確率也為 100%，資料如表 4.21 所示。我們比較原系統與新系統的訊號雜音比及正確率，如表 4.22 所示，不論全因子模式或是重要因子模式，在閾值設於 4~40 時都有 100% 之分類正確率。

觀察訓練樣本與測試樣本比率由 1：1 改變到 4：1，若將在閾值為 4 時，全因子模式判定的正確率分別為 98.66%、97.91%、100%，而重要因子模式判定的正確率分別為 97.91%、97.91%、100%。

表 4.19 實驗 4-5 以 Setoso 為馬氏空間之相關資料

Correlation Matrix	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)	Inverse matrices	A (花萼長)	B (花萼寬)	C (花瓣長)	D (花瓣寬)
A(花萼長)	1	0.696*	0.382	0.284	A(花萼長)	2.15	-1.357	-0.382	-0.21
B(花萼寬)	0.696*	1	0.264	0.187	B(花萼寬)	-1.357	1.941	-0.003	0.024
C(花瓣長)	0.382	0.264	1	0.387	C(花瓣長)	-0.382	-0.003	1.299	-0.394
D(花瓣寬)	0.284	0.187	0.387	1	D(花瓣寬)	-0.21	0.024	-0.394	1.207
(表中*者代表高度相關)									
Mean Value	49.75	33.95	14.625	2.475	Standard Deviation	3.284	3.701	1.798	1.118

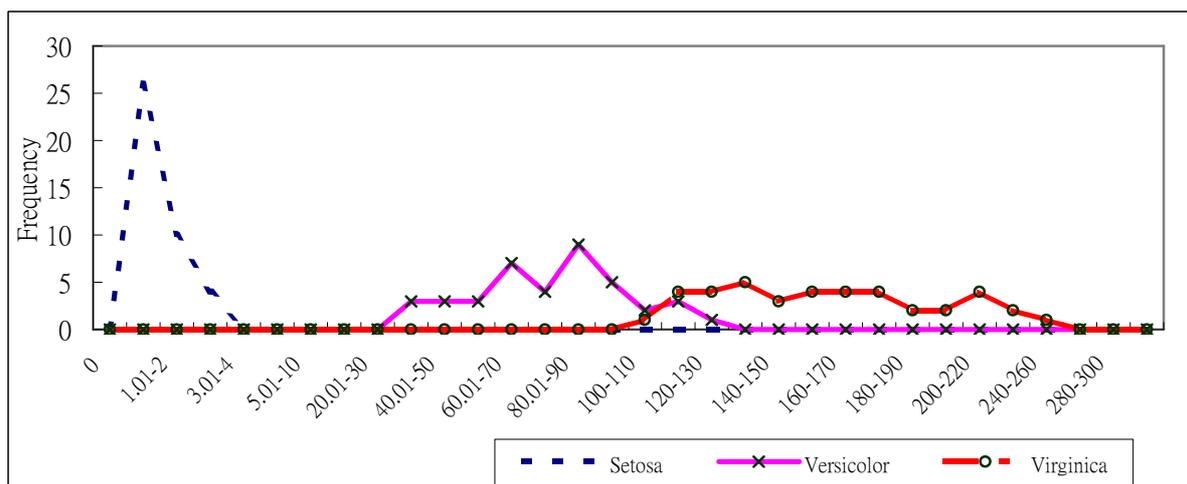


圖 4.27 實驗 4-5 全因子模式訓練樣本馬氏距離分配圖

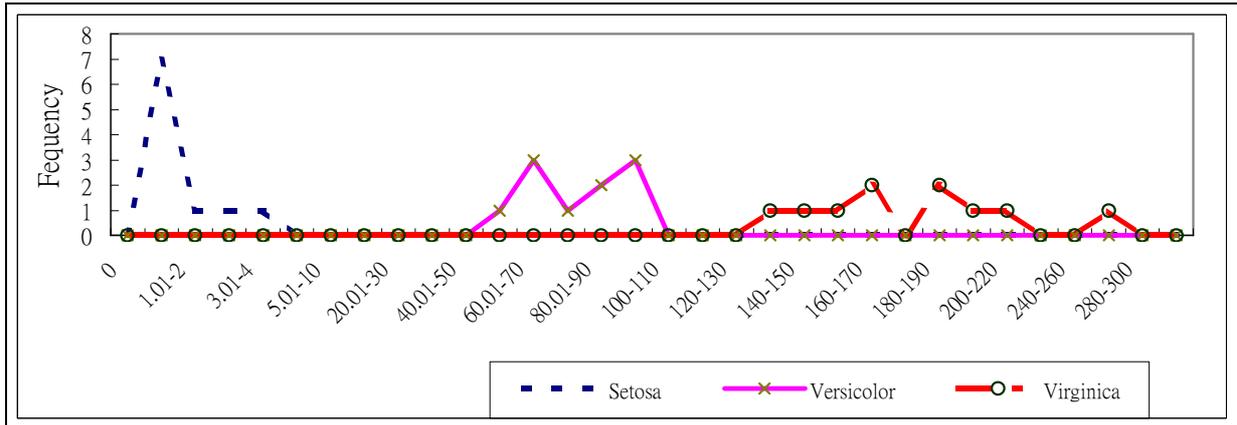


圖 4.28 實驗 4-5 全因子模式測試樣本馬氏距離分配圖

表 4.20 實驗 4-5 全因子模式訓練及測試樣本之分類正確率

閾值	實驗 4-5 全因子模式分類正確率			
	3	訓練	預測 \ 實際	Setosa(40)
Setosa			100%(40)	0%(0)
其他品種			0%(0)	100%(80)
測試		預測 \ 實際	Setosa (10)	其他品種(20)
		Setosa	100%(9)	0%(1)
		其他品種	0%(0)	100%(20)
閾值 4~40	訓練	預測 \ 實際	Setosa (40)	其他品種(80)
		Setosa	100%(40)	0%(0)
		其他品種	0%(0)	100%(80)
	測試	預測 \ 實際	Setosa (10)	其他品種(20)
		Setosa	100%(10)	0%(0)
		其他品種	0%(0)	100%(20)

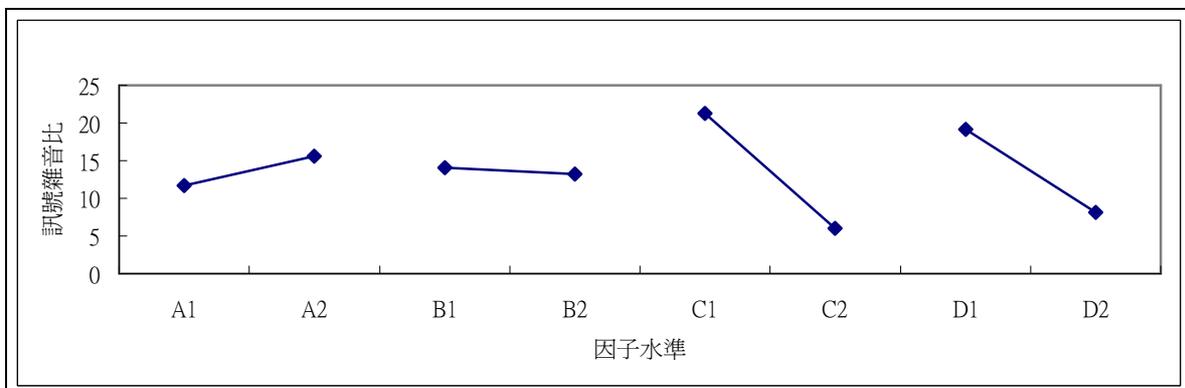


圖 4.29 實驗 4-5 因子效果圖

表 4.21 實驗 4-5 重要因子模式訓練及測試樣本之分類正確率

閾值	實驗 4-5 重要因子分類正確率			
	3.5~40	訓練	預測 \ 實際	Setosa (40)
Setosa			100%(40)	0%(0)
其他品種			0%(0)	100%(80)
測試		預測 \ 實際	Setosa (10)	其他品種(20)
		Setosa	100%(10)	0%(0)
		其他品種	0%(0)	100%(20)

表 4.22 實驗 4-5 全因子及重要因子模式之訊號雜音比及正確率比較表

	SN	總正確率 (Threshold=3)	總正確率 (Threshold=4~40)
原系統 (4 個變數)	19.3	96.66%	100%
新系統 (3 個變數)	20.5	100%	100%
增益	0.8		

4.3.2 結論

觀察以上的現象，發現訓練及測試樣本比的選擇對分類績效會有影響，為瞭解是否有統計上的顯著差異，我們進行幾組實驗，將每一種紫羅蘭均分割成 n 組樣本，利用其中 n-1 組樣本組合成訓練樣本，留下 1 組樣本成為測試樣本，分別進行幾組樣本比例的實驗，分別是 1:1, 2:1, 4:1，每一種狀況均重複上述之實驗，可獲得每一組實驗之測試樣本的分類正確率，如表 4.23 所示。利用無母數之變異數分析 Kruskal-Wallis 檢定以下之虛無假設：

H_0 : 不同抽樣策略對分類績效沒有影響。($H_0: \mu_{1:1} = \mu_{2:1} = \mu_{4:1}$)

H_1 : 不同抽樣策略對分類的績效有影響。($H_1: \mu_i$ 不完全相等)

分析結果發現，其卡方值為 1.470，其 p 值為 0.480 未達顯著水準，故判定在全因子模式中訓練與測試樣本比不影響分類的績效。由上述三組實驗的觀察，發現 MTS 的分類績效並未顯著的受到訓練與測試樣本比率之影響。

表 4.23 抽樣策略之全因子測試樣本正確率比較表

訓練測試樣本比 \ 次數	1	2	3	4	5
1 : 1	98.67%	98.67%			
2 : 1	100%	100%	97.91%		
4 : 1	100%	93%	100%	100%	100%

4.4 訊號雜音比之影響

4.4.1 實驗說明與過程

本節主要觀察不同的訊號雜音比的對分類績效之影響。本節共執行兩種實驗，均以 Setosa 為馬氏空間，實驗 4-1 採用望大特性，實驗 4-6 利用動態特性。

<實驗 4-1>採用望大特性之訊號雜音比，已於 4.2 節中執行，不再重述。

<實驗 4-6>採用動態特性之訊號雜音比

本實驗有關馬氏空間之建構結果與實驗 4-1 相同，接著為利用直交表與訊號雜音比篩選重要變數，實驗配置如表 4.6 所示。在本例中，由於無法獲得訊號因子之真值，所以利用 Working Average 的方式來進行。在此根據兩種不同品種，將資料分成兩組，每一組各選擇 10 個觀察值，令 $y_{ij} = \sqrt{MD_{ij}}$ ， $i=1,2$ ， M_i 為第 i 組之 Working Average，經由計算獲得 $M_1 = 9.169$ ， $M_2 = 13.488$ ，假設 $y_i = \beta M_i$ ，根據動態 $\eta = 10 * \log \frac{\beta^2}{\sigma_e^2}$ 可計算獲得重要因子，如圖 4.30 所示。

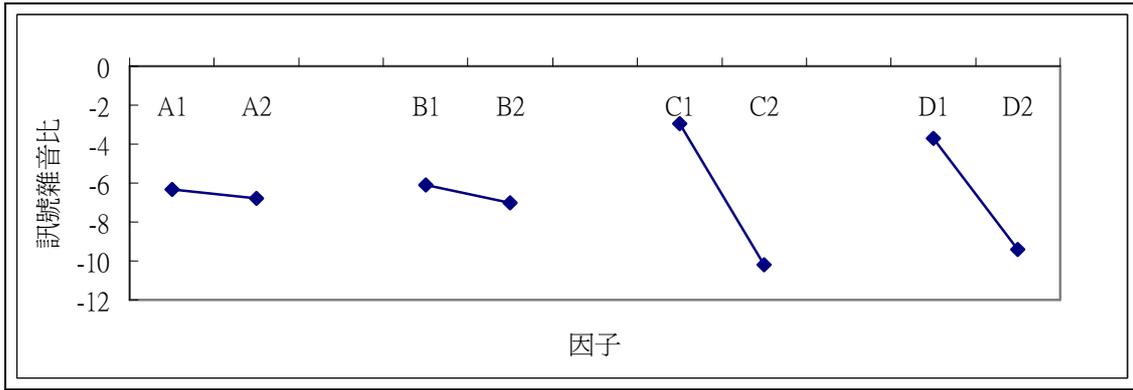


圖 4.30 實驗 4-6 之因子效果圖

本實驗訊號雜音比之 ANOVA 分析列於表 4.24 中，其中 $S_{\beta} = 190.56307$ ， $S_e = 120.28$ ， $\hat{\beta} = 1.672$ 。如果 $y_i = \beta M_i$ 成立，則可利用已知的 SMD 來預測訊號因子 $M = y / \hat{\beta} \pm 3(\sqrt{MSE} / \hat{\beta})$ ，這種利用配適的迴歸方程式中已知的因變數 (Dependent Variable) 反預測一個未知的自變數 (Independent Variable) 的過程，在統計學上稱為「口徑測定」 (Calibration)，或稱為「逆預測」 (Inverse Prediction)，而逆預測的優點是可利用觀察值建立一種便宜、快速但較為粗略的測定 y 來觀察另一種精確但可能昂貴耗時的 x 之間的關係。觀察實驗所篩選出的重要因子 C、D 所形成的新的量表的訊號雜音比及正確率，如表 4.25 所示，發現新的量表的表現並沒有比較好。

表 4.24 實驗 4-6 之 ANOVA 分析表

ANOVA	自由度	SS	MS	F	顯著值
迴歸	1	190.56307	190.5631	30.10144	3.28E-05
殘差	19	120.283241	6.330697		
總和	20	310.846311			

表 4.25 實驗 4-6 全因子與重要因子模式之訊號雜音比及正確率之比較表

	SN	總正確率 (Threshold=3~5)	總正確率
全因子(4 個變數)	-1.9273	93.75% ~ 97.91%	100%(Threshold=5~42)
重要因子(2 個變數)	-5.573	93.75%	100%(Threshold=6~54)
增益	-3.6457		

4.4.2 結論

表 4.26 為實驗 4-1 及 4-6 之比較表。本實驗之結果發現使用望大特性，不論在訊號雜音比的增益或是分類正確率上都比較好，這與田口博士的認知不同。在實驗 4-6 中是採用 Working Average 的方式來處理，是否因為所採用之 Working Average 不適當，才導致分類結果並不佳，值得再討論。

表 4.26 兩組實驗之訊號雜音比的增益及正確率比較

實驗	訊號雜音比	最佳化前的訊號雜音比	最佳化後的訊號雜音比	增益	測試樣本正確率 (重要因子) Threshold=4
實驗 4-1	望大特性	20.4	22.3	1.9	97.91%
實驗 4-6	動態特性	-1.9273	-5.573	-3.65	93.75%

4.5 閾值的討論

在多變量的系統中，要有效率的進行診斷，閾值的決定是很重要的。在 MTS 執行的步驟 2 中，提及可以利用異常觀察值之 SMD 來判定量表好壞，通常，異常觀察值的 SMD 不太可能全都大於正常觀察值的 SMD。田口建議利用品質損失函數選擇閾值，但實務上，這種閾值的決定方法似乎相當複雜，因為成本的評估具不確定性(Bovas,2003)。

在 MTS 中，田口博士不願用機率來決定閾值的原因，是因為他認為異常的觀察值都具有獨特性。田口(2002)引用了俄國作家托爾斯泰(Tolstory)在安娜卡列寧娜(Anna Karenina)的一句話「*All happy families look alike. Every unhappy family is unhappy after its own fasion*」，來強調異常的觀察值不會形成另一組群體的思維。但是將 MTS 應用在紫羅蘭的案例時，每一種紫羅蘭都由其共同的特性，所以在這種案例中，應該可以採用機率的概念來處理閾值，本節嘗試兩種做法並觀察其結果，其一為利用 Hotelling T^2 管制圖的概念來訂定閾值，第二是採用最小化型 I 及型 II 錯誤的方式來決定閾值。

第一個方法利用 Hotelling T^2 的管制界限來當成閾值以辨識異常值。Hawkins(2003)認為利用 SMD 來進行診斷的方法，本質上與利用正常群體來建立

Hotelling T^2 管制圖沒有差別。Hotelling $T^2 = (\bar{X} - \bar{\bar{X}})^T S^{-1} (\bar{X} - \bar{\bar{X}})$ 管制圖的管制上限為 $UCL = \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha, p, m-p}$ ，管制下限 $LCL=0$ 。SMD 與 $T^2 = (\bar{X} - \bar{\bar{X}})^T S^{-1} (\bar{X} - \bar{\bar{X}})$ 之差別除了變數標準化外，還多除以變數個數 p ，我們觀察 UCL 公式中 p, m, α 與 UCL 之關係， P 代表變數的個數， m 代表樣本大小， α 代表臨界值。以紫羅蘭為例，當辨識變數為 4 個，各種樣本量分別為 $m=25, 34, 40, 50$ 時，在型 I 錯誤的機率在 0.05 時，其管制的上限分別為 13.50263, 12.18245, 11.69728, 11.18697，轉換成馬氏空間的閾值分別為 3.375657, 3.045612, 2.92432, 2.796743。觀察前述實驗的結果，利用這些閾值進行判定，結果如表 4.27 所示。以 Hotelling T^2 所建立之閾值觀察各組實驗之正確率，如表 4.28 所示。

表 4.27 Hotelling T^2 與 MTS 閾值之轉換表

P	4	4	4	4
M	25	34	40	50
α	0.05	0.05	0.05	0.05
F	2.840096	2.689632	2.633534	2.574033
係數	4.754286	4.529412	4.441667	4.346087
UCL	13.50263	12.18245	11.69728	11.18697
MD 閾值	3.375657	3.045612	2.92432	2.796743

表 4.28 以 Hotelling T^2 所建立之閾值與正確率

實驗 4-1	馬氏空間樣本數=34	閾值=3.045	訓練樣本正確率=100%	測試樣本正確=93.75%
實驗 4-4	馬氏空間樣本數=25	閾值=3.375	訓練樣本正確率=100%	測試樣本正確=98.67%
實驗 4-5	馬氏空間樣本數=25	閾值=2.9243	訓練樣本正確率=96.67%	測試樣本正確=96.67%

以上的閾值決定的方法是根據 Hotelling T^2 的概念來決定的，利用 Hotelling T^2 來決定閾值時，其型 I 錯誤之機率 α 均為 0.05，當樣本量愈大，此閾值會愈小。

第二種方法，可採用最小化型總錯誤率的方式，針對分類問題，特別是分類的類別超過兩組以上時，除應建立閾值以區分目標群體與其他群體外，還應建立次閾值以協助

其他類別之間的區分，在訓練樣本中，透過最小化型 I 與型 II 錯誤之總合決定對應的閾值及次閾值，並以此閾值與次閾值決定測試樣本分類正確率。

4.6 MTS 與其他方法之比較

本節將 MTS 與其他傳統的方法進行比較，包括鑑別分析與決策樹。首先，將每一種紫羅蘭均分割成 5 等分，每一等分都有 10 個樣本，利用其中 4 組樣本組合成訓練樣本，留下 1 組樣本成為測試樣本，共有 $c_4^5 = 5$ 種狀況。在本研究中使用 See5 決策樹工具進行分類，它依據 C4.5 進行改良，適用於連續型資料及離散型資料。並利用鑑別分析處理相同的 5 筆資料，分類之結果如表 4.29 所示。利用無母數之變異數分析 Kruskal-Wallis 檢定決策樹、鑑別分析及 MTS 之分類績效。

建立以下之虛無假設：

H_0 : 不同的分類技術對分類績效沒有影響。 $(H_0: \mu_{DT} = \mu_{MTS} = \mu_{DA})$

H_1 : 不同的分類技術對分類的績效有影響。 $(H_1: \mu_i$ 不完全相等)

分析表 4.29 之資料發現，卡方值為 0.994，其 p 值為 0.608 未達顯著水準，故判定這三種技術在紫羅蘭分類問題上沒有顯著差異。

表 4.29 三種分類技術之分類結果比較表

方法 \ 實驗次數	1	2	3	4	5	平均值
MTS	100%	93.3%	100%	96.7%	100%	98%
決策樹	97.6%	83.3%	100%	96.7%	96.7%	94.86%
鑑別分析	100%	96.7%	93.3%	100%	100%	98%

實驗結果發現，MTS 的分類績效不輸傳統的鑑別分析與決策樹。本研究之結果認為，針對分類的問題，MTS 是一個值得推薦的新方法，MTS 除了可獲得不錯的分類績效之外，還有以下幾個優點：(1) 可以提供異常觀察異的嚴重程度；(2) 有能力檢查所選擇的辨識變數是否恰當；(3) 具穩健性；(4) 若有適當的訊號因子時，可以提供逆預測的功能。當然，MTS 也有些缺點：(1) 馬氏空間的選擇並不容易；(2) 針對多個類別的分類

問題時要提供多個適當的次閾值以利決策，例如分類的對象有 3 類時就必須有 2 個閾值，若分類的對象有 N 類時就必須有 N-1 個閾值。



第五章 MTS 在信用評分之應用

MTS 是針對多變量系統之綜合評價提出的一套方法，而信用評分的特性就是一個綜合評價的問題，所以本研究提出以 MTS 建構信用評分模式。本章共分成 5 節，5.1 節探討信用評分與國內目前常用的信用評分方式，5.2 節說明 MTS 信用評分模式之建構，5.3 節為 MTS 信用評分模式之實證研究，5.4 節為 MTS 之模式修正，5.5 節則為各種信用評分模式之比較。

5.1 信用評分及現況

要降低信用風險且不因冗長之徵信程序而流失商機，授信機構必須要準確且快速的完成信用的評估，而一般較為常用的信用評估方式就是「信用評分」。信用評分在商業活動中扮演相當重要的機制，透過信用評分可協助進行風險的控管。本節針對信用評分之模式及其評估之準則進行說明，並簡述國內目前常用的信用評分方式。

5.1.1 信用評分及相關研究

所謂「信用」是指債務人違約的可能性以及違約造成債權人損失的大小。在整個授信的過程中，授信機構必須評估申請者的風險水準並決定是否授信，授信之後必須監控其信用狀態。無論是篩選新客戶或是監控信用狀況，都希望能及早偵測出其違約的機率以進行適當之決策，此過程稱為風險評估(Risk Assessment)。而信用評分(Credit Scoring)是風險評估的一種，信用評分利用客戶信用的歷史資料評估客戶信用之風險，並根據此風險狀況將客戶分成「信用正常」及「信用異常」兩種。本質上，信用評分是一種「分類決策」(Thomas, 2000)。可分成(1)申請前的階段；(2)信用申請階段；(3)信用績效階段，如圖 5.1 所示。

早在 1960 年代後期，銀行及授信業者已察覺到信用評分的必要性，由於信用卡申請人數的遽增，若不採用自動授權決策系統則會造成人力及經濟上的負荷。在 1980 年代已成功地將信用評分應用在信用卡上，意味著銀行經開始利用信用評分技術在個人信

貸上。有關信貸的決策，信用評分不但能處理大量的交易資料，且其準確度比人為主觀的評估更高(Rosenberg, *et al.*, 1994；Hand, 1997)。

目前已有許多技術被應用到信用評分模式上，如：線性鑑別分析、羅吉斯迴歸、決策樹、類神經網路、線性規劃、馬可夫鏈、最鄰近法、基因演算法等(Thomas, 2000)。其中線性鑑別分析是最早應用在信用評分的統計技術，之後的相關研究也常以鑑別分析做為比較的標竿。由於鑑別分析的使用上有些假設條件的限制，雖然有學者認為違反常態的假設對線性鑑別分析的績效影響不大(Reichert, 1983)，但仍有不少研究針對此缺點進行改善。Henley(1995)利用羅吉斯迴歸處理信用評分問題，Tam(1992)與Henley等(1996)則利用最鄰近法來評估消費者的信用風險。Makowski(1985)，Carter(1987)與 Davis 等(1992)則將決策樹應用在信用卡上，而 Coffman(1986)則將決策樹與鑑別分析進行比較。Desai 等(1997)將基因演算法應用在信用機構中並將結果與類神經網路進行比較，West(2000)則利用五種類神經網路方法來建立信用評分模式，包括 MLP(Multi-Layer Perceptron)、MOE(Mixture-of-Experts)、RBF(Radial Basis Function)、LVQ(Learning Vector Quantization)及 FAR(Fuzzy Adaptive Resonance)等，並將結果與傳統的鑑別分析及羅吉斯迴歸、最鄰近法、核密度估計(Kernel Density Estimation)及決策樹進行比較，結果發現 MOE、RBF 是可行的信用評分模式，而羅吉斯迴歸則是傳統方法中最準確的。除上述的方式外，目前也有學者結合不同的演算法來改善信用模式的正確性，Chiu 等(2001)利用結合類神經網路及 CART 來改善信用評分的正確率。

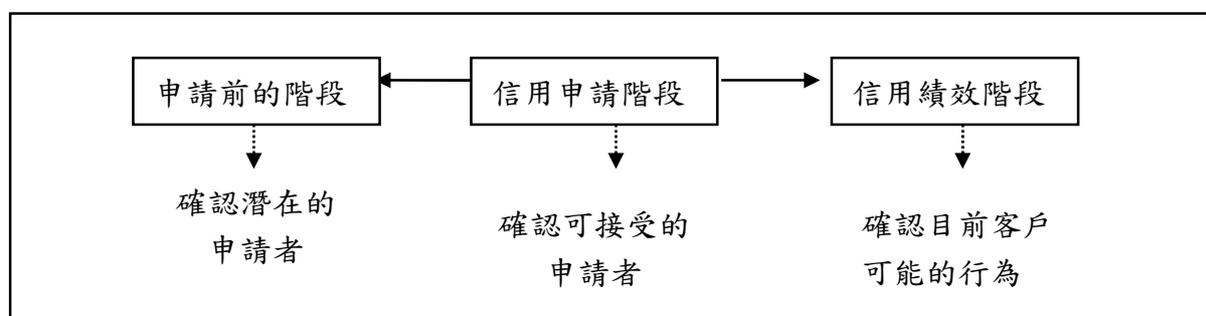


圖 5.1 信用評分模式之階段

5.1.2 評估信用評分模式的準則

有關模式的評估，包括兩種方式：混淆矩陣(Confusion Matrix)及兩種錯誤率之互償法。

一、混淆矩陣：

混淆矩陣提供真實類別與預測類別之矩陣表，如表 5.1 所示。對於信用評分問題而言，一般會考慮兩種錯誤：(1)型 I 錯誤代表「信用風險」，它是將「信用異常」的客戶判成「信用正常」的錯誤，如果犯型 I 錯誤之機率 α 很高時，則表示授信政策太寬鬆，這將使得信用風險擴張；(2)型 II 錯誤代表「商業風險」，它是將「信用正常」的客戶判成「信用異常」的錯誤，如果犯型 II 錯誤之機率 β 很高時，則表示授信政策過度嚴苛，將容易喪失市場佔有率。通常評估及比較分類模式的優劣時經常採用總錯誤率。但因為兩種風險的重要性不盡相同，所以用總錯誤率進行評估有時較不適合，而有關錯誤率之估計，將於後續討論。

表 5.1 混淆矩陣

預測類別	實際類別		實際類別	
	數量	百分比	數量	百分比
信用異常	A	C	B	$C/(C+D)=\beta$
信用正常	B	D	A	$B/(A+B)=\alpha$

二、兩種錯誤的互償：

兩種錯誤的互償，包括(1) ROC(Receiver Operating Characteristic)曲線；(2)成本函數；(3)閾值的決定等。

所謂的ROC曲線，是一種衡量分類績效的指標，它由一組 $(\beta, 1-\alpha)$ 所形成的軌跡，由於分類模式之型 I 及型 II 錯誤機率會隨著改變閾值的變動而有所改變，ROC曲線用以表現當閾值改變時，兩種錯誤的互償。使用ROC 分析主要的目的，是藉此分析所得的資訊做為決策者參考的依據(DeLeo, 2001)。ROC曲線愈高，則代表其分類之正確率愈高，當比較分類模式時會出現模式之ROC曲線相交的現象，此時也可用另一個指標

AUC(Area Under the Curve)來評估，AUC代表在ROC曲線下的面積，可以利用這個值的大小來比較分類模式的績效，AUC的值愈大，代表其分類績效愈好。

其次是成本函數，當使用分類模式進行分類時，最佳的模式是最小化期望的損失。在兩種分類的狀況下，其成本的函數為：期望成本= $P_b * \text{cost}(I) * \alpha + P_g * \text{cost}(II) * \beta$ ，其中 P_b 、 P_g 代表母體中信用「壞」「好」的比率， $\text{cost}(I)$ 、 $\text{cost}(II)$ 代表兩種錯誤的成本。不過這個成本函數隱藏著一個假設是每一個申請者的風險都相同，但實務上並非如此。實務上，分類錯誤成本的計算是相當複雜的(Liu, 2002)。

最後是閾值的決定，這是一個重要的問題，如何選擇最適合的閾值，Liu(2002)建議利用斜率 S 及 ROC 曲線之切點，可找出最佳之閾值範圍；其中斜率 S 為 PR/CR ，而 $PR = P_g/P_b$ ， $CR = \text{cost}(I) / \text{cost}(II)$ 。影響閾值的因素可能有總體環境、經濟景氣的改變；授信機構的風險政策、能忍受之最大風險等。

前述有關混淆矩陣中提及，通常評估及比較分類模式的優劣時經常採用總錯誤率，接下來說明有關錯誤率之估計。一般而言，模式之「真正錯誤率」(True Error Rate)是指模式實際應用時所產生的分類錯誤率，這個真正錯誤率在模式尚未應用前，是無法真正得知的。然而，實務上為評估模式的優劣，模式使用者不可能等到實際應用後才得知模式的能力如何，必須在模式使用前就估計此錯誤率。而所謂表面錯誤率(Apparent Error Rate)或稱訓練樣本錯誤率則是模式在訓練樣本中得到的分類錯誤率，然而，訓練樣本錯誤率並非一個好的估計式，因為訓練樣本錯誤率愈小，表示分類模式愈符合訓練樣本的資料之特性。這種情形意謂模式過度的學習，以至於太符合訓練樣本之獨特性，這種現象就稱為“Over-fitting”。估計一個分類模式之真正錯誤率所採用的方法稱為保留法(Holdout Method)或是訓練與測試法(Train and Test Method)，基本的想法是利用一組訓練樣本來建模，然後利用第二組樣本來測試。這組測試樣本所獲得的錯誤率稱為測試樣本錯誤率(Test Sample Error Rate)。Weiss(1991)提及當測試樣本的樣本量足夠大($n > 1000$)時，測試樣本錯誤率是一個正確的估計式，因此，通常測試樣本錯誤率被用來估計真正錯誤率。

以上討論有關分類錯誤率的問題，然而，分類錯誤率並非選擇模式的唯一指標，模式的選擇不僅根據分類錯誤率，同時也會考慮其他的準則。其他的準則如下所述(Liu, 2002)：

(1)分類模式的速度

這是指模式計算時間的長短，選擇模式時應同時考慮建模與實際應用時所需的時間，特別是模式應用時所需的時間之長短，因為即時、迅速的決策才會吸引更多潛在的客戶(Hand & Henley, 1997)。

(2)分類模式的透明度與解釋性

由模式所產生的信用分數可提供模式使用者決策參考的依據，但模式的解釋性及合理性應該要能讓模式使用者接受，模式的重要特性之一就是可觀察出每個輸入變數對輸出變數的影響。例如線性鑑別分析所產生的模式的解釋性比類神經網路所產生的模式來得高，因為類神經網路的模式其運算過程有如黑箱，對於分類的結果無法提供何以如此之原因，尤其當授信者必須對那些被拒絕的申請者說明理由時，類神經網路模式並無法提供一個道理。



(3)分類模式的單純性

除考慮上述兩個準則外，也應考慮模式的單純性，模式應愈簡單愈好。

5.1.3 訓練及測試樣本的抽樣策略

本節討論有關抽樣策略的問題，包括母體漂移(Population Drift)、訓練與測試樣本比率及比較分類模式正確率之問題。

為確保測試樣本錯誤率是真正錯誤率的可靠估計式，訓練與測試樣本應隨機、獨立自母體中抽出。故一般的抽樣策略是從所有可用之樣本中，同時隨機抽取出訓練及測試樣本。有時，為了為獲得較好的分類績效，必須對模式的結構及某些參數進行調整，例如：決策樹的修剪、類神經網路中隱藏層數量的調整等，必須有其另一組測試樣本用以比較及選擇不同模式結構的績效。因此需要三組樣本，一為建模的樣本，另一為選擇最佳結構及參數的樣本，最後為測試樣本。為了增加最終模式的訓練樣本，有時候會將建

模樣本及調整結構之樣本合併成一個訓練樣本。但測試樣本不應該用來調整模式的結構，以確保評估品質。

以測試樣本錯誤率估計真正錯誤率的基本假設是測試樣本具代表性，足以代替未來真正的母體。對於實務上的信用評分問題，母體會隨著時間而改變。這種母體的改變，稱為母體漂移，這違反了以樣本估計錯誤率的基本假設。故以漂移的母體所建立的模式時，模式的重建問題就很重要，因為當母體漂移時，模式就變得無效。針對漂移母體的問題想要獲得較可靠的模式，Weiss 等人(1991)提出一有效的抽樣方法來處理這個問題。

一般而言，訓練樣本愈多，其分類模式之績效愈好，但當訓練樣本超過某一個數量其效果就開始遞減。另一方面，測試樣本愈多，錯誤率的估計會愈準確，對同一組樣本而言，這會產生互償的效果。為處理這種互償的效果，學者 Weiss(1991)建議訓練樣本與測試樣本的比例為 2/3 與 1/3。

設有二個分類模式 A 與 B，欲了解此二個分類模式之績效是否具有統計上顯著的差異，假設 $\Delta\hat{P} = \hat{P}_{Aerr} - \hat{P}_{Berr}$ ，其中 \hat{P}_{Aerr} 代表 A 分類模式的測試錯誤率， \hat{P}_{Berr} 代表 B 分類模式的測試錯誤率。

根據中央極限定理，可獲得統計量 $Z = \frac{(\hat{P}_{Aerr} - \hat{P}_{Berr}) - \mu_{\hat{P}_{Aerr} - \hat{P}_{Berr}}}{\sqrt{\text{var}_{\hat{P}_{Aerr} - \hat{P}_{Berr}}}}$ ，用以

檢定二個分類模式之分類錯誤率相等之虛無假設。令 $R = (a + b) / N$ ，其中 a 代表若用分類模式 A 可正確分類但以分類模式 B 無法正確分類的觀察值個數，而 b 則為可由分類模式 B 正確分類但以分類模式 A 就無法正確分類之觀察值個數。經由公式的推導可以獲得公式(5.1)。

$$\text{當 } |\Delta\hat{P}| \geq \frac{4R + 2\sqrt{R^2 - R(4 - N)}}{N - 4} \text{，則 } Z \geq 2 \quad (5.1)$$

這公式代表著兩個分類模式在顯著水準設為 1%~5%時，也就是 Z 的值約為 2 時，錯誤率之最小顯著差距為 R 與 N 之函數關係(Liu, 2002)。除了利用較大的測試樣本之外，也可利用交叉驗證(Cross Validation)或拔靴法(Bootstrap)等方式來估計真正的錯誤率。

5.1.4 國內目前常用的信用評分方式

目前國內銀行所採用之信用評分方式大致分成(1)內部評分卡；(2)JCIC 分數；(3)廠商提供評分卡；(4)顧問服務；(5)依 Data Mining 產生評分規則，其優缺點如表 5.2 所示。有關傳統信用評分卡之缺點如下：(1)評分的調整頻繁，由於外在環境變動加劇，使得評分標準亦經常調整；(2)每家銀行的客戶屬性差距極大，所謂專家型評分卡(Generic Scorecard)亦無法完全正確評量不同銀行客戶的風險；(3)評分卡的產生需仰賴廠商，銀行自主性低；(4)每張評分卡收費高昂，銀行投資高。而以 Data Mining 建立評分模式則有以下之優點：(1)以各家銀行本身的資料進行學習，最符合各別銀行之特性；(2)具解釋性；(3)可自行更新模式；(4)可以擁有自家銀行的 Know-How；(5)可擴增到行為評分或其他相關評分功能。

表 5.2 國內目前常用的信用評分方法一覽表(台灣金融研訓院)

方法	說明	優點	缺點
內部評分卡	銀行自行訂定參考分數	簡易使用	1)標準不一 2)不夠客觀 3)處理速度慢
JCIC 分數	JCIC 依照客戶資料產生參考分數	具公信力	1)未參考個別銀行自身之客戶狀況 2)每筆皆要付費 3)未考慮不同產品之信用條件
廠商提供評分卡	廠商依其建置經驗產生	廠商累積全世界經驗信賴度高	1)未考慮到各別銀行之客戶狀況 2)收費昂貴且需仰賴廠商更新
顧問服務	廠商將其實際建置經驗提供評分建議		1)黑盒子，銀行無法自行解釋 2)高度依賴廠商 3)收費極高
依 Data Mining 產生評分規則	由銀行之客戶實際繳款紀錄，學習產生專屬規則	1)考量各別銀行之實際發生狀況 2)銀行可自行調整更新 3)客觀可排除人為判斷缺失	1)成果待為驗證

5.2 MTS 信用評分模式之建構

本節提出以 MTS 建構信用評分之模式，首先說明傳統信用評分與 MTS 信用評分模式之差別，其次說明如何建構 MTS 信用評分模式。

首先說明傳統信用評分的流程與 MTS 信用評分流程之差別，以進件信用評分為例如圖 5.2 所示。流程中最主要的差別在步驟 1，傳統的信用評分模式必須要定義出何謂「信用異常」的客戶，但在 MTS 中則必須定義「信用正常」的客戶。田口博士(Taguchi *et al*,2001)認為傳統信用評分模式無法順利運作的原因，是因為傳統模式以信用異常的客戶為研究對象所造成的。由於信用異常者其異常的原因各有不同，所以若以其為研究對象，則不易找出共通性。至於如何定義「信用正常」及如何建構馬氏空間則是信用評分模式的設計者必須決定的。



圖 5.2 傳統信用評分與 MTS 信用評分之比較

本研究利用資料探勘建構信用評分模式之程序來協助 MTS 之信用評分模式的建構，程序如圖 5.3 所示。其中第一階段是問題的定義及資料的準備，此階段是最重要也

最為繁複；第二階段是模式的建立，可利用各種演算法來建立所需的模式，第三階段則是模式的應用與評估，模式的建構者必須瞭解模式的能力並建立模式修正之機制。

本研究以 MTS 建構信用評分模式，其流程如圖 5.4 所示，包括(1)定義信用類別；(2)資料前處理；(3)抽樣計劃；(4)選擇辨識信用之控制因子與雜音因子；(5)形成馬氏空間；(6)計算馬氏距離以形成量表；(7)驗證量表能力；(8)利用直交表及訊號雜音比篩選重要因子；(9)建立新量表；(10)驗證量表能力；(11)協助進行信用決策。步驟(1)~(7)所建立的模式，在本研究中稱為全因子模式(Full Model)，步驟(8)~(10)所建立之模式，則稱為重要因子模式(Reduced Model)，以下將針對前述各項步驟加以說明。

(1)信用類別的定義

首先要定義信用好壞必須經過一段時間的觀察，由於傳統的研究都以信用異常者為對象，故一般銀行對於「信用異常」有清楚的定義，通常是以連續三個月未繳款者為信用異常。但對於信用正常則無明確之定義，一般認為若不是信用異常者，則為信用正常。但在 MTS 信用評分模式中，必須清楚定義何謂「信用正常」，定義的良窳與否關係著 MTS 模式的能力展現。



(2)資料的前處理

資料的前處理對任何分類技術而言都是相當重要的，透過信用領域專家的協助，可對於辨識信用的變數更為清楚，應如何處理遺漏值、如何判斷離群值、如何判定不合理之資料編碼以及如何進行合理之資料轉換，乾淨的資料對模式的建立有相當大的幫助。

(3)抽樣的規劃

建立模式所用的樣本是否有代表性，將會影響模式的效益及適用性，故以 MTS 建構信用評分模式時仍必須考量抽樣的計劃，包括訓練樣本與測試樣本的比例；信用好壞類別的比例及抽樣的方法等。

(4)控制因子與雜音因子的選擇

對 MTS 信用評分模式而言，所謂控制因子是指用來辨識信用好壞的變數，包括顧客基本資料、聯徵資料、繳款資料及相關信用的行為變數等，而雜音因子則為會影響信用分類績效但又不易控制的因子，例如：客戶進件的通路或是核准信用的審核者等。

(5)馬氏空間的形成

利用信用正常者之資料建構馬氏空間是整個流程中最關鍵的步驟，因為信用辨識與預測的良窳有賴這群正常值的選取。

(6)信用量表的建構

選擇正常觀察值並利用所有可用之信用辨識因子建立全因子模式之量表，透過公式(3.2)計算馬氏空間中每一個觀察值之 SMD 以建構量表。

(7)量表的驗證

在此階段必須先檢查馬氏空間中所有 SMD 之平均值為 1 是否成立，如果平均值不為 1，則必須尋找新的樣本重建馬氏空間。理論上，一個理想的量表其異常觀察值之 SMD 都會大於正常觀察值之 SMD。但實務上並非如此，故必須決定適當之閾值來評估量表的能力，但如果量表的能力仍不理想，則必須回到步驟(1)、(2)、(3)。

(8)重要變數的篩選

利用直交表與訊號雜音比篩選重要變數，依辨識變數之多寡及對信用問題之先驗知識選定適當之直交表及訊號雜音以篩選重要變數。

(9)新量表的建構

重要變數的篩選對於信用分析的時效性及經濟性有其助益。選擇正常觀察值並利用重要的信用辨識變數建立重要因子模式之量表，以協助未來之預測。

(10)新量表的驗證

重複步驟(7)的程序，以決定新量表之辨識信用之能力。

(11)信用決策的進行

利用所建構之量表進行信用評估，並協助信用決策。

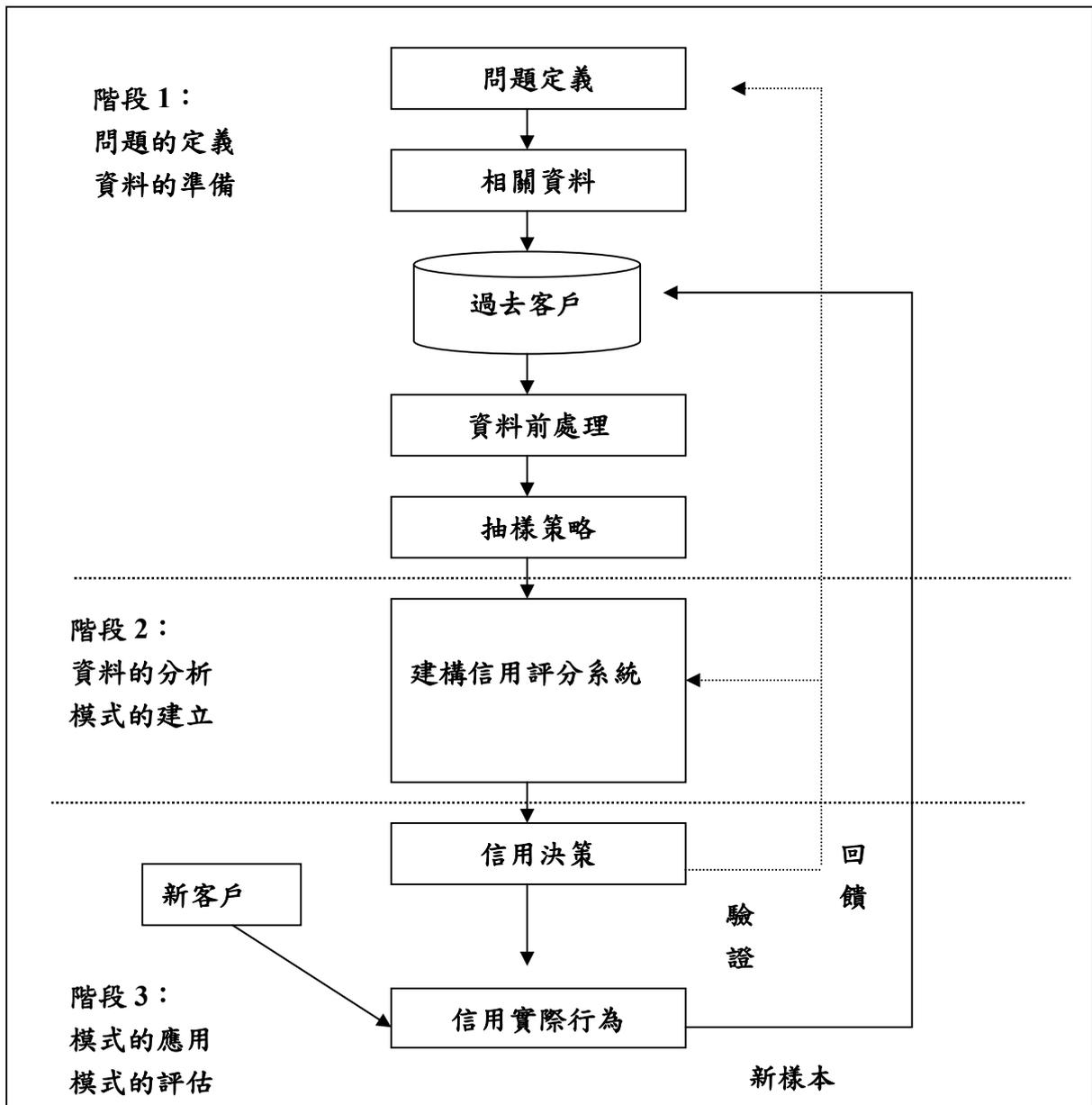


圖 5.3 以資料探勘流程建構信用評分的程序(Liu, 2000)

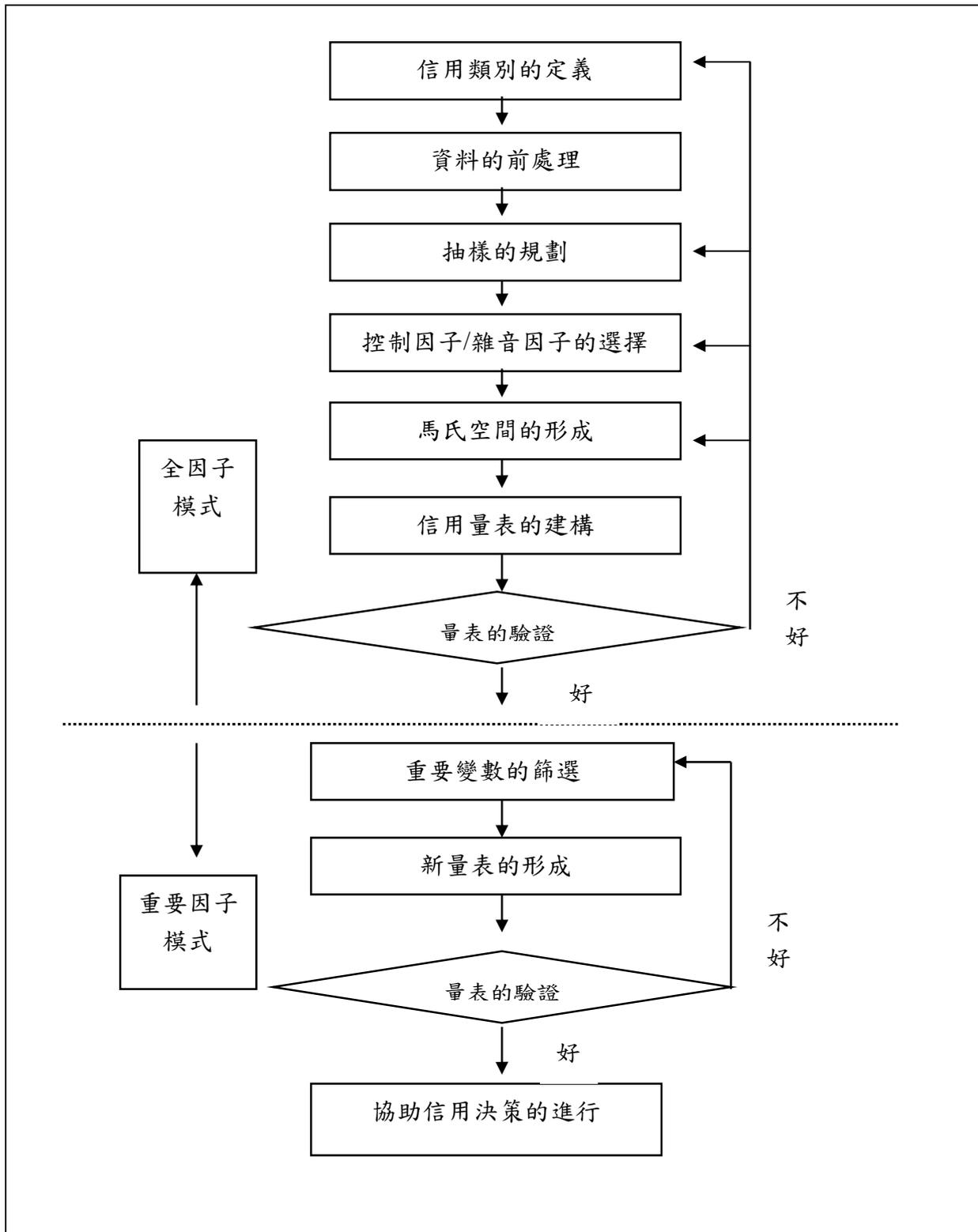


圖 5.4 MTS 信用評分模式流程圖

5.3 MTS 信用評分模式之實證研究

本研究利用國內某銀行的信用卡資料，驗證 MTS 信用評分模式之可行性。首先針對信用卡資料進行前處理及檢查資料間的關係；其次利用 MTS 建構各種信用評分模式，最後探討 MTS 信用評分模式之優缺點。

5.3.1 資料之前處理與檢查

本節針對資料的前處理及資料的檢查進行說明。此筆信用資料共有 15 個變數，其中有 5 個變數為名目尺度分別是性別、婚姻狀況、職業別及自有不動產、發卡通路，而教育程度、職位、條件差異為順序尺度，年齡由出生年月轉換而來、信用額度、年薪及預支現金金額及最高金額則為等比尺度，變數資料如表 5.3 所示。信用資料中類別變數必須重新編碼，以符合 MTS 運算之要求。故利用 1-of-(c-1) 編碼方式，以 c-1 個虛擬變數來代表 c 個類別，以避免線性相依(Linear Dependency)的問題。

而信用正常與否的定義，則根據銀行所提供的 Block Code 來定義，當客戶之 Block Code 為「空白」時，代表信用記錄良好則定義為「信用正常」；而當 Block Code 為「U」、「Y」時分別為「遲繳 180 天以上」及「遲繳 90 天以上」，代表已列入呆帳準備者則定義為「信用異常」。經定義及資料之篩選後可用的資料共 5925 筆，其中信用正常者為 5250 人，信用異常者為 675 人。

表 5.4 為客戶資料之次數分配表，包括性別、婚姻、教育程度、條件差異、職位、職業、居住地區、發卡通路及信用正常與否之分佈狀況。表 5.5 則提供有關年齡、信用額度、薪水、預支現金及最高金額之平均值及標準差等資訊。表 5.6 為交叉分析，以了解變數與信用正常與否之關係，觀察結果發現，除自有不動產之外，其餘變數與信用類別之交叉分析均具有統計的顯著性。表 5.7 針對信用好壞兩組資料進行 t 檢定，發現在預支現金、薪水、最高金額及信用額度上均有顯著的差異。表 5.8 則顯示連續變數之間的相關係數，其中信用額度與最高金額間的相關係數為 0.722，預支現金與最高金額間的相關係數為 0.649。

表 5.3 信用資料一覽表

	變數名稱	衡量尺度	說明
1	性別	名目尺度	0：男性，1：女性
2	年齡	等比尺度	此資料由出生年月日轉換而成
3	婚姻狀況	名目尺度	1：已婚，2：單身
4	教育程度	順序尺度	1：專上，2：高中，3：其他
5	條件差異	順序尺度	0：正常， 1：財力、工作、保人、年紀、信用有異常
6	職位	順序尺度	
7	職業別	名目尺度	
8	信用額度	等比尺度	
9	居住地區	名目尺度	
10	年薪	等比尺度	
11	自有不動產	名目尺度	Y：有；N：無
12	發卡通路	名目尺度	代表推廣信用卡的進件通路
13	預支現金金額	等比尺度	
14	最高金額	等比尺度	
15	主要碼	順序尺度	U：遲繳 180 天以上 Y：遲繳 90 天以上

表 5.4 客戶基本資料之次數分配表

變數	類別
性別	男性：4175(70.5%)；女性：1750(29.5%)
婚姻狀況	已婚：2767(46.7%)；未婚：3158(53.3%)
教育程度	專上：4143(69.9%)；高中：1242(21%)；其他：540(9.1%)
條件差異	條件正常：5347(90.2%)；條件異常：578(9.8%)
職位	代碼 0：15.7%(930)；代碼 1：13.5%(800)；代碼 2：8.2%(487)；代碼 3：7.7%(457)；代碼 4：33%(1955)；代碼 5：3%(178)；代碼 6：3%(176)；代碼 7：15.4%(913)；代碼 8：0.3%(19)；代碼 9：0.2%(10)
職業	代碼 1：15.1%(896)；代碼 2：0.8%(46)；代碼 3：25.5%(1513)；代碼 4：10.5%(623)；代碼 5：0.9%(53)；代碼 6：4.4%(258)；代碼 7：7.7%(458)；代碼 8：1.8%(104)；代碼 9：10.3%(611)；代碼 10：12.3%(729)；代碼 11：1.7%(99)；代碼 12：9%(535)
地區	北部：2002(33.8%)；中部：1794(30.3%)；南部：2129(35.93%)
發卡通路	代碼 1：5022(84.9%)；代碼 100：892(15.1%)
信用類別	信用正常：5250(88.6%)；信用異常：675(11.4%)

表 5.5 客戶基本資料之敘述統計表

變數	最小值	最大值	平均數	標準差
年齡	22	79	35.13	8.88
信用額度	10000	1000000	154014	111418.25
薪水	0	4148000	370063	272543.13
預支現金	-23927	692410	21710	71895.65
最高金額	0	1037163	83906	107081.57

表 5.6 基本資料與信用類別交叉分析結果一覽表

變數	Person 卡方值	顯著性檢定
性別	30.516	0.000*
婚姻	133.111	0.000*
教育程度	1062.270	0.000*
條件差異	1874.318	0.000*
職位	408.124	0.000*
職業	325.217	0.000*
居住地區	498.98	0.000*
發卡通路	51.26	0.000*
自有不動產	0.295	0.878

表中*代表具有統計顯著性

表 5.7 客戶基本資料之 t 檢定一覽表

變數	信用類別	平均數	標準差	變異數之 F 檢定		平均數之 t 檢定	
				統計量	顯著性	統計量	顯著性
預支現金	正常	1666.92	17643.57	9166.258	0.000*	-36.623	0.000*
	異常	177606.57	124653.75				
薪水	正常	352215.62	275538.89	122.422	0.000*	-18.279	0.000*
	異常	508881.48	199556.17				
最高現金	正常	58928.38	78475.12	85.965	0.000*	-54.417	0.000*
	異常	278176.87	100823.77				
信用額度	正常	138124.38	104040.18	42.233	0.000*	-38.328	0.000*
	異常	277607.41	86878.45				

表中*代表具有統計顯著性

表 5.8 相關分析一覽表

	年齡	信用額度	薪水	預支現金	最高現金
年齡	1	0.527	0.238	0.162	0.317
信用額度	0.527	1	0.424	0.419	0.722*
薪水	0.238	0.424	1	0.187	0.248
預支現金	0.162	0.419	0.187	1	0.649*
最高金額	0.317	0.722*	0.248	0.649*	1
表中*代表具有統計顯著性					

5.3.2 各種 MTS 信用評分模式之建構

本節建構 4 組實驗，以觀察 MTS 應用於信用評分之績效。表 5.9 為實驗規劃一覽表，表中所謂基本資料為客戶的性別、年齡、婚姻狀況、教育程度、條件差異、職位、職業、居住地區、薪資、有無不動產等 10 個變數。行為變數則為客戶在一段時間內巷銀行預支之現金及最高金額，雜音因子設定為發卡通路。在動態模式中則以信用額度為訊號因子。所謂標準配置是指田口博士針對類別變數所提出的配置法，而新的配置則指本研究所提出的配置法。

實驗 5-1 觀察利用基本資料與利用標準配置及望大特性來建構 MTS 信用評分模式。實驗 5-2 與實驗 5-1 的差別在於直交表的配置，本研究針對類別資料提出一個配置方法，透過實驗 5-2 觀察此種配置方法的優缺點。實驗 5-3 則是增加辨識的變數，加入客戶的行為變數以觀察 MTS 信用評分模式績效之改變。實驗 5-4 則嘗試利用動態特性來處理信用評分的問題。

表 5.9 各種信用評分模式一覽表

名稱	變數	訊號雜音比	雜音	訊號	直交表配置
實驗 5-1	基本資料	望大特性	通路	無	標準配置 L12
實驗 5-2	基本資料	望大特性	通路	無	新的配置 L32
實驗 5-3	基本資料+行為變數	望大特性	通路	無	L32
實驗 5-4	基本變數+行為變數	動態特性	通路	信用額度	L32

<實驗 5-1>

實驗 5-1 利用客戶的基本資料來建構模式。利用 1-of-(c-1)的編碼方式，將職業重新編成 11 個變數，分別為代號職業-1、職業-2、職業-3、職業-4、職業-5、職業-6、職業-7、職業-8、職業-9、職業-10、職業-11。而居住地區則編成 2 個變數，分別為地區-1、地區-2。重新編碼共有 21 個變數。資料共 5925 筆，其中信用正常者共 5250 人，信用異常者共 675 人。為了解 MTS 的可行性，將整筆資料劃分成 3 等份，取其中 2 份組成訓練樣本，另 1 份為測試樣本，也就是訓練及測試樣本之比率約為 2：1，共組成 3 組樣本。此外為避免母體漂移(Population Drift)的問題，本研究採用 Weiss 等人(1991)所建議之方法來抽樣。

表 5.10 為實驗 5-1 三組樣本之馬氏距離分配表，從表 5.10 中可發現這三組樣本之 SMD 之分配有重疊的現象，故要利用 SMD 來辨識信用的好壞並不容易。

表 5.10 實驗 5-1 之馬氏距離一覽表

樣本	平均值'	訓練樣本		測試樣本	
		信用正常	信用異常	信用正常	信用異常
1	1.0002	0.259~12.258	0.26~12.263	0.344~8.919	0.344~15.224
2	1.0002	0.251~11.006	0.329~12.913	0.25~11.004	0.334~7.942
3	1.0006	0.255~10.905	0.338~12.665	0.255~0.543	0.408~11.006

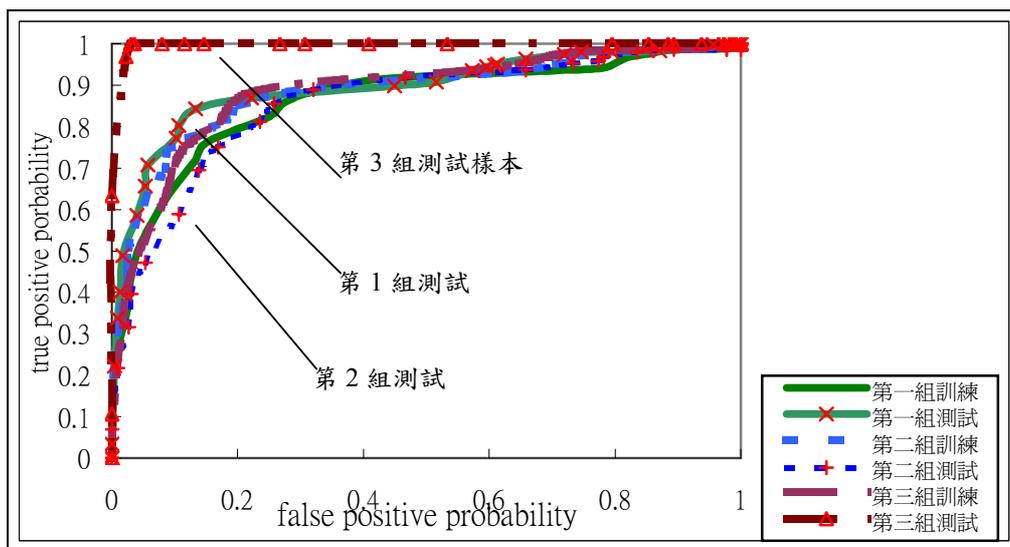


圖 5.5 實驗 5-1 三組訓練及測試樣本分類正確率之 ROC 曲線

三組訓練及測試樣本分類正確率之 ROC 曲線，如圖 5.5 所示。在各種不同的閾值下，觀察 MTS 信用評分模式之績效，由圖中發現測試樣本之分類績效依序為第三組、第一組及第二組。表 5.11 代表三組資料在不同閾值下，訓練樣本分類之正確率。利用訓練樣本分類績效最佳時所對應的閾值，當作未來判定測試樣本正確率之閾值，故第一組在閾值為 6 時測試樣本之分類正確率為 89.15%；第二組在閾值為 5 時測試樣本之分類正確率為 88.86%；第三組在閾值為 4 時測試樣本之正確率為 89.82%。

由於此筆信用資料中，信用正常與否之比例為 7.78 比 1，在不使用任何預測變數下，將所有客戶都判定成信用正常時，此時模式之正確率為 88.6%(3500/3950)，這個值稱為自然(Native)正確率。我們觀察這三組樣本所得之正確率只比自然正確率略高一些，表示實驗 5-1 所建構之 MTS 信用模式績效不盡理想，雖然如此，本研究仍利用此資料進行後續之變數篩選。

表 5.11 各組的訓練樣本在不同閾值下分類正確率一覽表

組別	正確率 Threshold=1	正確率 Threshold=2	正確率 Threshold=3	正確率 Threshold=4	正確率 Threshold=5	正確率 Threshold=6
1	77.29%	87.74%	85.97%	88.05%	88.10%	88.35%
2	76.57%	87.06%	87.62%	88.88%	88.96%	88.22%
3	76.78%	88.20%	88.40%	89.69%	89.34%	88.63%

有關重要變數的篩選，實驗 5-1 根據田口博士對類別資料配置的建議，利用原始變數的個數找出所對應的直交表並進行配置，由於實驗 5-1 共 10 個辨識變數，故本研究採用 L_{12} ，其配置方式，如表 5.12 所示。表中的數字代表直交表中的行，英文字母為變數之代號，本研究將變數配置在前 10 行。任選一組資料，經 12 次實驗結果獲得之因子效果圖，如圖 5.6 所示。圖中因子 A(性別)、B(年齡)、C(婚姻)、D(教育程度)、E(條件差異)、G(職位)、H(居住地區)為重要的辨識因子。

最後，利用重要的信用辨識因子來形成的量表以進行信用的分類，其分類之正確率 86.32%，這個值低於自然正確率。觀察實驗 5-1，不論全因子模式或是重要因子模式，其分類之正確率都不盡理想。

表 5.12 實驗 5-1 之直交表配置

1	2	3	4	5	6	7										8		9	10		
A	B	C	D	E	F	G										H		I	J		
性 別	年 齡	婚 姻	教 育 程 度	條 件 差 異	職 位	X0 職 業	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1 地 區	Y2	薪 資	房 屋	

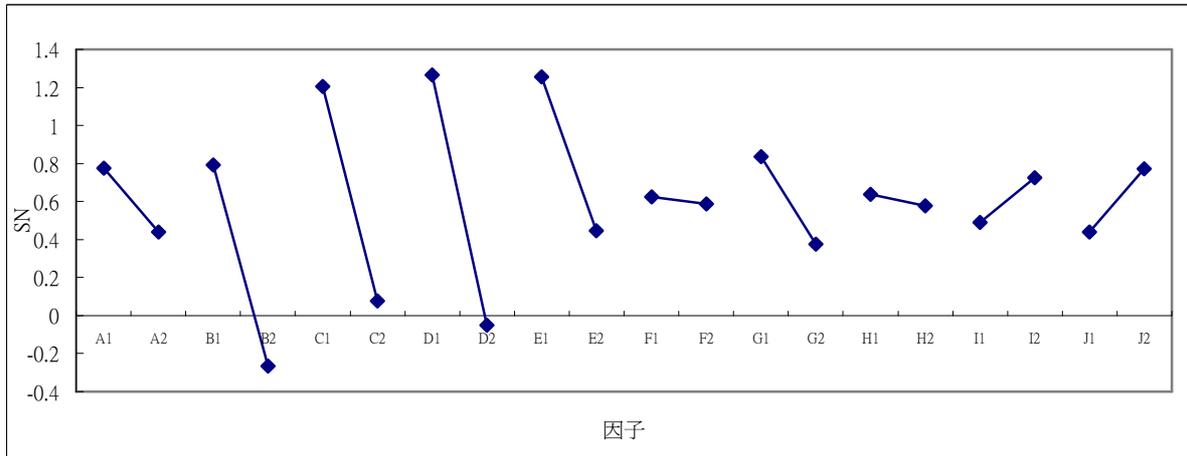


圖 5.6 實驗 5-1 重要因子效果圖

<實驗 5-2>

實驗 5-1 所採用的標準配置，雖然將類別變數透過 1-of-(c-1)的編碼方式分成幾類，但在直交表配置時仍視為一個變數來處理，這種方法的缺點是無法得知該變數的哪一個類別對信用的辨識較為重要。本研究規劃實驗 5-2，以改善上述標準配置的缺點。由於實驗 5-2 與 5-1 的差異僅在直交表的配置上，故實驗 5-2 所建構出的全因子模式之結果與實驗 5-1 相同，結果如表 5.10 所示。

在重要因子篩選的部份，本研究建議的新的配置法是將每一個編碼後的變數都視為一個變數來進行配置，在本例中，基本資料共 10 個變數經編碼後共有 21 個變數，故採用 L₃₂ 直交表進行配置，配置方法如表 5.13 所示。利用同樣的資料，經 32 次實驗獲得之因子效果圖，如圖 5.7 所示。由圖中可知因子 A(性別)、B(年齡)、C(婚姻)、D(教育程度)、E(差異條件)、F(職位)、I(職業-3)、J(職業-4)、L(職業-6)、R(地區-1)、S(地區-2)為重要因子。最後利用這些重要因子所形成之新量表進行信用的評估，得知此模式分類之

正確率為 88.6%。

從實驗 5-2 中可觀察以下的結果，利用本研究所提議之直交配置方法得出之重要因子模式之分類正確率比實驗 5-1 略高出 2.37%，且由此實驗可知職業-3、職業-4 及職業-6 對信用正常與否的辨識是重要的。所以雖然此種配置所用的實驗次數較多，但可獲得較多的資訊。

實驗 5-1 及 5-2 所建構之 MTS 信用評分模式之分類績效都不盡理想，從 MTS 的理論來看，這個現象反映所選擇的辨識變數不恰當，但是在 5.3.1 節中，卡方檢定及 t 檢定都顯示所用的變數與信用正常與否有關。觀察表 5.13，其中只有 2 個辨識變數是連續變數，其餘均為類別變數。是否因為類別變數過多且不符合馬氏距離之假設所造成的，使得 MTS 在信用評分模式的表現不理想，值得觀察。

表 5.13 實驗 5-2 之直交表配置

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U		
性別	年齡	婚姻	教育程度	條件差異	職位	職業 1	職業 2	職業 3	職業 4	職業 5	職業 6	職業 7	職業 8	職業 9	職業 10	職業 11	地區 1	地區 2	薪資	房屋		

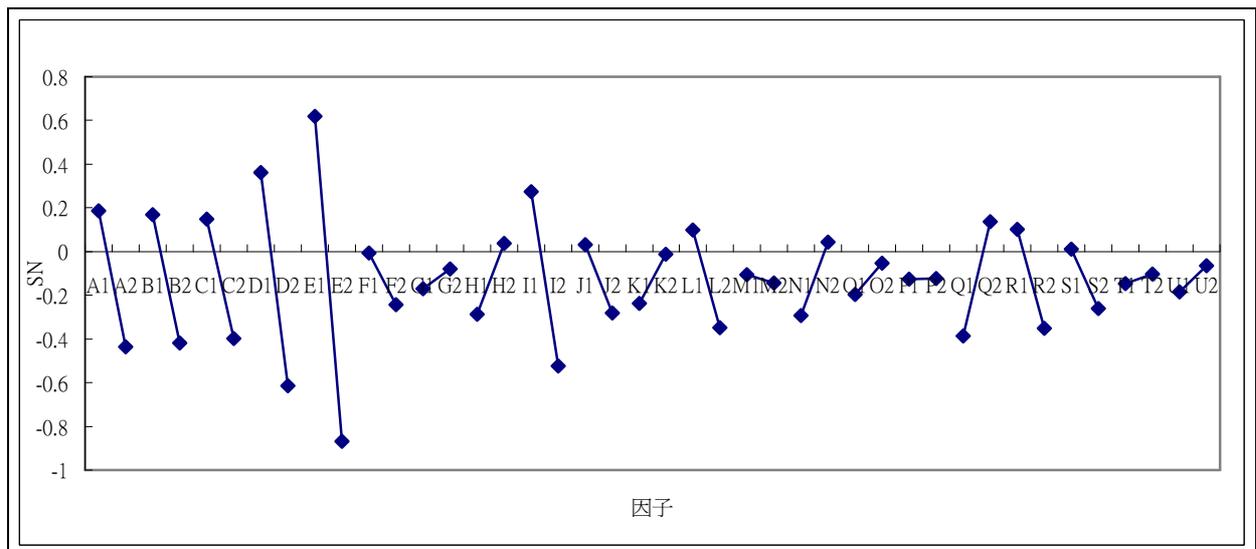


圖 5.7 實驗 5-2 之重要因子效果圖

<實驗 5-3>

在實驗 5-1 與實驗 5-2 中僅利用客戶之基本資料來進行信用之預測，其分類績效均不理想。建構實驗 5-3，除顧客基本資料外加入客戶之行為變數，包括最高金額及預支現金。將這些變數經編碼後，共得 23 個變數。

首先，利用這 23 個變數所形成之馬氏空間建構進行全因子模式，並評估模式之分類績效。表 5.14 列出三組樣本之 SMD 平均值及 SMD 之範圍。實驗 5-3 三組樣本之分類正確率之 ROC 曲線，如圖 5-8 所示。研究發現實驗 5-3 所建構之 MTS 模式之 ROC 曲線均比實驗 5-1 所建構之模式來的好。針對每一組資料，找出其訓練樣本最佳分類正確率所對應的閾值，分別為 5.5(第一組)、4.5(第二組)、5.5(第三組)。再根據此閾值來決定測試樣本的分類正確率，分別為 96.25%(第一組)、98.63%(第二組)、90.12%(第三組)。由實驗 5-3 之結果發現，由 MTS 所建構之信用評分模式之分類正確率比自然正確率 88.6%高出許多，實驗 5-3 之平均正確率 95%也比實驗 5-1 的平均正確率 89.27%改善了 5.73%。

表 5.14 實驗 5-3 之馬氏距離一覽表

樣本	平均值	訓練樣本		測試樣本	
		信用好	信用不好	信用好	信用不好
1	1.000005	0.237~16.447	0.175~15.976	0.238~16.698	4.184~17.006
2	1.000000	0.229~13.067	0.717~72.710	0.229~11.250	4.707~12.773
3	0.999997	0.234~17.6	4.184~97.851	0.234~0.524	0.685~9.51

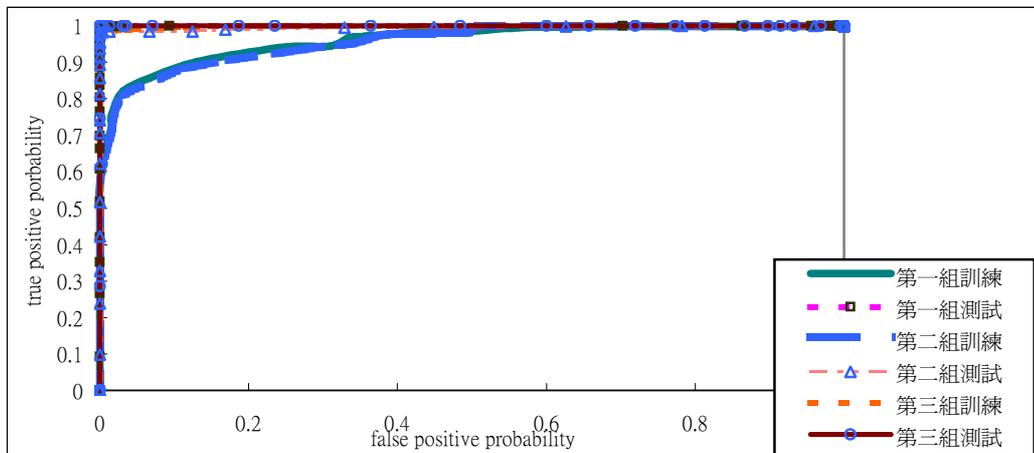


圖 5.8 實驗 5-3 全因子模式訓練及測試樣本分類正確率之 ROC 曲線

表 5.15 實驗 5-3 之直交表配置

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
性 別	年 齡	婚 姻	教 育 程 度	條 件 差 異	職 位	職 業 1	職 業 2	職 業 3	職 業 4	職 業 5	職 業 6	職 業 7	職 業 8	職 業 9	職 業 10	職 業 11	地 區 1	地 區 2	薪 資	房 屋	最 高 金 額	預 支 現 金

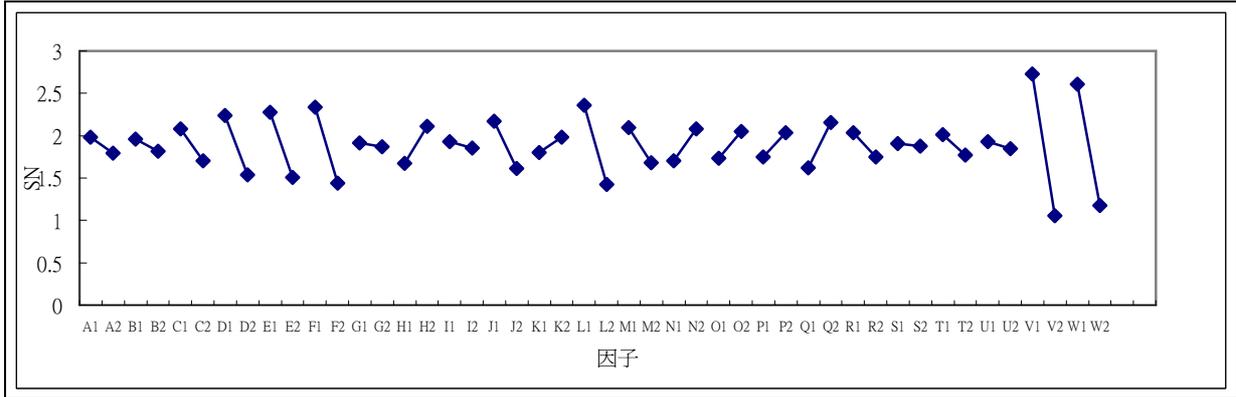


圖 5.9 實驗 5-3 第一組資料之重要因子效果圖

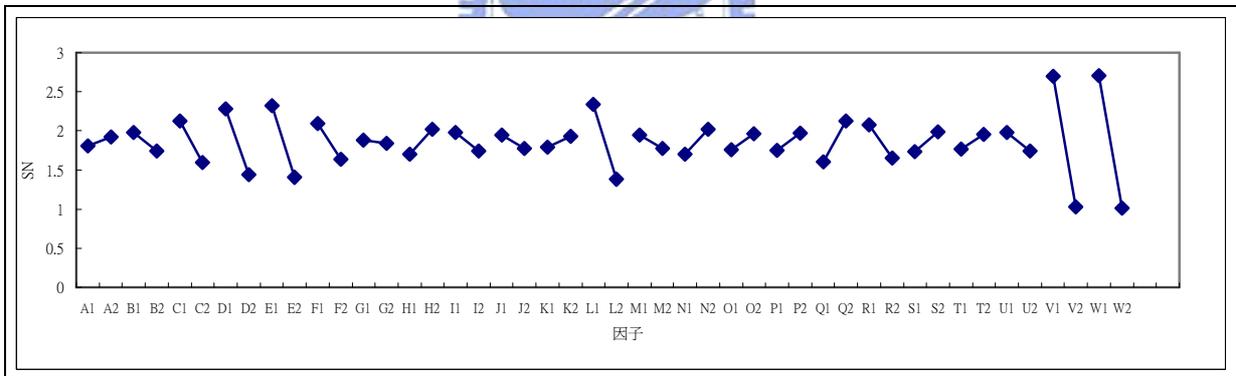


圖 5.10 實驗 5-3 第二組資料之重要因子效果圖

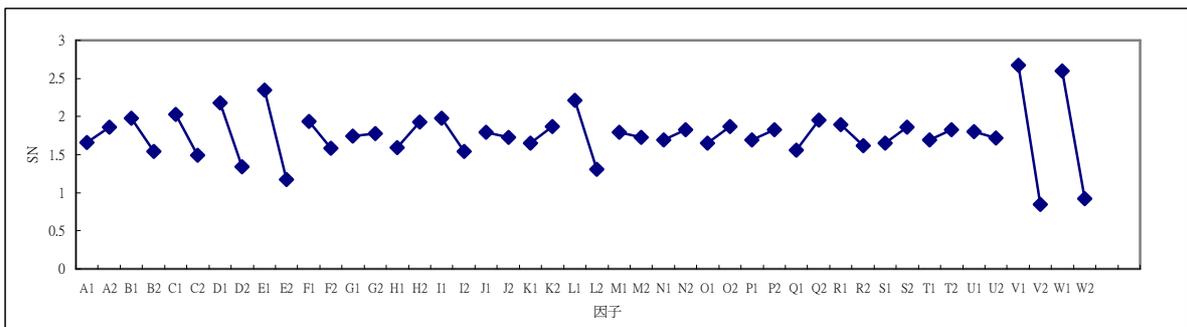


圖 5.11 實驗 5-3 第三組資料之重要因子效果圖

表 5.16 實驗 5-3 之三組資料重要因子一覽表

組別	重要因子 (依序)
1	V(1.69),W(1.43),L(0.93),F(0.89),E(0.77),D(0.7),J(0.55),M(0.4),C(0.37),R(0.29), T(0.24),A(0.18),B(0.14),I(0.07),G(0.04),S(0.03)
2	W(1.69),V(1.67),L(0.95),E(0.9),D(0.83),C(0.53),F(0.45),R(0.42),I(0.24),U(0.23), B(0.23),J(0.17),M(0.17),G(0.03)
3	V(1.82),W(1.67),E(1.17),L(0.9),D(0.83),C(0.53),B(0.43),I(0.43),F(0.35),R(0.27), U(0.08),J(0.07),M(0.06),

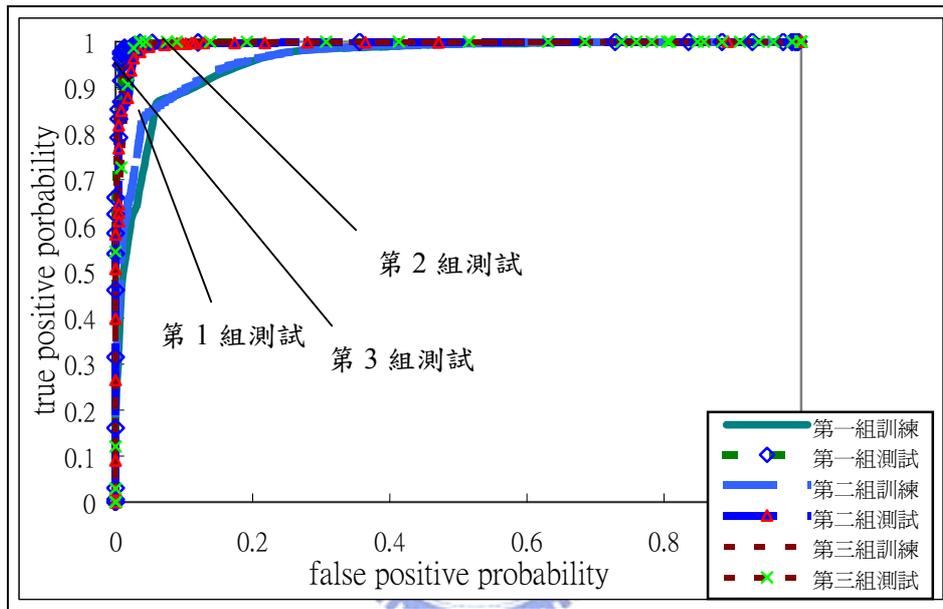


圖 5.12 實驗 5-3 重要因子模式訓練及測試樣本分類正確率之 ROC 曲線

表 5.17 實驗 5-3 之三組資料重要因子模式之正確率

組數	訓練樣本正確率	測試樣本正確率
1	94.96% (閾值=4)	97.59%
2	95.06% (閾值=4)	98.48%
3	99.49% (閾值=5)	91.58%
平均數		95.88%

利用 L₃₂ 直交表進行重要變數篩選的實驗，直交表之配置如表 5.15 所示，將三組資料分別經過直交表之配置共執行 96 次實驗，其因子效果圖如圖 5.9~5.11 所示。

茲將本實驗三組資料所篩選出的重要因子整理如 5.16 所示，表中英文字母代表因子，而括號中的數值則代表訊號雜音比的增益。儘管每一組實驗所篩選出的重要因子不盡相同，但變數 V(最高金額)、W(預支現金)、L(職業-6)、F(職位)、E(條件差異)、D(教

育程度)、J(職業-4)、M(職業-7)、R(地區-1)、B(年齡)、I(職業-3)等變數都是三組樣本的重要變數。

為了解重要因子所形成之新量表其分類之正確率，透過 ROC 曲線觀察三組樣本在不同閾值下的分類正確率，如圖 5.12 所示。實驗結果顯示，當加入行為變數之後，由 MTS 所形成之信用評分模式之績效大幅提升。表 5.17 為三組資料重要因子模式之分類正確率，可知在閾值為 4~5 時，三組測試樣本之平均正確率為 95.88%，比實驗 5-2 之 88.6%高出 7.28%。

<實驗 5-4>

實驗 5-4 嘗試利用動態訊號雜音比來處理信用評分問題，在 MTS 中，假設訊號因子 M_i 代表因客戶不付款而造成公司損失的金額，利用 $Y_j = \beta M_j$ ，其中 $Y_i = \sqrt{MD_j}$ ， $j=1,2,\dots,t$ 來觀察馬氏距離與訊號因子之間的關係。假設 $Y_j = \beta M_j$ 成立，則馬氏距離為 D^2 的客戶，其所對應的損失則為 $M = D / \hat{\beta} \pm 3\sqrt{MSE / (\hat{\beta})}$ 。這種利用已知依變數 (Dependent Variable) 的值反預測一個未知的獨立變數 (Independent Variable) 的過程，稱為「口徑測定」(Calibration)。本實驗嘗試利用動態特性的方式，找出以馬氏距離與信用額度之關係。

本實驗之實驗條件與實驗 5-3 相同，故其全因子模式之分類正確率與實驗 5-3 之結果相同。在篩選重要因子的部份，利用 L_{32} 直交表並蒐集 100 筆信用異常的資料進行實驗，利用動態的訊號雜音比來處理。針對每一組資料的 SMD 之平方根與信用額度，找出一關係式，但以本例所找出的 β 值幾乎為 0，這個模式對於逆預測而言沒有幫助。所以這個實驗可說是失敗的。

5.3.3 以 MTS 建構信用評分之優缺點探討

利用 MTS 建構信用評分模式有以下幾個優點：(1)透過 MTS 所建構之信用量表，可以衡量客戶信用異常的嚴重程度。(2)具穩健性，傳統的信用評分模式常因外在環境的變動，使得評分標準需經常調整。若能使信用評分模式具穩健性，則可以減少調整之次數。利用 MTS 建構信用量表時，可以透過在直交表內加入雜音因子的方式，來找出對雜音不敏感的重要因子。而雜音可以是不同的通路、時間或是經濟環境等。(3)具經濟及時效性，假設能夠獲得適當的訊號因子，則可利用 SMD 與訊號之間的關係建立一個反預測的模式，利用客戶信用相關的資料所建立的 SMD 來反預測訊號因子，這種模式將具有經濟及時效性。(4)MTS 可以改善傳統的信用評分的方式，傳統的作法是將客戶在每一個項目上的個別得分直接加總，這種缺點是會失去相關性的資訊，利用 MTS 則可以解決這個問題。(5)提供較為彈性之決策參考，決策者可隨著總體經濟環境、授信政策或是願意冒的風險的變化來決定 MTS 信用評分模式之閾值。(6)MTS 可提供一種檢查的機制，當所選用的辨識變數不恰當時，MTS 模式將會反映這個現象。

然而，MTS 也有以下的缺點：(1) MTS 在處理類別型資料時必須重新編碼，而客戶的信用資料中有相當多的類別變數，這個過程較為繁瑣。(2)受共線性之影響，如果變數間存在高度共線性時，則會影響評分之績效。(3)由於信用資料中有許多類別變數，類別變數往往無法滿足常態之假設，故無法符合馬氏距離之假設條件，這使得 MTS 在信用評分的應用上產生了限制。

5.4 MTS 之修正模式

MTS 是一種結合「馬氏距離」與「田口方法」的技術，其中「馬氏距離」用以進行型態辨識，而「田口方法」則用以縮減變數。雖然田口博士宣稱使用這個方法，不需對資料進行任何假設，但馬氏距離的使用前提就是要假設資料為常態，有關這個部份仍有不少爭議(Woodall, *et al.*, 2003; Hawkins, 2003; Bovas, 2003)。

由於信用資料通常是類別的型態，類別資料往往無法滿足符合常態之假設，本研究提出一修正之方法以處理類別資料的問題。根據前述文獻可知，羅吉斯迴歸(LR)適用於類別或連續資料，不需要常態的假設。羅吉斯迴歸除提供預測分類之組別外，還可計算出事件發生的機率。此外，羅吉斯迴歸的透明度也較類神經網路佳。故本研究採用整合羅吉斯迴歸與田口方法的方式來建構一個更適合信用評分性質之模式，以 LRTM 表示此整合模式。圖 5.13 為本研究所提出之 LRTM 信用模式之流程。透過羅吉斯迴歸計算出每一個客戶的違約機率以此衡量客戶之信用狀況，以此取代 MTS 中之 SMD，並利用田口方法之直交表與訊號雜音比及 Ω 轉換來篩選重要的辨識因子，最後利用重要變數衡量客戶之信用狀態以提供信用決策之參考。

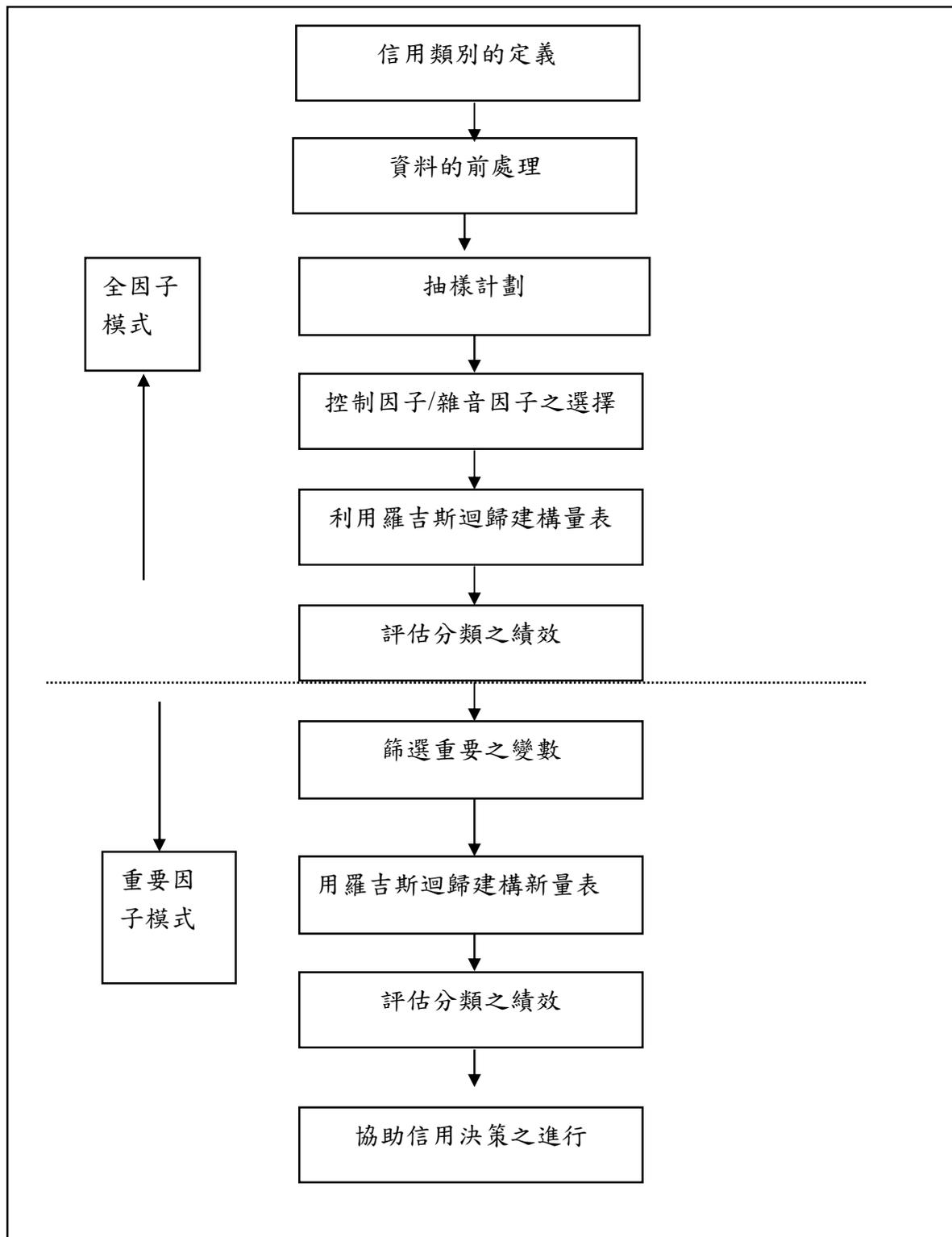


圖 5.13 LRTM 信用模式

本節規劃並執行兩組實驗如表 5.18 所示，透過這兩組實驗觀察所提之 LRTM 信用評分模式之可行性，並將其結果與 MTS 信用評分模式進行比較。

表 5.18 LRTM 實驗規劃一覽表

名稱	變數	訊號雜音比	雜音	直交表配置
實驗 5-5	基本資料	望大特性	通路	L32
實驗 5-6	基本資料+行為變數	望大特性	通路	L32

<實驗 5-5>

實驗 5-5 之實驗條件與實驗 5-2 相同，首先利用羅吉斯迴歸來建立全因子之信用評分模式，所建立之模式其分類之結果，如表 5.19 所示。比較表 5.19 與實驗 5-2 之結果，發現同樣的實驗條件下所建立之全因子信用模式，LRTM 模式的平均正確率為 92.67% 優於 MTS 之 89.19%。

表 5.19 實驗 5-5 訓練及測試樣本分類正確率

		訓練樣本 (預測值)		測試樣本 (預測值)	
		0	1	0	1
1(實際值)	0	3432	68	1700	50
	1	213	237	79	146
		總正確率：92.9%		總正確率：93.46%	
2(實際值)	0	3429	71	1748	2
	1	189	26	163	62
		總正確率：93.4%		總正確率：91.64%	
3(實際值)	0	3441	59	1750	0
	1	143	307	140	85
		總正確率：94.9%		總正確率：92.91%	

為說明如何運用 LRTM 來篩選重要的因子，以第一組資料資料進行重要因子的篩選為例，利用表 5.13 之直交表 L₃₂ 進行實驗的配置，每一組實驗根據所指定的變數進行羅吉斯迴歸模式的建立，並透過公式(5.1)來預估信用異常的客戶之違約機率。

$$p(Y) = \frac{1}{1 + e^{-(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_n X_n)}} \quad (5.1)$$

在(5.1)中 Y 為依變數，代表信用正常與否；而 X_i 則表示預測變數， β_i 則為迴歸係數。表 5.20 為執行第一組資料 32 次實驗所獲得羅吉斯迴歸之迴歸係數一覽表。將每一組實驗所獲得的客戶違約機率進行轉換，凡違約機率 $P(Y)$ 大於 0.5 者則定義為「信用異常」，再計算出每一組實驗中信用異常之比例 P 。當 P 很接近 0 或 100% 時，可加性非常差，必須透過 Ω 轉換(蘇朝墩, 2002)，如公式(5.2)所示，將不具加法性的數據轉換成可具加法性的數值。

$$\Omega = 10 * \log_{10} \left(\frac{1-p}{p} \right) \quad (5.2)$$

由於本實驗選擇 100 位信用異常的客戶進行重要變數的篩選，故計算訊號雜音比時採用望大特性訊號雜音比，如公式(5.3)所示，並根據此訊號雜音比篩選出重要變數。

$$SN_{P-LTB} = 10 * \log_{10} \left(\frac{p}{1-p} \right) \quad (5.3)$$

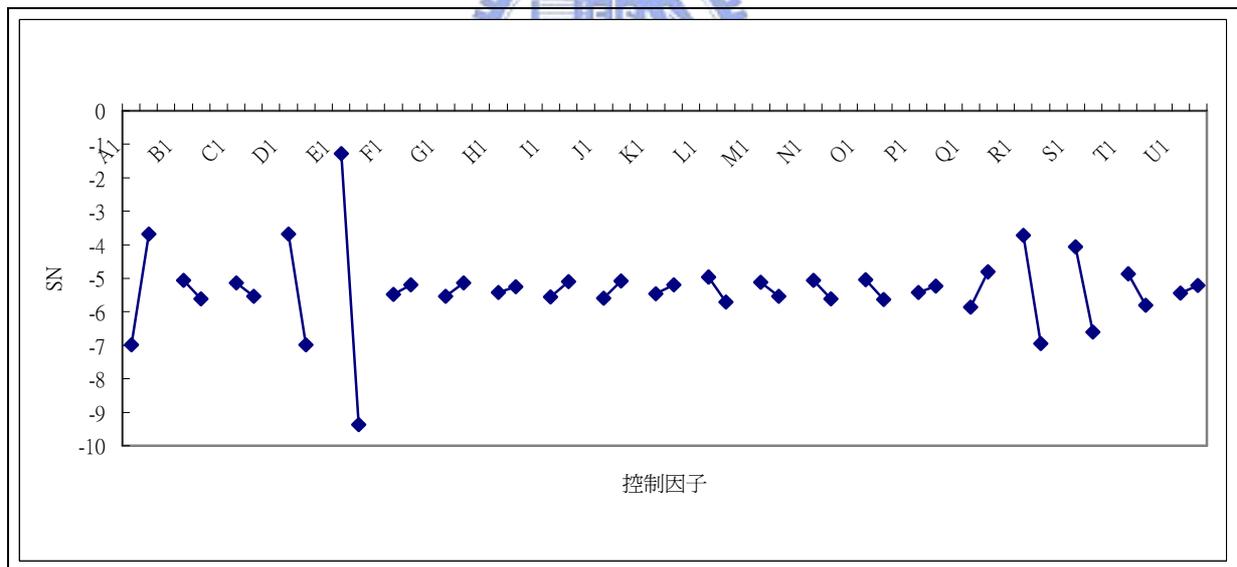


圖 5.14 實驗 5-5 重要因子效果圖

表 5.20 實驗 5-5 羅吉斯迴歸模式迴歸係數一覽表

實驗組別	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
常數	-23.165	-5.19	-15.34	-5.44	-18.904	-6.276	-10.288	-0.257	-17.906	-12.773	-8.472	-6.093	-17.261	-9.555	-8.435	-2.927	
A	0.251	0.312	0.2	0.288	0.342	0.417	0.437	0.497	0.203	0.257	0.207	0.207	0.341	0.412	0.411	0.455	
B	-0.012	-0.013	-0.012	-0.012	0.035	0.013	0.035	0.032									
C	-0.308	-0.326	-0.276	-0.345	-0.624	-0.485	-0.575	-0.465									
D	1.513	1.491	1.506	1.486					1.481	1.556	1.481	1.540					
E	3.545	3.541	3.493	3.484					3.436	3.638	3.410	3.599					
F	0.014	-0.027	-0.018	0.000									-0.041	-0.088	-0.073	-0.097	
G	1.449	0.689	1.049	1.111									1.678	0.947	1.141	0.661	
H	7.169	5.163			5.891	4.475			6.942	5.334			5.711	4.572			
I	0.692	0.039			0.421	0.138			0.422	0.184			0.682	0.203			
J	0.305	-0.544			-0.151	-0.547					-0.29	-0.343			-0.153	-0.666	
K	0.004	-0.577			0.133	-0.024					-0.25	-0.58			-0.073	-0.442	
L	-0.852	-1.383						-1.122	-1.287	-1.039	-1.306					-1.010	-1.392
M	0.284	-0.651						-0.045	-0.116	-0.051	-0.392					-0.0065	-0.163
N	-0.734	-1.666						-0.822	-0.959			-1.283	-1.426	-0.373	-0.739		
O	0.255	-0.490						-0.235	-0.250			-0.153	-0.479	0.203	-0.151		
P	1.784		1.505		1.878			1.633	1.265			1.105		2.054		1.624	
Q	5.192		5.026		5.412			5.167	5.195			4.03		5.803		5.476	
R	3.363		3.303		2.826			2.866			3.167		3.188		2.736		2.719
S	3.323		3.299		2.899			2.968			3.254		3.194		2.731		2.714
T	-0.314		0.305			0.293		0.296	0.343			0.325			0.325		0.347
U	-0.195		-0.219			-0.344		-0.287	-0.336			-0.321			-0.057		-0.171
實驗組別	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
常數	-9.163	-12.274	-6.966	-9.273	-12.386	-17.385	-5.427	-7.208	-11.972	-14.256	-7.087	-8.168	-7.088	-14.025	-1.1	-4.201	
A																	
B	0.002	0.01	0.002	0.009	0.034	0.027	0.032	0.027									
C									-0.174	-0.331	-0.16	-0.315	-0.789	-0.637	-0.748	-0.572	
D	1.436	1.405	1.467	1.455					1.432	1.405	1.456	1.422					
E																	
F	-0.058	-0.061	-0.060	-0.069										-0.083	-0.104	-0.074	-0.091
G					1.097	0.820	1.202	0.836	0.982	0.975	1.231	1.065					
H	5.663	5.622			6.037	7.295			5.743	5.851			5.634	7.625			

I			0.569	0.363			0.467	0.191			0.675	0.44			0.184	-0.042
J	-0.413	-0.541			-0.304	-0.435					0.006	-0.314			-0.462	-0.675
K			0.51	0.176			-0.111	-0.473	0.269	0.215			-0.483	-0.379		
L	-0.896	-0.952					-0.977	-1.301	-0.74	-0.79					-1.269	-1.524
M			-0.114	-0.48	0.088	0.047					0.019	-0.399	-0.021	-0.051		
N	-1.050	-1.244					-0.772	-1.004			-0.621	-0.937	-0.968	-1.035		
O			0.099	-0.22	-0.196	-0.273			-0.013	-0.258					-0.363	-0.513
P	-1.331		1.64		1.688		1.789		1.488		1.719		1.491		1.355	
Q		4.465		4.665		6.046		5.070		4.360		4.411		6.081		4.792
R	0.347		0.349		0.477		0.468			0.193		0.189		0.481		0.49
S		0.397		0.388		0.127		0.170	0.263		0.25		0.228		0.228	
T	0.428		0.423			0.252		0.261	0.365		0.345			0.294		0.287
U		-0.297		-0.26	-0.34		-0.366			-0.161		-0.204	-0.444		-0.487	

圖 5.14 為重要的因子效果圖，重要因子依序為 E (8.09)、D (3.31)、R (3.22)、S (2.5)、T(0.9)、L (0.74)、O (0.59)、N(0.56)、B (0.54)、M(0.42)、C(0.39)，其中 E 代表條件差異、D 為教育程度、R 為地區-1、S 地區-2、T 是薪資、L 是職業-6、O 為職業-9、N 是職業-8、B 為年齡、M 為職業-7、而 C 是婚姻，括號中的數值則為其訊號雜音比之增益。為驗證所找出之重要因子的辨識能力，利用上述的因子建立羅吉斯迴歸模式，並檢視其模式之正確性，其訓練模式之分類正確率為 92.8%，而測試模式之分類正確率也達 93.56%，這個結果高於實驗 5-2 之 88.6%。

為了解由 LRTM、MTS 所建構之模式正確率之差異，是否具有統計的顯著性。將三組樣本分別利用 LRTM 及 MTS 模式進行分析。表 5.21 為全因子 LRTM 及 MTS 模式分類正確率之比較表，進行無母數之 Mann-Whitney 檢定兩組資料之差異，在全因子模式中得 Mann-Whitney U 統計量為 0.00 所對應之 p 值為 0.05，故判定這兩組平均正確率的差異具統計之顯著性。

表 5.22 為重要因子 LRTM 及 MTS 模式分類正確率之比較表，其為 Mann-Whitney U 統計量為 0.00 對應之 p 值為 0.05，判定這兩組平均正確率的差異具統計上的顯著性。由

上述的結果得知，本研究所提之 LRTM 模式與 MTS 相比，在分類正確率上 LRTM 的效果較佳。

表 5.21 LRTM 模式與 MTS 模式分類正確率之比較表--全因子模式

正確率 \ 實驗次數	1	2	3
LRTM	93.46%	91.64%	92.91%
MTS	89.15%	88.86%	89.82%

表 5.22 LRTM 模式與 MTS 模式分類正確率之比較表--重要因子模式

正確率 \ 實驗次數	1	2	3
LRTM	93.56%	92.4%	93.97%
MTS	88.6%	90.68%	90.58%

<實驗 5-6>

實驗 5-6 之實驗條件與實驗 5-3 相同，利用羅吉斯迴歸來建立全因子之信用評分模式，其分類之結果如表 5.23 所示。表 5.24 比較 LRTM 與 MTS 在全因子模式之分類正確率，可以觀察出 LRTM 的平均分類正確率為 95.93%，而 MTS 的平均分類正確率則為 95%。為了解是否具有統計上的顯著性，透過無母數之 Mann-Whitney 對兩組資料進行比較，在顯著水準為 0.05 下，得 Mann-Whitney U 統計量為 3.00 對應之 p 值為 0.513，故並無顯著的差異。

表 5.25 比較 LRTM 與 MTS 重要因子模式之分類正確率，可以觀察出 LRTM 的平均分類正確率為 95.93%，而 MTS 的平均分類正確率則為 95.88%。透過無母數之 Mann-Whitney 對兩組資料進行比較，在顯著水準為 0.05 下，得 Mann-Whitney U 統計量為 4.00 對應之 p 值為 0.827，故 LRTM 與 MTS 在此筆資料上並無顯著的差異。

觀察實驗 5-5 及實驗 5-6 得知，當我們只用客戶基本資料來建立信用評分模式時，無論是全因子或是重要因子模式，本研究所提出之 LRTM 的表現會比 MTS 的表現好。但是當我們加入行為變數時，LRTM 與 MTS 的表現就無顯著之差異。

回顧圖 5.1 可知信用評分模式可分成「信用申請階段」及「信用績效的階段」，在申請階段所獲得的資料通常是客戶的基本資料，而顧客的基本資料中多數為類別資料，所以研究認為 LRTM 模式適合建構此階段的信用模式。而由於 MTS 有較不易受到樣本比例變化之影響的特性，所以研究認為當信用變數有較多連續資料時，MTS 也是一個值得考慮的模式。

表 5.23 實驗 5-6 訓練及測試樣本分類正確率

		訓練樣本 (預測值)		測試樣本 (預測值)	
		0	1	0	1
1(實際值)	0	3455	45	1701	49
	1	107	343	11	214
		總正確率：96.2%		總正確率：96.96%	
2(實際值)	0	3455	45	1735	15
	1	112	338	11	214
		總正確率：96%		總正確率：98.68%	
3(實際值)	0	3489	11	1750	0
	1	7	443	155	70
		總正確率：99.5%		總正確率：92.15%	

表 5.24 實驗 5-6 與實驗 5-3 正確率之比較表--全因子模式

正確率	1	2	3
實驗 5-6(LRTM)	96.96%	98.68%	92.15%
實驗 5-3(MTS)	96.25%	98.63%	90.12%

表 5.25 實驗 5-6 與實驗 5-3 正確率之比較表--重要因子模式

正確率	1	2	3
實驗 5-6(LRTM)	98.78%	96.40%	94.12%
實驗 5-3(MTS)	97.59%	98.48%	91.58%

5.5 各種信用評分模式比較

本節將比較 MTS、LRTM、鑑別分析與決策樹在信用評分模式之正確率。雖然類神經網路應用在信用評分模式中也有不錯的表現，但由於類神經網路運算的過程宛如黑箱，對信用評分模式的使用者而言，並不容易理解，故在本研究不討論類神經網路應用於信用評分的部份。

利用實驗 5-2 的資料進行各種模式之比較。利用 See5 建構決策樹模式，其測試樣本分類正確率分別為 94.2%、94.4%、92.9%。利用鑑別分析所獲得之測試樣本分類正確率分別為 88.6%、88.6%、88.6%。資料如表 5.26 所示。為了解各種方法之平均正確率是否具有統計的顯著差異，利用無母數 Kruskal-Wallis 檢定，其卡方值為 9.207，對應之 p 值為 0.027，故在顯著水準 α 為 0.05 下，這四種方法具有顯著性之差異。由於平均之正確率來看，依序為決策樹(94.2%)、LRTM(93.31%)、MTS(89.95%)、鑑別分析(88.6%)。由此結果可知，當類別變數較多時，決策樹與 LRTM 之結果會優於 MTS 與鑑別分析。

表 5.26 四種分類技術之分類結果比較表

實驗 5-2 之 資料	方法	實驗次數	1	2	3	平均值
		MTS		88.6%	90.68%	90.58%
	LRTM		93.56%	92.4%	93.97%	93.31%
	決策樹		94.2%	94.4%	92.9%	94.2%
	鑑別分析		88.6%	88.6%	88.6%	88.6%
實驗 5-3 之 資料	方法	實驗次數	1	2	3	平均值
	MTS		97.59%	98.48%	91.58%	95.88%
	LRTM		98.78%	96.40%	94.12%	96.43%
	決策樹		95.8%	97.3%	90.9%	94.67%
	鑑別分析		97.8%	97.92%	98.86%	97.84%

另外，利用實驗 5-3 的資料進行各種模式之比較。利用 See5 建構決策樹模式，其分類正確率分別為 95.8%、97.3%、90.9%。利用鑑別分析所獲得之分類正確率分別為 97.8%、97.92%、98.86%。資料如表 5.26 所示。利用無母數 Kruskal-Wallis 檢定，其卡方值為 4.179，其對應之 p 值為 0.243，故在顯著水準 α 為 0.05 下，這幾種方法並沒有顯

著性之差異。

由上述的資料發現，MTS 及鑑別分析都較適合於實驗 5-3 的資料，LRTM 模式及決策樹則不論在實驗 5-2 的資料上或是實驗 5-3 的資料上都有不錯的表現。



第六章 結論

本研究探討 MTS 的特性，並將 MTS 應用於信用評分中評估其績效及可行性，最後並針對 MTS 處理類別資料的問題提出一改善方法 LRTM。本章針對研究之結果及未來可行之研究提出建議。

6.1 研究結果與結論

本論文研究之結果如下：

一、MTS 特性之探討：

經過實驗的觀察發現，利用 MTS 處理分類問題時，有三項重要特性：(1)慎選馬氏空間，因為實驗發現馬氏空間的選擇會顯著地影響 MTS 分類之績效。(2)MTS 之分類效果不易受樣本比例之影響，這個特性使得使用者不須擔心樣本的議題。(3)有關訊號雜音比的部份，經實驗之觀察仍建議採用望大特性。

二、在信用評分的應用方面：

本研究提出以 MTS 建構信用評分模式，除利用實際的信用資料實證 MTS 之可行性之外，並針對 MTS 在處理類別資料時的缺點，提出一改善的模式 LRTM，並驗證 LRTM 在信用評分模式之可行性。透過實驗觀察，利用 MTS 建構信用評分模式時，若僅利用客戶之基本資料，其分類結果是不理想的，主要原因是類別資料過多所造成。但當模式加入行為變數時，則 MTS 信用評分模式就可以呈現較佳的結果。針對 MTS 缺點所提出的模式 LRTM 的分類績效，不論在僅有客戶基本資料的狀況或是加入行為變數的狀況下都有不錯的表現。

針對 MTS，本研究有以下之看法：有關 MTS 之優點如下：(1)藉由閾值的建立，MTS 除了可以進行分類之外，亦可衡量異常的嚴重程度，可提供較多的訊息；(2)馬氏距離有考慮變數之間的相關性，這使得 MTS 在進行綜合評價時更為適合；(3)MTS 透過直交表與訊號雜音比可篩選重要變數，縮減變數個數會節省系統進行預測或辨識所需的時間，故具有成本之效益。(4)具穩健性，延續著田口品質工程的精神，利用 MTS 所設

計出來的模式也具有對雜音不敏感之特性。至於 MTS 之缺點如下：(1)缺乏圖形分析，視覺化的分析對使用者來說是相當方便且易於了解，MTS 在這個部份仍有改善的空間。(2)由於馬氏空間的建立只以正常群體為對象，田口建議正常群體盡量具有均質性，但過度的均質性以致該變數之標準差為 0 的狀況時會出現無法標準化的情形，此時就無法計算馬氏距離。

應用 MTS 時最重要的議題就是馬氏空間的製作，這個階段可以說幾乎決定 MTS 應用之成敗。以本研究在信用評分應用上為例，銀行及相關的法規只針對信用異常的客戶進行定義，並無對「信用好」下定義，類別定義的不確定性影響 MTS 之分類績效，實務上，有時不見得可以對類別做出正確的定義，如何制定良好的定義將是一重要的前提。第二個議題是訊號因子之設定，「口徑測定」的好處是它可透過馬氏距離與訊號因子的關係來達成逆預測的效果。在信用評分的案例中，假設可以找出客戶之馬氏距離及其因違約造成公司損失金額間的關係時，則可根據馬氏距離來及早預測可能損失的金額，以提供及時之決策參考。第三個議題是誤差因素的設定，就是雜音的設定。以信用評分為例，授信人員其評估的基準可能不盡相同，所以授信人員間的差異就是誤差因素，實務上要找出或模擬出適合的雜音並不容易。

6.2 未來研究建議

當面對屬性(變數)過多的情況，利用直交表進行篩選變數就顯得複雜，而 MTS 的作法是利用多重馬氏距離法將屬性分割成好幾個群體，最佳化的進行都是針對個別的群體，然而這種作法卻忽略了每一群體間的關係，所以考慮群體間的關係是未來可探討之議題。由研究發現馬氏空間選擇對於 MTS 分類績效有重要的影響，如何進行馬氏空間的選擇，是使用 MTS 中最重要的一件事，而目前並無適當的方法來決定馬氏空間，所以尋找出適當決定馬氏空間的方法，是未來 MTS 研究中可著眼的方向。

參考文獻

英文部分

1. Bovas, A. and Asokan Mulayath, V., 2003, "Discussion-A Review and Analysis of the Mahalanobis-Taguchi System," *Technometrics*, Vol. 45, NO.1, pp. 22-25
2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C. J., 1984, *Classifications and Regression Trees*, Wadsworth, Pacific Grove, California, USA.
3. Carter, C. and Catlett, J., 1987, "Assessing Credit Card Applications Using Machine Learning," *IEEE expert*, Vol. 2, pp. 71-79
4. Carven, M. W. and Shavlik, J. W., 1997, "Using Neural Networks for Data Mining," *Future Generation Computer Systems*, Vol. 13, pp. 221-229
5. Chiu, C. C. and Kao, L. J., 2001, "Mining the Customer Credit by Using the Neural Network Model with Classification and Regression Tree Approach," IFSA World Congress and 20th NAFIPS Intern Conference, Joint 9th, Vol. 2, pp. 923-928
6. Cios, K. J., Pedrycz, W. and Swiniarski, R. W., 1998, *Data Mining : Methods for Knowledge Discovery*, Kluwer Academic, Boston.
7. Coffman, J. Y., 1986, "The Proper Role of Tree Analysis in Forecasting the Risk Behavior of Borrowers," *Management Decision Systems*, Atlanta, MDS Reports, 3,4,7,and 9
8. Curt, H., 1995, "The Devil's in The Detail: Techniques, Tools, and Applications for Database Mining and Knowledge Discovery-Part 1," *Intelligent Software Strategies*, Vol. 6, pp. 1-15
9. Davis R. H., Edelman D. B. and Gammerman A. J., 1992, "Machine Learning Algorithms for Credit-Card Applications," *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 4, pp. 43-51
10. DeLeo, M. and Rosenfeld, S. J., 2001, "Essential roles for receiver operating characteristic methodology in classifier neural network applications," *Proceeding*

International Joint Conference Neural Networks, Vol. 4, pp. 2730-2731

11. Desai, V. S., Conway, D. G., Crook, J. N. and Overstreet, G. A., 1997, "Credit Scoring Models in the Credit Union Environment Using Neural Networks and Genetic Algorithms," *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 8, pp. 323-346
12. Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P., 1996, "The KDD Process for Extracting Useful Knowledge from Volumes of Data," *Communications of the ACM*, Vol. 39, pp. 27-34
13. Fish, K. E., Barnes, J. H. and Ailen, M. W., 1995, "Artificial Neural Networks: A New Methodology for Industrial Market Segmentation," *Industrial Marketing Management*, Vol. 24, pp. 431-438
14. Freeman, J. A. and Skapura, D. M., 1992, *Neural Networks Algorithms, Applications and Programming Techniques*, Addison-Wesley, Michigan.
15. Galitz, L. C., 1983, "Consumer Credit Analysis," *Management Finance*, Vol. 9, pp. 27-33
16. Hand, D. and Henley, W. E., 1997, "Statistical Classification Methods in Consumer Credit Scoring: a Review," *Journal of the Royal Statistical Society*, Series A, Part 3, pp.523-542
17. Hawkins, D. M., 2003, "Discussion- A Review and Analysis of the Mahalanobis-Taguchi System," *Technometrics*, Vol. 45, pp. 25-29
18. Henley W. E., 1995, *Statistical Aspects of Credit Scoring*. Dissertation. The Open University, Milton Keynes, UK.
19. Henley W. E. and Hand, D. J., 1996, "A K-Nearest Neighbor Classifier for Assessing Consumer Credit," *Statistician*, Vol. 44, pp.77-95
20. Hosmer, D. W. and Lemeshow, S, 1989, *Applied statistics Regression*, John Wiley & Sons, New York
21. Johnson, R. A. and Wichern, D. W., 1998, *Applied Multivariate Statistical Analysis*, Prentice-Hall

22. Kass, G., 1980, "An Exploratory Technique for Investigating Large Quantities of Categorical Data," *Applied Statistics*, Vol. 29:2, pp. 119-127
23. Liu, Y., 2002, A Framework of Data Mining Application Process for Credit Scoring, Institute of Information Systems, University of Göttingen, Nr.01/2002, Göttingen.
24. Liu, Y., 2002, The Evaluation of Classification Models for Credit Scoring, Institute of Information Systems, University of Goettingen, Nr.01/2002, Göttingen.
25. Lu, H., Setiono, R. and Liu, H., 1996, "Effective Data Mining Using Neural networks," *IEEE Transaction on Knowledge and Data Engineering*, Vol. 8, No. 6, pp. 957-961
26. Lunani, M., Nair, V. N. and Wasserman, G. S., 1997, "Graphical Methods for Robust Design with Dynamic Characteristics," *Journal of Quality Technology*, Vol. 29, pp. 327-338
27. Mahalanobis, P. C., 1936, "On the generalized distance in statistics," *Proceedings National Institute of Science of India*, pp. 49-55
28. Makowski, P., 1985, "Credit Scoring Branches Out," *The credit world*, Vol. 75, pp. 30-37
29. Martell, T. F. and Fitts, R. L., 1981, "A Quadratic Discriminant Analysis of Bank Credit Card User Characteristics," *Journal of Economics and Business*, Vol. 33, pp. 153-159
30. Moxon, B., 1996, "Defining Data Mining", DBMS Data Warehouse Supplement. Downloaded from: <http://www.dbmsmag.com/9608d53.html>, 05.12.2000.
31. Quinlan, J. R., 1993, *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Francisco, CA.
32. Rajesh, J., Taguchi, G., and Taguchi, S., 2003, "Discussion-A Review and Analysis of the Mahalanobis-Taguchi System," *Technometrics*, Vol. 45, NO. 1, pp. 16-21
33. Reichert A. K., Cho, C. C. and Wager, G. M., 1983, "A Examination of the Conceptual Issues Involved in Developing Credit-Scoring Models," *Journal of Business and Economics Statistics*, Vol. 1, pp. 101-114
34. Rosenberg, E. and Gleit, A., 1994, "Quantitative Methods in Credit Management: A

- Survey,” *Operations Research*, Vol. 42, No. 4, pp.589-613
35. Rumelhart, D. E., Hinton, D. E. and Williams, R. J., 1986, *Learning Internal Representations by Error Propagation in Parallel Distributed Processing*, MIT Press, Cambridge, MA, pp.318-362
36. Sharma, S., 1996, *Applied Multivariate Techniques*, John Wiley & Sons
37. Tam, K. Y., and Kiang, M. Y., 1992, “Managerial Applications of Neural Networks: the Case of Bank Failure Predictions,” *Management Science*, Vol. 38 :7, pp. 926-947
38. Taguchi, G., Chowdhury, S. and Wu, Y., 2001, *The Mahalanobis-Taguchi System*, McGraw-Hill
39. Taguchi, G. and Jugulum, R., 2000, “New Trends in multivariate diagnosis,” *Sankhyā: The Indian Journal of Statistics*, Vol. 62, Series B, pp. 233-248
40. Taguchi, G. and Jugulum, R., 2002, *The Mahalanobis-Taguchi Strategy*, John Wiley & Sons, New York
41. Thomas, L. C., 2000, “A Survey of Credit and Behavioral Scoring: Forecasting Financial Risk of Lending to Consumers,” *International Journal of Forecasting*, Vol. 16, pp. 149-172
42. Tsui, K.L., 1999, “Modeling and Analysis of Dynamic Robust Design Experiments,” *IIE Transactions*, Vol. 31, pp. 1113-112
43. Weiss, S. H. and Indurkha, N., 1998, *Predictive Data Mining: A Practical Guide*, Morgan Kaufmann, San Francisco, CA.
44. West, D., 2000, “Neural Network Credit Scoring Models,” *Computers & Operations Research*, Vol. 27, pp. 1131-1152
45. Woodall, W. H., Koudelik, R., Tsui, K. L., Kim, S.B., Stoumbos, Z. G., and Carvounis, C. P., 2003, “A Review and Analysis of the Mahalanobis-Taguchi System,” *Technometrics*, Vol. 45, pp. 1-15

中文部份

1. 王濟川、郭志剛著，2003，Logistic 迴歸模型—方法及應用，五南出版
2. 蘇朝墩著，2002，品質工程，中華民國品質學會出版

