

應用模糊局部類神經網路於演化型最佳化方法之研究

莊志勇

余志成*

國立台灣科技大學機械工程系

*國立高雄第一科技大學機械與自動化工程系

106 台北市大安區基隆路四段43號

811 高雄市楠梓區卓越路2號

Tel : (07)6011000 ext.2267

Tel : (07)6011000 ext.2228

E-mail : D9003102@mail.ntust.edu.tw

E-mail : jcyu@cems.nkfust.edu.tw

摘要

本文主要探討當類神經網路初始學習樣本稀少時，對網路模型準確度的影響，並提出神經網路模糊信賴度的觀念，以及應用於最佳化搜尋的策略。類神經網路常用來模擬複雜系統之響應曲面，搭配基因演算法進行設計最佳化，然而類神經網路模擬曲面需採用大量的學習樣本，方能達到全域模擬的準確度。但實際的工程設計經常受限成本及時效性的考量，或是以現有的現場資料來建立模型，未能提供足夠與均佈的實驗樣本，往往造成模擬網路不具有全域的準確性，若直接應用於全域最佳化搜尋，其搜尋結果往往具有相當大的誤差而缺乏實用性。鑑於類神經網路模型的特性，即當模擬點離學習樣本越近，其準確度越高，而內插設計的預測值準確度也會高於外插設計的準確度。本文利用模糊規則規範基因演算法於網路模型的有效搜尋區間，藉以提昇預測結果的有效性，再將所得的暫態最佳設計加入學習樣本，重新訓練網路模型，使得學習樣本增加在最佳點附近的採樣，提升類神經網路於該區域的模擬準確度與學習樣本的效率。文中利用隨機實驗樣本結合典型的數值案例，論證此方法於演化型最佳參數設計之優點。

關鍵字：類神經網路、基因演算法、模糊信賴度、最佳化設計

Abstract

This study addresses the effect on the prediction accuracy of the neural network if the model has limited initial training samples. An inference engine of prediction reliability using fuzzy inference is proposed to characterize the searching domains for design optimization. Neural networks are often applied to simulate the response of complex systems. To ensure the generality of the simulate model, well-spread and abundant training samples are often required. However, practical applications are subject to cost and time constraints. And, sometimes, existed field data, which might be not well designed, are applied to establish the model. The initial network model usually

lacks the global accuracy and the optimum search from the model is liable to significant errors compared with the verification results. The heuristic rules of the network model suggest that the prediction accuracy of the model is closely related to the distances between sampling points and the evolved designs, and the interpolation designs usually have better prediction accuracy than the extrapolation. This study describes a fuzzy inference engine to restrict the searching domains in the optimum search using genetic algorithm to ensure the searching reliability. The generated local optimum from GA search is introduced to the training samples to retrain the neural network. The searching process iterates until the convergence of the optimum. The sampling points will then congregate on the most probable regions of global optimum, and increase the sample efficiency. A numerical example with small initial random samples is presented to illustrate the advantage of the proposed scheme.

Keywords: Neural Network, Genetic Algorithms, Fuzzy reliability, Optimization

1. 前言

目前常用的最佳化方法大致可分為兩大類：一為以斜率為基的搜尋法(Gradient based techniques)，如最陡梯度法(Steepest Descent Method)及共軛梯度法(Conjugated Descent Method)等，只能獲得區域的最佳解；另一為模擬演化型最佳化(Simulated Evolutionary Optimization)，如基因演算法(Genetic Algorithms)及模擬退火法(Simulated Annealing)等[1]，能得到全域最佳解的搜尋方法。對於全域最佳化方法，由於需要大量的世代搜尋，時常無形間造成實驗成本快速攀升，因而多適用於數值模擬的最佳化，往往不適用需以實驗獲得系統參數與輸出的工程問題。

類神經網路常用來模擬複雜系統之響應曲面，搭配全域搜尋法進行最佳化設計的搜尋。Cios等人[2]使用舊有的實驗資料作為倒傳遞類神經網路的訓練樣本，藉以模擬系統輸出入關係，若當訓練樣本數量不夠或不具代表性時，再增加新的實驗

* 聯絡作者

資料來彌補其不足。Nandi 則結合遺傳演算法於類神經網路的最佳化搜尋，並應用於製程參數設計與最佳容差設計[5]。再者，Sitarama 等人[3]利用法國數學家 Jacques Hadamard (1987)發明的直交表，以全因子規劃之參數及輸出值模糊化，經過學習產生新的模糊規則庫。藉由模糊近似推論預測系統輸出，再以基因演算法配合權重計算，進行多目標全域最佳值搜尋。但為求得模擬系統的全域準確性，往往需採用大量的學習樣本，藉以模擬複雜的系統，然而當學習樣本數目受限時，尤其是實際的工程問題常受限於成本與時間因素，無法提供足夠與均佈的樣本時，模擬系統的準確性可能侷限於局部的設計空間，對於結合最佳設計搜尋法的影響與策略，在文獻中則無詳細的探討。

本文提出模糊信賴度的觀念規範類神經網路的有效區間，結合基因演算法於局部網路模型的最佳設計搜尋，提昇預測結果的有效性，以反覆學習與搜尋的方式，求得實際最佳解，並作為實驗資料數量不足與實驗成本高時的設計最佳化的策略。

2. 稀少樣本之類神經網路預測特性

2.1. 倒傳遞類神經網路

類神經網路係利用仿生物神經網路內部傳遞，所衍生出的資訊處理系統，屬於軟計(Soft Computing)的一種。主要用來處理當系統數學型態不明、非線性及複雜任務要求的問題。其種類甚多，本文採用多層回饋式類神經網路中，最著名的監督式倒傳遞類神經網路(Back-propagation Neural Networks, BPNNs)作為系統模擬工具。

倒傳遞類神經網路是將 Window-Hoff 學習規則，廣義化到多層具非線性可微分轉移函數網路中創造出來的。基本架構包含三層：第一、三層分別代表輸入及輸出層；中間層則為隱藏層，主要用來處理參數間的交互作用、關聯性與敏感度等[7]。

因屬於監督式學習網路，所以需經過樣本學習與目標值比對過程，藉以提昇網路的預測能力。利用學習樣本配合最陡梯度法訓練網路單元間的權重矩陣，用來增進類神經網路預測的準確性。

2.2. 類神經網路模型的預測準確度特性

當利用大量的實驗樣本去訓練類神經網路模型，一般皆可獲得極佳的預測結果。然現今多數的實際工程設計問題，往往因成本及時效的考量下限制了實驗樣本的數量。而若應用於最佳設計的搜尋，全域的精度提昇也並非必要，只要在有機會存在最佳化設計的附近增加訓練樣本，提昇局部的準確性，即可符合工程最佳解的應用需求。

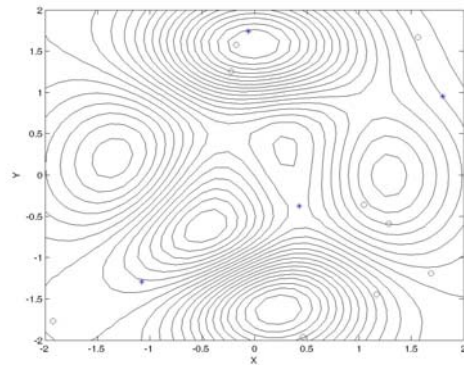
類神經網路的模擬準確度一般將隨著實驗樣本數目的減少而降低，尤其對於複雜的工程系統。

而由擬合模型的特性可知類神經網路可能存在著當預測點距離學習樣本點越近其準確度將會提高；反之，距離學習樣本點越遠其準確度將隨之降低；而內插設計的準確性一般也會高於外插設計的固有特性，本章將以一數值例說明此特性的存在。

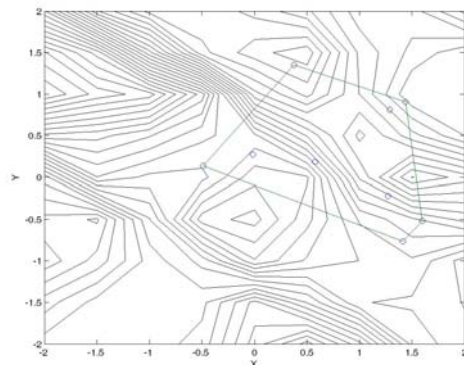
圖 1(a)為典型的二維多極值 Peaks 非線性函數例等高線圖，其方程式如下式：

$$z = 3(1-x)^2 e^{-x^2-(y+1)^2} - 10\left(\frac{x}{5} - x^3 - y^5\right) e^{-x^2-y^2} - \frac{1}{3} e^{-(x+1)^2-y^2} \quad (1)$$

本研究以九個學習樣本點如圖 1(b)中的空心圓點訓練一個類神經網路其模擬模型的預測響應曲面如圖 1(b)等高線圖所示，可看出與實際系統存在相當的誤差。為了瞭解倒傳遞類神經網路在此情況下，預測模型的準確度的分佈特性，我們利用隨機取樣方式在等高線圖面範圍內隨機採樣 180 點，並計算預測設計與學習樣本點間之最小距離，並統計其網路模型預測值與實際值之誤差列於表 1。



(a) 實際 Peaks 函數的等高線圖



(b) 類神經網路模擬之響應曲面

圖 1 實際與類神經網路模擬之 Peaks 函數等高線圖

圖 1(b)中將學習樣本的外圍以直線連接成一凸多邊型，藉以區分內插與外插設計點。從表 1 中可以清楚的驗證之前的經驗假設：

- (1) 當預測點距離學習樣本點越遠其預測準確度將會越低。

(2) 內插設計的模擬準確性高於外插設計的準確度。

表 1 內外部隨機抽樣點預測值與實際值誤差

內外設計區域別	距離學習樣本的最小距離	預測誤差平均值
內插設計	0.2 以下	0.1085
	0.2-0.5	1.1488
外插設計	0.2 以下	1.8661
	0.2-0.5	2.1595
	0.5-1.0	2.8444
	1.0-2.0	3.5359
	2.0 以上	4.0394

2.3. 內外插設計的定義

在二維空間中我們可將學習樣本點相連，找出包圍所有樣本點的凸多邊形，在此凸多邊形內部（含邊界）的點稱為內插點，而在外部的點稱為外插點，如圖 1(b)所示。但為區分廣義 n 維度變數空間中內插與外插設計，本文提出超幾何凸多邊體 (Convex Hyper Polyhedron, CHP) 概念，CHP 將包圍所有學習樣本點的最小凸多邊體，藉由判斷預測設計座落於超幾何封閉空間之內部或外部，來區分內插與外部設計。

CHP 的邊界係由位於外部的 n 個樣本點所擬合的超平面，其方程式如式(2)，

$$x_n + \sum_{j=1}^{n-1} a_{ij}x_j - c_i = 0 \quad (2)$$

$$j = 1, 2, 3, \dots, n$$

以圖 2 的二維變數設計為例，圖中共有三個樣本點（圖中空心圓點），欲構成一個二維參數問題的設計邊界，須有兩個的實驗樣本點，構成一直線方程式。而在這個例子中 CHP 由三條直線方程式所圍成，同理可觀察出圖 1(b)的 CHP 則由 5 條邊界直線方程式所圍成。

為便於判斷任一預測點所在區域，本文提出一套判斷法則，首先由 m 個學習樣本點中任選 $n-1$ 個點配合預測點可擬合出一 n 維空間的超平面如式(2)，共可找出 m 差集 n 個超平面，再將其餘的樣本點一一代入該超平面方程式可得 $y_i(X)$ 如下

$$y_i(X) = x_n + \sum_{j=1}^{n-1} a_{ij}x_j - c_i \quad (3)$$

$$i = m-n+1$$

若存在一個超平面其所有的 $y_i(X)$ 皆同號，則代表該預測點位於 CHP 的外側，即為外插點。否則則為內插點。

以圖 2 為例，假若設計初期有三個學習樣本點： S_A 、 S_B 與 S_C ，這三點剛好構成學習樣本的 CHP。

若要判斷 P_D 四為內插或外插設計，則選擇 P_D 與另外一點透過公式(2)計算其直線方程式，共構成三條方程式。對每一條方程式分別代入剩餘之座標點，其所得的函數值 y_j 並不存在相同正負值，代表這些直線都會切割樣本點空間，使樣本點分別位於直線的兩側，即 P_D 一定位於 CHP 的內部點。而以另一預測點 P_E 為例，如同先前的方式計算可知，因 $P_E S_B$ 所構成的直線方程式，代入剩餘之樣本點所得的函數值 y_j 皆同號，故表示為一外插點。

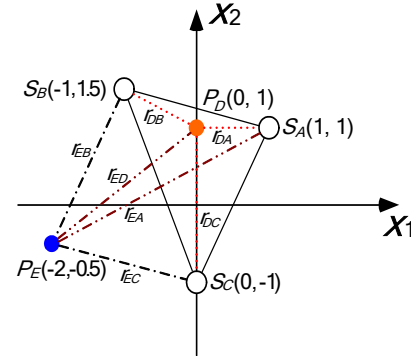


圖 2 預測點與學習樣本空間內外插的判斷例

3. 外插距離與預測信賴度

對於稀少學習樣本所訓練的類神經網路，其預測信賴度由前一節的討論中可知道將與預測設計離學習樣本的距離有關，本節將定義外插距離 (Extrapolation Distance, ED)，並提出模糊推論規則配合 Ed 作為預測信賴度之指標，以規範未來基因演算法於模擬曲面搜尋之有效搜尋區域，增加搜尋結果的實用性。

3.1. 樣本外插距離

為了計算預測點與每個樣本的距離，必須先對輸出變數的座標位置進行正規化，以式(4)將類神經網路的輸入層參數對映至 $[-1, 1]$ 的區間，以便於推論法則的運用並增進網路的學習效果。另一方面也可避免計算結果受少數數值較大的參數所主導。

$$z_{kl} = \frac{\left(x_{kl} - \frac{(\max(x_k) + \min(x_k))}{2} \right)}{\left(\frac{(\max(x_k) - \min(x_k))}{2} \right)} \quad (4)$$

本文使用平均歐氏 (Mean Euclid Distance) 作為計算類神經網路的預測點與學習樣本點間的距離指標如(5)式

$$r_{ij} = \left[\frac{1}{n} \sum_{k=1}^n (P_{ik} - S_{jk})^2 \right]^{0.5} \quad (5)$$

其中 n 表示有 n 個變數 (n 個維度)。

當預測點距離學習樣本越近，其準確度相對會較高。因此計算預測點之平均距離後，採用最短的距

離作為輸出預測信賴度的評斷準則如式(6)。本文將此一最小樣本距離稱為外插距離(Extrapolation Distance) [6]，如圖 2 中預測點 P_D 與樣本點 A 的距離 r_{DA} 最短，因此預測點 P_D 的外插距離表示為 $ED_D = r_{DA}$ 。

$$ED_i = \min(r_{ij}) \quad (6)$$

3.3. 模糊規則與歸屬函數設定

對於多遠的外插距離該給多少的信賴度是屬於模糊性的概念。本文將依兩個準則訂定模糊規則與歸屬函數，一為預測點與學習樣本的外插距離越遠給予較低的信賴度；另一則為預測點屬於外插設計時，相對亦給予較內插設計低的信賴度。文中所有的歸屬函數將利用典型的三角形及梯形函數表示。

外插距離歸屬函數的設定，對預測點屬於外插的部份依遠近分為三種語意述詞：PB (Positive Big)、PM (Positive Medium) 及 PS (Positive Small)；至於外插距離剛好等於零時，則以 ZE (Zero) 表示。對預測點屬於內插部份則分為：NB (Negative Big)、NM (Negative Medium) 及 NS (Negative Small)。每一語意述詞的設定數值，係以正規化參數空間隨機採樣計算出距離的平均值約在 0.7，而最大值約在 1.4。因此定義出外插距離歸屬函數如圖 3。

而初始信賴度的歸屬函數，係將全集合對應到 [0,1] 區間，其設定依信賴度高低分為 Excellent、Good、Fair、Poor 及 Bad，共 5 個等級，如圖 4 所示。

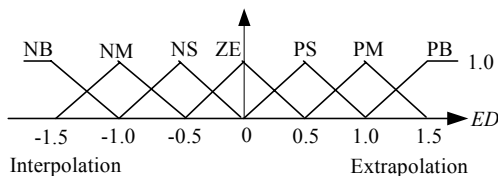


圖 3 外插距離歸屬函數

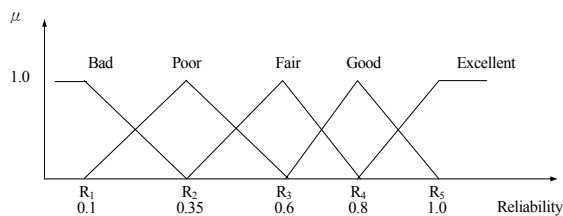


圖 4 初始信賴度歸屬函數

為了改善最佳化方法的搜尋效率及漸進收斂程度，本文於是提出自我學習信賴度法(Self-Learning Reliability Method, SLRM)。此方法係隨著每一迭代過程中所預測之新的設計參數誤差程度及內外插進行動態調整信賴度值。SLRM 調

整信賴度從屬函數之中央峰值為 R_i^a ， a 為修正後次方數，初始值設 $a=1$ 。

若 GA 所搜尋的最佳設計其預測值與實驗值誤差越小時，則下一迭代的 a 值相較於前一代之 a 值會遞減，相對信賴度將會被放寬，對於較為簡單系統模式，雖是稀少樣本點，但以足以訓練準確的全域模型時，可在基因演算法的最佳設計搜尋時逐漸擴大搜尋範圍；反之，若其預測值與實驗值誤差越大時，則下一迭代的 a 值相較於前一代之 a 值會遞增，相對會緊縮信賴度，間接表示基因演算法的搜尋範圍逐步縮小。所以，對於整體而言利用此法將使基因演算法於搜尋過程中，逐漸往最佳值區域接近，方不致於受類神經網路於外插模擬誤差過大的影響，進而造成迭代過程發散。

3.4. 預測信賴度推論法則

由之前所提出關於網路模型預測準確度的經驗法則，進一步根據上述所提及的 7 個外差距離，搭配 5 個不同等級的初始信賴度歸屬函數語意述詞，設定 7 條模糊規則，如表 3 所示。最後經由模糊推論與解模糊化後，方能計算出該預測點的信賴度。

表 2 預測信賴度模糊規則

編號	模糊規則
R1	IF RD is [PB] THEN prediction reliability is [Bad]
R2	IF RD is [PM] THEN prediction reliability is [Poor]
R3	IF RD is [PS] THEN prediction reliability is [Fair]
R4	IF RD is [ZE] THEN prediction reliability is [Excellent]
R5	IF RD is [NS] THEN prediction reliability is [Good]
R6	IF RD is [NM] THEN prediction reliability is [Fair]
R7	IF RD is [NB] THEN prediction reliability is [Poor]

4. 基因演算法最佳設計搜尋

本文結合模糊局部類神經網路與遺傳演算法，期望在較短的時間內，以少量的實驗計算成本，獲得最佳(較佳)的設計。由學習樣本所建立的類神經網路，作為一個虛擬系統響應產生器(Virtual Response Generator)，藉由類神經網路回想快速的效率，以基因演算法搜尋模擬曲面上的最佳值。而另一方面，為避免因學習樣本的減少，造成預測誤差影響最佳設計的搜尋，我們結合網路模型的信賴度預測，限制 GA 的有效搜尋範圍，信賴距離，配合模糊系統推論網路預測值的信賴度，以控制 GA 的搜尋範圍，再將所得的暫態最佳設計加入學習樣本，重新訓練網路模型，重複訓練與搜尋直至預測設計收斂為止。如此使得學習樣本增加在最佳點附近的採樣，提升類神經網路於該區域的模擬準確度與學習樣本的效率。整個演算流程如圖 5 所示。

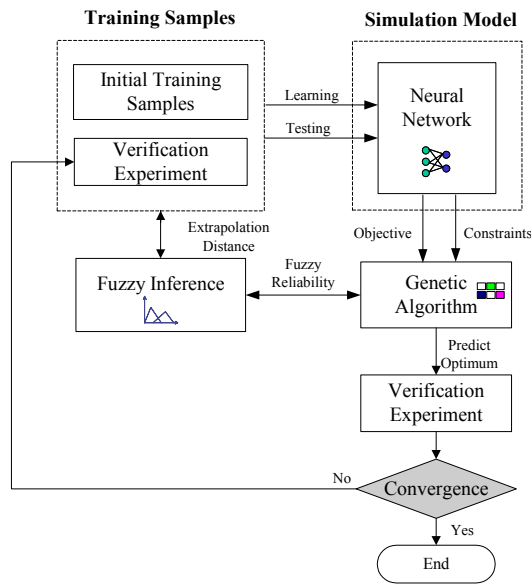


圖 5 最佳化設計流程

5. 典型數值案例分析

本節利用 Chen 等人[4]所發表論文中之 Peaks 函數，如式(1)；說明所提出最佳化策略的應用，並與一般類神經網路結合遺傳演算法所得結果比較。此函數為一雙變數之三維空間圖形如圖 6，其極值分佈等高線如圖 1(a)所示。應用本文所提出之模糊類神經網路搭配基因演算法於 $x, y \in [-\infty, +\infty]$ ，進行全域極大值之搜尋。

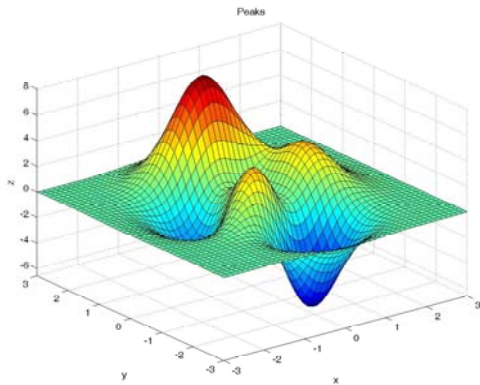


圖 6 Peaks 函數之三維空間分佈圖

5.1. 隨機樣本取樣

利用隨機抽樣方式於初始設計範圍內，隨機抽取一組樣本，每組包含 13 個樣本點；其中 4 個作為測試樣本，另外 9 個作為學習樣本，其位置分佈關係，如圖 1(b)所示。

5.2. 倒傳遞類神經網路模擬

將每一組隨機樣本點依照測試與學習區分，批次輸入倒傳遞類神經網路進行曲面模擬，當學習誤

差達預先設定之均方誤差($MSE = [\text{容許誤差}]^2$)或最大循環數時，隨即完成曲面模擬學習。類神經網路主要參數設定，如表 3 所示。

表 3 倒傳遞類神經網路主要參數設定

隨機樣本編號	隱藏層神經元個數	學習率	遞減學習率	遞增學習率	均方誤差
Fuzzy NNGA	17	0.50	0.80	1.05	10^{-8}
NN+GA	8	0.80	0.70	1.05	2.5×10^{-3}

5.3. 模糊推論

因為類神經網路對於外插曲面模擬本身就存在有較內插區域差的準確度，再加上本案例僅運用少量的隨機樣本點進行曲面模擬之因素，進而使得遠離學習樣本點之預測值更添堪憂。於是，加入模糊信賴度的概念，藉由自我學習信賴度法調整模糊推論信賴度，對於類神經網路所模擬之曲面進行修正，以增加隨後基因演算法於搜尋全域極大值之效率。

5.4. 基因演算法搜尋最佳值

首先，利用先前類神經網路學習與測試之 13 個樣本點，作為基因演算法之初始設計族群，如此可以減少再進行實驗所產生之不經濟及耗時等問題。隨後經由選擇、交配及突變過程篩選出較佳之新設計族群。計算適應函數值；另一方面則以公式(1)進行實驗獲得實際值，若實驗值與適應函數值間誤差大於容許誤差，再重新將此組參數組合加入當作類神經網路之學習樣本點，逐至誤差收斂小於容許誤差，方可獲得一組全域極大值參數。基因演算法之參數設定如表 4 所示。

表 4 基因演算法之參數設定

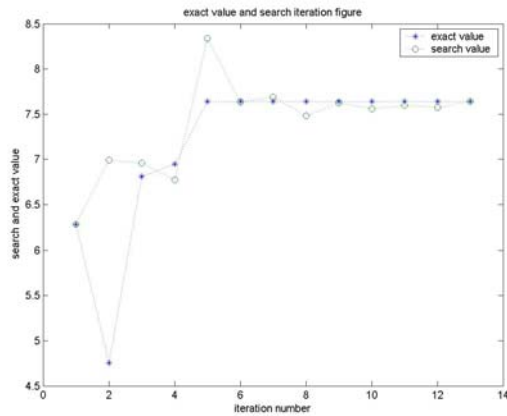
隨機樣本編號	容許誤差	交配率	突變率	最大世代數
Fuzzy NNGA	0.0001	1.0	0.080	300
NN+GA	0.05	0.8	0.080	300

5.5. 結果比較與討論

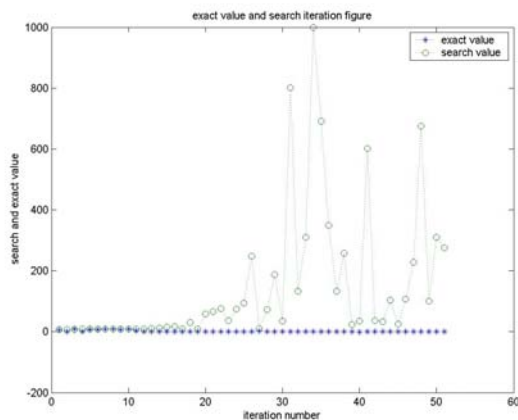
將類神經基因演算法與自我學習信賴度法進行分析與比較。由圖 7 及表 5 可很清楚的看出，傳統類神經基因演算法之迭代過程會有發散的現象，難以收斂於一定值。是因為類神經網路於外插區域模擬誤差往往過大，使 GA 誤以為找到最佳值，為了改善此一棘手的問題，我們提出模糊自我學習信賴度法進行調整信賴度值，使搜尋侷限於學習樣本點的附近，藉由迭代反覆演化學習，逐漸逼近最佳值，即使因樣本點數目的限制無法搜尋至收斂，甚至中途停止，運用模糊類神經網路仍可提供合理的較佳設計。

表 5 Fuzzy NNGA 與傳統 NNGA 最佳化結果

編號	理論值	最佳值	誤差 (%)
Fuzzy NNGA	8.1062	7.6147	6.06
NN+GA	8.1062	NG	NG



(a) 模糊類神經基因演算法(FUNGA)



(b) 類神經基因演算法(NNGA)

圖 7 FUNGA 與 NNGA 最佳化迭代收斂結果

6. 結論

本文所提出模糊預測信賴度推論法則，利用超幾何凸多邊體概念動態調整設計邊界進行內外插判讀與自我學習信賴度法調整信賴度，成功的應用於稀少隨機學習樣本點情況下，限制倒傳遞類神經網路所模擬出之響應曲面的有效區間，藉以增加基因演算法於全域最佳值之搜尋結果可行性。避免因過度信賴網路模型的全域準確性，造成應用最佳化設計搜尋時的誤差造成搜尋過程的發散。

7. 誌謝

本研究承蒙國科會支持，計劃編號 NSC 91-2212-E-155-007，特誌謝意。

8. 參考文獻

- [1]. D.B. Fogel, "A Introduction to Simulated Evolutionary Optimization", *IEEE transactions on Neural Networks*, vol. 5, No. 1, Jan. 1994.
- [2]. K. J. Cios, G. Y. Baakliini and A. Vary, "Soft Computing in Design and Manufacturing of Advanced Materials", *Journal of engineering for gas turbines and power* Vol. 177 161-165, 1995.
- [3]. P. Sitarama Chakravarthy and N. Ramesh Babu, "A New Approach for Selection of Optimal Process Parameters in Abrasive Water Jet Cutting", *Materials and manufacturing processes* Vol. 14, No. 4, 581-600, 1999.
- [4]. J. Chen, D. S. H. Wong, S. S. Jang and S. L. Yang, "Product and Process Development Using Artificial Neural-Network Model and Information Analysis", *Journal of the AIChE*, Vol. 44, No. 4, 876-887, April 1998.
- [5]. S. Nandi, S. Ghosh, S. S. Tambe, B.D. Kulkarni, "Artificial neural-network-assisted stochastic process optimization strategies", *AIChE Journal*,; v.47, n.1, pp. 126-141, Jan, 2001.
- [6]. J. Yu, X. Chen, T.R. Hung and F. Thibault, "Optimization of Extrusion Blow Molding Processes Using Soft Computing and Taguchi Method", *Journal of Intelligent Manufacturing*, 15, 625-634, 2004.
- [7]. J. S. Jang, C. T. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence*, Prentice-Hall, 1997.