

以資料變換方式建立之模糊推論系統

陳鴻進* 黃有評**

壹、計畫緣由及目的
貳、研究方法
參、模擬結果
肆、結論
伍、參考文獻

摘要

昔日的模糊推論系統，均以調整歸屬函數的中心點及左右寬度，去滿足每一個輸入點。因此為了滿足某一條件而去調整一條模糊法則，可能破壞其他法則已建構出的模型；如此將造成耗時。對於已處理過的資料，本文提出另一觀點的調整方法。每個前提變數的歸屬函數是相等分佈且固定在整個轉換值域中。為了得到較佳的建構模型，任一個轉換函數，結論部分或上述二者均須被最佳化。因此本文將每個三角形的歸屬函數視為二個沒有共同接點的三角形，使得每一條法則將被拆為二條無關的法則。因此改善了法則調整中相互破壞的情形。本文採用最陡坡降法來調整模糊推論中的參數。經由模擬結果，我們驗證本研究所提方法較傳統其它方法更能得到更佳的结果。

關鍵字：最陡坡降法，模糊模型，轉換函數

* 陳鴻進：電子工程科專任講師

** 黃有評：大葉大學資訊工程系教授

壹、計畫緣由及目的

本計畫之目的就是要提出一套簡單實用且有系統的作法來改進模糊模型的建立與調整。我們的做法是在原始資料輸入模糊模型前先做前處理動作。資料之前處理是透過所設定之轉換函數，將資料轉換後映射至另一空間如圖三所示，再將轉換後之資料代入模糊模型中，以執行模糊化及推論工作。

貳、研究方法

我們規劃之研究方法及進行步驟如下：

- (1)、建立資料前處理之轉換函數：設定一連續、可微分之轉換函數。一般可調式模糊模型是在調整歸屬函數之區間、形狀，但是往往調整步驟繁瑣，調整參數眾多，其效率不佳且成效有限。因此，我們利用轉換函數將資料轉換，使原本對應到歸屬函數之程度值得以因而改變。表面上看似乎沒什麼意義，但是深究其內涵可知，一般是調整歸屬函數使資料能最佳地對應至某一歸屬函數，而我們轉換函數的目的地是相同的，只是作法恰好相反，是調整轉換函數使資料能因轉換後對應至某一歸屬函數，所以其目的與傳統方法相同均可達成調整模糊模型的優點，只是作法不同，即我們的作法只須調整轉換函數之參數即可，而不必修改歸屬函數之區間、形狀。調整轉換函數的方法是採用最陡坡降的方式(因轉換函數是連續、可微分的)。其理論作法如下：若第 i 條法則如下，

Rule i : If x_1 is A_{1i} and x_2 is A_{2i} and ...and x_n is A_{ni} , then y is w_i .

法則被驅動之強度：

$$\mu_i = \prod_{j=1}^n A_{ji}(k(x_j)),$$

其中 $A_{ji}(k(x_j))$ 表示轉換後之變數 x_j 在模糊集 A_{ji} 上之歸屬度。因系統可能有多條法則同時被驅動，經推論後可得如下之輸出：

$$y_c = (\sum_{i=1}^m \mu_i w_i) / \sum_{i=1}^m \mu_i,$$

若定義誤差函數為：

$$E = \frac{1}{2}(y_c - y_d)^2,$$

其中 y_d 為期望輸出值。藉此，即可利用微分方式求取預調之參數 p_j 下一步調整之方向 $\partial E / \partial p_j$ ，即：

$$\frac{\partial E}{\partial p_j} = \frac{\partial E}{\partial y_c} \frac{\partial y_c}{\partial \mu_i} \frac{\partial \mu_i}{\partial A_{ji}(k(x_j))} \frac{\partial A_{ji}(k(x_j))}{\partial k(x_j)} \frac{\partial k(x_j)}{\partial p_j}$$

而新的參數 p_j 即可以下式獲得：

$$p_j(t+1) = p_j(t) - \gamma_p \frac{\partial E}{\partial p_j},$$

其中 γ_p 代表學習速率。同樣地，我們亦可求取模糊法則結論部分之調整方式：

$$\begin{aligned} w_i(t+1) &= w_i(t) - \gamma_w \frac{\partial E}{\partial w_i} \\ &= w_i(t) - \gamma_w \frac{\partial E}{\partial y_c} \frac{\partial y_c}{\partial w_i}, \end{aligned}$$

其中 γ_w 代表結論部分之學習速率。因此，只要定義好轉換函數即可進行調整並進行推論工作。我們目前提出一種可行之轉換函數型態為：

$$k(x_i) = R / [1 + e^{-n(x_i - c)}], i = 1, 2, \dots, n.$$

其中 R 決定轉換後之值域範圍， c 代表轉換資料之最小值，而 n 則代表變數個數。

- (2)、簡化模糊系統之建立：待選定系統所需之變數並建構一套適用之轉換函數後，即可依已知資料進行參數調整、分析模擬結果並與其它方法進行比較，從而朝向架構簡單、實用效果佳之設計目標前進。簡言之，我們提出之改良方式的作法與傳統模糊模型作法比較如圖三所示，即本計畫將原始資料經由函數轉換至另一定義域後再映射至固定型之歸屬函數執行模糊化工作，而傳統做法是以動態調整歸屬函數來達成最佳化執行模糊化之工作。

參、模擬結果

本文模擬函數採用單輸入單輸出(SISO)函數與雙輸入單輸出(DISO)函數 ;7 或 14 個歸屬函數。

參、一、

單輸入單輸出(SISO)函數：

待測函數：

$$y = 3x(x-1)(x-1.9)(x+0.7)(x+1.8), -2.1 \leq x \leq 2.1$$

轉換函數為 $k(x) = 4.2/[1 + e^{-px}]$.

參數 p 調整偏量:

$$\frac{\partial E}{\partial p} = (y_c - y_d) \left(\frac{w_i - y_c}{\sum \mu_i} \right) \left(\frac{\mu_i}{A_i(k(x))} \right) \left(\frac{-\text{sgn}(k(x) - a_i)}{b_i} \right) \cdot x \cdot k(x) \left(1 - \frac{k(x)}{4.2} \right)$$

其中 a_i 表三角型歸屬函數的中心點， $2b_i$ 表三角型歸屬函數的底邊寬度。

結論調整偏量為 $\frac{\partial E}{\partial w_i} = (y_c - y_d) \frac{\mu_i}{\sum \mu_i}$.

p 的初值為 0.00000005, 學習速率 γ_p 為 0.0005, 學習速率 γ_w 為 0.0005

下一個新偏量值為 $p(t+1) = p(t) - \gamma_p \frac{\partial E}{\partial p}$.

參、二.

雙輸入單輸出(DISO)函數:

待測函數:

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2, 1 \leq x_1, x_2 \leq 5.$$

轉換函數為

$$k(x_i) = 6/[1 + e^{-p_i(x_i-1)}], i = 1, 2.$$

參數 p 調整偏量:

$$\frac{\partial E}{\partial p_j} = (y_c - y_d) \left(\frac{w_j - y_c}{\sum \mu_i} \right) \left(\frac{\mu_i}{A_{ji}(k(x_j))} \right) \left(\frac{-\text{sgn}(k(x_j) - a_{ji})}{b_{ji}} \right) (x_j - 1) k(x_j) \left(1 - \frac{k(x_j)}{6} \right)$$

因雙輸入, 故 $\partial E / \partial p_j$ 取平均值 i.e., 下一個新偏量值為

$$p_i(t+1) = p_i(t) - \gamma_p \left(\frac{\partial E}{\partial p_1} + \frac{\partial E}{\partial p_2} \right) / 2.$$

表一顯示出模擬結果, 可以看出 14 個歸屬函數優於 7 個歸屬函數。

表二表示 mean squared errors (MSE) 在不同方法上的比較表。

圖一表示單輸入單輸出函數於 14 個歸屬函數以 simple function 所得之圖型。

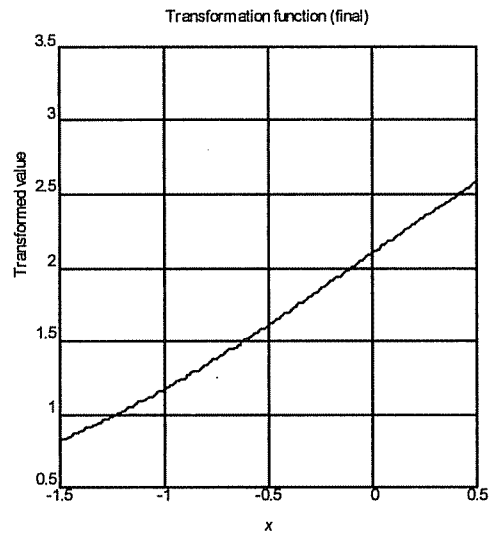
圖二表示雙輸入單輸出函數於 14 個歸屬函數以 simple function 所得之圖型。

肆、結論

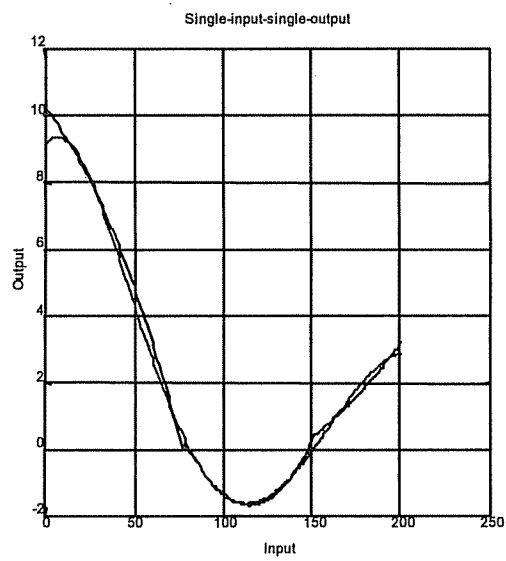
本文採用灰關聯函數的關聯分析方法來篩選出重要的前題變數，簡化結構辨識之過程，以建立模糊模型，降低系統之複雜度。並且在原始資料輸入模糊模型前先藉由轉換函數將資料轉換映射至另一定義域，類比於調整歸屬函數的中心點及其左右寬度。此一轉換函數為一連續且可微分的S型函數，故可由連鎖法則及微分方程來推導出系統的誤差函數，再以最陡坡降法來調整系統的參數以降低系統的推論誤差。模擬結果並與它法比較(表二)，驗證了此法優於其他方法。此建模未重大修改又兼備可調，學習的特性，是此一領域罕見獨創的方法。

伍、參考文獻

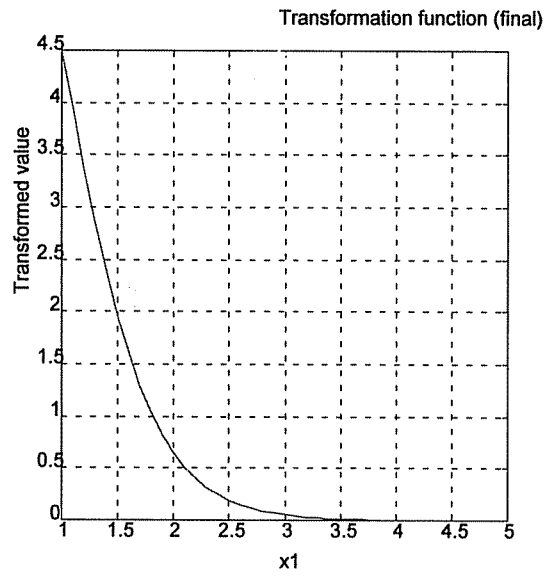
1. J. L. Deng. Control problems of grey systems, *Systems and Control Letters*, 5 (1982) 228-294.
2. . Nomura, I. Hayashi, and N. Wakami, A learning method of fuzzy inference rules by descent method, *Proc. FUZZ-IEEE* (1992) 203-210.
3. F. Guely and P. Siarry, Gradient descent method for optimizing various fuzzy rule bases, *Proc. FUZZ-IEEE* (1993) 1241-1246.
4. . Homaifar and E. McCormick, Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms, *IEEE Trans. Fuzzy Systems* 3(2) (1995) 129-139.
5. . Sugeno and T. Yasukawa, A fuzzy-logic-based approach to qualitative modeling, *IEEE Trans. Fuzzy Systems* 1(1) (1993) 7-31.



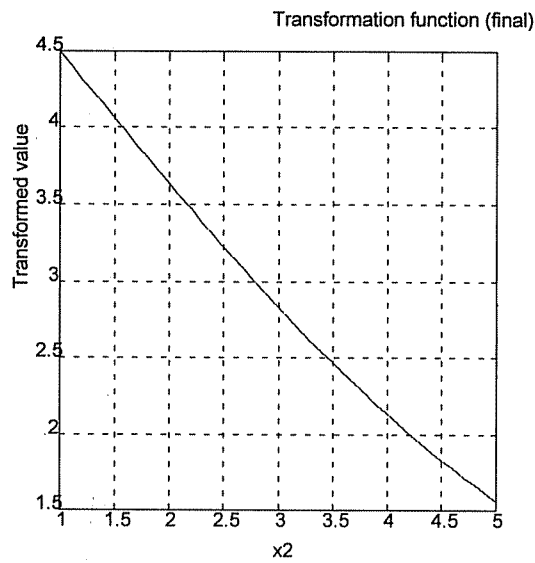
圖一.(A)表示單輸入單輸出函數的轉換函數圖型.



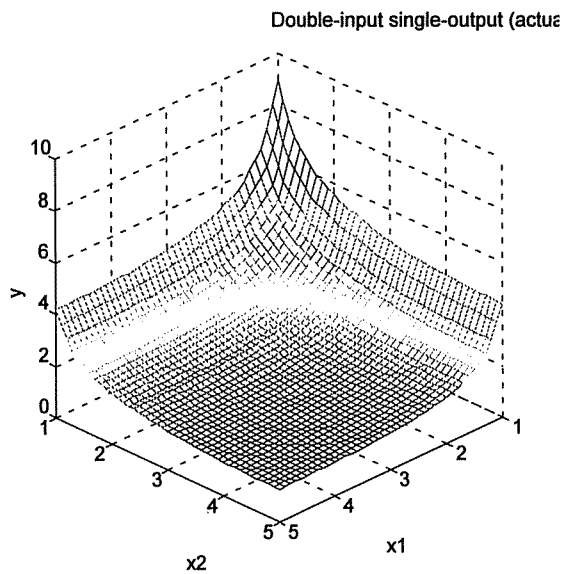
圖一.(B)表示單輸入單輸出函數之實際輸出(實線)與推論輸出(虛線)之圖型.



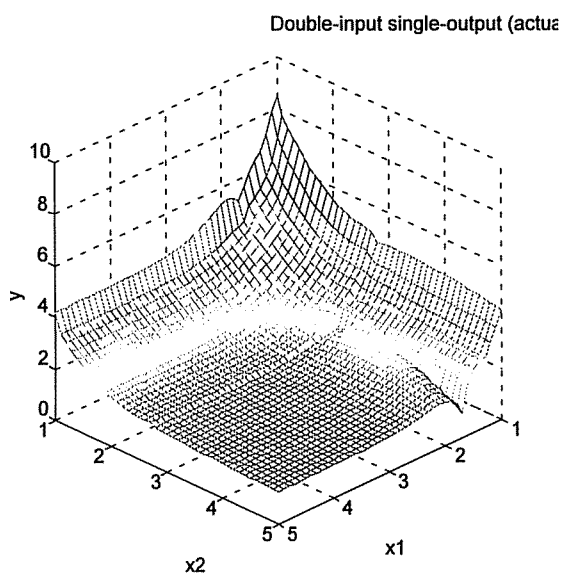
圖二.(A)表示雙輸入單輸出函數之 x_1 的轉換函數圖型.



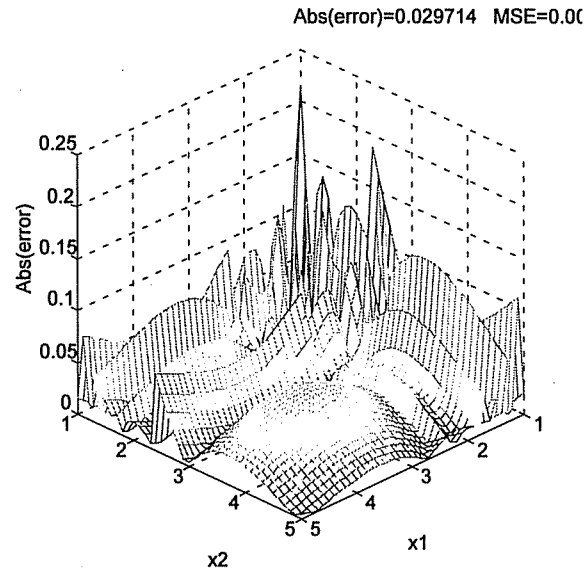
圖二.(B)表示雙輸入單輸出函數之 x_2 的轉換函數圖型.



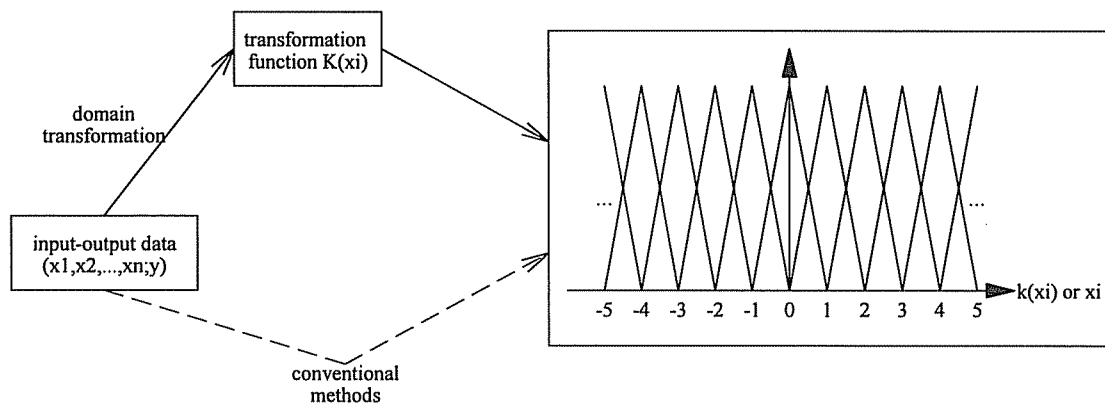
圖二.(C)表示雙輸入單輸出函數之實際輸出之圖型



圖二.(D) 表示雙輸入單輸出函數之推論輸出圖型, 平均誤差為 0.029714, MSE 為 0.001491



圖二(E) 示雙輸入單輸出函數之絕對誤差輸出圖型



圖三. 改良方式的作法與傳統模糊模型作法比較

表一 模擬結果之比較。CP 表結論部份的型式。No. MF 表歸屬函數的個數。Avg. 表平均絕對誤差。MSE is the mean squared error. Avg. D is the average absolute differences.

	Single-input-single-output system				
	CP	No. MF	Avg.	MSE	Avg. D
	real numbers	7	0.3054	0.1209	1.391134
	real numbers	14	0.1758	0.0551	0.220473
	simple functions	7	0.2379	0.1115	
	simple functions	14	0.1767	0.0653	
	Double-input-single-output system				
	CP	No. MF	Avg.	MSE	Avg. D
	real numbers	7	0.040032	0.002470	0.0385440
	real numbers	14	0.029714	0.001491	0.0297030
	simple functions	7	0.105136	0.019700	
	simple functions	14	0.081677	0.013900	

表二 mean squared errors (MSE)在不同方法上的比較表。

Method	MSE
our model (3-label for each input)	0.04701
our model (5-label for each input)	0.02257
our model (7-label for each input)	0.008385
gradient descent method	0.01574
from [5]- without parameter identification	0.318
from [5]- with parameter identification	0.079
from [5]- position-gradient type	0.010
from [5]	0.073

The design of a fuzzy system based on data transformation methods

Hung-Jin Chen Yo-Ping Hung

Abstract

In the past day of a inference fuzzy system, we always refine the central point and width of a membership function in order to fitting the input of a system We use the gradient descent method in this proposed system. Since adjusting a rule to satisfy one pattern may deteriorate the others performance and result in a lengthy tuning process. This paper approaches from another viewpoint to adjusting the fuzzy model built for the preprocessed data. The membership functions defined for each premise variable are equally distributed and fixed in the transformed domain. To better identify the fuzzy model, either the transformation functions or consequent parts of fuzzy rules or both need to be optimized. We then treat each triangular membership functions as two disjoint ones such that each fuzzy rule is divided into mutually independent rules. This in turn benefits the refinement of consequent parts in the fuzzy rules since adjusting a rule will not be demonstrated but also the results from the conventional approaches will be givens for comparisons. We use the least squared method to calculates the desired consequent real numbers for the data located in the same region of transformed domain. The conformity of the after-tuned consequence parts in the fuzzy rules with the desired values further verifies the effectiveness of the presented methodology.

Keywords: Gradient descent methods; Fuzzy modeling; Transformed function

