

結合獨立成份分析與類神經網路於財務時間序列預測模式之建構

Financial time series forecasting using independent component analysis and artificial neural network

呂奇傑 Chi-Jie Lu

清雲科技大學 工業工程與管理系

Department of Industrial Engineering and Management, Ching Yun University

李天行* Tian-Shyug Lee

輔仁大學 管理學研究所

Graduate Institute of Management, Fu-Jen Catholic University

高人龍 Jen-Lung Kao

輔仁大學 企業管理學系

Department of Management, Fu-Jen Catholic University

陳學群 Hsueh-Chun Chen

輔仁大學 應用統計研究所

Graduate Institute of Applied Statistics, Fu-Jen Catholic University

摘要:財務時間序列預測一直是財務決策之重要議題，但由於財務時間序列資料具有高頻率、雜訊、非定態與混沌等性質，使得其在時間序列預測領域中，向來被視為是一極具挑戰性的應用領域。本研究提出一結合獨立成份分析(independent component analysis, ICA)與倒傳遞類神經網路(back-propagation neural network, BPN)之二階段模式建構程序，先利用 ICA 具有將混合訊號分離出個別獨立來源訊號之能力，從預測變數中估計出獨立成份，在使用測試接受法找出獨立成分中代表資料中雜訊的獨立成分後，將代表雜訊的獨立成份去除，並利用其餘不代表雜訊的獨立成份重建時間序列資料，得到濾除雜

* 通訊作者：台北縣新莊市，輔仁大學管理學研究所， E-Mail: 036665@mail.fju.edu.tw

訊後之預測變數，最後再使用 BPN 以濾除雜訊後之預測變數建構預測模式，期望可以使 BPN 於建構模式時不受雜訊影響，進而提升預測結果的準確度。為驗證所提方法之有效性，本研究以日經 225(Nikkei 225)現貨開盤指數進行實證研究，並與直接使用 BPN、隨機漫步(random walk)模式及其他常用之去除雜訊方法-簡單移動平均(simple moving average)及小波框架(wavelet frame)之預測結果進行比較。實證結果顯示，所提之方法不論在預測誤差或是趨勢預測準確度的表現上均較直接使用 BPN、隨機漫步、簡單移動平均及小波框架為佳。

關鍵詞：財務時間序列預測；獨立成份分析；類神經網路；股價指數

Abstract: The characteristics of time series financial data are inherently high frequency, noisy, non-stationary and deterministically chaotic, which render the time series financial forecast extremely challenging. Owing to advantages in building non-parametric and non-linear models, artificial neural networks (ANN) have also been applied to time series predictions, especially for modeling financial time series forecasting. In ANN based financial time series modeling, one of the primary issues is the inherent high noise. It is an important but difficult task to identify and alleviate the noise in order to build a reliable ANN forecasting model. To minimize the influence of noise, we propose to conduct financial time-series forecasting by combining the approaches of both independent component analysis (ICA) and back-propagation neural network (BPN). The combined approach first applies ICA to deduce independent components (ICs) from the forecasting variables. After identifying and removing each IC contained noise, the filtered IC signals are then used to reconstruct forecasting variables and applied to the BPN forecasting model. In order to validate the performance advantage of the proposed approach, the Nikkei 225 opening cash price index has been used as an illustrative example. The experimental results show that the proposed model outperforms in forecasting accuracy the conventional BPN model with non-filtered forecasting variables, the random walk model, the simple moving average model, and the wavelet frame model.

Keywords: Financial time series forecasting; Independent component analysis; Artificial neural network; Stock index

1. 緒論

證券市場一直是投資理財最常用的投資標的之一，但證券市場的高波動性，常造成投資人的重大損失，因此無論是專業分析師的預測、市井小民的口耳相傳或是各種基本分析與技術分析，都顯示了投資者為了獲取報酬，嘗試去預測股價的慾望，因而造就了財務時間序列預測研究議題的蓬勃發展。然而由於金融商品的交易特性以及價格容易受政府政策、景氣循環、國際市場的漲跌及物價膨脹等眾多因素影響，使得財務時間序列具有高頻率(high frequency)、雜訊(inherently noisy)、非定態(non-stationary)與混沌(chaotic)等性質(Cao, 2003; Deboeck, 1994; Yaser and Atiya, 1996)。其中雜訊會使得建構之預測模式產生過度配適(over-fitting)或配適不足(under-fitting)的問題；而非定態的特性則使得時間序列的統計分配將隨著時間而改變，增加建構預測模式之困難度。因此，財務時間序列是極為複雜且難以預測的時間序列資料，為一極具挑戰性的應用領域。

在針對財務時間序列建構預測模式所使用的工具方面，可以分為定性與定量模式兩大類。定性預測模式主要是在沒有足夠量化資料可運用、或資料無法量化時使用，常見的有德爾菲法(Delphi technique)、專家意見法等；而定量預測模式則是相信歷史會重現，也就是過去的數據走勢可能會在未來重演，因此可以透過以往的歷史資料加以預測，其中又可分為以統計為基礎的傳統預測模式以及人工智慧(artificial intelligence, AI)方兩大類。常見的傳統預測模式有指數平滑法(exponential smoothing)、時間序列分析(time series analysis)及迴歸分析(regression analysis)等，由於傳統的預測模式常需符合特定之假設，因而限制了其實用性(Lee and Chen, 2002; Lee and Chiu, 2002)，因此不需過多假設及擁有學習能力特性的人工智慧方法便逐漸受到重視及使用。

在眾多人工智慧工具中，類神經網路(artificial neural network, ANN)由於擁有捕抓變數間非線性關係、不需假設資料的形態與學習動態系統的特性，且能找出資料中所隱含的模式或特徵(pattern)，使得其在建構變數間為非線性資料型態關係的預測問題中有優越的表現(Zhang, Patuwo and Hu, 1998)，並且廣泛的應用在各預測領域中。而在各種模式中，又以倒傳遞類神經網路(back-propagation neural network, BPN)最常被討論與使用(Vellido, Lisboa and

Vaughan, 1999; Zhang, Patuwo and Hu, 1998)。由於財務時間序列預測問題的複雜性，以及類神經網路於建構預測模式的優異展現，因此本研究將使用 BPN 建構財務時間序列預測模式。

而在建構財務時間序列預測模式時，最常面臨之問題就是財務時間序列資料包含許多的雜訊(Deboeck, 1994; Yaser and Atiya, 1996)。由於資料中存在雜訊時，會使得預測模式在訓練過程中學習到包含雜訊的資訊，而產生過度配適或配適不足的問題，進而在測試階段喪失預測績效的可靠性與準確度，因此如何偵測以及去除財務時間序列資料中的雜訊，以降低其對 BPN 預測結果的影響，進而提升預測績效，就成為重要的研究議題。

獨立成份分析(independent components analysis, ICA)是一個新穎的訊號處理技術，主要用於處理未知來源分離(blind source separation)的問題(Hyvärinen and Oja, 2000; Lee, 1998)，其目的是在沒有任何有關訊號混合機制的事前資訊下，從觀察到的混合訊號(mixing signals)中找出潛在的來源訊號(latent source signals) 這些潛在來源訊號在 ICA 中稱為獨立成份(independent component, IC)(Hyvärinen, Karhunen and Oja, 2001)。由於時間序列的雜訊資訊往往無法由觀察到的混合訊號中獲得，因此利用 ICA 將混合訊號分離出潛在來源訊號後，透過觀察來源訊號之特徵擷取表現，將有助於找出資料中的雜訊。由於 ICA 具有前述特性，因此本研究將利用 ICA 做為濾除時間序列中雜訊的工具，期望藉此增加預測模式之精準度。

綜上所述，本研究導入 ICA 訊號分離技術於財務時間序列預測問題上，提出一結合獨立成份分析與倒傳遞類神經網路之兩階段時間序列預測模式之建構程序，首先利用 ICA 分離出個別獨立的潛在來源訊號(即獨立成份)，接者去除代表資料中雜訊的獨立成份之後，將保留的獨立成份重建時間序列資料，得到濾除雜訊後之預測變數，再利用 BPN 以濾除雜訊後的預測變數建構預測模式。本研究所提方法能減少 BPN 在模式建構時，因受雜訊影響而產生過度配適或配適不足的問題，提升 BPN 模式預測結果的精確度。為驗證所提方法之有效性，本研究將以日經 225(Nikkei 225)現貨開盤指數進行實證研究，並與直接使用 BPN 模式、隨機漫步(random walk)及其他常用濾除雜訊方法-簡單移動平均(simple moving average)及小波框架(wavelet frame)之預測結果進行比較。

本文各節區分如下：第二節將回顧獨立成份分析及類神經網路之相關文獻；第三節針對獨立成份分析及倒傳遞類神經網路進行簡單的介紹；第四節

則說明所提之兩階段預測模式；第五節將本研究所提出之預測模式透過 Nikkei 225 指數開盤指數進行實證分析，並說明其結果；最後第六節則為本研究之結論與建議。

2. 文獻回顧

2.1 獨立成份分析

獨立成份分析(ICA)模式已廣泛應用於醫學訊號處理、語音訊號處理、特徵萃取及人臉辨識的問題中(Bartlett, Movellan and Sejnowski, 2002; Beckmann and Smith, 2004; James and Gibson, 2003; Jang, Lee and Oh, 2002; Jung et al., 2001; Kim et al., 2004; Lin et al., 2004; Vigario et al., 2000)。Ikeda and Toyama (2000)應用 ICA 於磁振造影(MEG)資料，先利用因素分析(factor analysis)做前處理，使磁振造影訊號非相關，再利用 ICA 於經過前處理後之訊號，得到磁振造影資料的獨立成份。如此真實腦波訊號、機械上感應器之雜訊及人工訊號的特徵將個別的被凸顯在獨立成份中，進而提供觀察者更有意義的訊息。Déniz, Castrillón and Hernández (2003)結合 ICA 與支援向量機(support vector machine, SVM)進行人臉辨識，先利用 ICA 對訓練影像估計出獨立成份，並以獨立成份作為人臉特徵，之後利用 SVM 建立分類模式以進行人臉分類。Visser and Lee (2003)利用 ICA 於語音增強上，假設觀察訊號是由背景噪音(background noise)、干擾來源(interfering point source)及發音來源訊號(speaker signal)所混合而成，先透過 ICA 於觀察訊號中找出獨立成份，並於獨立成份中將干擾來源去除，之後再利用說話者活動偵測(speaker activity detection)技術去除背景噪音，以達到語音增強的目的。

由前述文獻可知，ICA 主要用於訊號處理及特徵辨識等領域，而於財務預測領域之文獻則相對不足，仍有極大之發展空間。Back and Weigend (1997)應用 ICA 模式分析股價指數走勢的特徵，使用 ICA 模式針對日本股票交易所前 28 大公司每日收盤指數的時間序列資料擷取獨立成份，之後將所得獨立成份進行排序，實驗結果顯示，排序在前的獨立成份能抓住股價走勢的特徵。Kiviluoto and Oja (1998)應用 ICA 模式分析影響某零售商現金流量的潛在因素，經由對獨立成份的分析可知此零售商現金流量的多寡主要受假日、季節

及其他零售商競爭的影響。Oja, Kiviluoto and Malaroiu (2000)利用 ICA 模式進行匯率預測，先利用 ICA 針對匯率資料獲得獨立成份，再透過線性與非線性轉換過濾獨立成份中的雜訊，之後利用自我迴歸(autoregressive, AR)模式預測下一期各獨立成份值，最後這些預測之獨立成份被用來產生下一期的匯率預測值，結果顯示結合 AR 與 ICA 之預測結果較直接使用 AR 為佳。

2.2 類神經網路

類神經網路是目前發展極為迅速的一門學科，目前已廣泛的應用於各領域，包括品質管制、資料探勘、信用評等、財務預警及股市、期貨指數預測等 (Andrisevic et al., 2005; Lee and Chen, 2002; Lee and Chiu, 2002; Lee and Chen, 2005; Tsaih, Hsu and Lai, 1998; Yagci et al., 2005; Zhang, Patuwo and Hu, 1998)。而在類神經網路眾多的模式中，倒傳遞類神經網路(BPN)為應用最普遍的模式(Zhang, Patuwo and Hu, 1998)，並且根據 Vellido, Lisboa and Vaughan (1999)指出自 1992 至 1998 年間，使用類神經網路於商業相關問題研究中，BPN 之比率高達 78%，因此本研究亦將使用 BPN 建構財務時間序列預測模式。Grudnitski and Osburn (1993)以貨幣供給月成長率、期貨價格波動度，以及大型投機、避險者與小型交易者佔月底市場淨部位之比率作為 BPN 之輸入變數，對 S&P 500 與黃金期貨未來每個月的價格波動進行預測。研究結果發現，BPN 對於預測下個月 S&P 500 和黃金期貨的價格變動，分別達到 75%和 61%的準確率，且每期平均報酬率為 17.04%和 16.36%。Greg (2001) 使用美國與加拿大利率收益率差(yield spread)、90 天商業本票利率(90-day corporate paper rate)、廣義與狹義貨幣供給額(M2 and M1)及多倫多交易所 300 股票指數(Toronto stock exchange 300 index)作為輸入變數，利用 BPN、線性迴歸、隨機漫步、AR 及指數平滑法預測加拿大國內生產毛額(GDP)的成長率，並以 1968 年第一季至 1999 年第二季之資料進行實證，結果顯示，預測每季 GDP 成長率時，隨機漫步預測結果最佳，但在預測每年 GDP 成長率時，BPN 之預測結果明顯優於線性迴歸與時間序列模式。

Ilan, Min and Sadowski (2001)以前一期(one-step)或前數期(multi-step)資料作為輸入變數；利用 BPN、溫特斯指數平滑法(Winters exponential smoothing)、ARIMA 及多元迴歸模式預測 1978 年 1 月至 1995 年 4 月美國零售業之月營業額。結果顯示，在不穩定的經濟條件下，以前數期資料建構之

BPN 模式預測表現最佳。Kanas and Yannopoulos (2001)以股利及交易量作為輸入變數，使用 BPN 及線性迴歸模式，預測 1980 年 1 月至 2000 年 12 月道瓊指數(Dow Jones index)與金融時報指數(Financial times index)，結果顯示 BPN 在不同的預測標的物上皆有較佳的預測結果。Cao, Leggio and Schniederjans (2005)使用公司報酬、公司市值、風險指標(beta)及帳面-市值比(book-to-market)作為輸入變數，利用迴歸與 BPN 模式進行新興市場股價指數預測，並以 1999 年 1 月至 2002 年 12 月 367 家上市公司於上海股票交易所之每日收盤指數進行實證研究，結果發現 BPN 在各變數組合中皆有較好的預測效果。

由於預測變數存在雜訊或離群值時，往往對於 BPN 預測模式之精確度及可靠性造成負面影響，因此在建構 BPN 預測模式時，預測變數中如果存在雜訊或離群值，將大幅影響該預測模型的預測效果，因此也吸引學者們研究的興趣(Chen and Romagnoli, 1998; Chen, Bandoni and Romagnoli, 1998; Christensen and Pearson, 1992; Muñoz and Muruzábal, 1998)。Yu, Wang and Jiang (1996)指出在 BPN 的學習過程中，殘差的線性轉換及過度學習的特性，使得資料離群值對於預測模式的準確率有很大的影響，因此利用閾值(threshold)提出針對殘差的非線性轉換，並將學習過程分為正常樣本及離群值樣本的兩階段學習過程，降低離群值對預測模式的影響，並提高整體預測準確率；Park, Joo and Choi (2000)則利用信賴區間法濾除預測變數內的雜訊，先對訓練資料建立信賴區間，再透過去除位於預測值信賴區間外的觀測值後，得到過濾後之預測變數，最後則利用 BPN 於濾除雜訊後的預測變數建構預測模式，並以凝血劑劑量時間序列資料進行實證研究，結果發現經過前處理的預測變數能提高 BPN 的預測準確率。

3. 研究方法

3.1 獨立成份分析

令 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]^T$ 為由 \mathbf{x}_i 所組成維度為 $M \times N$ ($M \leq N$) 的混合訊號矩陣，其中 \mathbf{x}_i 為維度 $1 \times N$ 的時間序列訊號，一個基本的 ICA 模式可以表示

為(Hyvärinen and Oja, 2000)：

$$\mathbf{X} = \mathbf{AS} = \sum_{i=1}^M \mathbf{a}_i \mathbf{s}_i \quad (1)$$

其中 \mathbf{a}_i 是維度 $M \times M$ 的未知混合矩陣(mixing matrix) \mathbf{A} 的第 i 個行向量(column vector)，向量 \mathbf{s}_i 是維度 $M \times N$ 的未知來源矩陣(source matrix) \mathbf{S} 的第 i 個列向量(row vector)，亦即無法由混合訊號矩陣 \mathbf{X} 所直接觀測的潛在來源訊號。在假設來源訊號 \mathbf{s}_i 互為統計獨立的情況下，為了估計 \mathbf{s}_i ，ICA 的方法是找到一個維度 $M \times M$ 的解混合矩陣(demixing matrix) \mathbf{W} ，將所觀察到的混合訊號矩陣 \mathbf{X} 進行轉換，以產生維度為 $M \times N$ 的矩陣 \mathbf{Y} ，亦即

$$\mathbf{Y} = [\mathbf{y}_i] = \mathbf{WX} \quad (2)$$

其中， \mathbf{y}_i 是矩陣 \mathbf{Y} 的第 i 個列向量。向量 \mathbf{y}_i 之間必須盡可能的互為統計獨立，並且被稱為獨立成份(IC)。當解混合矩陣 \mathbf{W} 為混合矩陣 \mathbf{A} 的反矩陣時，即 $\mathbf{W} = \mathbf{A}^{-1}$ ，這些獨立成份將可用於估計潛在來源訊號 \mathbf{s}_i 。

主成份分析 (principal component analysis, PCA) 常被用來與 ICA 做比較。ICA 被視為是 PCA 的延伸，這兩者主要不同處在於主成份分析使用的是二階統計量(second order statistics)，利用共變異矩陣來找出主成分(principal component)，主成分的投影方向為為資料變異最大的方向。而 ICA 則是使用高階統計量 (high order statistics)，利用人工智慧或最佳化的方法來找獨立成份，獨立成份的投影方向是資料獨立性最大的方向。想像一個構成菱形分佈的二維混合訊號，PCA 所投影的方向為此菱形兩個對角線的方向(變異最大)，而 ICA 之投影方向則為此菱形兩個邊界的方向(獨立性最大)。有關 PCA 與 ICA 關係的說明，可參考 Hyvärinen, Karhunen and Oja (2001)。

ICA 模式有一個基本的假設或限制，即各獨立成份間必須統計獨立(Hyvärinen and Oja, 2000)，請注意前述假設並不代表觀察到的訊號或隨機變數資料互為統計獨立。因此，ICA 可以被描述成一個最佳化問題，將各獨立成份的獨立性量測當成目標函數，經由最佳化的技術找出解混合矩陣 \mathbf{W} 。亦即，解混合矩陣 \mathbf{W} 可以由一個非監督式(unsupervised)的演算法決定，該演算法的目標即是使 IC 間的統計獨立性最大。由於 IC 具有非高斯分配(non-Gaussian distribution)，隱含各 IC 之間互為統計獨立(Hyvärinen, Karhunen

and Oja, 2001)，因此 IC 之非高斯特性(non-Gaussianity)可做為 ICA 演算法的目標函數。而在可用於量測獨立成份非高斯特性之方法部份，常見的有高階統計量法(high order cumulants)，共同資訊法(mutual information)及負熵法(negentropy)，其中又以負熵法最常被討論與使用(David and Sanchez, 2002; Hyvärinen, Karhunen and Oja, 2001)。

負熵法是延伸資訊理論中的熵(entropy)的定義所提出的方法，在資訊理論中熵為一資訊量的測量單位，用以表示隨機變數之觀察值的資訊程度，當隨機變數的結構性越差越亂(即越無法預測)時，其熵會越大。令隨機向量 \mathbf{y} 之機率密度函數為 $p_{\mathbf{y}}(\eta)$ ，則 \mathbf{y} 的熵 H 可定義為

$$H(\mathbf{y}) = -\int p_{\mathbf{y}}(\eta) \log p_{\mathbf{y}}(\eta) d\eta$$
 (Cover and Thomas, 1991)。根據資訊理論，在所有具有相同變異數的隨機分配中，呈高斯分配的變數會具有最大的熵。因此若變數具非高斯分配，則熵值會較其具有高斯分配時來的小(Hyvärinen, Karhunen and Oja, 2001)。由此延伸，可以對熵 H 的定義做一修正當作非高斯的測量值。熵 H 修改後的定義稱為負熵 J (negentropy)，定義如下(Hyvärinen, Karhunen and Oja, 2001)：

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}) \quad (3)$$

其中 \mathbf{y}_{gauss} 表示和 \mathbf{y} 具有相同變異數的高斯變數向量。負熵的值永不為負值，即 $J(\mathbf{y}) \geq 0$ ，且只有 \mathbf{y} 為高斯分配時 $J(\mathbf{y})$ 其值才會為零，所以 ICA 的目標函數為最大化隨機向量 \mathbf{y} 的負熵，即 *Maximize* $J(\mathbf{y})$ 。

然而，負熵最大的問題就是計算太過複雜，因為根據定義，要估計出負熵就必須先估計 \mathbf{y} 的機率密度函數 $f(\mathbf{y})$ 。為了解決這這個問題，Hyvärinen (1999)發展出負熵的近似函數

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2 \quad (4)$$

其中 v 是平均數為零且變異數為 1 的高斯分配之隨機變數。 G 可為任何非二次方函數(non-quadratic function)，因為 G 若為二次方函數， $J(y)$ 必為零。

在眾多用來求解 ICA 模式的演算法中，Hyvärinen (1999)所提出 FastICA 演算法由於具有良好的運算效率，並且是處理 ICA 問題上最常使用的演算法之一，因此本研究將使用該演算法處理從財務時間序列資料中分離出獨立成份的工作。

3.2 倒傳遞類神經網路

類神經網路的原始想法及基本構造皆和神經生物學中的神經元(neuron)構造相似。一般來說，類神經網路中神經元的作用和真實的神經元相仿，是將外界的輸入值依相對重要性的不同給予權數(weights)，加總後再經由神經元中的激發函數(activation function)轉換，而得到一輸出值，其構造如圖 1 所示：

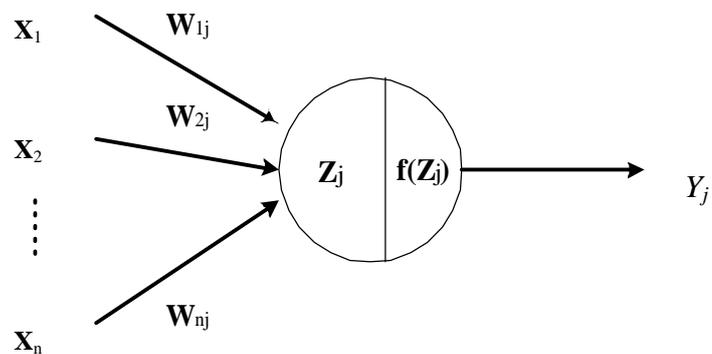


圖 1 神經元之構造

在圖 1 中， $X_i, i=1, 2, \dots, n$ 代表輸入值； W_{ij} 代表連接鍵的權數；

$Z_j = W_{ij} \times X_i$ 代表加權和； $f(Z_j)$ 則是代表一激發函數；最後 Y_j 則為神經元的輸出值。

類神經網路有許多不同的網路型態，其中倒傳遞類神經網路為應用最廣的模式之一(Vellido, Lisboa and Vaughan, 1999)。倒傳遞類神經網路乃隸屬於監督式學習(supervised learning)網路模式的一種，其資料是以順向(forward)

之方式向前傳遞。一般而言，倒傳遞類神經網路之結構包含三層：輸入層(input layer)、隱藏層(hidden layer)及輸出層(output layer)，其中隱藏層之數目可以是一層或多層。在輸入層部分，神經元的數目即為我們所欲輸入的變數個數，而輸出層中神經元的輸出結果則為網路最後的輸出值。一個簡單倒傳遞類神經網路模式的結構圖示可整理如圖 2。

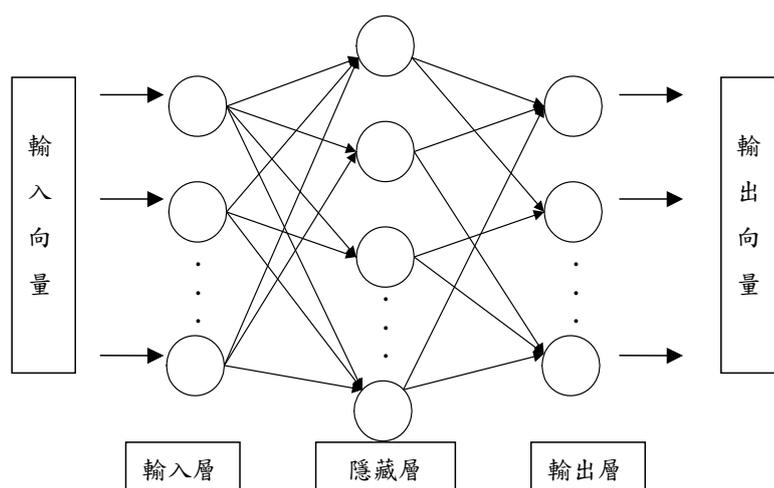


圖 2 倒傳遞類神經網路之構造

在圖 2 中，類神經網路的第一層為輸入層；第二層為隱藏層；第三層為輸出層，層與層間有連接鍵相互連接，同層間的神經元則沒有連接鍵。每一連接鍵均含一相對應之權數(W_{ij})，代表輸入訊息的重要程度。

倒傳遞類神經網路的演算流程是由輸入層將資料傳給隱藏層，經計算與轉換後送到輸出層，計算與實際值的差距後，再回傳給隱藏層去修正連接鍵的權數，使得輸出層之輸出值與真實值的均方誤差最小，此訓練過程會重複至誤差收斂至假設之條件為止。倒傳遞類神經網路模式使用的激發函數是 Sigmoid function，其型式為 $f(x) = (1 + \exp(-x))^{-1}$ ，而最常用的訓練過程則是最陡坡降法 (gradient steepest descent method)，是用來調整權數變動的幅度 ΔW_{ij} ，可以式 5 表示：

$$\Delta W_{ij} = -\eta(\partial E / \partial W_{ij}) \quad (5)$$

其中 η 為學習率； $E = \frac{1}{2} \sum (T_j - A_j)^2$ 為誤差函數； T_j 表實際值； A_j 為網路輸出值。

類神經網路對於變數的選取有較大的自由度，沒有如迴歸分析般的限制，研究者須以文獻探討、專家意見判斷或經由統計方法處理，來選取輸入層之輸入變數。而在輸入層之變數決定後，對於網路結構中的隱藏層數目、隱藏層中神經元數目、訓練的學習率大小等，都需要以主觀邏輯判斷，或以不同組合加以測試，並沒有一定理可資遵循，有關上述參數如何決定及其餘神經網路模式之種類與應用可參考 Anderson and Rosenfeld (1988) 與 Tang and Fishwick (1993) 等人之著作。

4. 結合 ICA 與 BPN 預測模式之建構

本研究提出一結合獨立成份分析與倒傳遞類神經網路之財務時間序列預測模式。此預測模式之建構可以分成利用 ICA 對預測變數去除雜訊的前處理，及使用 BPN 以去除雜訊後的預測變數建構預測模式等兩部分。首先在 ICA 之前處理部分，在將 M 個維度為 $1 \times N$ 的預測變數 \mathbf{x}_i 組成混合矩陣後，利用 ICA 模式從矩陣 \mathbf{X} 中估計出 M 個維度為 $1 \times N$ 的獨立成份 (\mathbf{y}_i) 及維度為 $M \times M$ 的解混合矩陣 \mathbf{W} 。接著，為找出代表雜訊的獨立成份，本研究利用 Cheung and Xu (2001) 提出之測試接受法 (testing-and-acceptance, TnA) 對獨立成份在趨勢特徵擷取之表現進行評估。TnA 法是基於資料還原 (data reconstruction) 的排序方法 (ICA 中還原的公式如下面式 (6) 所示)，使用 relative hamming distance (RHD) 還原誤差值為指標，衡量還原後的序列資料與原始序列間的趨勢相似程度。RHD 值越小代表兩個序列資料之趨勢越相似，當趨勢完全一致時其值為 0，完全相反時則為 4。

使用 TnA 法排序的步驟為先計算每一個獨立成份各自進行還原後與原始資料之間的 RHD 值，有最小 RHD 值的獨立成份是最能夠捕捉到原始序列資料中主要趨勢結構的獨立成份，因此被排序為第一個；接著將排序第一的

獨立成份與其餘的獨立成份兩兩組合成一組後，再次進行還原並計算 RHD 值，同樣有最小 RHD 值的獨立成份組合是表現最佳的組合，組合內排序第一的獨立成份外的另一個獨立成份將是排序第二的獨立成份，以此類推，重複此組合還原及比較選擇的程序，一次排序一個獨立成份，直到所有的獨立成份被排序完成。

舉例說明，圖 3 為 4 個維度均為 1×794 的財務時間序列訊號，將其組成維度為 4×794 的混合訊號矩陣後，利用 ICA 模式於此混合訊號矩陣後，可以估計 4 個維度為 1×794 的獨立成份，此 4 個獨立成份(IC1-IC4)的波形圖列示於圖 4。其次利用 TnA 法對 4 個獨立成份的特徵擷取表現進行排序，表 1 為 TnA 法排序過程的 RHD 還原誤差值，排序過程為先對 4 個 IC 各別還原，由表 1 中可知，4 個 IC 中以 IC1 的 RHD 值 1.7531 最低，因此被排序為第一個。接著將 IC1 與其餘的獨立成份兩兩組合成一組後再進行還原，同樣由表 1 中可以發現，各組合的還原結果中以 IC1 與 IC3 組合的 RHD 值 1.3346 最低，因此除了 IC1 為排序第一外，IC3 將排序第二。重複此一操作，最後之排序結果依序為 IC1、IC3、IC2 及 IC4，其中以 IC1 最能夠捕捉原始資料結構中的趨勢特徵，IC4 的特徵擷取表現則較差。而在 4 個獨立成份的排序中，決定了排名前三的獨立成份後，剩下的必定排名最後，因此第 4 次之比較在實際操作中是不需要的。

在獨立成份排序後，我們將選擇排序在前的數個獨立成份，代表預測變數資料中的主要趨勢結構，其餘的獨立成份則代表預測變數資料中的雜訊。在保留代表主要趨勢結構的獨立成份及去除(將其值設為 0)代表雜訊的獨立成份並且還原後，還原的預測變數資料中將存在最少的雜訊資訊。還原的式子如下：

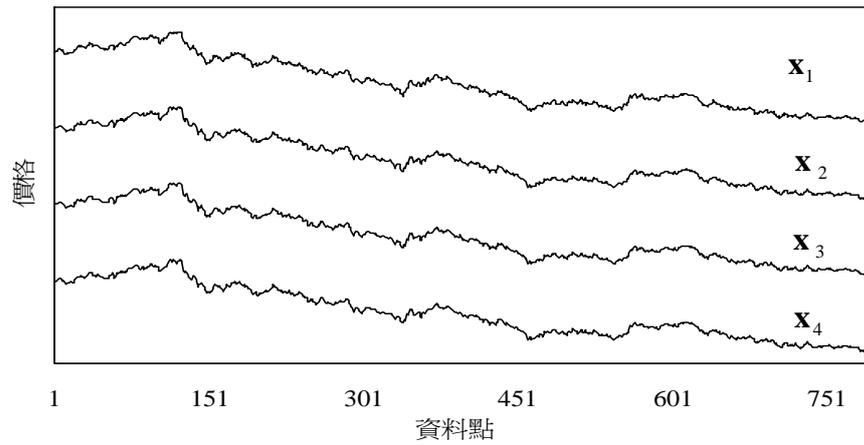


圖 3 四個維度均為 1×794 的財務時間序列訊號

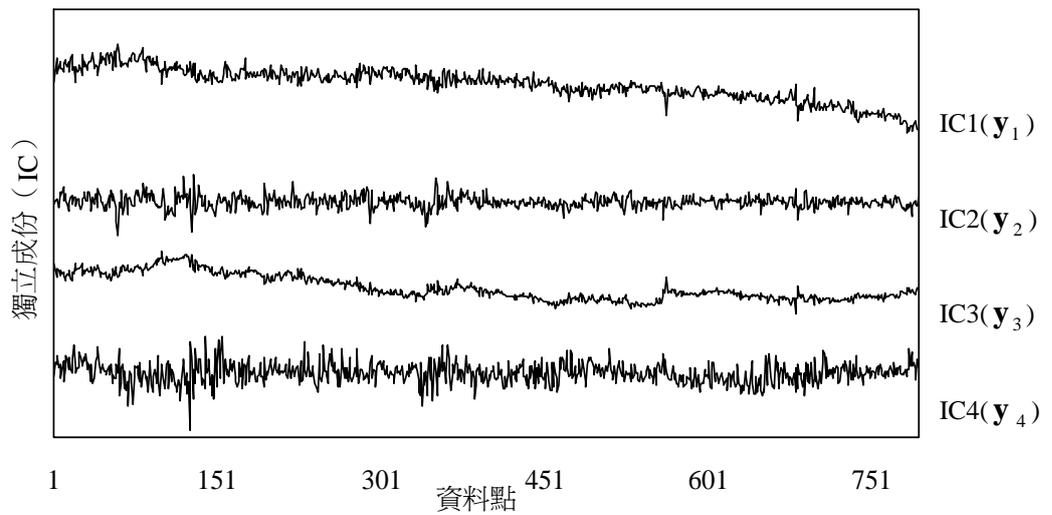


圖 4 圖 3 之財務時間序列訊號的 4 個獨立成份

$$\hat{\mathbf{X}} = \sum_{i=1, i \neq k}^M \mathbf{a}_i \mathbf{y}_i, 1 \leq k \leq M \quad (6)$$

其中 $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_M]^T$ 為維度 $M \times N$ 的還原預測變數矩陣， $\hat{\mathbf{x}}_i$ 是還原後的預測變數， \mathbf{a}_i 為混合矩陣 \mathbf{A} 的第 i 個行向量(column vector)， $\mathbf{A} = \mathbf{W}^{-1}$ ， \mathbf{y}_i 是第 i 個獨立成份， k 是代表雜訊的獨立成份個數。

表1 測試接受法排序過程的RHD還原誤差值

比較次數	還原之獨立成份 (IC)	RHD 還原誤差值
第 1 次	IC1	1.7531
	IC2	1.8528
	IC3	1.9779
	IC4	1.9044
第 2 次	IC1、IC2	1.7137
	IC1、IC3	1.3346
	IC1、IC4	1.7373
第 3 次	IC1、IC3、IC2	0.7662
	IC1、IC3、IC4	1.4294*
第 4 次	IC1、IC3、IC2、IC4	0

*在第三次比較時，加入 IC4 還原所得 RHD 還原誤差值反較第二次比較所得結果(IC1、IC3)之誤差為高(1.3346->1.4294)；反之，在加入 IC2 後所得 RHD 還原誤差值則降低(1.3346->0.7662)，可知 IC4 為雜訊，因其加入後 RHD 還原誤差值不降反升。

在還原的過程中，重要的工作是選擇適當的獨立成份個數，達到最佳的雜訊去除效果，因選擇過多 IC 會使得雜訊的濾除效果不佳，反之選擇太少則可能會喪失預測變數的趨勢結構。本研究根據依序還原排序第一的獨立成份、前兩個獨立成份及所有獨立成份的 RHD 值之變化，選擇適當的獨立成份個數。由於還原的獨立成份越多，還原的訊號會與原始訊號越相似，所以

此一 RHD 值的變化將是由大到小的排序。而選擇的準則為若 RHD 值在某個獨立成份加入還原組合後之下降幅度不大，表示此一獨立成份包含較少的趨勢特徵資訊，加入還原後並無法使還原後訊號更近似原始訊號，所以此一及排序在其之後的獨立成份將被用來代表原始資料中的雜訊；反之，若 RHD 值在某個獨立成份加入還原組合後顯著下降，則此一獨立成份代表預測變數中的主要趨勢結構。除此之外，若有 M 個獨立成份，則至多選擇 $M - 1$ 個獨立成份代表預測變數中的主要趨勢結構，或是最少會選擇一個獨立成份(亦即排序最後獨立成份)代表預測變數中的雜訊，因為財務時間序列資料存在雜訊(Deboeck, 1994; Yaser and Atiya, 1996)。

以圖 4 中 4 個獨立成份為例，圖 5 為還原排序第一的獨立成份(即 IC1)、前二個獨立成份(即 IC1 及 IC3)到前四個獨立成份(即 IC1-IC4)的 RHD 值，這些 RHD 值可以從 TnA 法的排序過程中(即表 1 中)整理而得。由圖 5 可知 RHD 值從還原第一個到還原前 4 個 IC 均明顯下降，雖然如此，由於最少會選擇排序最後的獨立成份代表預測變數資料中的雜訊，因此前三個獨立成份(即 IC1、IC3 及 IC2)將被用來代表預測變數中的主要趨勢結構，排序最後的 IC4 則代表預測變數中的雜訊部分。為說明代表主要趨勢 IC 與代表雜訊之 IC 在還原結果上之差異，圖 6 為以圖 2 中 x_1 前 201 個資料點為例，還原代表主要趨勢結構的 IC(即 IC1、IC3 及 IC2)以及代表雜訊的 IC(即 IC4)之結果。由圖 6 可以發現，還原 IC1、IC3 及 IC2 的結果會與原始訊號(即 x_1)在主要趨勢及結構(shape)上非常相近，而代表雜訊的 IC 之還原結果除了在趨勢上與原始訊號不符外，並且在整體結構的資訊貢獻度上明顯不佳。

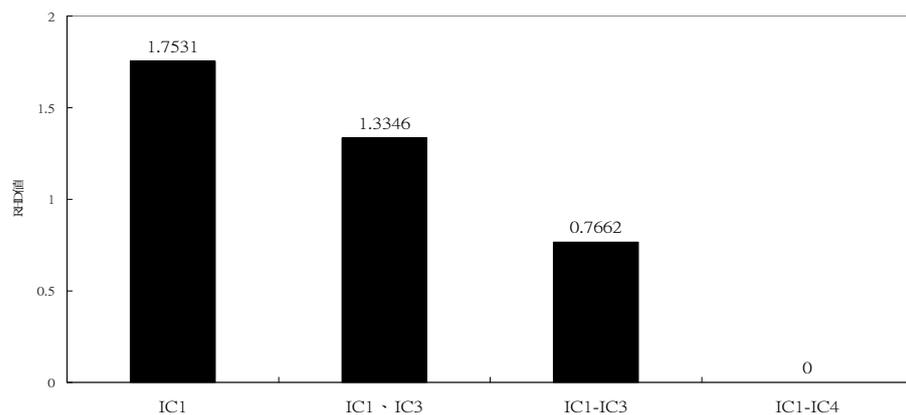


圖 5 獨立成份依排序前後組合並還原的 RHD 值

在利用 ICA 模式得到濾除雜訊之預測變數後，接著將利用 BPN 以前述步驟所得之變數建構預測模式。有關 BPN 在建構預測模式需決定的網路架構及參數，由於目前尚無特定的定理可以遵循，研究者通常以試誤法 (trial-and-error) 決定，因此本研究參考 Lee and Chiu (2002) 及 Lee and Chen (2002) 之研究，使用試誤法決定本研究中 BPN 參數之設定，最後本研究所提之結合 ICA 與 BPN 的實證結果，將說明於第五節中。

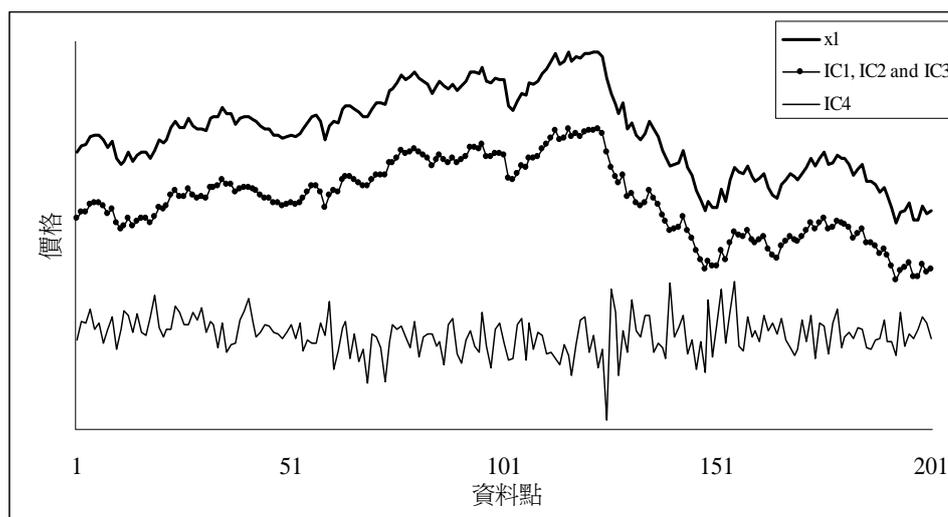


圖 6 還原代表主要趨勢結構的 IC (即 IC1、IC3 及 IC2) 以及代表雜訊的 IC (即 IC4) 之結果 - 以圖 3 中 x_1 前 201 個資料點為例

5. 實證結果

5.1 資料收集及績效衡量準則

為驗證本研究所提結合 ICA 與 BPN 預測模式之預測能力，將以 Nikkei 225 為研究標的，對 Nikkei 225 之開盤指數進行預測，並將所提方法之預測結果與直接使用 BPN 及隨機漫步模式的預測結果進行比較，以驗證所提方法的有效性。Nikkei 225 實證資料採用之期間為 1999 年 10 月 4 日至 2004 年 9 月 30 日，共計五年 1144 個交易日之資料，資料中訓練樣本的部分為 1999

年 10 月 4 日至 2003 年 3 月 27 日共 794 筆資料(佔實證資料之 69.41%)，另保留 2003 年 3 月 28 日至 2004 年 9 月 30 日共 350 筆資料(佔資料期間 30.59%) 作為測試樣本，圖 7 為實證資料之走勢圖。

而在 Nikkei 225 開盤指數使用之預測變數方面，由於期貨市場反映市場資訊的速度較現貨為快，並且其收盤成交指數會率先反映盤後訊息(Lee and Chen, 2002, Lee and Chiu, 2002)，因此本研究選擇 3 個以 Nikkei 225 指數為標的期貨指數，分別為大阪證券交易所(Osaka securities exchange, OSE)、新加坡衍生性商品交易所(Singapore exchange derivatives trading, SGX)和芝加哥商業交易所(Chicago mercantile exchange, CME)之 Nikkei 225 期貨前一日收盤指數，及前一日現貨收盤指數等 4 個預測變數，4 個預測變數的走勢圖如圖 8 所示。

在預測績效的評估指標方面，以均方根誤差(root mean squared error, RMSE)、平均絕對離差(mean absolute deviation, MAD)、平均絕對百分比誤差(mean absolute percentage error, MAPE)、方向對稱性(directional symmetry, DS)、正確上漲趨勢率(correct up-trend, CP)及正確下跌趨勢率(correct down-trend, CD)等 6 個常用之指標，做好評估預測模式精確度之指標，指標之計算式如表 2 所示。在上述指標中，RMSE 與 MAD 為衡量預測值與實際值誤差之指標，為了避免正負值會相互抵銷，因此分別為平方和再開根號及絕對值的型態；MAPE 則為用於衡量預測值與實際值的相對誤差，不受實際值與預測值之大小與單位影響，能夠客觀獲得實際值與預測值間之差異程度；此 3 指標(即 RMSE、MAD 及 MAPE)均是越小越代表預測值與實際值越接近。而 DS、CP 及 CD 則是用來衡量預測的漲跌方向與實際方向的準確率，其中 DS 衡量的是整體(亦即不管上漲或下跌)的趨勢正確率指標，而 CP 與 CD 則是分別用來評估上漲及下跌時的趨勢正確率指標，此 3 指標之值越大越代表預測值與實際值越接近。

5.2 結合 ICA 與 BPN 模式之預測結果

本節將驗證結合 ICA 與 BPN 模式之預測能力，並以 Nikkei 225 開盤指數進行實證研究。為了比較 Nikkei 225 開盤指數之預測結果，本研究將結合 ICA 與 BPN 模式(簡稱：結合模式)之預測結果與直接使用 BPN(直接 BPN 模式)及隨機漫步模式之結果進行比較，其中隨機漫步模式為直接利用前一日現

貨收盤指數為今日現貨開盤指數之預測值，直接 BPN 模式為直接使用 BPN 模式進行預測，不加入 ICA 作為前處理，而利用 ICA 對 Nikkei 225 開盤價預測變數去除雜訊的前處理部分也已於第四節中說明，不再重複。不過由 ICA 實證之結果可知(詳見表 1)，ICA 確實能濾除財務時間序列中之雜訊，適合作為類神經網路預測模式建構前，濾除資料中雜訊之前處理工具。

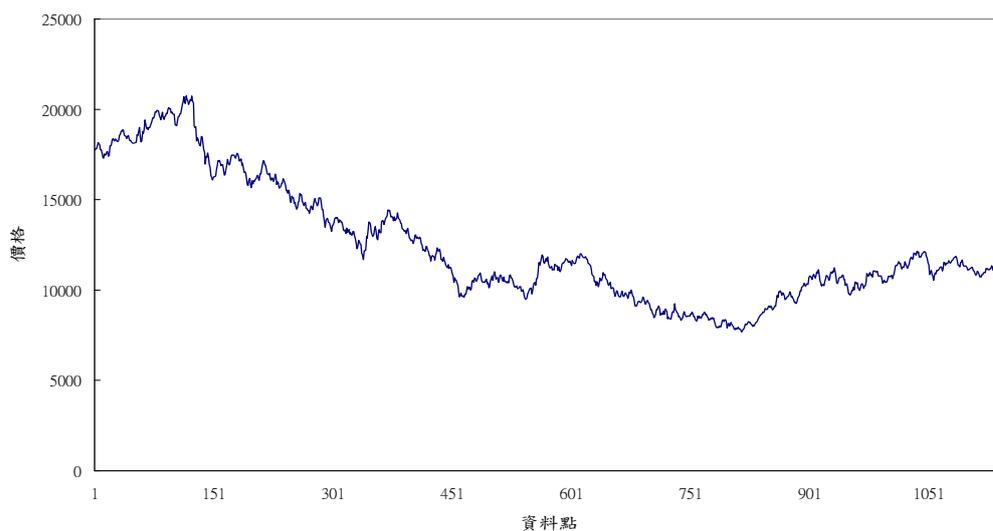


圖 7 Nikkei 225 開盤價走勢圖

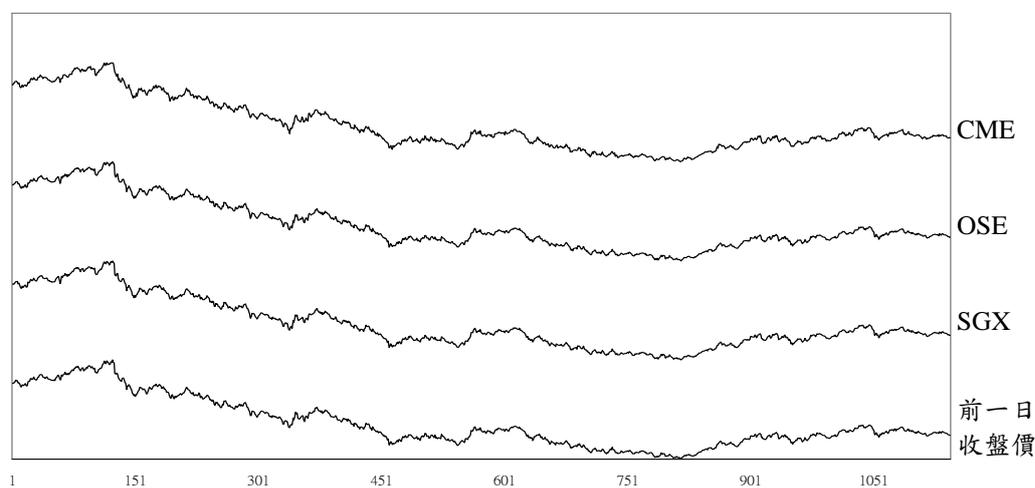


圖 8 Nikkei 225 開盤價預測變數走勢圖

表 2 績效衡量指標及其公式

指標	公式
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^N (T_i - A_i)^2}{N}}$
MAPE	$MAPE = \frac{\sum_{i=1}^N \left \frac{T_i - A_i}{T_i} \right }{N}$
MAD	$MAD = \frac{\sum_{i=1}^N T_i - A_i }{N}$
DS	$DS = \frac{100}{N} \times \sum_{i=1}^N d_i, \text{ 其中 } d_i = \begin{cases} 1 & (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$
CP	$CP = \frac{100}{N_1} \times \sum_{i=1}^N d_i, \text{ 其中 } d_i = \begin{cases} 1 & (T_i - T_{i-1}) > 0 \text{ and } (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$
CD	$CD = \frac{100}{N_2} \times \sum_{i=1}^N d_i, \text{ 其中 } d_i = \begin{cases} 1 & (T_i - T_{i-1}) < 0 \text{ and } (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$

*上面所有式子中， T 代表實際值， A 代表預測值， N 代表資料筆數， N_1 代表資料上漲筆數， N_2 代表資料下跌筆數。

在預測結果部分，首先在直接 BPN 模式方面，由於本研究使用 4 個預測變數，因此 BPN 的輸入層節點數為 4。而在隱藏層部分，因為單一隱藏層已足夠描述複雜的非線性系統 (Hornik, Stinchcombe and White 1989; Zhang, Patuwo and Hu, 1998)，因此採用單一隱藏層設定，隱藏層內之節點數無一定的選用標準，本研究根據 Lee and Chen (2002) 的建議，使用 $2n \pm 2$ 的方式選擇隱藏層節點數，其中 n 為輸入層節點數，是以本研究選取 5 種不同隱藏層節點數(5, 6, 7, 8 及 9)。在學習率部分，學習率介於 0 與 1 之間，學習率較小，會造成緩慢的收斂過程；學習率較大，雖然會提高搜尋的效率，但目標函數會產生不穩定的震盪現象，因此學習率的設定需要配合資料類型以試誤方式決定，本研究使用 3 種學習率(0.020, 0.025 及 0.030)進行測試。整體而言，本

研究使用使用 5 種隱藏層節點數搭配 3 種學習率 5 種共 15 種參數組合來測試最佳的 BPN 模式。在網路學習法則方面，參考 Lee and Chen (2002) 的建議，激發函數與訓練過程採用 Sigmoid function 與最陡坡降法，迭代次數為 10000 次。

表 3 直接 BPN 在不同參數組合下的預測結果

隱藏層節點數	學習率	訓練樣本 RMSE*	測試樣本 RMSE*
6	0.020	0.004215	0.004155
	0.025	0.004160	0.004029
	0.030	0.004122	0.004220
7	0.020	0.006262	0.005269
	0.025	0.004311	0.004241
	0.030	0.004195	0.004390
8	0.020	0.008794	0.008820
	0.025	0.004306	0.004237
	0.030	0.004140	0.004156
9	0.020	0.004297	0.004186
	0.025	0.004141	0.004052
	0.030	0.004142	0.004152
10	0.020	0.009236	0.009428
	0.025	0.004163	0.004171
	0.030	0.004161	0.004218

*訓練與測試樣本 RMSE 皆為目標變數重新尺度化至[0.15, 0.85]間之結果

直接 BPN 模式在不同參數組合下的訓練及測試結果列示於表 3。由表 3 可知當隱藏層神經元數為 9 且學習率為 0.025 時有最小的測試資料 RMSE 值，圖 9 為 BPN 模式最佳模式訓練樣本之 RMSE 趨勢圖，由圖中可知，此 BPN 模式之收斂情況良好，並無過度學習的情形，是直接 BPN 模式的最佳參數組合。而在結合模式方面，其所使用的 BPN 網路架構與參數測定與直接使用 BPN 模式相同，表 4 為結合模式在不同參數組合下的訓練及測試結果，由表 4 中可知當隱藏層神經元數為 8 且學習率為 0.025 時有最小的測試資料

RMSE 值，圖 10 為結合 ICA 與 BPN 模式最佳模式訓練樣本之 RMSE 趨勢圖，由圖中可知，此 BPN 模式之收斂情況良好，亦無過度學習的情形，是結合 ICA 與 BPN 模式的最佳參數組合。

表 4 結合模式在不同參數組合下的預測結果

隱藏層節點數	學習率	訓練樣本 RMSE	測試樣本 RMSE
6	0.020	0.005507	0.004437
	0.025	0.005718	0.003778
	0.030	0.005481	0.003956
7	0.020	0.007640	0.006525
	0.025	0.005499	0.004383
	0.030	0.005465	0.004415
8	0.020	0.006438	0.005008
	0.025	0.005495	0.003635
	0.030	0.005483	0.004481
9	0.020	0.005621	0.003788
	0.025	0.005771	0.003905
	0.030	0.005465	0.004177
10	0.020	0.007353	0.005576
	0.025	0.005485	0.004364
	0.030	0.005452	0.004281

*訓練與測試樣本 RMSE 皆為目標變數重新尺度化至[0.15, 0.85]間之結果

表 5 隨機漫步、直接 BPN 及結合模式的預測結果彙總表

預測模型	指標					
	RMSE	MAD	MAPE	DS	CD	CP
結合 ICA 與 BPN 模式	67.86	53.05	0.52%	86.25%	87.77%	84.47%
直接 BPN	75.64	57.45	0.56%	78.51%	80.32%	76.40%
隨機漫步	137.85	105.77	1.02%	50.43%	54.26%	45.96%

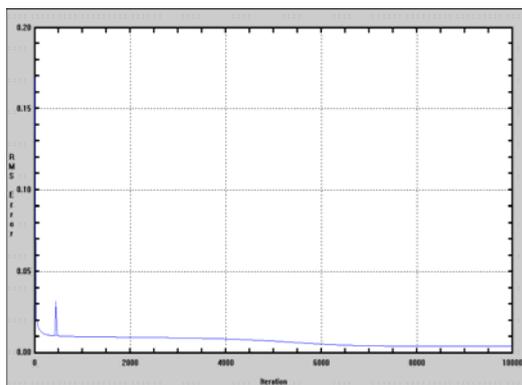


圖 9 直接使用 BPN 最佳模式訓練樣本 RMSE 趨勢圖

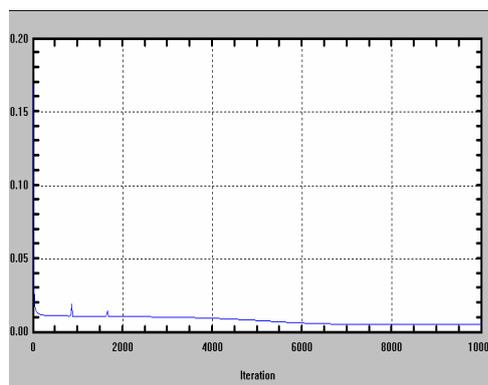


圖 10 結合 ICA 與 BPN 最佳模式訓練樣本之 RMSE 趨勢圖

表 5 為結合模式、直接使用 BPN 與隨機漫步模式在 Nikkei 225 開盤指數之預測結果。由表 5 各評估指標之值可以得知，所提之結合模式在離差衡量指標 RMSE、MAD 及 MAPE 分別為 67.86，53.05 及 0.52%，均小於直接 BPN 模式及隨機漫步模式，代表結合模式所預測的數值與實際值最接近，預測誤差最小，此外，相對直接使用 BPN 及隨機漫步模式而言，結合模式亦有最高的 DS、CD 及 CP 值，分別為 86.25%、87.77% 及 84.47%，代表其在股價趨勢的預測上亦較其他兩個方法準確。整體而言，本研究所提之結合 ICA 與 BPN 之兩階段預測模式，相較於直接 BPN 模式及隨機漫步模式，能提供最小的預測誤差以及最高的趨勢預測準確率，為一值得推薦之模式建構程序。

5.3 ICA 與其他去除雜訊方法的比較

為評估 ICA 濾除雜訊的有效性，本研究亦將 ICA 去除雜訊後的預測結果與其他常用於去除雜訊方法之預測結果進行比較。本研究考慮兩種常用的濾除雜訊方法，一為簡單移動平均(simple moving average, MA)，另一則為小波框架(wavelet frame, WF) (Zheng et al., 1999; Gonzalez and Woods 2002, Quatieri, 2002, Stevenson, 2002)。首先在簡單移動平均方面，簡單移動平均是時間序列方法中最常用於濾除隨機雜訊及趨勢的方法，其目的為使用前 n 期的資料消除隨機變化，其計算是根據時間移動，每次計算將加上一個最新紀錄，同時減少一個最舊紀錄。由於 MA 可以消除隨機成份(使誤差減小)，所

以能被用來濾除資料中雜訊的影響，進而表示時間序列的趨勢。整體而言，期數 n 越小，對趨勢之預測的落後會越小，但是對消除隨機變化的結果愈不明顯；反之，期數 n 越大，對消除隨機變化的結果愈明顯，但是對趨勢預測的落後會越大。由於並無一明確規則決定 n ，一般通常經由試誤 (trial and error) 決定。本研究考慮期數為 1 至 10 的簡單移動平均，經試誤後決定最終之期數為 3。

而在小波框架方面，小波框架是小波轉換(wavelet transform)中的一種形式，是訊號處理中常用於濾除雜訊的方法之一，基本概念為利用低通濾波器(low pass filters)及高通濾波器(high pass filters)將原始訊號分解為近似訊號(approximation signal)與細部訊號(detailed signals)。由於近似訊號包含原始訊號的整體趨勢，是消除隨機變化的結果，因此可以用於濾除時間序列資料中的雜訊。本研究考慮支援長度為 1 至 10 之 Haar 小波函數，進行小波框架的計算，經試誤後決定最佳之小波框架長度為 3。有關小波轉換的詳細內容可參考 Bjorn (1995)、Assume, Campbell and Murtage (1998)及 Zheng et al. (1999) 等人之著作。

在 ICA、MA 及 WF 預測績效比較部分，以 RMSE、MAPE 及 DS 做為評估指標，各模式的預測結果列示於表 6。由表 6 各評估指標之值可以得知，不論在離差衡量指標或是趨勢正確率指標中，皆以結合 ICA 與 BPN 的預測模式表現最佳。代表 ICA 去除雜訊之效果較簡單移動平均與小波框架為佳，能有效的提升預測工具的預測績效。

表 6 不同去雜訊方法之預測結果彙總表

預測模型 \ 指標	RMSE	MAPE	DS
結合 ICA 與 BPN 模式	67.86	0.52%	86.25%
結合 WF 與 BPN 模式	101.72	0.75%	65.62%
結合 MA 與 BPN 模式	192.40	1.46%	48.42%
直接 BPN	75.64	0.56%	78.51%
隨機漫步	137.85	1.02%	50.43%

6. 結論

不論在金融界或學術界，財務時間序列預測一直都是重要的議題。但由於財務時間序列容易受到如政府政策或其他金融市場漲跌等因素影響，使得財務時間序列資料具有非定態及包含許多雜訊之特性。本研究提出一結合獨立成份分析(ICA)與倒傳遞類神經網路(BPN)之兩階段財務時間序列預測模式，先利用 ICA 具有將混合訊號分離出個別獨立來源訊號之能力，從預測變數中估計出獨立成份，並在去除代表雜訊的獨立成份及保留剩餘的獨立成份重建時間序列後，利用 BPN 將濾除雜訊後的預測變數建構財務時間序列預測模式。期望所提之方法能減少 BPN 在模式架構時，因受雜訊影響而產生過度配適或配適不足的問題，進而提升 BPN 模式的預測精準度。

為驗證所提兩階段預測模式之有效性，本研究以 Nikkei 225 開盤指數之財務時間序列資料進行實證研究。實證結果顯示，ICA 確實能濾除財務時間序列中之雜訊，適合作為濾除資料中雜訊之前處理工具。此外，本研究所提之結合 ICA 與 BPN 模式，由於已經利用 ICA 將資料中之雜訊去除，因此不論是在預測誤差或是預測精確度的表現上，均較直接使用 BPN 及隨機漫步模式為佳。而在評估利用 ICA 濾除雜訊的效果方面，本研究亦將使用 ICA 作為濾除雜訊工具之預測結果，與其他濾除雜訊工具，如簡單移動平均及小波框架之結果進行比較，結果亦顯示所提之結合 ICA 與 BPN 兩階段模式有最佳之表現。在後續研究建議方面，由於本研究所使用之 ICA 模式，為假設觀察訊號為原始訊號線性組合之情況，並不一定適用於非線性組合之情況，因此在未來研究上，可以利用非線性 ICA(non-linear ICA)進行資料去除雜訊之工作；此外，可以本研究之研究成果及預測模式擬定投資操作策略，並以投資績效作為預測結果之衡量，進一步增加研究結果之實用性。

致謝

本研究感謝國科會專題研究計畫支持，計畫編號 NSC 95 2221-E-030 -002。

7. 參考文獻

- Anderson, J. A. and Rosenfeld, E. (1998), *Neurocomputing: Foundation of Research*, Cambridge, MA: MIT Press.
- Andrisevic, N., Ejaz, K., Rios-Gutierrez, F., Alba-Flores, R., Nordehn, G., and Burns, S. (2005), “Detection of Heart Murmurs Using Wavelet Analysis and Artificial Neural Network,” *Journal of Biomechanical Engineering-Transaction of the ASME*, 127(6), 899-904.
- Assume, A., Campbell, J., and Murtagh, F. (1998), “Wavelet-Based Feature Extraction and Decomposition Strategies for Financial Forecasting,” *Journal of Computational Intelligence in Finance*, 6(2), 5-12
- Back, A. and Weigend, A. (1997), “Discovering Structure in Finance Using Independent Component Analysis,” in *Proceeding of 5th International Conference on Neural Networks in Capital Market*, 15-17.
- Bartlett, M. S., Movellan, J. R., and Sejnowski, T. J. (2002), “Face Recognition by Independent Component Analysis,” *IEEE Transactions on Neural Networks*, 13(6), 1450–1464.
- Beckmann, C. F. and Smith, S. M. (2004), “Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging,” *IEEE Transactions on Medical Imaging*, 23(2), 137-152.
- Bjorn, V. (1995), “Multiresolution Methods for Financial Time Series Prediction,” in *Proceedings of the IEEE/IAFE 1995 Conference on Computational Intelligence for Financial Engineering*, 97.
- Cao, L. (2003), “Support Vector Machines Experts for Time Series Forecasting,” *Neurocomputing*, 51(April), 321-339.
- Cao, Q., Leggio, K. B., and Schniederjans, M. J. (2005), “A Comparison Between Fama and French’s Model and Artificial Neural Networks in Predicting the Chinese Stock Market,” *Computers and Operations Research*, 32(10), 2499-2512.
- Chen, J., Bandoni, A., and Romagnoli, J. A. (1998), “Outlier Detection in Process Plant Data,” *Computers and Chemical Engineering*, 22(4/5), 641-646.

- Chen, J. and Romagnoli, J. A. (1998), "Strategy for Simultaneous Dynamic Data Reconciliation and Outliers Detection," *Computers and Chemical Engineering*, 22(4-5), 559-562.
- Cheung, Y. M. and Xu, L. (2001), "Independent Component Ordering in ICA Time Series Analysis," *Neurocomputing*, 41(1-4), 145-152.
- Christensen, R. and Pearson, L. M. (1992), "Case-Deletion Diagnostics for Mixed Models," *Technometrics*, 34(1), 38-45.
- Cover, T. M. and Thomas, J. A. (1991), *Elements of Information Theory*, New York, NY: John Wiley and Sons.
- David, V. and Sanchez, A. (2002), "Frontiers of Research in BSS/ICA," *Neurocomputing*, 49(1-4), 7-23.
- Deboeck, G. J. (1994), *Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets*, New York, NY: Wiley.
- Déniz, O., Castrillón, M., and Hernández, M. (2003), "Face Recognition Using Independent Component Analysis and Support Vector Machines," *Pattern Recognition Letters*, 24(13), 2153-2157.
- Gonzalez, R. C., and Woods, R. E. (2002), *Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall.
- Greg, T. (2001), "Neural Network Forecasting of Canadian GDP Growth," *International Journal of Forecasting*, 17(1), 57-69.
- Grudnitski G. and Osburn, L. (1993), "Forecasting S&P and Gold Futures Prices: An Application of Neural Networks," *The Journal of Future Markets*, 13(6), 631-643.
- Hornik, K., Stinchcombe, M., and White, H. (1989), "Multilayer Feedforward Networks are Universal Approximation," *Neural Networks*, 2, 336-359.
- Hyvärinen, A. (1999), "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks*, 10(3), 626-634.
- Hyvärinen, A., Karhunen, J., and Oja, E. (2001), *Independent Component Analysis*, New York, NY: Wiley.
- Hyvärinen, A. and Oja, E. (2000), "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, 13(4-5), 411-430.

- Ikeda S. and Toyama, K. (2000), "Independent Component Analysis for Noisy Data-MEG Data Analysis," *Neural Networks*, 13(10), 1063-1074.
- Ilan, A., Min, Q., and Sadowski, R. J. (2001), "Forecasting Aggregate Retail Sales: A Comparison of Artificial Neural Networks and Traditional Methods," *Journal of Retailing and Consumer Services*, 8(3), 147-156.
- James, C. J. and Gibson, O. J. (2003), "Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis," *IEEE Transactions on Biomedical Engineering*, 50(9), 1108-1116.
- Jang, G. J., Lee, T. W., and Oh, Y. H. (2002), "Learning Statistically Efficient Features for Speaker Recognition," *Neurocomputing*, 49(1-4), 329-348.
- Jung, T. P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T. W., and Sejnowski, T. J. (2001), "Imaging Brain Dynamics Using Independent Component Analysis," *Proceedings of the IEEE*, 89(7), 1107-1122.
- Kanas, A. and Yannopoulos, A. (2001), "Comparing Linear and Nonlinear Forecasts for Stock Returns," *International Review of Economics and Finance*, 10(4), 383-398.
- Kim, T. K., Kim, H., Hwang, W., and Kittler, J. (2004), "Independent Component Analysis in a Local Facial Residue Space for Face Recognition," *Pattern Recognition*, 37(9), 1873-1885.
- Kiviluoto, K. and Oja, E. (1998), "Independent Component Analysis for Parallel Financial Time Series," In *Proceedings of the Fifth International Conference on Neural Information*, 895-898.
- Lee, T. S. and Chen, I. F. (2005), "A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks and Multivariate Adaptive Regression Splines," *Expert Systems with Applications*, 28(4), 743-752.
- Lee, T. S. and Chen, N. J. (2002), "Investigating the Information Content of Non-Cash-Trading Index Futures Using Neural Networks," *Expert Systems with Applications*, 22(3), 225-234.
- Lee, T. S. and Chiu, C. C. (2002), "Neural Network Forecasting of an Opening Cash Price Index," *International Journal of Systems Science*, 33(3), 229-237.

- Lee, T. W. (1998), *Independent Component Analysis: Theory and Application*, Boston, MA: Kluwer Academic Publishers.
- Lin, Q. H., Zheng, Y. R., Yin, F., and Liang, H. L. (2004), "Speech Segregation Using Constrained ICA," *Lecture Notes in Computer Science*, 3173, 755-760.
- Muñoz, A. and Muruzábal, J. (1998), "Self-Organizing Maps for Outlier Detection," *Neurocomputing*, 18(1-3), 33-60.
- Oja, E., Kiviluoto, K., and Malaroiu, S. (2000), "Independent Component Analysis for Financial Time Series," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 111-116.
- Park, H., Joo, D. S., and Choi, D. J. (2000), "The Effects of Data Preprocessing in the Determination of Coagulant Dosing Rate," *Water Research*, 34(13), 3295-3302.
- Quatieri, T. F. (2002), *Discrete-Time Speech Signal Processing-Principles and Practice*, Englewood Cliffs, NJ: Prentice Hall.
- Stevenson, W. J. (2002), *Production Operations Management*, New York, NY: McGraw-Hill.
- Tang, Z. and Fishwick, P. A. (1993), "Feed Forward Neural Nets as Models for Time Series Forecasting," *ORSA Journal on Computing*, 5(4), 374-385.
- Tsaih, R., Hsu, Y. and Lai, C. C. (1998), "Forecasting S&P 500 Stock Index Futures with a Hybrid AI System," *Decision Support Systems*, 23(2), 161-174.
- Vellido, V. N., Lisboa, P. J. G., and Vaughan, J. (1999), "Neural Networks in Business: A Survey of Applications (1992-1998)," *Expert Systems with Applications*, 17(1), 51-70.
- Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M., and Oja, E. (2000), "Independent Component Approach to the Analysis of EEG and MEG Recordings," *IEEE Transactions on Biomedical Engineering*, 47(5), 589-593.
- Visser, E. and Lee, T. W. (2003), "Speech Enhancement Using Blind Source Separation and Two-Channel Energy Based Speaker Detection," in

Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal, 884-887.

- Yagci, O., Mercan, D. E., Cigizoglu, H. K., and Kabdasli, M. S. (2005), "Artificial Intelligence Methods in Breakwater Damage Ratio Estimation," *Ocean Engineering*, 32(17-18), 2088-2106.
- Yaser, S. and Atiya, A. F. (1996), "Introduction to Financial Forecasting," *Applied Intelligence*, 6(3), 205-213.
- Yu, R. Q., Wang, J. H., and Jiang, J. H. (1996), "Robust Back Propagation Algorithm as Chemometric Tool to Prevent the Overfitting to Outliers," *Chemometrics and Intelligent Laboratory Systems*, 34(1), 109-115.
- Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998), "Forecasting with Artificial Neural Networks: The State of the Art," *International Journal of Forecasting*, 14(1), 35-62.
- Zheng, G., Starck, J. L., Campbell, J. G., and Murtagh, F. (1999), "The Wavelet Transform for Filtering Financial Data Streams", *Journal of Computational Intelligence in Finance*, 7(3), 18-35.