國立政治大學應用數學系碩士學位論文

以類神經網路建構風險值模型 Constructing a VaR Model by Artificial Neural Network

碩士班學生: 許力夫 撰

指導教授:林士貴 教授

張宜武 教授

中華民國 107 年 12 月 24 日

摘 要

本研究利用2008/01/04至2018/09/28台灣加權指數進行分析與評估風險值之預測效果。研究成果與貢獻如下:一、延伸 GARCH 模型,假設即期波動度與非預期變動、歷史波動度呈非線性關係,透過類神經網路來捕捉更多非線性槓桿、波動叢集等效果。二、針對不同模型,分別利用最大概似法、反向傳播演算法等進行參數估計與訓練網路。實證結果顯示新模型對於波動度具有較好之配適能力。三、與變異數法、歷史模擬法、GARCH 模型比較,在95%信心水準下新模型所計算之風險值具有較低之穿透率。因此新模型所計算之風險值應可有效改善企業投資時所需提撥之準備金,主管機關亦可透過此模型來訂定投資人所需付出之保證金,皆可降低信用風險與穩定金融市場。本研究可提供上述產業評估風險時較為精準、客觀與較有效率之工具。



關鍵詞: 風險值、類神經網路、GARCH 模型、反向傳播演算法、共變異數法、歷史模擬法

Abstract

We analyze the VaR prediction by using the TWII data from 2008/01/04 to 2018/09/28. The contribution and results are as following: First, to extend the GARCH model, we assume spot volatility, unexpected volatility and historical volatility have non-linear relationship. By training neural network, we capture more non-linear lever effects and cluster volatility effects. Second, compared with different VaR models, we use Max Likelihood method to estimate the parameters and Backpropagation to train the neural network. The results show that the new model fits the volatility better than others. Third, compare the new model with other methods, VaR values predicted by new model have lower ABLF value. Therefore, the VaR values evaluated by new model can improve the reserve fund when the enterprise invests. The financial authority also can set the security deposit by using new model. This study can provide the abovementioned industrial a precise and objective tool to evaluate the risk.



Keywords: Value-at-Risk(VaR), Neural Network, Variance-covariance method, Historical simulation method, GARCH(1,1) model, Backpropagation

目 錄

摘	要		2
Al	BSTRA	ACT	3
目	錄		4
圖	目錄		5
表	目錄		6
第	一章	緒論	7
	第一節	· 研究背景與動機	7
	第二節	i 研究目的	8
第	二章	文獻探討	9
•	, 		
	第一節		9
	第二節	5 類神經網路應用於風險值模型	11
第	三章	研究方法	13
	第一節	5 風險值介紹	13
	第二節		14
	第三節		
	第四節		16
	第五節	5 GARCH(1,1)模型 5 GARCH(1,1)概念結合神經網路	17
	第六節	76110611	
第	四章	實證資料	24
	第一節	5 資料期間	24
	第二節	5 日資料敘述性統計	24
	第三節		
	第四節	· 神經網路超參數設定	
第	五章	實證結果	29
第	六章	結論	32
參	考文劇	ξ	34

圖目錄

圖3-1:風險值介紹圖	13
圖3-2:一個神經元運算圖	18
圖3-3:Sigmoid 函數圖	19
圖3-4:Tanh 函數圖	19
圖3-5:神經網路圖	20
圖4-1:台灣加權股價指數價格圖	24
圖4-2:台灣加權股價指數報酬圖	24
圖4-3:模型建構示意圖	25
圖4-4:2個神經元的激活函數選取圖(µ̃)	26
圖4-5:3個神經元的激活函數選取圖(µ̃)	26
圖4-6:6個神經元的激活函數選取圖(µ)	26
圖4-7:μ̃超參數選取圖	27
圖4-8:2個神經元的激活函數選取圖 $(\tilde{\sigma}^2)$	/ 27
圖4-9:3個神經元的激活函數選取圖 $(\tilde{\sigma}^2)$	28
圖4-10:6個神經元的激活函數選取圖 $(\tilde{\sigma}^2)$	28
圖4-11:ỡ²超參數選取圖	28
圖5-1:移動視窗500日之穿透圖	29
圖5-2:移動視窗1000日之穿透圖	30
圖5-3:移動視窗1500日之穿透圖	31
圖5-4:移動視窗2000日之穿透圖	32

表目錄

表2-1:	國內傳統風險值模型文獻整理	10
表2-2:	國外傳統風險值模型文獻整理	. 11
表2-3:	神經網路風險值模型文獻整理	12
表4-1:	台灣加權股價指數敘述性統計	24
表5-1:	移動視窗500日之各風險值模型比較	.29
表5-2:	移動視窗1000日之各風險值模型比較	30
表5-3:	移動視窗1500日之各風險值模型比較	31
表5-4:	移動視窗2000日之各風險值模型比較	.32

第一章 緒論

第一節 研究背景與動機

當今全球化的發展,各國企業競爭越來越激烈,也因此帶動了金融市場的蓬勃發展,金融商品推陳出新。隨著衍生性金融商品不斷創新,金融機構所承擔的風險也隨之增高,故除了追求高額的報酬之外,也不能忽視風險管理這區塊。例如:1995年英國霸菱銀行事件因為交易員越權進行衍生性金融商品的操作不當導致虧損約14億美元,銀行甚至因此倒閉;1997年的亞洲金融風暴;2008年的二次房貸金融風暴,由房地產蔓延到信貸市場所引起的經濟大衰退正是因為忽略了市場風險所帶來的後果才會導致風暴的產生。過去的金融事件更是警惕了人們更應該了解金融資產的極端事件以規避金融市場波動的風險,所以風險管理也因此成了一大課題。

風險管理的目的在於如何正確衡量不同來源的風險並加以有效監督與控制,而如何降低風險所帶來的影響並確認其來源,皆是在擬定風管目標與決策時須加以考量的。一般風管多用於企業風險管理、資產配置管理、銀行信用風險等等。巴賽爾銀行監督管理委員會於1996年允許各銀行利用風險值以衡量金融市場風險,可對券商之風險性資產進行風險揭露、風險控管與規劃、風險資源分配下之策略性資產配置。

而風險值(Value-at-Risk, VaR),是當前最廣受使用且有效的風險控管工具,以一簡單明確的數值,來描述投資商品在具波動的金融市場下所承擔的風險。 在不同模型與假設下,估算出不同的風險值。該模型除了要有準確度,也要被政府、法人機構、投資者所接受。估算風險值的準確度在於其採用之機率分配是否能有效捕捉到金融市場發生的極端事件,因此如何將金融商品報酬之機率分配的尾端特性真實的呈現出來,才是風險值模型所應考量的問題。

早期評估股票市場風險的方式以變異數作為指標,但因投資機構僅關心投資損失的部分,故此指標不符合普遍所關切的損失風險之需求。爾後有了歷史

模擬法 (Historical Simulation Method)、共變異數法 (Covariance Method)、蒙地卡羅模擬法 (Monte Carlo Simulation)等方法估算風險值,以利投資機構做好資產配置達到最佳效率。

而近年來人工智慧在科技發展的地位舉足輕重,其中的機器學習更是大量被廣泛應用於各大領域,舉凡影像辨識、市場預測分析,尤其是當紅 AlphaGo 的圍棋策略。有鑑於此,希望能藉由神經網路模型,訓練出良好的預測函數,估算出較為準確的風險值,有效改善企業投資時所需提撥之準備金。

第二節 研究目的

本篇論文希望將當紅的神經網路運用在預測市場波動度,透過類神經網路 訓練出函數預測報酬未來的可能波動度,進而估算出風險值。由於過往對報酬 變異度僅有固定常數的假設較不能捕捉到報酬率的真實狀況,故藉由神經網路 捕捉非線性槓桿和波動叢聚的效果,以達到波動度的準確性。

本文的後續章節如下:第一章為初步介紹本篇論文的研究動機與目的;第二章為參考文獻,第一節為國內外傳統風險值模型文獻探討,第二節為類神經網路應用於風險值模型;第三章為研究方法,第一節風險值的介紹,第二節為共變異數法,第三節為歷史模擬法,第四節為 GARCH(1,1)模型與第五節為以 GARCH 概念結合神經網路之模型,第六節為風險值模型準確性評估方法;第四章為實證資料,第一節為資料期間,第二節為日資料敘述性統計,第三節為模型建構,第四節為神經網路超參數設定;第五章為實證結果;第六章為本篇結論。

第二章 文獻探討

本論文著重在風險值模型之比較,故參考了傳統風險值模型以及類神經網路的風險值模型之比較的相關文獻,從中選取欲比較的傳統風險值模型,以下將依序介紹國內外的傳統風險值模型文獻和類神經網路風險值模型文獻並將結論彙整於表中。

第一節 國內外傳統風險值模型比較之文獻

一、國內傳統風險值模型比較之文獻

古永嘉(2006),本文以ARMA-GARCH風險值模型與最常被使用的歷史模擬法、蒙地卡羅法與變異數-共變異數法三法比較模型的預測效果。經過實證比較,發現ARMA-GARCH模型的不管在象徵性指標與比較性驗證指標的表現上比其他三個風險值模型好,且通過LR檢定下的穿透次數為最低。

蘇榮斌(2008),研究使用辛普森數值方法求標準化偏態t分配在不同厚尾、偏態參數值下之下方百分位值,再使用GARCH-N(N(GARCH model with normal error distribution)、GARCH-TGARCH model with student t error distribution) 與GARCH-ST (GARCH model with skewed student t error distribution)模型來進行95%種信賴水準一天期風險值之估計與預測。由GARCH-N模型產生之ABLF、AQLF及非預期平均損失值是三個模型中最小,為最佳風險值模型。

柯博倫(2011),本研究評估GARCH、EGARCH(Exponential GARCH)與TGARCH(Threshold GARCH)等估計類股風險值之績效。實證結果以ABLF值、AQLF值和非預期損失比較發現,在95%信心水準下GARCH模型為較佳的風險值估測模型。

表2-1:國內傳統風險值模型文獻整理

作者	方法	對象	結果
古永嘉(2006)	ARMA-GARCH 、	彰化銀行、新竹商	ARMA-GARCH 通過
	歷史模擬法、共變	銀、台北商銀、農	LR 檢定,且有效降低
	異數法、蒙地卡羅	民銀行、高雄銀	失敗次數,為較佳風險
	模擬	行、聯邦銀行等上	值模型。
		市商業銀行之股價	
蘇榮斌(2008)	GARCH	台灣、南韓、新加	以 ABLF、AQLF 及
	GARCH-T 與	坡與馬來西亞等四	非預期平均損失值作
	GARCH-ST	國股票市場指數	比較,在95%信心水
//	IEX .		下,GARCH 為最佳
100		7	風險值模型。
柯博倫 (2011)	GARCH	食品類、金融類、	結果發現回溯測試在
	EGARCH	機電類、造紙類、	95%信心水準下,金
-	TGARCH	紡織類、營建類、	融類股與塑膠類產業
\\		塑膠類、水泥類股	以 GARCH 為最佳估
	tio	價指數	測風險值模型。

Andjelić G. (2010),本篇論文比較了歷史模擬法與共變異數法,研究對象為歐洲國家的股價指數,回溯測試指出在95%信心水準下,共變異數法有較低的穿透次數,為較佳風險值模型。

Čorkalo Š. (2011), 本篇論文比較了共變異數法、歷史模擬法及拔靴法在股票投資組合的風險值評估,發現歷史模擬法的預測結果有別於共變異數法,顯示報酬分配明顯異於常態,而拔靴法也是從歷史資料抽取,故和歷史模擬法有相近的結果。

表2-2:國外傳統風險值模型文獻整理

作者	研究方法	研究對象	結果
Andjelić G.	歷史模擬法、共	歐洲國家股價指	回溯測試在 95%
(2010)	變異數法	數	信心水準下,共
			變異數法為較佳
			的風險值模型
Čorkalo Š(2011)	歷史模擬法、共	德意志電信, 葡	報酬分配有異於
	變異數法、拔靴	萄牙里斯本、克	常態分配,故歷
	法	羅地亞、格勒	史模擬法與拔靴
	-4	布、Ericsson	法為較好的風險
/	正义	Nikola Tesla 股票	值模型

第二節 類神經網路應用於風險值模型

神經網路近年來已廣泛運用於金融時間序列,其原理是模擬神經運作的方式,利用建立的訓練資料(training data)多次的自我學習與修正錯誤以達到目標的輸出值。其最大的特色在於非線性輸出以及多層結構的預測能力。以下兩篇參考類神經網路建立模型的文獻,並簡述其實證結果且彙整於表中。

Liu (2005),此研究利用標準普爾500指數(S&P 500)、道瓊工業平均指數 (DJI)、福特汽車公司(Ford)與國際商業機器公司(IBM)四種股價指數,比較歷史模擬法、GARCH模型和類神經網路風險值模型法(輸入變數為歷史模擬法與GARCH估算之風險值,輸出值為 combined VaR forecast),由回溯測試與概似比檢定發現,類神經網路為最佳風險值估測模型。

CHEN C.T. (2009),此研究利用那斯達克和道瓊工業期貨指數資料,結合了GARCH模型和類神經網路(輸入變數為5-day Bias、10-dayMA (Moving Average), 12-day RSI (Relative Strength Index), and KD indicator,輸出值為 Close Index 收盤指數),數據結果通過概似比檢定,且較傳統模型(歷史模擬法、共變異數法、蒙地卡羅法)BLF值明顯較小,為最佳風險值估測模型。

Chen X.L. (2009),此研究利用2003/5/3~2009/12/29的恆生指數(Hang Seng Index, HSI)資料,比較 ARMA-GARCH 與類神經網路兩種風險值模型,結果顯示神經網路所建立的風險值模型通過 LR 檢定且穿透率最小,為較佳風險值模型。

表2-3: 類神經網路風險值模型文獻整理

作者	方法	研究對象	結果
Liu (2005)	歷史模擬法、	標 準 普 爾 500 指	類神經網路通過
	GARCH、類神經	數、	LR 檢定,穿透率
	網路	道瓊工業指數、	最低且非預期損
	7	福特汽車公司、	失最小,為較佳
		國際商業機器公司	風險值模型。
CHEN C.T. (2009)	歷史模擬法、共變	那斯達克100指數	數據結果類神經
-	異數法、蒙地卡羅	, 道瓊指數	網路通過概似比
Z	法、神經網路	//	檢定,且較傳統
		5	模型更為準確
	2/	Init [®]	(BLF 明顯較小)
Chen X.L. (2009)	GARCH、神經網		兩方法皆通過 LR
	路	Index(HSI)	檢定,而神經網
			路有較小的穿透
			率,為較佳預測
			模型。

第三章 研究方法

第一節 風險值介紹

風險值為一常見的風險指標,簡稱 VaR(Value at Risk)。廣義來說,市場波動度導致投資金融商品可能的最大損失即為風險值。狹義來說,在特定期間與信心水準下,所投資的金融商品預期可能產生的最大損失。例如,某資產報酬率服從標準常態分配,則有5%的機會,最大損失幅度為1.645。抑或是在95%信心水準(confidence level)下,風險值為1.645。

假設風險值的評估期間為 T 天,信心水準為 $1-\alpha$,可以確定有 α 的機會, 投資金融商品未來 T 天的預期最大損失金額(ΔX_T)不會超過該 VaR。

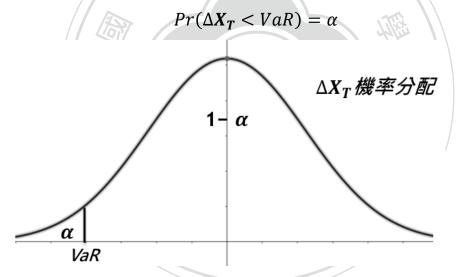


圖3-1:風險值介紹圖

由於本篇論文將以報酬率代表風險值,故作以下推導

$$Pr(\Delta X_T < VaR)$$

$$= Pr(X_T - X_0 < VaR)$$

$$= Pr\left(\frac{X_T - X_0}{X_0} < \frac{VaR}{X_0}\right)$$

$$= Pr\left(R < \frac{VaR}{X_0}\right) = \alpha$$

$$\Rightarrow Pr(R < VaR^r) = \alpha$$

其中 X_0 , X_T 為第0期與第 T 期的資產價格, VaR^r 為報酬率風險值,R 為投資報酬率。其中風險值又可分為絕對風險值與相對風險值,絕對風險值為相對於損益為0元的估計值;而相對風險值則是相對於投資期望報酬的估計值。本文皆以相對風險值來表示。

第二節 共變異數法(Variance-Covariance Approach)

假設資產報酬率為常態分配,由於常態分配線性組合仍為常態,且投資組合報酬率為個別資產報酬率的線性組合,可以直接估算出投資在不同期間與信心水準下的 VaR^{r-N} 。主要用於線性商品。

 $在(1-\alpha)$ 的信心水準下,

$$P(R < VaR^{r-N}) = \int_{-\infty}^{VaR^{r-N}} f(R) dR = \alpha$$

標準化下 Z~N(0,1),

$$P(R < VaR^{r-N}) = P(Z < \frac{VaR^{r-N} - \mu}{\sigma}) = \alpha$$
$$\frac{VaR^{r-N} - \mu}{\sigma} = Z_{\alpha} \to VaR^{r-N} = \mu + Z_{\alpha} \cdot \sigma$$

此模型優點是由於統計分配假設為常態分配,故在計算上容易且效率高。 此模型缺點也因為是常態分配,故無法處理厚尾問題(極端事件)。

第三節 歷史模擬法(Historical Simulation)

假設資產過去的報酬率在未來的評估期間重現,於是擷取過去某段期間的歷史報酬資料,以模擬方式重建資產未來的報酬分配圖,進而估算出 VaR^{r-H} 。

首先模擬個別資產的報酬分配,令 X_t 為資產在時間第t期的價格,如果投資者處於t=0,選取過去N天的價格為模擬資料。再將相鄰兩天的價格取對數後相減,則求出過去N天該資產的日報酬,表為

$$R_{1} = \frac{X_{0} - X_{-1}}{X_{-1}}$$

$$R_{2} = \frac{X_{-1} - X_{-2}}{X_{-2}}$$

$$\vdots$$

$$R_{N} = \frac{X_{-N+1} - X_{-N}}{X_{-N}}$$

其中 R_i , i=1,2...,N為在未來一日可能發生的模擬報酬。再將 R_i 由小至大排序描繪出投資商品的一天報酬分配圖 $(R_{(i)},i=1,2...N)$,依據第五百分位數算出 VaR^{r-H} 。如圖示,若 $N\alpha$ 不為整數,採用線性內插法算出其預測報酬率風險值。

$$\begin{array}{ccc}
R_1 & R_{(1)} \\
R_2 & R_{(2)} \\
\vdots & R_{(N)}
\end{array}$$

$$VaR^{r-H} \\
= R_{([N\alpha])} + (R_{([N\alpha]+1)} - R_{([N\alpha])}) * (N\alpha - [N\alpha])$$

$$\pm \alpha = 0.05$$

此模型的優點其一是計算簡單且易理解;其二是以過去實際報酬作為未來之依據,不需資產報酬分配假設,可以反映過去資產報酬的分配特性(如厚尾現象);其三為具有無母數方法的特性,不需假設統計分配;歷史序列資料已反映過去報酬變化之間的相關性及波動性,此特性使得歷史模擬法不受模型風險的影響。

此模型的缺點其一是歷史資料取得可能存疑,抑或是存在時間不長,歷史資料過少。歷史資料發生極端狀況不常見,因此經常缺乏足夠的資料來模擬分配極端狀況;其二是市場結構改變,使得過去資料不能完全反映未來市場。直覺上,越接近現在的資料越能夠反映當時的市場狀況,若選取資料期間過長,則太過久遠的資料會稀釋較近資料所提供的訊息,使得估計發生偏誤。

第四節 GARCH (1,1)模型

由於股票市場的指數為一複雜且不穩定的時間序列,在期望值上不容易預測而波動度相對容易,故 George-Box & Jenkins(1976)提出了自我整合回歸(Auto Regression Moving Average; ARMA)時序模型加以修正。但這樣的時間序列模型無法獲得有效的估計,由於前提假設干擾項的變異為常數。而 Engle(1982)發現金融報酬時間序列具有波動叢聚性,考慮到條件變異數隨時間而改變的現象,因此將模型設定了時間序列資料條件變異數受到前期殘差項的影響,提出了自我回歸條件變異數模型(Auto Regression Conditional Heteroskedastic; ARCH)。但ARCH的波動度只考慮了前期的殘差平方項, Bollerslev(1986)根據傳統 ARMA模型的移動平均(即落後期的條件變異數)加入 ARCH 模型,提出了一般化自我回歸條件變異數模型(Generalized Auto Regression Conditional Heteroskedastic; GARCH)修正了 ARCH模型。

一報酬的時間序列 R_t ,使 $\varepsilon_t = R_t - \mu_t$ 為一殘差序列,滿足以下三式則為 GARCH(1,1)模型:

$$R_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \sqrt{\sigma_t^2} Z_t$$

$$\sigma_t^2 = \omega + \alpha_0 \sigma_{t-1}^2 + \beta_0 \varepsilon_{t-1}^2$$

其中 μ_t 為第 t 期的條件期望報酬率, Z_t 為一序列獨立且服從相同分配的隨機變數 (白噪音,平均數為0,標準差為1), σ_t^2 為第 t 期的條件變異數, $0 \le \alpha_0 + \beta_0 < 1$ 且 $\alpha_0 > 0$, $\beta_0 > 0$ 以滿足共變異穩定的條件。假設 $\mu_t = \mu$,採用最大概似函數估計 法求出所有未知參數 $(\mu, \omega, \alpha_0, \beta_0)$,並設 σ_1^2 波動度起始值與 ε_1^2 殘差起始值進行波動方程的迭代,估計出t+1期波動度,於是

$$VaR_{t+1}^{r-G} = \hat{\mu}_{t+1} + \sqrt{\hat{\sigma}_{t+1}^2} \cdot Z_{0.05}$$

其中 Z 為 N(0,1)。

此模型的優點是考慮到波動度會隨時間變化,而非過去僅僅只有固定常數的假設,如此也更貼近實際的金融市場狀況。

此模型的缺點則是無法解釋非對稱現象。殘差項的符號不影響波動度,也 就是說不管殘差項是正號或是負號對波動的影響都會是正向的。在實務上,若 市場上放出利空消息,對波動度的影響反而會是增大的;反之,若是利多消息, 波動度的影響卻是小的。故模型無法解釋報酬與波動度的負相關現象。

第五節 GARCH(1,1)概念結合神經網路

本研究採用落後一期的 GARCH(1,1)模型概念,報酬的平均數、波動度方 程式為

$$R_t = \widetilde{\mu_t} + \varepsilon_t$$

$$\varepsilon_t = \sqrt{\widetilde{\sigma_t^2}} Z_t$$

$$\widetilde{\sigma_t^2} = \omega + \alpha_0 \sigma_{t-1}^2 + \beta_0 \varepsilon_{t-1}^2$$

μt: 第t日的期望報酬率(取第t日與前20日的報酬取平均數)

 σ_t^2 : 第t日的報酬波動度(取第t日與前20日的報酬取變異數)

首先透過神經網路幫我們訓練出 σ_r^2 與 μ_r 的預測函數,當輸入值為前1期的奇 期望值 μ_{t-1} 與報酬率 R_{t-1} 時,目標輸出值將會是下一期的報酬波動度 $\widetilde{\mu}_t$,函數 式如下,

$$f_1(\mu_{t-1}, R_{t-1}) = \widetilde{\mu_t}$$

式如下, $f_1(\mu_{t-1},R_{t-1})=\widetilde{\mu_t}$ 以及當輸入值為前1期的殘差平方 ϵ_{t-1}^2 與波動度 σ_{t-1}^2 時,目標輸出值將會是下一期的報酬波動度 $\widetilde{\sigma_t^2}$,函數式如下,

$$f_2(\varepsilon_{t-1}^2, \sigma_{t-1}^2) = \widetilde{\sigma_t^2}$$

第二步將平均數方程式與波動度方程式移項得出 Z_t ,再將 Z_t 由小到大排序 取出第5分位數值。(取法同歷史模擬法)

$$\frac{R_t - \mu_t}{\sqrt{\sigma_t^2}} = Z_t$$

最後將預測出的 $ilde{\mu}_t$ 、 $ilde{\sigma}_t^2$ 與歷史模擬法求出的 $Z_{t,0.05}$ 值和標準常態 $Z_{0.05}$ 值代入

$$VaR_{t+1}^{r-NGH} = \tilde{\mu}_{t+1} + \sqrt{\tilde{\sigma}_{t+1}^2} \cdot Z_{t,0.05}$$

$$VaR_{t+1}^{r-NGN} = \tilde{\mu}_{t+1} + \sqrt{\tilde{\sigma}_{t+1}^2} \cdot Z_{0.05}$$

一、 類神經網路

基本上需要訓練的神經網路可分成三大步驟:建立網路模型(Building Model)、損失函數(Loss Function)、訓練(Training)。第一步就像是人類大腦運作,建立一個函數對應網路的架構;接著第二步決定使用何種損失函數;最後一步即是讓神經網路自行訓練以找出最佳解。

二、 一個神經元運算

類神經網路的概念是由人類大腦的神經網路而來(又稱人工神經網路),是 一種模仿生物神經網路的結構和功能所產生的數學模型。以下我們將介紹類神 經網路中神經元的架構與操作。

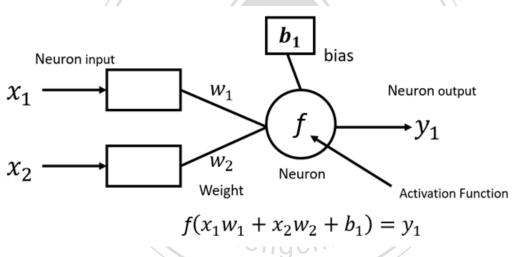


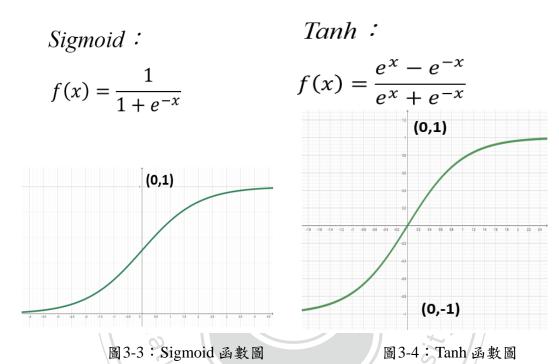
圖3-2:一個神經元運算示意圖

如圖3-2一個神經元的運算,每個神經元都有個簡單的函數:左側神經元的輸入值 x_1,x_2 和右側的輸出值 y_1 。而 w_1,w_2 則是對應每個輸入值的權重且激活函數(f)為非線性函數,常用的激活函數如 Sigmoid function、Rectified linear unit(ReLU)、Tanh function。一個神經元的運作是輸入值乘上對應的權重且加總後再加上誤差值(bias),所得出的數值也是激活函數的輸入值。舉個例子來說, x_1,x_2 分別是1、2和 w_1,w_2 分別是2、5,誤差值 b_1 是-1,而激活函數用 ReLU 函數,則激活函數的輸入值即是 $1 \times 2 + 2 \times 5 - 1 = 11$ 。因此最後的輸出值將會是11。我們稱權重 w_1,w_2 和誤差值 b_1 為參數,這些參數將決定神經網路的運算。

三、 激活函數

激活函數在類神經網路的運作上扮演重要的角色,如果沒有它,不管幾層的學習都會只能有線性的變化。因此,激活函數的功能就是使運算過程引入非線性的特徵。以下我們將介紹本篇選取的兩種激活函數:

Tanh function . Sigmoid function .



四、 損失函數

當類神經網路的架構以及激活函數確定時,可調整的參數就是權重(weights)和誤差值(bias)。首先說明何謂損失函數,它本身是個函數被用來衡量真實質與預測值的近似程度,所以其值越小越好。以下為本篇論文採用的損失函數: 均方根誤差(Root Mean Square Error)

Chengchi Unive

$$L(w,b) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\tilde{y}_i - y_i)^2}$$

其中 L 為損失函數,w 為權重(weight),b 為誤差值(bias), \tilde{y} 為神經網路輸出值,y 為目標輸出值。

五、 梯度下降法

現實生活中我們對問題的最終目標永遠都是最佳解,基本上就是在問題中得出最佳的輸出值。在1986年,Rumelhart & Hinton 提出了反向傳播算法(Back Propagation)來訓練神經網路,訓練的目標即是找到最小化損失函數的最佳解,此演算法就是常用的梯度下降法。此法顧名思義就是透過梯度來調整參數以找出最佳解,而梯度在數學上稱作微分。也就是說,在神經網路找出最佳解採用的梯度下降法,網路輸出與目標輸出之間的差值,用迭代的算法,反向地去調整前面各層的參數(也就是反向傳播 Backpropagation),直到整個網路收斂穩定。舉例來說,站在山的頂端,目標是抵達山的最低點,但我們被蒙眼且能見度是零。所以要達到目的地最好的辦法即是察覺出哪個方向是往下坡走的趨勢,一步步跟隨著向下的趨勢,進而抵達目的地。本篇論文的神經網路是兩個輸入變數以及一個輸出變數,隱藏層為一個,故模型圖如下:

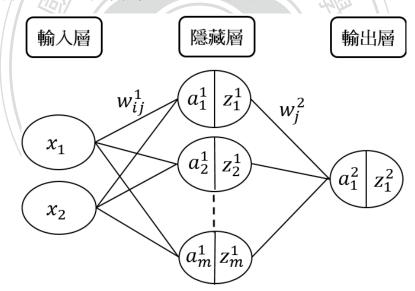


圖3-5:神經網路圖

$$a_j^1 = \sum_{1}^{2} x_i * w_{ij}^1 + b_j^1, \quad i = 1,2; j = 1,2 \dots m$$

$$a_1^2 = \sum_{1}^{m} z_j^1 * w_j^2 + b_1^2$$

$$z_j^h = f(a_j^h), h = 1,2$$
where $f(x) = \frac{1}{1 + e^{-x}}$

我們的目標即是最小化損失函數。假設損失函數 $L = \frac{1}{2}(Y_1 - Z_1^2)^2$, Y_1 為目標 輸出值, z_1^2 為神經網路輸出值,平方是為確保誤差為正,係數 $\frac{1}{2}$ 則為微分計算方 便。

步驟一:

給定初始權重。

步驟二:

利用微分中的鏈鎖率(Chain rule)算出每個權重對 L 的影響($\frac{\partial L}{\partial w_i^p}$):

從輸出層往回一層一層微分。

隱藏層與輸出層間的權重:

$$\frac{\partial L}{\partial w_i^2} = \frac{\partial L}{\partial z_1^2} \times \frac{\partial z_1^2}{\partial a_1^2} \times \frac{\partial a_1^2}{\partial w_i^2} = -(Y_1 - z_1^2) \times \frac{d}{dx} f(a_1^2) \times z_j^1$$

隱藏層與輸出層間的 bias:

$$\frac{\partial L}{\partial b_1^2} = \frac{\partial L}{\partial z_1^2} \times \frac{\partial z_1^2}{\partial a_1^2} \times \frac{\partial a_1^2}{\partial b_1^2} = -(Y_1 - Z_1^2) \times \frac{d}{dx} f(a_1^2) \times 1$$

輸入層與隱藏層間的權重:

$$\frac{\partial L}{\partial w_{ij}^1} = \frac{\partial L}{\partial z_m^1} \times \frac{\partial z_m^1}{\partial a_m^1} \times \frac{\partial a_m^1}{\partial w_{ij}^1} = -(Y_1 - Z_1^2) \times \frac{d}{dx} f(a_m^1) \times Z_j^0$$

輸入層與隱藏層間的 bias:

所入層與隱藏層間的 bias:
$$\frac{\partial L}{\partial b_1^2} = \frac{\partial L}{\partial z_1^2} \times \frac{\partial z_1^2}{\partial a_1^2} \times \frac{\partial a_1^2}{\partial b_1^2} = -(Y_1 - z_1^2) \times \frac{d}{dx} f(a_1^2) \times 1$$

步驟三: 利用**梯度下降法**調整各個權重
$$w^{New} = w^{Old} - \eta \times \frac{\partial L}{\partial w}$$

其中η為學習率,值越大則收斂速度越快。

步驟四: 重複步驟二、三至損失函數(Loss function)收斂為止

第六節 風險值模型準確性評估方法

為了比較以上五種風險值預測模型的預測能力,本研究使用參考文獻中常 用的三種準確性評估方法,分別為二元損失函數(Binary loss function ,BLF)、二 次損失函數(Quadratic loss function, QLF)及非條件涵蓋概似比檢定 (LR test of unconditional coverage, LRuc)或簡稱回溯測試。

一、 二元損失函數(Binary loss function, BLF)

當實際報酬率大於所預測的風險估計值則稱為失敗,而每次失敗發生 時則二元損失函數給定數值1,反之為0。二次函數定義式如下:

$$BL_{t+1} = \begin{cases} 1, & if \ r_{t+1} < VaR_{t+1}^r \\ 0, & if \ r_{t+1} \ge VaR_{t+1}^r \end{cases}$$

其中 r_{t+1} 為 t+1期的實際報酬率, VaR_{t+1}^r 為 t+1期的預測風險估計值, BL_{t+1} 為二元損失函數。接著將失敗次數作加總,並除以預測的總天數(本篇為 1141天)即可算出平均二元損失函數(Average Binary loss function ,ABLF)。若整個樣本期間每日實際發生報酬值小於風險值估計值的平均失敗率 $\hat{\sigma}$ 愈接近於模型所設定報酬超過風險值的比率,則表示該風險值模型預測風險愈準確。

二、 二次損失函數(Quadratic loss function, QLF)

本篇使用了 Lopez(1999)提出的二次損失函數,有別於二元損失函數只考慮次數,此函數不僅考慮失敗次數也考慮了失敗發生時,其實際報酬值 與風險估計值的差距大小。二元損失函數定義式如下:

$$QL_{t+1} = \left\{ \begin{array}{ll} 1 + (r_{t+1} - VaR_{t+1}^r)^2, & if \ r_{t+1} < VaR_{t+1}^r \\ 0 & , & if \ r_{t+1} \ge VaR_{t+1}^r \end{array} \right.$$

其中 $(r_{t+1}-VaR_{t+1}^r)^2$ 為實際報酬值與風險估計值的差值大小, QL_{t+1} 為二元損失函數。以此值確保當失敗發生時所偏離的值較大時給予較大的加權,反之給予較小的加權。再來將 QL_{t+1} 作加總,再除以測試期天數即可算出平均二次損失函數(Average Quadratic loss function, AQLF)。此法比二元損失函數在模型更有準確性。

三、 非條件涵蓋概似比檢定(LR test of unconditional coverage, LR_{uc})

Kupiec(1995)提出了概似比檢定(Log-likelihood Ratio test),是基於二項式分配所求出的一個概似比率統計量 LR,檢定實際失敗率是否符合預設的理論失敗比率(α)。令 x 為實際報酬超過風險估計值(即為失敗)的次數,且此隨機變數服從二項分配,於測試期 n 天理論上的失敗機率為 α 的情況下,其機率質量函數為:

$$f(x) = C_x^n \alpha^x (1 - \alpha)^{n-x}$$

令實際失敗率 $p=\frac{x}{n}$,虛無假設為 H_0 : $p=\alpha$,其概似比檢定統計量 LR_{uc} 服從自由度為1的卡方分配,其式如下:

お下方が配、共文如下、
$$LR_{uc} = -2\ln[(1-\alpha)^{n-x}\alpha^x] + 2\ln\left[\left[1-\left(\frac{x}{n}\right)\right]^{n-x}\left(\frac{x}{n}\right)^x\right] \sim x^2(1)$$

在顯著水準為 α 下,其 LR_{uc} 值落入拒絕域則拒絕虛無假設,表示模型不合適,反之,則不拒絕虛無假設,表示模型的準確性高。

本文有關風險值模型的準確性的衡量準則包括平均失敗率(ABLF)、平均二次損失函數(AQLF)與回溯測試所用之非條件涵蓋概似比率檢定(LRuc),其中以非條件涵蓋概似比率檢定(LRuc)作為準確性之主要衡量準則。若模型皆通過回溯測試時,則再以ABLF、AQLF值大小來判定模型的準確性,其值愈小代表該模型愈準確。另外非預期平均損失(unexpected loss)則表示當實際資產報酬率超過模型所估計之風險值時之平均損失,亦即模型預測失敗時的平均損失,其值愈小代表該模型愈準確,藉以提供一般風險管理者作為另一個參考的依據。

第四章 實證資料

第一節 資料期間

本研究所取的資料期間為2008/1/4至2018/9/28,以每日收盤價當作每日價格率。其資料來源為 Yahoo Finance 網站。本研究以台灣加權股價指數為實證對象,總共2642日收盤價,將價格取一階差分乘上100得出每日報酬,共有2641筆日報酬。

第二節 日資料敘述性統計

台灣加權指數總樣本期間的每日收盤價如下圖4-1,共有2642筆資料,偏態係數為-0.4318,峰態係數為0.7690;日報酬率如圖4-2,分配具有左偏現象,也就是多數在平均之上,峰態係數為4.36988,表示分配具有高狹現象,較常態分配來得高瘦。

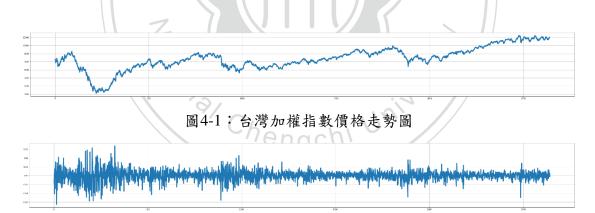


圖4-2:台灣加權指數報酬走勢圖

表4-1:台灣加權指數日報酬之敘述性統計

日報酬之敘述性統	日報酬之敘述性統計					
平均數	0.01736	偏態係數	-0.24073(左偏)			
標準差	1.16378	峰態係數	4.36988(高狹)			
最小值	-6.51328	最大值	6.74217			

第三節 模型建構:移動視窗法

本研究的五種風險值模型將以過去一段時間的報酬資料來建構模型,再將預測日報酬資料輸入模型以比較其值是否穿透。所以本研究將選擇以第一日~第 N 日報酬資料建立模型,測試第 N+1日的是否發生穿透。每一日模型將向右平移一日,新加的日報酬資料將更貼近市場的變化,預測也會更準確,如下圖 4.3.1所示。N 將設為500日、1000日、1500日、2000日四種作比較,以測試各風險值模型的準確性。

第1~N日報酬

(N+1) 日

第2~(N+1)日報酬

(N+2)∃

第3~(N+2)日報酬

(N+3) ⊟

圖4-3:模型建構示意圖

第四節 神經網路超參數設定

類神經網路的超參數(Hyperparameter)設定有許多種,在本研究的類神經網路須設定的部分為:神經元個數、激活函數的選取。

隱藏層的層數與神經元的個數參照葉怡成(2003),將隱藏層設為一個,神經元的個數設為輸入層與輸出層之神經元總數的1/2、1、2倍來比較。故以本研究的輸入層2個神經元以及輸出層1個神經元來說,分別為2、3、6個神經元。然而在激活函數的選取,將由 sigmoid 函數與 tanh 函數比較取其訓練結果優者。學習率將設為0.1,迭代次數設為200次。最終選擇上述6種組合,利用所有樣本的報酬率資料進行超參數的設定以均方根誤差(Root Mean Square Error, RMSE)小且穩定的神經元個數與激活函數作為本研究的神經網路超參數設定。

圖4-4~圖4-7為µ神經網路的超參數設定。由圖4-4和圖4-5知,在隱藏層為2和3個神經元的狀況下,皆是 sigmoid 函數收斂較快;而在圖4-6中,隱藏層為3個神經元的狀況下,則是 tanh 函數收斂較快且收斂值較小;在圖4-7中,由隱藏

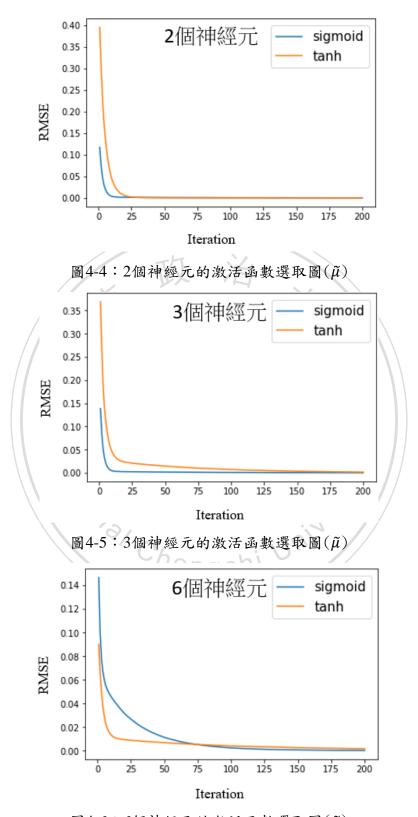


圖4-6:6個神經元的激活函數選取圖(ũ)

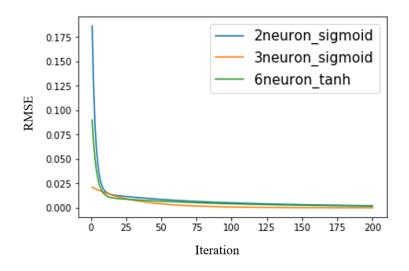


圖4-7: µ的激活函數選取圖

圖 $4-8\sim$ 圖4-11為 $\tilde{\sigma}^2$ 神經網路的超參數設定。由圖 $4-8\sim$ 圖4-10知,無論隱藏層為2,3,6個神經元的狀況下,皆是 tanh 函數收斂較快且值較小;在圖4-11中,由隱藏層6個神經元及 tanh 激活函數為我們最後決定使用的超參數設定。

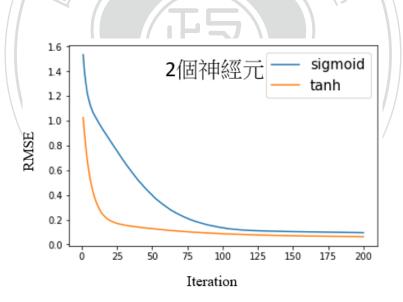


圖4-8:2個神經元的激活函數選取圖 $(\tilde{\sigma}^2)$

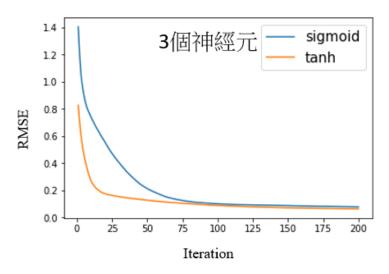


圖4-9:3個神經元的激活函數選取圖 $(\tilde{\sigma}^2)$

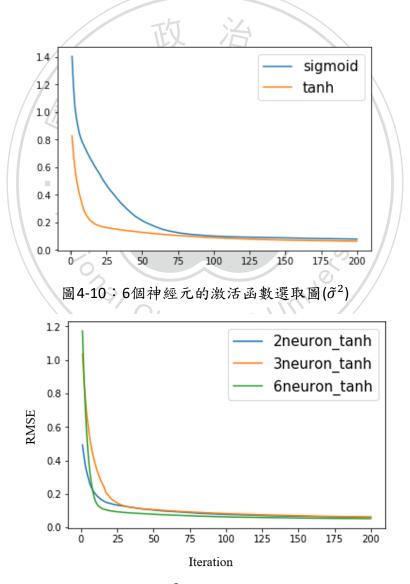


圖4-11: $\tilde{\sigma}^2$ 的激活函數選取圖

第五章 實證結果

本章實證結果將以平均失敗率(ABLF)、平均二次損失函數(AQLF)與回溯測試所用之非條件涵蓋概似比率檢定(LRuc)評估模型的準確度,並求出預測報酬率風險值的平均數與標準差以下依序為移動視窗500日~2000日的各風險值模型的實證結果。

圖5-1為移動視窗500日之穿透率結果。由表5-1:移動視窗500日之實證結果知所有方法皆通過概似比檢定,其中以NNGARCH-history的AQLF值為最小,表示其風險值模型最為準確。而GARCH的預測報酬風險標準差最小,顯示其穩定性最高。

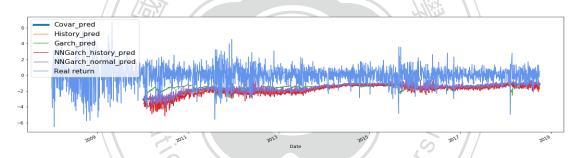


圖5-1:移動視窗500日之穿透率結果

表5-1:移動視窗500日之各風險值模型比較(**表示通過概似比檢定)

	共變異數法	歴史模擬 法	GARCH	NNGARCH- normal	NNGARCH- history
VaR ^r	-1.6540	-1.7111	-1.4876	-1.5744	-1.7639
$\sigma_{\overline{VaR^r}}$	0.5045	0.6163	0.2625	0.5444	0.6874
ABLF	0.0443	0.0443	0.0490	0.0504	0.0439
AQLF	0.0919	0.0914	0.0970	0.1069	0.0900
LR _{uc}	0.6434**	0.4257**	0.6434**	0.0038**	0.7571**

圖 5-2 為移動視窗 1000 日之穿透率結果。表 5-2:移動視窗 1000 日之實證結果只有 GARCH 和 NNGARCH-normal 通過概似比檢定,GARCH 的預測報酬風險標準差最小,顯示其穩定性較高。其中以 NNGARCH-normal 的 AQLF 值為較小,表示其預測報酬率風險較為準確。

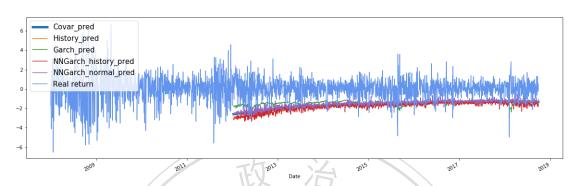


圖5-2:移動視窗1000日之穿透率結果

表5-2:移動視窗1000日之各風險值模型比較(**表示通過概似比檢定)

	共變異數法	歷 史 模 擬法	GARCH	NNGARCH- normal	NNGARCH- history
$\overline{VaR^r}$	-1.6337	-1.6863	-1.4406	-1.5219	-1.7801
$\sigma_{\overline{VaR^r}}$	0.3717	0.4630	0.1859	0.3682	0.4521
ABLF	0.0304	0.0304	0.0390	0.0353	0.0274
AQLF	0.0570	0.0565	0.0690	0.0659	0.0486
LR _{uc}	6.6116	5.7158	1.9574**	3.5744**	12.0826

圖5-3為移動視窗1500日之穿透率結果。所有方法皆通過概似比檢定, GARCH 的預測報酬風險標準差最小,顯示其穩定性較高,其中以 NNGARCHhistory 的 AQLF 值為最小,表示其預測報酬率風險最準確。

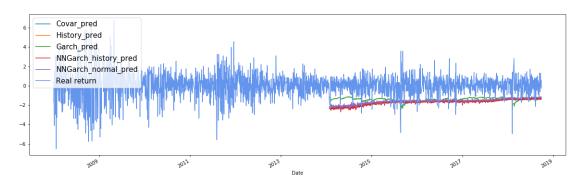


圖5-3:移動視窗1500日之穿透率結果

表5-3:移動視窗1500日之各風險值模型比較(**表示通過概似比檢定)

	共變異數	歷史模擬	GARCH	NNGARCH-	NNGARCH-
	法	法		normal	history
VaR ^r	-1.6464	-1.6990	-1.4094	-1.5750	-1.7206
$\sigma_{\overline{VaR^r}}$	0.2626	0.3083	0.1830	0.2558	0.2824
ABLF	0.0324	0.0306	0.0394	0.0368	0.0324
AQLF	0.0647	0.0616	0.0767	0.0691	0.0614
LR _{uc}	3.6593**	3.6090**	1.2505**	1.9900**	3.6593**

圖5-4為移動視窗2000日之穿透率結果。只有 GARCH 和 NNGARCH-normal 通過概似比檢定,其中以 NNGARCH-normal 的 AQLF 值較小,表示其預測報酬風險較準確。GARCH 的預測報酬率風險標準差最小,表示其穩定性較高。共變異數法、歷史模擬法與 NNGARCH-history 的 ABLF 值偏低,預測的報酬率風險值有偏低的現象。

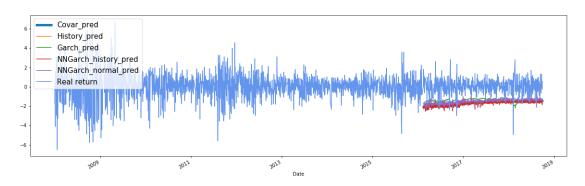


圖5-4:移動視窗2000日之穿透率結果

表5-4:移動視窗2000日之各風險值模型比較(**表示通過概似比檢定)

	共變異數法	歴 史 模 擬法	GARCH	NNGARCH- normal	NNGARCH- history
VaR ^r	-1.6937	-1.7189	-1.3817	-1.4947	-1.7339
$\sigma_{\overline{VaR}^r}$	0.2126	0.2351	0.1544	0.2133	0.2360
ABLF	0.0202	0.0187	0.0343	0.0358	0.0202
AQLF	0.0421	0.0403	0.0633	0.0587	0.0398
LR _{uc}	6.6137	7.4592	1.6111**	1.2902**	6.6137

Chengchi wid 第六章 結論

實證結果顯示不管視窗的大小, GARCH 與 NNGARCH-normal 都通過概似比檢定,而 NNGARCH-normal 在視窗1000日與2000日的預測是最為準確。 NNGARCH-history 則是在視窗500日與1500日中通過概似比檢定,且預測也是最準確。 GARCH(1,1)模型的 AQLF 值整體上高於其他模型的結果,顯示預測報酬風險似乎有些低估的現象,但不管視窗大小,我們發現其預測報酬風險的變動卻是所有模型中最穩定的。共變異數法、歷史模擬法在視窗1000日與2000日有偏低的 ABLF 值,顯示預測報酬風險有較保守的現象。整體上,神經網路的

風險值模型有較佳的預測結果。

未來研究可朝深度學習發展,尤其是遞歸神經網路(Recurrent Neural Network; RNN)中的長短期記憶模型(Long Short-Term Memory; LSTM),因為其神經網路不僅考慮了時間性,也考慮近期與遠期的記憶性影響的程度,適合運用在預測一長串金融時間序列的相關研究上。



参考文獻

英文文獻

- [1] Chen C.T. (2009), "Forecasting Value at Risk (VAR) in the futures market using Hybrid method of Neural Networks and GARCH model", 2009 International Joint Conference on Computational Sciences and Optimization p17-21.
- [2] Moreno J.J.(2011),"Artificial neural networks applied to forecasting time series", Psicothema 2011. Vol. 23 n ° 2, p322-329.
- [3] Chen X.L(2009), "A Statistical Neural Network Approach for Value-at-Risk.", 2009 International Joint Conference on Computational Sciences and Optimization p17-21.
- [4] Engle(1982), "Autoregressive Conditional Heteroscedasticity with Estimation of the Variance in U.K. Inflation," Econometrica, Vol. 50, 1982, p987-1008...
- [5] Čorkalo Š (2011), "Comparison of a Value at Risk approaches on a stock portfolio", Croatian Operational Research Review (CRORR) Vol. 2, 2011 p81-90.
- [6] Arnerić1 J.(2014), "GARCH based artificial neural networks in forecasting conditional variance of stock returns", Croatian Operational Research Review, p329-343.
- [7] Andjelićj G.(2010), "Application of VaR in emerging markets: A case of selected Central and Eastern European Countries", African Journal of Business Management Vol. 4(17), p3666-3680.
- [8] Tsay R.S. (2002)," Analysis of Financial Time Series," ISBN-13: 978-0470414354
- [9] Kupiec(1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," The Journal of Derivatives, winter, p73-84.
- [10] Rumelhart & Hinton(1986), "Learning representations by back-propagating errors", Nature volume 323, p533–536.
- [11] Yan Liu(2005), "Value-at-Risk Model Combination Using Artificial Neural Networks", Emory University. August 2005.

.

中文文獻

- [1] 洪儒瑶、古永嘉、康健廷(2006), ARMA-GARCH 風險值模型預測績效實證, Journal of China Institute of Technology Vov.34-2006.6, p13-35。
- [2] 柯博倫、雷立芬(2010), GARCH 估測風險值績效之探討,臺灣銀行季刊第 六十二卷第四期(p234-243)。
- [3] 黄華山、邱一薰(2005),類神經網路預測台灣50 股價指數之研究, 資訊、科技與社會學報,第5卷第2期,p19-42。
- [4] 蔡玉娟、林家妃、張修明(2010),應用倒傳遞類神經網路及時間序列法建構股價報酬率預測模型-以台灣股市為例,國立屏東科技大學資訊管理系碩士班論文。
- [5] 蘇榮斌、蔡孟祥(2008), 風險值之預測:以台灣、韓國、新加坡及馬來西亞 等國家股票市場為例, Journal of China Institute of Technology Vol.39-2008.12, p179-196

Za Chengchi Unive

[6] 葉怡成(2003),類神經網路模式應用與實作,臺北市:儒林。