

參考文獻

1. Chou F.N.F., Wu C.W., ” Reducing the impacts of flood-induced reservoir turbidity on a regional water supply system”, Advance in Water Resources, 33(2), 146-157, 2010.
2. Einstein H., The Bed-Load Function for Sediment Transportation in Open Channel Flow, U.S. Department of Agriculture, Soil Conservation Service, Technical Bulletin 71, Washington, D.C.,1950.
3. Fleming G., ”Suspended Solids Monitoring : A Comparison Between Three Instruments”, Water and Waste Engineering, 73(883), 377-382,1969.
4. Kynch G.J., ”Theory of sedimentation”, Transactions of the Faraday Society, 48, 166-175, 1952.
5. Meyer-Peter E. and Muller R., ”Formulas for Bed-Load Transport”, Proceedings on Second Meeting of International Association for Hydraulic Research, Stockholm, 39-64,Sweden, 1948.
6. Metropolis N. and Ulam S., “The Monte Carlo Method”, Journal of the American Statistical Association, 44(247), 335-341, 1949.
7. Nagy H.M., Watanabe K., and Hirano M., ”Prediction of Sediment Load Concentration in Rivers Using Artificial Neural Network Model”, Journal of Hydraulic Engineering-ASCE, 128(6), 588-595, 2002.
8. Sarangi A. and Bhattacharya A.K., ”Comparison of Artificial Neural Network And Regression Models for Sediment Loss Prediction from Banha Watershed in India”, Agricultural Water Management, 78(3),

- 195-208, 2005.
9. Sun H., Cornish P.S. and Daniell T.M., "Turbidity-Based Erosion Estimation in A Catchment in South Australia", Journal of Hydrology, 253(1-4), 227-238, 2001.
 10. Toniolo H., Parker G. and Voller V., "Role of Poned Turbidity Current in Reservior Trap Efficiency", Journal of Hydrology Engineering, 133(6), 579-595,2007.
 11. 王英銘，「自組非線性系統應用於濁度預測」，國立成功大學水利及海洋工程學系土木工程系，碩士論文，民國 94 年。
 12. 李念勳，「應用類神經網路於非點源污染預測模式及預測採樣之研究」，國立成功大學環境工程學系，碩士論文，民國 96 年。
 13. 柯建州，「直線型高分子線性鬆弛之蒙地卡羅模擬」，國立中正大學化學工程系，碩士論文，民國 90 年。
 14. 程勝璟，「甲仙攔河堰懸移質輸運率研究」，國立屏東科技大學土木工程系，碩士論文，民國 92 年。
 15. 黃振傑，「沉砂池泥砂沉降特性之研究」，逢甲大學水利工程學系，碩士論文，民國 96 年。
 16. 趙信嘉，「沉砂池對離槽水庫營養負荷之研究」，國立成功大學環境工程系，碩士論文，民國 89 年。
 17. 楊錦釗，「旗山溪河床下降及防治之研究」，曹公農業水利研究發展基金會，民國 86 年。
 18. 經濟部水利署，高屏溪攔河堰檢測資料研判分析委託研究計畫，經濟部水利署，民國 91 年。
 19. 經濟部水利署水利規劃試驗所，濁水溪河槽水庫可行性規劃(一)

河道水理及輸砂分析，5-17 頁，民國 91 年。

20. 經濟部水利署水利規劃試驗所，通用性區域水資源調度與供需分析模式建立(3/3)，財團法人成大水利海洋研究發展文教基金會，民國 95 年。
21. 經濟部水利署南區水資源局，高屏溪攔河堰水質（濁度）及泥砂監測計畫，經濟部水利署南區水資源局，民國 96 年。
22. 經濟部水利署南區水資源局，高屏溪攔河堰高濁度缺水分析及改善策略檢討計畫，經濟部水利署南區水資源局，民國 96 年。
23. 逢甲大學，渾水水庫沉降特性與排渾減淤操作研究，經濟部水利署，民國 94 年。
24. 逢甲大學，渾水水庫沉降渾水水庫沉降特性與排渾減淤操作研究(2/2)，經濟部水利署，民國 95 年。
25. 易任與王如意編著，應用水文學，國立編譯館，民國 68 年。
26. 楊萬全，水文學，地理研究叢書，民國 71 年。
27. 楊朝仲，張良正，葉欣誠，陳昶憲，葉昭憲編著，系統動力學-思維與應用，五南圖書出版股份有限公司，台灣，民國 96 年。

附錄一 遺傳演算法

一、 傳統搜尋法的介紹

所謂搜尋，係指在一既定之問題條件下，於有限的可行解中找出一特定的答案，此答案可滿足所定之目標函數。而由可行解所構成的集合，稱為解答空間(Solution Space)。目標函數的訂定，則可能為搜尋最大或最小值(Davidor,1991)。

傳統中常見的搜尋方法可分為三大類，分別是微分(Calculus)，列舉(Enumerative)與隨機(Random)

(Goldberg,1989;Davidor,1991；張斐章、陳莉,1993)，在此略述如下：

(一)微分方法：

大致上，此法可分為間接與直接兩種方法，間接法是藉由目標函數之導數為零，解出區域極值(Local Extrema)，通常必須給定一平滑且無限制式的目標函數，在任一方向搜尋斜率為0的點，以發現可能的峰值(Peak)。而直接法是利用目標函數搜尋局部最佳值，且依局部的坡降(Gradient)來指引所搜尋的方向，亦即利用登山(Hill-Climbing)的簡單觀念，在最陡的允許方向向上攀登函數，以找出局部最佳值。

整體而言，微積分方法之主要缺點有二：(1)僅能搜尋某一局部的範圍，初始值影響結果甚鉅，易被區域最佳值(Local Optimum)所誤導。(2)微積分法依賴導數之存在，然而自然界中情況充滿著不連續多峰與擾動的空間，不易以傳統的微分方法求解。

(二)列舉法：

此法之理念極為簡易直接，即在有限空間內逐點搜尋其目標函數值，然其主要缺點為缺乏效率。甚至於動態規劃(Dynamic Programming)，也有此種維度的困擾(Curse of Dimensionality)。(Bellman,1961)

(三)隨機(Random)搜尋

此法無固定模式，僅於空間中任意選取座標點並求其目標函數值，缺點亦為缺乏效率，然其表現較列舉法略佳。故可知傳統的搜尋法不易求得問題之整體最佳解，否則即是缺乏效率，而遺傳演算法(Genetic Algorithms)則可克服以上的缺點，對於多峰函數的最佳化問題，遺傳演算法一次搜尋多點的群集，具同時平行攀登許多個峰值及逐次改良搜尋結果的能力，可避免被區域最佳值(Local Optimum)誤導，並且具有不錯之效率(蔡威平，1999)。

二、遺傳演算法概述

遺傳演算法的理論基礎可回溯自1859年達爾文(Charles Darwin)的「物種演化」(On the Origin of Species by Means of Nature Selection)書中的「物競天擇，適者生存」的演化及淘汰觀念。在這種由自然選擇的演化機制中，生物界中的每個個體會把它們的特徵傳遞到下一代，而生物的特徵是由生物細胞內的染色體來決定的(染色體即是由基因所組成的基因鏈)，由於每個個體的特徵都不大相同，因此不同特徵的個體對環境的適應力也不大一樣，同時生物的突變及交配也會使得上下代個體之特徵不相同，而適應力較高的個體，即它們的特徵較適合於目前的環境，在後代的數目上由於競爭的緣故，適應力較高的個體的後代數目會比適應力較低的後代數目多，因此這會把整個族群的特徵引導向更適合生存於自然環境的方向發展，在長時間中，這

種引導所發生的變異會越來越累積，最後演變至產生一整個特徵能適應於特別生態環境下的種族。

將這種自然界的選擇方法系統化並發展一可用之模式最早是由密西根大學的John Holland教授在1975年於Adaption in Natural and Artificial System文中所提出，發展出遺傳演算法搜尋技術的基本架構，並且由其學生David Goldberg成功地運用在工程問題上。之後，有許多研究亦證實了遺傳演算法在最佳化問題的求解上是十分有效率的，其有以下幾個優點：

1. 可優選連續(continuous)及不連續(discrete)的參數。
2. 在優選的過程中，無須求得目標函數的導數。
3. 搜尋的方式不同於以往的單點搜尋方式，而是採用多點搜尋，因此不易掉入局部最佳解(local optimum)。
4. 可以處理多參數的優選問題。
5. 具有隱平行運算的能力，若在平行電腦中，可大量解省運算的時間。
6. 在優選複雜非線性的問題中，其演算機制可跳脫局部最佳解(local optimum)。
7. 演算優選的結果，可提供一組最佳解，而非只有單一最佳解。

8. 參數優選需經由解碼的過程，而整個演算的機制是在解碼後的參數集合中進行，不是在參數集合本身，因此演算的機制不受函數型能的影響。

以上的優點，使得我們發現當傳統的最佳化方法無法解決一個問題或得到令人滿意的優選結果時，遺傳演算法便是一個很有趣且擁有很大潛力去替代傳統的優選法。

對所有的問題而言，遺傳演算法並非都是一個最佳的方法，例如當在處理一具有凸函數型態，且僅有少量變數之問題時，一般傳統以微積分為基礎的搜尋法，即可比遺傳演算法快速的找到最佳解，別外一些簡單優選的問題，傳統的演算法亦都能很快的解決，然而當我們在處理實際的問題時，經常會遇見的是非凸函數且多變數型態或更複雜的問題，這是一般演算法不易解決的，而遺傳演算法就有解決此類問題的能力，且可得到近以全域最佳解。

三、 遺傳演算法的架構

遺傳演算法將欲求解的問題變數或參數以一種類似染色體的資料結構(Chromosome-Like Data Structure)來編碼，並應用一些遺傳運算元(Operators)如交換(Crossover)、突變(Mutation)對大量的染色體作運算，運算後產生的子代除了保存親代中具優勢的特質外，也有可能因為基因的交換與突變而比親代的表現更佳。基本的遺傳演算法包含下列幾個步驟（如附圖1-1所示）：

步驟一：將問題的變數編碼：

例如可以二進位字串(Binary String)的形式來表示變數，其間的轉

換為二進位與十進位的對應，如將二進位字串1001解碼，則可對應於十進位的變數值9，而1100對應於12，1001與1100可看作是兩條染色體。

步驟二：產生初始群集(Initial Population)

以隨機的方式產生多條染色體作為初始解。

步驟三：計算目標函數值(Evaluation)

將初始群集大量的染色體解碼後對應的變數值一一代入問題模式中，計算函數或目標函數值。

步驟四：計算適合度

適合度愈高表示該染色體具有較優的特質，將來被複製(Reproduction)的機會也較大。

步驟五：複製(Reproduction)或選取(Selection)

為演化出更優良的個體，必須從原來族群中篩選出較佳的個體，組成下一代的族群，這就是複製。因此，擁有較高適應值的染色體，便有較高的機率被選擇出來進行複製。一般而言，有下列兩種方式：

(一)輪盤法(Roulette wheel):

所謂輪盤法是假設一個可轉動的輪盤，在輪盤上劃分許多扇形區塊，區塊的面積大小正比於個體被複製的機率。因此，個體的適應值越高，適應值佔有族群適應值總和的比例也越高，在輪盤上所佔的面積也越大，而被選上的機率也越大。輪盤法的示意圖參考附圖1-2所示。詳細步驟則如下所示；

第一步：計算總適應值；為第一個個體適應值至最後一個個

體適應值之總和。

第二步：求個體適應值的份量；為個體適應值除以總適應值。

第三步：繪製輪盤；在盤上劃分扇形區塊，區塊大小正比於個體適應值的份量。

第四步：射靶；隨機亂數產生 0 至 1 的數字，此數落於何區，則該區個體被複製一次。

(二)比較選取法(Tournament selection):

比較選取法，即是模仿自然界的生物彼此競爭情形，當某一個體的適應值愈高，其經由比較選取後，存活下來而被複製的機會愈高，此選取法有一好處，即是染色體被複製下來的機率與染色體間適合度的相對值大小無關而是取決於相對大小，因此較適合於個體間適合度值相對變化很大之問題，一般做法如下：

第一步：依每代總數，設定一個合理的比較個數(N)。

第二步：每次從母代隨機選取 N 個染色體，比較其適合度最優者複製至子代。

步驟六：基因交配(Crossover)或重組(Recombination)

進行完複製的步驟之後，便接著要進行交配的程序。在基因演算法運作的過程中，程式運用輪盤法或比較選取法等方法挑選兩條染色體，將其基因排列作重新的組合，以產生新的兩個染色體，這就是交配的過程。基本上，交配的方式有三種，分別是單點交配、雙點交配和均一化交配。下面分別對這三種方式作介紹：

(一)單點交配：

在進行單點交配時，程式會先依亂數決定一個切斷點，利用這個切斷點，將原先挑選出欲進行交配兩個的染色體切成兩部分，再將切

開的部分重新組合成一對新的染色體。

(二)雙點交配：

雙點交配的步驟與單點交配類似，唯一的不同處是在進行雙點交配時，程式會先依亂數決定兩個切斷點，利用這兩個切斷點，將原先挑選出欲進行交配的染色體切成三部分，再將切開的部分重新組合成新的染色體。

(三)均一化交配：

使用均一化方式進行交配時，首先會產生一個和染色體長度相等的二進位陣列，在這個二進位陣列中，每一個位元均會依亂數決定此位元的值為0或1，如此一個完整的二進位陣列稱之為面具(mask)。利用這個面具，可以決定染色體交配的位置。若是在面具中的位元值為1時，就進行染色體在此位元的互換；反之面具中的位元值為0時，則不進行互換。和單點交配、雙點交配不同處是，利用單雙點交配時，每次交配切斷點的位置不是固定的，而均一化交配則是利用面具的方式，使得同一世代之所有個體在固定的位置交配。三種交配方式的示意圖如附圖1-3~附1-5所示。

步驟七：突變

演化過程最後一個步驟是突變。在遺傳演算法的運作過程中，程式會隨機產生一個突變的機率值，若是此值低於事先所定義的突變率，染色體便會進行突變的程序。所謂突變，是隨機選定染色體的某些位元，將此些位元的值作0與1的互換，此步驟對染色體上的每個位元皆存在有突變的機率，可對染色體上各分段所代表的參數進行突變，將染色體的體質作一立即的改變，使搜尋的點更為零散，以防止過早收斂於局部最佳值，而無法獲得整體最佳解。突變示意圖則如附

圖1-6所示。

四、 模板理論及隱平行性質

在理論上遺傳演算法通常以處理模板(schema)的觀點來進行分析，所謂模板代表了染色體空間上的次集合。在一個模板中，某些基因是沒有限制的，而其他的基因則包括特定的值。以二位元編碼而言，模板的定義是由字母{0,1,*}所形成的字串(string)，其中*表示在字串中相對應的位置上並無限制，即為1或為0均可，舉例，字串01*1*0所描述的模板為；{010100,010110,011100,011110}。模板理論解釋了遺傳演算法的高效率，因為較短的，表現較好的(適合度較高的)一些次字串(substring)將呈指數次方型式的快速成長。在推導模板理論(Schema Theorem)及隱平行性質(Implicit Parallelism)之前，須先介紹一些基本觀念：

當我們在表現一個字串內的二位元址時，通常包含了二種表示法：即(0,1)。現在再加入一種新的符號(*)，其意義為二位元址可能為0或1任意一種。舉例說明如下：如有一字串為{11*0*}，則此字串有四種可能的組合狀態" {11000,11001,11100,11101}。由這些觀念我們可以清楚地知道，模版{01*0*10}比{0*****}更明確地被定義；而模板{1****0*}比{1**1***}跨越(span)更多的字串部份。為了明確定義以上這些觀念，便導入兩個性質，模板的階(Schema Order),及定義長度(Defining Length)。

所謂模板的階，係指模板已經被固定的位址數目，以 $O(H)$ 代表之(H 為某一模板)。如10*1*1*，其階為4，而**1**1*，其階為2。而定義長度，則為模板內第一個定址至最後一個定址間之距離，記為

$\delta(H)$ ，如： $\{0**1*1*\}$ ，其定義長度為 $7-2=5$ ，而 $\{1*****\}$ ，其定義長度為 $7-7=0$ 。

設在時刻 t ，有模板 H 的範例 m 個，記為 $m(H,t)$ 。母群體 $A(t)$ ，其中一字串 A 的表現值為 f_i ，則存活機率定義如下：

$$p_i = \frac{f_i}{\sum f_i}$$

則於時刻 $t+1$ ，當選取 n 個子代，模板 H 的出現個數之期望值為：

$$m(H,t+1) = m(H,t) \cdot \frac{f(H)}{\sum f_i} \cdot n = m(H,t) \cdot \frac{f(H)}{\bar{f}}$$

..... (附 1-1 式)

上式之 \bar{f} 表示模板的平均表現值。由此式得到一個觀念，當一個模板的表現值(f_i)越好，則在次一代時出現的個數會增加；反之，則會減少。亦即，較平均值高的模板，較容易繼續存活，而低於平均值的模板則會逐漸減少而被淘汰。

若將上式中之 f_i 改定為 $\bar{f} + cf = (1+c) \cdot \bar{f}$ ，則(4.1)式可改為：

$$m(H,t+1) = m(H,t) \cdot \frac{(\bar{f} + cf)}{\bar{f}} = m(H,t) \cdot (1+c)$$

..... (附 1-2 式)

假設 t 由0時刻開始，則可得：

$$m(H,t+1) = m(H,0) \cdot (1+c)^t \text{ (附 1-3 式)}$$

上式在數學上的意義為，當模板複製時，其數目會隨著某一比例，而呈指數的增加(或減少)。

然而，如果只是不斷地複製字串或模板，則只會在原有的範圍內不停地搜尋而已，並不會對搜尋的結果有所助益，因為並沒有新的點產生，這也就是為什麼要運用遺傳演算元的原因。

當我們在使用交換(Crossover)的時候，假定一字串長為 L ，定義長度為 $\delta(H)$ ，而交換的位置為隨機選取，則對一個模板造成破壞的機率 P_d 為：

$$P_d = \frac{\delta(H)}{L-1} \dots\dots\dots \text{(附 1-4 式)}$$

再加上對交換過程 P_c 發生的機率的考量，則模板 H 存活的機率，至少有 $1 - P_c \cdot P_d$ ，也就是：

$$P_s \geq 1 - P_c \cdot \frac{\delta(H)}{L-1} \dots\dots\dots \text{(附 1-5 式)}$$

若假設複製和交換為獨立事件，彼此並無互相影響，則模板 H 在第 $t+1$ 時刻的個數期望值為：

$$m(H, t+1) = m(H, t) \cdot \frac{f(H)}{\bar{f}} \cdot [1 - P_c \cdot \frac{\delta(H)}{L-1}] \dots \text{(附 1-6 式)}$$

接下來再考慮突變，假設一個字元(Bit)突變的機率為 P_m ，而模板 H 的階次(Schema Order)為 $O(H)$ ，則模板 H 經過突變程序後，不被破壞的機率，應為： $(1 - P_m)^{O(H)}$ ，假設 P_m 的值遠小於1，則以泰勒展開式展開後，二階(含)以上的項皆可被忽略，故可簡化成： $1 - O(H) \cdot P_m$ ，最後，我們可以認定一個模板 H ，在經過複製交換，及突變的程序後，由第 t 代(時刻)至第 $t+1$ (時刻)的期望出現個數 m ，會有如下之關係：

$$m(H, t+1) = m(H, t) \cdot \frac{f(H)}{\bar{f}} \cdot \left[1 - P_c \cdot \frac{\delta(H)}{L-1} - O(H) \cdot P_m \right]$$

..... (附 1-7 式)

從這個式子，我們得到一個很重要的結論：定義長度較短的，階次較低的，以及對目標函數 f 表現較好的模板，存活的機會越大。此一結論甚為重要，故我們給予一特別的名稱"模板理論"(Schema Theorem)。

"模板理論"證明遺傳演算法確實能夠自主性地選擇較好的模板，從而增進模擬的精度，接著討論此一演算法的效率問題。

遺傳演算法何以是一有效率的優選方法，這由Holland所提出的隱平行性質(Implicit Parellelism)可以得到理解(Goldberg,1985)。其中思想，在於假設一代若有 n 個字串，則此一演算法不只搜尋此 n 個字串而已，其總共的模板搜尋數共有 $n3$ 個，且並沒有利用到額外的記憶體，現將此一性質推導如下：

假設一群體(Population)由 n 個字串長度為 L 的字串所構成，定模板的存活機率必須大於 P_s 始能存活下來，且假定只運用簡易交換及突變二運算元。則最後存活下來之字串或模板，其誤差率(Error Rate) ϵ 會低於 $1 - P_s$ 。這些模板長度 L_s ，則會小於 $\epsilon(L-1) + 1$ 。假設模板中有一位址為固定，其他的位置可為固定或不定的情形，再考量在字串中模板位置的可能種類共有 $L - L_s + 1$ ，則一代內所搜尋的模板數目 n_s 為：

$$n_s = 2^{(L_s-1)} \cdot (L - L_s + 1) \cdot n$$

..... (附 1-8 式)

再假設群體規模(Size)為 $2^{L_s/2}$ ，模板的二位元址分布係二項式分

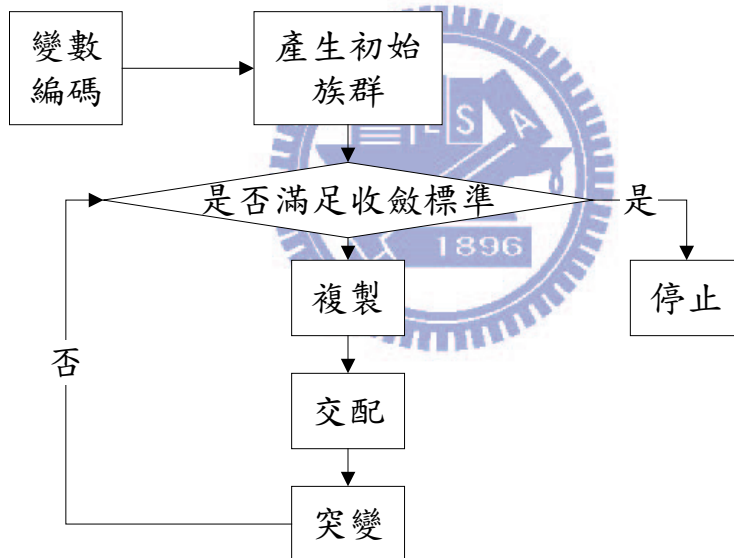
布(Binomial Distribution)，則會有一半的模板屬於較低階(Smaller Order)的存活率較高的模板，故得出一低限值：

$$n_s \geq 2^{(L_s-2)} \cdot (L - L_s + 1) \cdot n \dots\dots\dots (附 1-9 式)$$

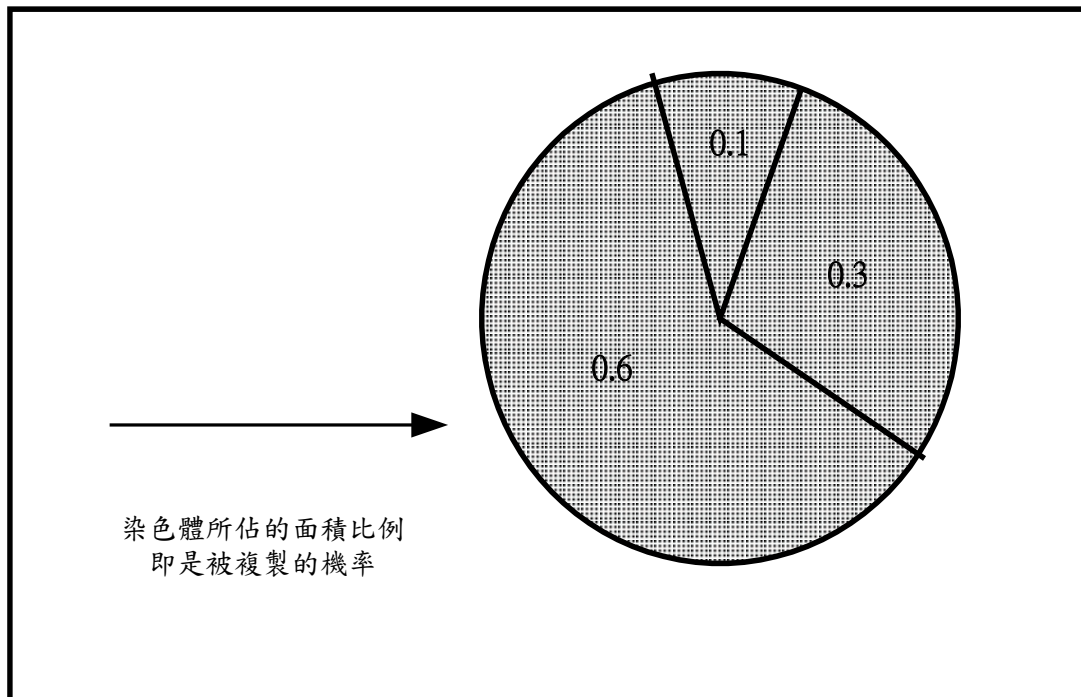
再將 n 值代入 ($2^{L_s} = n^2$) 便得：

$$n_s = \frac{(L - L_s + 1) \cdot n^3}{4} \dots\dots\dots (附 1-10 式)$$

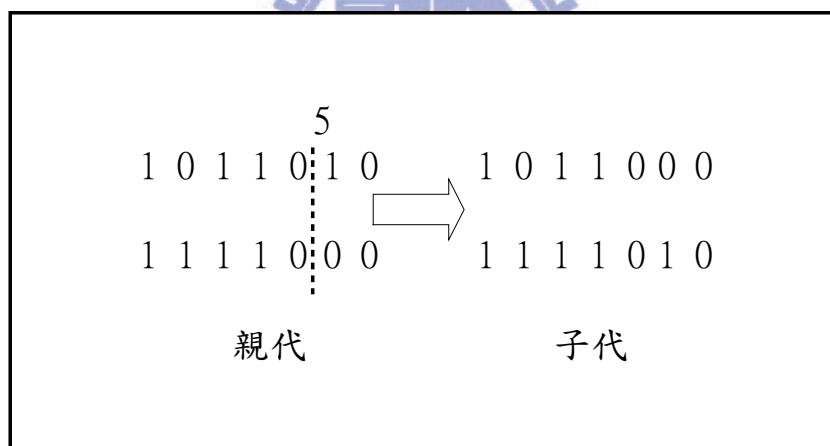
由此式可知，在遺傳演算法中，一代內所搜尋的模板數目，至少為其群體規模 n 的三次方倍。所以我們可以清楚地知道，遺傳演算法確實為一非常有效率的優選方法。



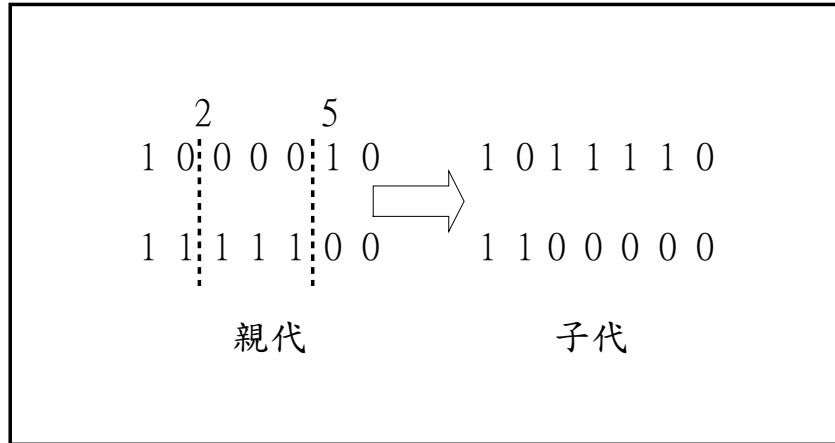
附圖 1-1 遺傳演算法流程圖



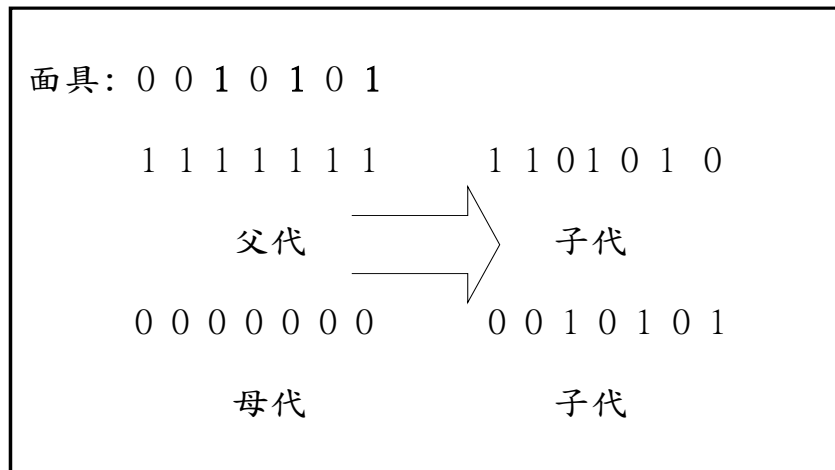
附圖 1-2 輪盤法示意圖



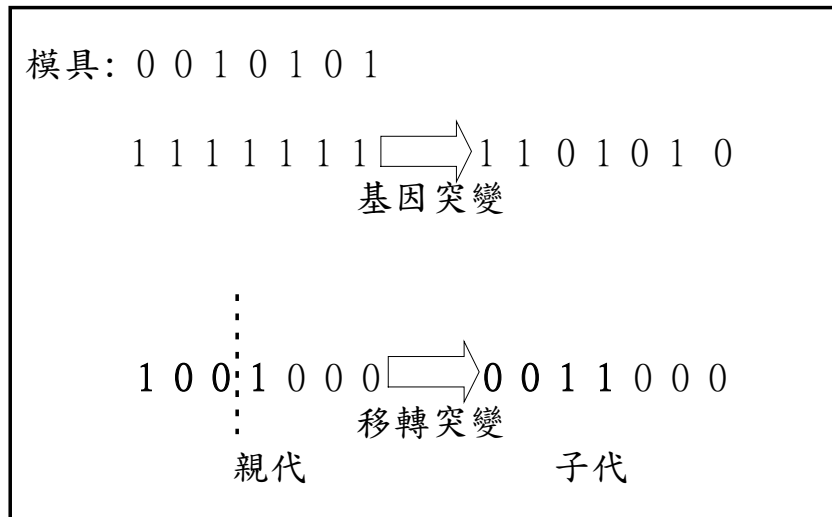
附圖 1-3 單點交配



附圖 1-4 雙點交配



附圖 1-5 均一化交配



附圖 1-6 突變示意圖



附錄二 類神經網路

類神經網路是一個可計算之多層網路，它使用大量簡單而具有平行處理能力之人工神經元來模擬人類學習行為，因此本章節將針對生物及人工神經元的構造、類神經網路的組成架構和基本理論，以及類神經網路中之倒傳遞網路(Backpropagation Network, BP)詳加介紹。

一、 類神經網路簡介

(一)生物神經元模型

生物神經網路是由巨量的神經細胞(neuron, 又稱神經元)所組成，形成一個高度連結網狀的神經網路，資訊的處理工作即透過上述之連結來進行。以人腦而言，人腦大約由1011個神經元所組成，而每一個神經元約有103根連結與其他神經元相連，所以人腦中約有1014根連結，因此人腦可以儲存大量而複雜的知識。神經元構造如附圖2-1所示，其主要構造如下：

1. 神經核(soma):神經細胞的核心，為一呈核狀的處理機構。
2. 神經軸(axon):神經細胞呈軸索狀的輸送機構。
3. 神經樹(dendrites):神經細胞呈樹枝狀的輸出入機構。
4. 神經節(synapse):神經細胞神經樹上呈點狀的連結機構。

當神經細胞透過神經節與神經樹從其它神經元輸入脈波訊號後，經神經核處理，產生一個新的脈波訊號，這個訊號再經過神經軸傳送到神經樹，再透過神經節與神經樹成為其它神經元的輸入脈波訊號，如果脈波訊號是經過興奮神經節(excitatory synapse)，則會增加脈波訊號的速率(pulse rate)，如果脈波訊號是經過抑制神經節(inhibitory synapse)，則會減少脈波訊號的速率。因此，脈波訊號的速率是同時

取決於輸入脈波訊號的速率，以及神經節的強度。而神經節的強度可視為神經網路儲存資訊之所在，神經網路的學習即在調整神經節的強度(葉怡成，1994)。

(二)人工神經元模型

根據神經細胞的結構與功能，從40年代開始，先後提出的神經元模型有數百種之多，其中對於腦模型、自動機、人工智慧有重大影響的是1943年由美國心理學家McCulloch和數學家Pitts共同提出的形式神經元模型，同常稱之為MP模型。(謝明富，1999)

此一模型（見附圖2-2所示）具有將輸入變數與輸出變數間，複雜的內在對映關係充份呈現的功能，其為解決非線性動態問題的最佳工具之一。此模型由許多人工神經元所組成，神經元又稱為節點(node)或臨界值元件(threshold element)，其輸入端輸入各自之訊息，藉由各自權重加權總和後傳入節點，透過閾值的過濾，繼而經由轉換函數轉換後輸出，其數學表示式如下：

$$Y_j = f(\text{net}_j), \text{net}_j = \sum_i W_{ij} X_i - b_j, i = 1, 2, \dots, n \text{ (附 2-1 式)}$$

其中： Y_j :為模仿生物神經元模型的輸出訊號； f :為模仿生物神經元模型的轉換函數(transfer function)，將輸入值之加權乘積和轉換成處理單元輸出值； W_{ij} :為模仿生物神經元模型的神經節強度，又稱連結加權值； X_i :為模仿生物神經元模型的輸入訊號； b_j :為模仿生物神經元模型的閾值(bias)； net_j :為輸入值之加權乘積和； n :為輸入訊號個數。

類神經網路常用之轉換函數有下列四種（見附圖2-3～附2-7）：

1. 位階臨界轉換函數(Step Threshold Transfer Function)：臨

界函數的輸出只隨輸入值的正副號所改變。

2. 線性轉換函數(Linear Transfer Function)：函數輸入值與輸出值呈線性關係。
3. 非線性轉換函數(Nonlinear Transfer Function)：函數輸出的最大值與最小值限制在一個特定的範圍內。
4. 臨界 S 型轉換函數(Sigmoid Threshold Transfer Function)：是最常被應用之函數，因為此種函數型態具有可微分且連續等性質，此特質使網路可以應用到非線性的學習領域中。常用函數有雙彎曲函數(附 2-2 式)與雙曲線正切函數(附 2-3 式)兩種。

$$f(x) = \frac{1}{1 + e^{-x}} \dots\dots\dots (附 2-2 式)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \dots\dots\dots (附 2-3 式)$$



二、類神經網路基本理論

類神經網路的架構乃仿照人類腦部神經網路的模型而發展出來的，一個類神經網路是由許多個人工神經元連結組成，並可以組成各種網路模式。(葉怡成，1994)

類神經網路的總體運作模式有兩種：(1)學習過程(Learning)－網路依學習演算法，從範例中學習，以調整網路權重值的過程，與(2)回想過程(Recalling)－網路依學習演算法，以輸入資料決定網路輸出資料的過程。

此外，學習過程的學習演算法又可分為三類：

1. 監督式學習

從問題領域中取得訓練範例(有輸入變數值，也有輸出變數)，並從中學習輸入變數與輸出變數的內在對映規則，以應用於新的案例(只有輸入變數值，而需推論輸出變數值的應用)。感知機網路、倒傳遞網路、機率神經網路、學習向量量化網路與反傳遞網路等五種類神經網路皆屬監督式學習。

2. 無監督式學習

從問題領域中取得訓練範例(只有輸入變數值)，並從中學習範例的內在集群規則，以應用於新的案例(有輸入變數值，而需推論它與那些訓練範例屬同一集群的應用)。自組織映射圖網路、自適應共振理論網路等兩種類神經網路皆屬無監督式學習。

3. 聯想式學習

從問題領域中取得訓練範例(狀態變數值)，並從中學習範例的內在記憶規則，以應用於新的案例(只有不完整的狀態變數值，而需推論其完整的狀態變數值的應用)。霍普菲爾網路、雙向聯想記憶網路等兩種類神經網路皆屬聯想式學習。

三、 倒傳遞神經網路 (BP)

在監督式學習模式中，由Rumelhart等學者於1986年所發展的誤差向後推導或稱倒傳遞學習演算法（簡稱BP）是被廣泛使用的一種學習演算法，因為其具有學習及回想的功能，故可進行定率預測。一

般倒傳遞網路可分為三部份(如附圖2-8所示)，輸入層用以接受外在環境的訊息，其神經元數目則依問題而定；輸出層用以輸出訊息給外在環境，其神經元數目同樣依問題而定；隱藏層將輸入與輸出層各處理單元間的相互關係充份地表現出來，其神經元數目並無標準可決定。倒傳遞網路模式學習訓練方式由所探討問題中取得相當數量之訓練樣本，並從樣本中應用向前餽入與誤差向後推導兩步驟推求輸入變數與輸出變數的內在對映規則，再應用回想功能，進行新案例之輸出變數值推估。下列為倒傳遞網路 (Bp) 之學習與回想過程建立步驟：

1. 倒傳遞網路學習過程的建立：

(1) 令輸入層、隱藏層與輸出層節點分別以 i 、 j 、 k 為下標符號，建立一輸入層、隱藏層與輸出層節點數分別為 m 、 n 、 o 個，以均勻分佈隨機亂數設定加權值矩陣 W_{ij} 、 W_{jk} 與閾值 b_j 、 b_k 的初始值。

(2) 輸入一個訓練案例的輸入向量 $\vec{X} = [x_i] \quad i = 1, 2, \dots, m$ 與目標輸出向量 $\vec{T} = [T_k] \quad k = 1, 2, \dots, o$ 。

(3) 計算推估輸出向量 \vec{O} 。

A. 輸入層向量

$$\vec{X} = [x_i] \quad i = 1, 2, \dots, m$$

B. 計算隱藏層輸出向量 \vec{H}

$$H_j = f_h(net_j), \quad net_j = \sum_{i=1}^m x_i w_{ij} - b_j, \quad j = 1, 2, \dots, n \quad (\text{附$$

2-4 式)

其中 $f_h(x)$ 為隱藏層轉換函數，可依問題型態挑選適當轉換函數。

C. 計算輸出層輸出向量 \vec{O}

$$O_k = f_o(net_k), \quad net_k = \sum_{j=1}^n H_j w_{jk} - b_k, \quad k=1,2,\dots,o \quad (\text{附 2-5 式})$$

2-5 式)

其中 $f_o(x)$ 為輸出層轉換函數，可依問題型態挑選適當轉換函數。

(4) 計算加權值矩陣修正量 Δw ，及閾值修正量 Δb 。

因為監督式學習的目的在降低網路輸出單元目標輸出值與推論輸出值之差距，所以一般以能量函數（又稱誤差函數）

表示學習的品質：

$$E = \frac{1}{2} \sum_{k=1}^o (O_k - T_k)^2 \quad \dots\dots\dots (\text{附 2-6 式})$$

因此網路的學習過程即為使能量函數最小化的過程，通常以最陡坡降法來使能量函數最小化，即每輸入一個訓練案例，網路即小幅調整加權值的大小，調整的幅度和誤差函數對該加權值的敏感程度成正比，即與誤差函數對加權值的偏微分值大小成正比：

$$\Delta w = -\eta \frac{\partial E}{\partial w} \quad \dots\dots\dots (\text{附 2-7 式})$$

式中：

Δw : 加權值調整的幅度。

η : 學習速率(learning rate)，控制每次以最陡坡降法最小化誤差函數的步幅，其範圍為 $0 \leq \eta \leq 1$ ，可視所需狀況自行設定範圍容許內之值。

經由一連串的代入及演算可推得以下各值：

A. 輸出層

$$\Delta w_{jk} = \eta(T_k - O_k) \cdot df_o(net_k) \cdot H_j \dots\dots\dots (附 2-8 式)$$

$$\Delta b_k = -\eta(T_k - O_k) \cdot df_o(net_k) \dots\dots\dots (附 2-9 式)$$

其中 $df_o(x)$ 為輸出層轉換函數之一階導函數，舉例說明：

若以雙彎曲函數 $f_o(x) = \frac{1}{1+e^{-x}}$ 為輸出層之轉換函數，其一階導

函數則為 $df_o(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ 。

B. 隱藏層

$$\Delta w_{ij} = \eta(T_k - O_k) \cdot df_o(net_k) \cdot w_{jk} \cdot df_h(net_j) \cdot x_i \dots\dots\dots (附 2-10 式)$$

$$\Delta b_j = -\eta(T_k - O_k) \cdot df_o(net_k) \cdot w_{jk} \cdot df_h(net_j) \dots\dots\dots (附 2-11 式)$$

其中 $df_h(x)$ 為隱藏層轉換函數之一階導函數。

(5) 更新加權值矩陣與閾值

$$w_{new} = w_{old} + \Delta w \dots\dots\dots (附 2-12 式)$$

$$b_{new} = b_{old} + \Delta b \dots\dots\dots (附 2-13 式)$$

(6) 重覆步驟 2 至步驟 6，直至收斂，即誤差不再有明顯變化且符合所要求的精度。

2. 倒傳遞網路回想過程的建立：

(1) 讀入加權值矩陣與閾值矩陣。

(2) 輸入一個欲作檢定或驗證的案例之輸入向量

$$\vec{X} = [x_i] \quad i = 1, 2, \dots, m。$$

(3) 計算推估輸出向量 \vec{O} 。

A. 輸入層向量

$$\vec{X} = [x_i] \quad i = 1, 2, \dots, m$$

B. 計算隱藏層輸出向量 \vec{H}

$$H_j = f_h(\text{net}_j), \quad \text{net}_j = \sum_{i=1}^m x_i w_{ij} - b_j \quad j = 1, 2, 3, \dots, n \quad (\text{附$$

2-14 式)

其中 $f_h(x)$ 為隱藏層轉換函數，可依問題型態挑選適當轉換函數。

C. 計算輸出層輸出向量 \vec{O}

$$O_k = f_o(\text{net}_k), \quad \text{net}_k = \sum_{j=1}^n H_j w_{jk} - b_k \quad k = 1, 2, 3, \dots, o$$

..... (附 2-15 式)

其中 $f_o(x)$ 為輸出層轉換函數，可依問題型態挑選適當轉換函數。

四、 敏感度分析

對任一模式之敏感度可定義為：模式輸出值受輸入變數變化影響程度。舉例說明：若模式對某一輸入變數極為敏感，則表示當此一變數有些微改變時，將導致模式輸出值產生巨幅變化。

以數學說明表示某輸入變數之敏感度，即為模式輸出變數對該輸入變數之一階偏微分值。舉例說明，假設輸入變數為維度 n 之向量 \vec{x} ，模式輸出值為維度 m 之向量 \vec{y} ，輸入變數與輸出值之函數關係為 $\vec{y} = f(\vec{x})$ ，若想知道各輸入變數對第 k 個輸出值之敏感程度，我們可以

計算第k個模式輸出值對各輸入變數之維度n一階偏微分值向量

$$(\bar{y}'_k = \left[\left[\frac{\partial y_k}{\partial x_i} \right] \right] \quad i=1,2,3,\dots,n)$$

對微分值向量之任一分量而言，正值代表輸出值隨該輸入變數增大而增大，負值則反向變化，其絕對值越大則代表該輸入變數對模式輸出值之影響程度越大，越趨近於零則表示模式輸出值對該變數變化並不敏感。因此，吾人可藉由模式之一階偏微分值向量，瞭解任一變數對模式之影響程度。

可是許多時候難以得到輸入變數對模式輸出值之微分解析式，通常都以差分式近似之。而類神經網路之倒傳遞神經網路由人工神經元所組成，其組成函數為可微分函數，因此可推導得倒傳遞神經網路輸出值對各輸入變數之一階偏微分解析式。

由前節倒傳遞神經網路(BP)的介紹可知，一單隱藏層之倒傳遞神經網路之輸出值 O_k 可表示如下：

$$O_k = f_o(net_k), \quad net_k = \sum_{j=1}^n H_j w_{jk} - b_k \quad k=1,2,3,\dots,o$$

$$H_j = f_h(net_j), \quad net_j = \sum_{i=1}^m x_i w_{ij} - b_j \quad j=1,2,3,\dots,n$$

其中， O_k 、 H_j 分別為輸出層、隱藏層之輸出單元， x_i 則為輸入層之輸入單元。此外，W為權重值，b為閾值， $f_o(x)$ 、 $f_h(x)$ 分別為輸出層與隱藏層之轉換函數。因此各輸入變數對輸出層輸出單元之一階偏微分解析式如下：

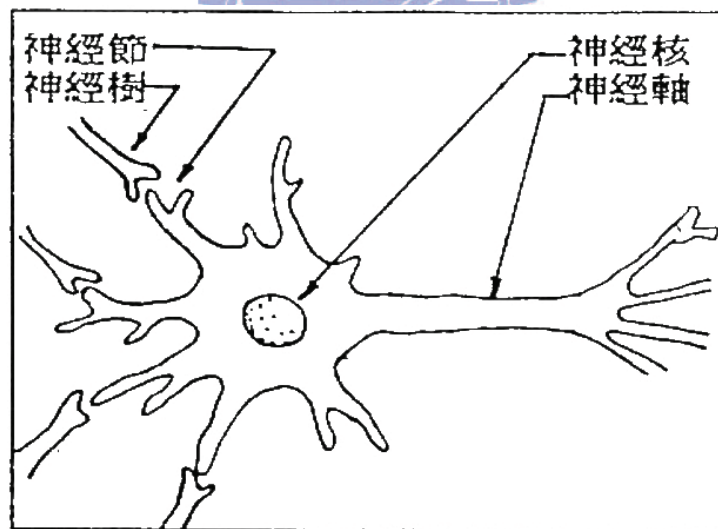
$$\frac{\partial O_k}{\partial x_i} = \sum_{k=1}^n df_o(net_k) \cdot w_{jk} \cdot df_h(net_j) \cdot w_{ij} \dots\dots\dots \text{(附 2-16 式)}$$

$$\text{其中：} df_o(x) = \frac{\partial f_o(x)}{\partial x}, \quad df_h(x) = \frac{\partial f_h(x)}{\partial x}$$

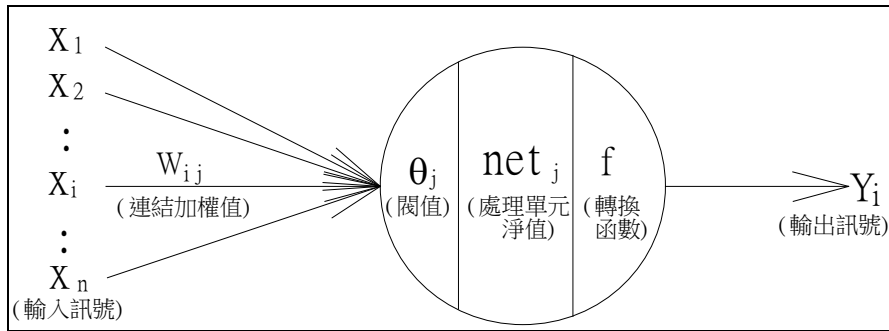
五、 類神經網路敏感度分析實例驗證

本案例選擇一簡單可微分函數以作為類神經網路敏感度分析之實例驗證，在此以 $y = 0.45 \sin(\pi x) + 0.5, -1 \leq x \leq 1$ 為訓練範本，建立一個單一隱藏層之單輸入單輸出類神經網路，隱藏層節點數為9個，輸出層與隱藏層之轉換函數均為雙彎曲函數 $f(x) = \frac{1}{1+e^{-x}}$ 。收斂標準訂為誤差值小於0.03。

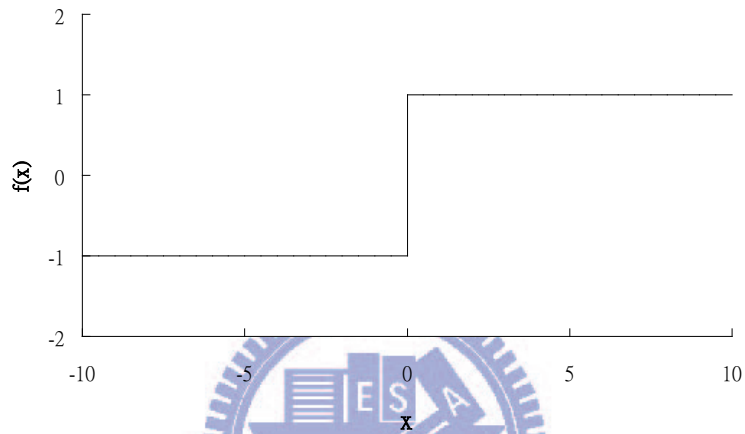
類神經網路訓練成功後（見附圖2-9），應用其微分解析式與該函數之微分式 $y' = 0.45\pi \times \cos(\pi x)$ 比較（如附圖2-10所示），可發現僅在輸入變數之上下限附近之微分值誤差較大，於中央部份較能抓住實際函數微分值變化趨勢。



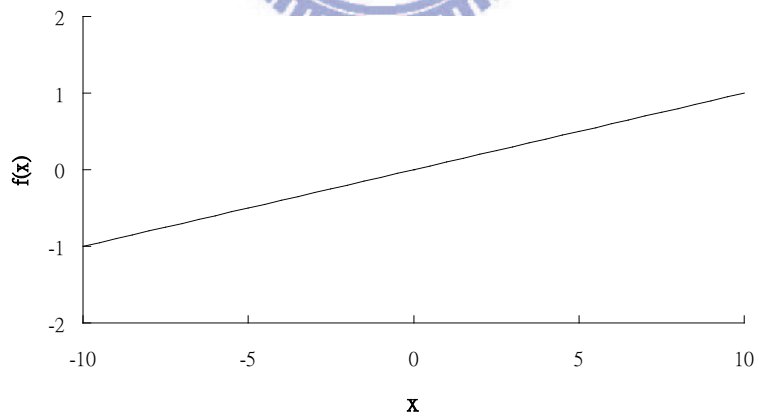
附圖 2-1 物神經元模型（葉怡成，1993）



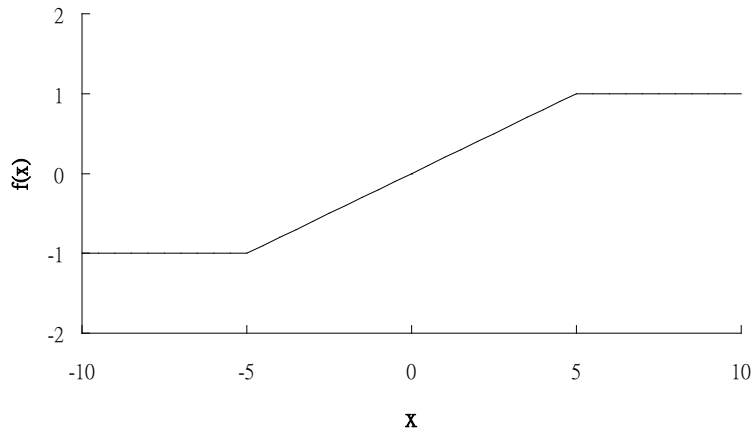
附圖 2-2 人工神經元模型



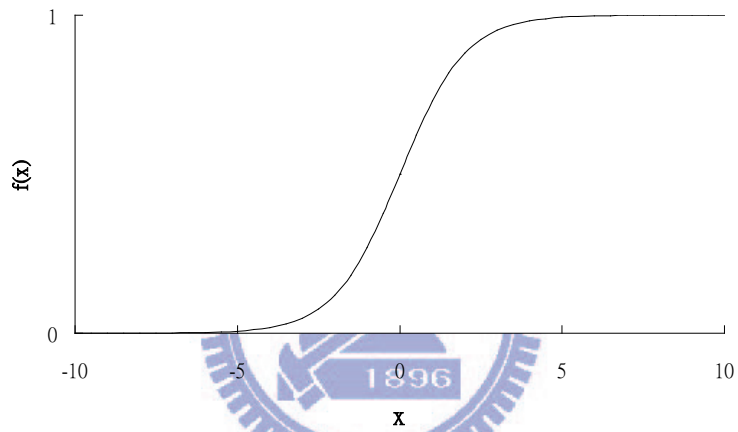
附圖 2-3 位階臨界轉換函數圖



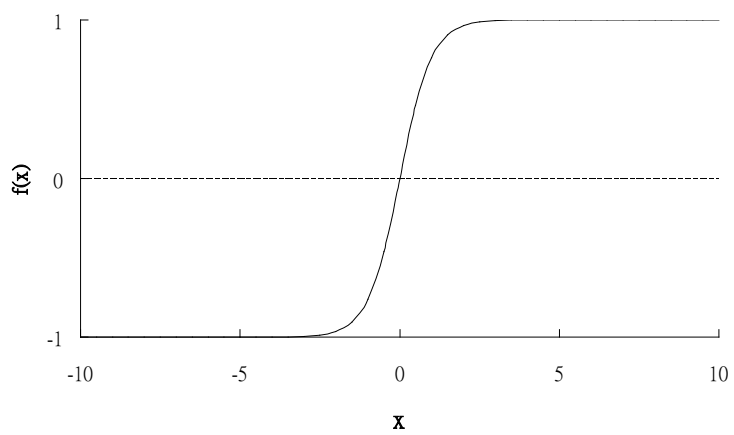
附圖 2-4 線性轉換函數圖



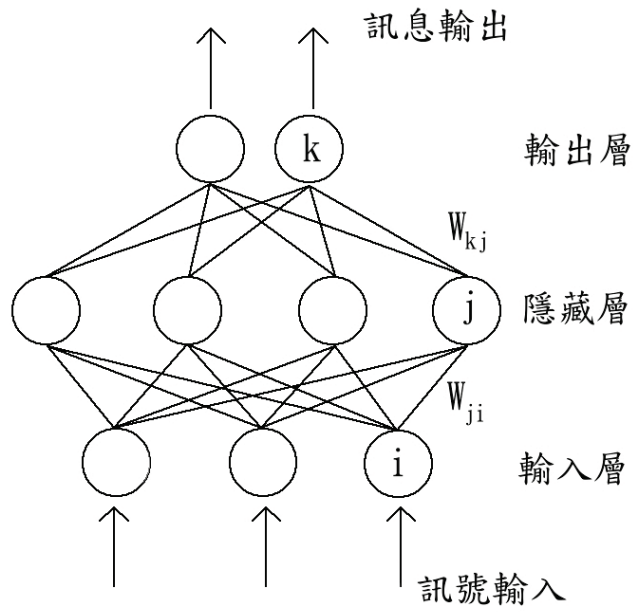
附圖 2-5 非線性轉換函數圖



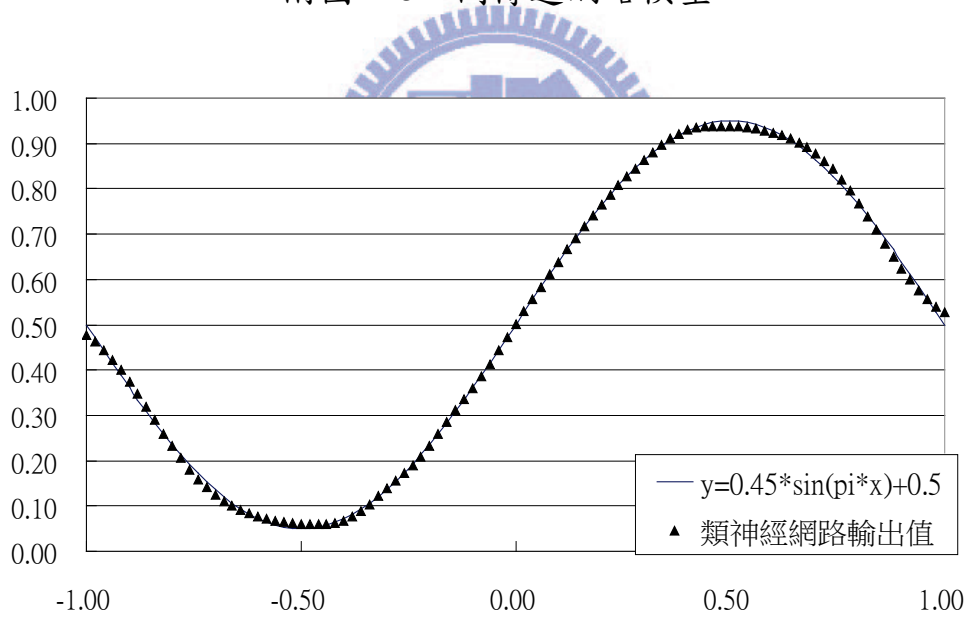
附圖 2-6 雙彎曲轉換函數圖



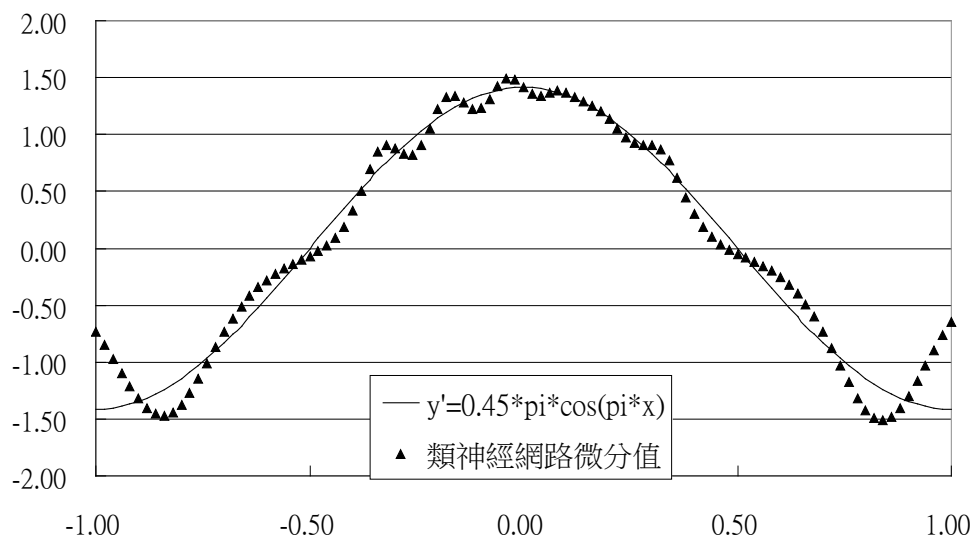
附圖 2-7 雙曲線正切轉換函數圖



附圖 2-8 倒傳遞網路模型



附圖 2-9 實際函數與類神經網路預測值軌跡圖



附圖 2-10 實際微分函數與類神經網路微分預測值軌跡圖

