Mathware & Soft Computing 5 (1998) 141-149

Fuzzy Inference Using a Least Square Model

H. Bustince, M. Calderón and V. Mohedano
Departamento de Automática y Computación.
Universidad Pública de Navarra.
Campus de Arrosadía. 31006 Pamplona. España.
e-mail: bustince@si.upna.es, vmohedano@si.upna.es

Abstract

In this paper, the method of least squares is applied to the fuzzy inference rules. We begin studying the conditions in which from a fuzzy set we can build another through the method of least squares. Then we apply this technique in order to evaluate the conclusions of the generalized modus ponens. We present different theorems and examples that demonstrate the fundamental advantages of the method studied.
Keywords: Approximate reasoning; fuzzy inference rules; generalized modus ponens; least squares model; fuzzy logic; power function

1 Introduction

The classical modus ponens is expressed by

$$\begin{array}{l} A \longrightarrow B \\ \underline{A \qquad\quad} \\ \qquad\quad B \end{array}$$

this means that if both:

A implies B and A are true, then
B is also true.

This reasoning scheme was extended to fuzzy reasoning by Zadeh $[13 - 16]$ as follows:

A) The implication $A \longrightarrow B$ is replaced by the fuzzy inference rule

if x is A then y is B

where A and B may be fuzzy sets, A in a universe of discourse X and B in a universe of discourse Y; and x is a variable which takes values in X, and y is a variable which takes values in Y. The fuzzy rule represents a relation between two variables x and y.

B) Similarly, the premise A is replaced by a fuzzy premise:

x is A'

where A' is a fuzzy set, expressing the knowledge we have about the value of x. A' is a fuzzy set in a universe of discourse X.

Combining the rule and the premise, it is possible to deduce a new piece of information, written

$$y \text{ is } B'$$

where B' is a fuzzy set in a universe of discourse Y.

The membership functions of A and A' are written as $\mu_A(x)$ and $\mu_{A'}(x)$ respectively; the membership functions of B and B' are denoted as $\mu_B(y)$ and $\mu_{B'}(y)$ respectively. The other well known classic method of inference is the modus tollens which is expressed as follows:

$$\begin{array}{l} A \longrightarrow B \\ \underline{\quad \text{not } B} \\ \text{not } A \end{array}$$

Therefore, exact reasoning modus ponens and modus tollens can be extended to approximate reasoning which deals with the inherent vagueness of human language. Thus, the generalized modus ponens (GMP) is introduced to reach a conclusion from fuzzy premises. These rules can be expressed in standard as follows

$$\begin{array}{ll} \text{If } x \text{ is } A \text{ then } y \text{ is } B & \\ \hline x \text{ is } A' & \\ \hline y \text{ is } B' & \text{(GMP)} \end{array}$$

The main advantage of these extensions to fuzzy reasoning is being able to deduce new information, even when the knowledge is not exactly identical to the condition of the rule or when the information we consider is not crisp. It is noted that in (GMP) when $A = A'$, then the generalized modus ponens reduce to the case of the modus ponens.

For the fuzzy inference rules, different methods have been suggested by various authors such as Zadeh [9, 12], Fukami [4], Mizumoto and Zimmermann [6], Ezawa and Kandel [3].

In this paper we are going to study the generalized modus ponens following a different reasoning from the one used by above-named authors. The basis will be the ideas established by J.F. Baldwin [1], A. Nafarie [7], M.B. Gorzalczany [5], etc., which can be summarized in the two following steps:

1) First relate A to A'
2) Build the consequence B' using the result of the comparison above and B.

We will carry out step 1) following the method presented by T.D. Pham and S. Valliappan [8], analysing and afterwards formalizing its main consequences.

We begin the paper studying the way to obtain from a fuzzy set another which is fuzzy as well, by means of the method of least squares. We focus above all on the conditions in which said construction can be carried out. We conclude this section analysing the main consequences of the method of construction presented. We must make clear that in this paper we at no time consider whether the approximation we make is appropiate or not, that is, we at no time use Pearson's coefficient, etc., for as we will see in the following section of the paper, in fact what we are interested in are solely and exclusively the coefficients calculated from the approximation carried out.

In section three we present a method for obtaining the consequence of the GMP. In thise case we study the main characteristics of said method, highlighting, through different examples, its great applicability when the membership functions in the premises are characterized by the fuzzy linguistic hedge operators as defined by Zadeh [9 − 11].

Like T. C. Chang, K. Hasegawa and C.W. Ibbs [2], for convenience we take the universes of discourse in this paper as follows:

1) $X = \{x_1, \cdots, x_n\}$, $Y = \{y_1, \cdots, y_m\}$ and $Z = \{z_1, \cdots, x_p\}$
2) $x_i, y_j, z_k \in (0, 1]$ for all $i = 1, \cdots, n,\ j = 1, \cdots, m,\ k = 1, \cdots, p$
3) $x_i < x_{i+1},\ \ y_j < y_{j+1}$ and $z_k < z_{k+1}$ for all i, j, k

We will denote as $FS(X)$ the set of all fuzzy sets on X, we will denote as $FS^*(X)$ the set of all fuzzy sets on X such that for all $x_i \in X$ it is verified that $\mu(x_i) \neq 0$. We will denote as A_c the complementary of the fuzzy set A and for $p \in \mathbb{R}^+ \cup \{0\}$, we will write $A^p = \{< x_i, \mu_A^p(x_i) > | x_i \in X\}$, evidently, p has to be greater than or equal to zero.

2 Construction of a $FS(X)$ from a $FS^*(X)$

Our objective is to approximate the membership functions of a fuzzy set $A \in FS^*(X)$ to ax_i^b type functions, where for each set $A \in FS^*(X)$ constants a and b are obtained by means of the following algebraic lineal equations:

$$(1)\left\{\begin{array}{l} n \text{ Ln } a + (\sum_{i=1}^{n} \mathrm{lnx_i})b = \sum_{i=1}^{n} \text{Ln } \mu_A(x_i) \\ (\sum_{i=1}^{n} \text{Ln } x_i) \text{ Ln } a + (\sum_{i=1}^{n} \text{Ln}^2 x_i)b = \sum_{i=1}^{n} \text{Ln } x_i \text{Ln } \mu_A(x_i) \end{array}\right.$$

obtained from taking logarithms in the expression $\mu_A(x_i) = ax_i^b$ and applying the method of least squares. Given the nature of the problem we are dealing with, it is clear that *the constant a has to be greater than zero* and it also must always be: for all $i = 1, \cdots, n,\ ax_i^b \leq 1$. The following theorem establishes the conditions in which the above-mentioned condition holds.

Theorem 1. *Let $A \in FS^*(X)$ and let a and b be the solutions of (1).*
$ax_i^b \leq 1$ if and only if

$$\left\{\begin{array}{l} 1)\ \ \textit{if}\ \ b \geq 0, \ \ \textit{then}\ |\sum_{i=1}^{n} Ln\ \mu_A(x_i)| \geq (nLn\ x_{max} - \sum_{i=1}^{n} Ln\ x_i)b \\ 2)\ \ \textit{if}\ \ b < 0, \ \ \textit{then}\ |\sum_{i=1}^{n} Ln\ \mu_A(x_i)| \geq (nLn\ x_{min} - \sum_{i=1}^{n} Ln\ x_i)b. \end{array}\right.$$

Proof.
$\Rightarrow$) $ax_i^b \leq 1$, then Ln $a + b$Ln $x_i \leq 0$ for all $i = 1, \cdots, n$, therefore

$$\text{Ln } a + b\text{Ln } x_{\max} \leq 0$$
$$\text{Ln } a + b\text{Ln } x_{\min} \leq 0.$$

1) If $b \geq 0$, we have that

$$\text{Ln } a + b \text{ Ln } x_{\max} = \\ = \frac{\sum_{i=1}^{n} \text{ Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{ Ln } x_i)b}{n} + b \text{ Ln } x_{\max} = \\ = \frac{\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{ Ln } x_i)b + bn \text{ Ln } x_{\max}}{n} \leq 0, \text{ therefore}$$
$$\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{Ln } x_i)b + bn \text{ n } x_{\max} \leq 0.$$

Since $x_{\min} = x_1 < x_2 \cdots < x_n = x_{\max}$, it follows that Ln $x_1 <$ Ln $x_2 < \cdots <$ Ln $x_n \leq 0$, therefore $\sum_{i=1}^{n}$ Ln $x_i \leq n$ Ln x_n, so $|\sum_{i=1}^{n}$ Ln $x_i| \geq |n$ Ln $x_n|$ from which we have that n Ln $x_n - \sum_{i=1}^{n}$ Ln $x_i \geq 0$, and since

$$\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{Ln } x_i)b + bn\text{Ln } x_{\max} \leq 0$$

is

$$|\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)| \geq (n\text{Ln } x_{\max} - \sum_{i=1}^{n} \text{Ln } x_i)b.$$

2) If $b < 0$, Ln $a + b$Ln $x_i \leq 0$ for all $i = 1, \cdots, n$, then Ln $a + b$Ln $x_1 \leq 0$, therefore

$$\frac{\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{Ln } x_i)b + bn\text{Ln } x_1}{n} \leq 0$$
$$\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{Ln } x_i)b + bn\text{Ln } x_1 \leq 0$$

since nLn $x_1 \leq \sum_{i=1}^{n}$ Ln x_i, it follows that $|n$Ln $x_1| \geq |\sum_{i=1}^{n}$ Ln $x_i|$, therefore nLn $x_1 - \sum_{i=1}^{n}$ Ln $x_i \leq 0$, so $(n$Ln $x_1 - \sum_{i=1}^{n}$ Ln $x_i)b \geq 0$, from which

$$|\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)| \geq (n\text{Ln } x_{\min} - \sum_{i=1}^{n} x_i)b.$$

$\Leftarrow$) If $b \geq 0$, then $|\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)| \geq (n\text{Ln } x_{\max} - \sum_{i=1}^{n} x_i)b$, through the same reasoning as before we have $(n\text{Ln } x_n - \sum_{i=1}^{n} \text{Ln } x_i)b \geq 0$; since
$|\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)| \geq (n\text{Ln } x_{\max} - \sum_{i=1}^{n} x_i)b \geq 0$, we have

$$\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) + (n\text{Ln } x_n - \sum_{i=1}^{n} \text{Ln } x_i)b \leq 0, \text{ then}$$
$$\frac{\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) + (n\text{Ln } x_n - \sum_{i=1}^{n} \text{Ln } x_i)b}{n} \leq 0, \text{ then}$$
$$\frac{\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{Ln } x_i)b}{n} + b\text{Ln } x_n \leq 0$$

therefore $\text{Ln } a + b\text{Ln } x_n \leq 0$, that is $ax_n^b \leq 1$, since $b \geq 0$ we have that $x_i^b \leq x_n^b$, then $ax_i^b \leq ax_n^b$ since $a > 0$, therefore $ax_i^b \leq ax_n^b \leq 1$, then $ax_i^b \leq 1$ for all $i = 1, \cdots, n$.
2) If $b < 0$, $|\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)| \geq (n\text{Ln } x_{\min} - \sum_{i=1}^{n} \text{Ln } x_i)b$, since $x_1 < x_2 < \cdots < x_n$ following a reasoning analogous to the above we have that $|n\text{Ln } x_1| \geq |\sum_{i=1}^{n} \text{Ln } x_i|$, then $n\text{Ln } x_1 - \sum_{i=1}^{n} \text{Ln } x_i \leq 0$ and since $b < 0$ we have $(n\text{Ln } x_1 - \sum_{i=1}^{n} \text{Ln } x_i)b \geq 0$, therefore $\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) + (n\text{Ln } x_1 - \sum_{i=1}^{n} \text{Ln } x_i)b \leq 0$, so

$$\frac{\sum_{i=1}^{n} \text{Ln } \mu_A(x_i) - (\sum_{i=1}^{n} \text{L n} x_i)b}{n} + b\text{Ln } x_1 \leq 0$$

that is, $\text{Ln } a + b\text{Ln } x_1 \leq 0$, from which $ax_1^b \leq 1$ and since $b < 0$ we have that $x_1^b \geq x_i^b$, taking into account that $a > 0$ it follows that $ax_1^b \geq ax_i^b$, then $ax_i^b \leq ax_1^b \leq 1$ for all $i = 1, \cdots, n$. □

Corollary 1. *In the same conditions as in theorem 1, the set* $\overset{\frown}{A} = \{< x_i, \mu_{\overset{\frown}{A}}(x_i) > |x_i \in X\}$ *with* $\mu_{\overset{\frown}{A}}(x_i) = ax_i^b$ *is a fuzzy set on* X.

Remark.

Up to now we have considered $FS^*(X)$ fuzzy sets, that is, sets such that for all $x_i \in X$ it is verified that $\mu(x_i) \neq 0$. Nevertheless, the construction analysed before can be generalized to any fuzzy set as indicated below.

If in the set $A \in FS(X)$ considered there exists an x_i with null membership function, $\overset{\frown}{A}$ is constructed by means of the following two steps:

1) With the method above we build the new memberships of the elements with membership value different from zero.
2) For the elements with null membership, this is maintained.

Theorem 2. *Let $A \in FS^*(X)$.*
$b = 0$ *if and only if* $(\sum_{i=1}^{n} \text{Ln } x_i) \cdot (\sum_{i=1}^{n} \text{Ln } \mu_A(x_i)) = n \cdot (\sum_{i=1}^{n} \text{Ln } x_i \text{Ln } \mu_A(x_i))$.

Proof. Is easily deduced clearing b in (1). □

Theorem 3. *Let $A, A' \in FS(X)$ and let $\overset{\frown}{A}, \overset{\frown}{A'} \in FS(X)$ the sets built from A and A' respectively in accordance with the corollary 1. Let $p \in \mathbb{R}^+ \cup \{0\}$, if $A' = A^p$, then $\overset{\frown}{A'} = (\overset{\frown}{A})^p$, holds.*

Proof.
Let $\mu_{\overset{\frown}{A}}(x_i) = a x_i^b$ and $\mu_{\overset{\frown}{A'}}(x_i) = a' x_i^{b'}$. From (1) we deduce that:

$$n \text{ Ln } a' + (\sum_{i=1}^{n} \text{Ln } x_i) b' = \sum_{i=1}^{n} \text{Ln } \mu_{A'}(x_i) = \sum_{i=1}^{n} \text{Ln } \mu_A^p(x_i)$$
$$= p \sum_{i=1}^{n} \text{Ln } \mu_A(x_i) = pn \text{ Ln } a + p(\sum_{i=1}^{n} \text{ Ln } x_i) b \quad (2)$$

$$(\sum_{i=1}^{n} \text{Ln } x_i) \text{ Ln } a' + (\sum_{i=1}^{n} \text{Ln}^2 x_i) b' = \sum_{i=1}^{n} \text{Ln } x_i \text{Ln } \mu_{A'}(x_i) =$$
$$= \sum_{i=1}^{n} \text{Ln } x_i \text{ Ln } \mu_A^p(x_i) = p \sum_{i=1}^{n} \text{Ln } x_i \text{ Ln } \mu_A(x_i) =$$
$$= p(\sum_{i=1}^{n} \text{Ln } x_i) \text{ Ln } a + p(\sum_{i=1}^{n} \text{Ln}^2 x_i) b \quad (3)$$

solving (2) we have

$$\text{Ln } \frac{a'}{a^p} = \frac{(\sum_{i=1}^{n} \text{Ln } x_i)(p \cdot b - b')}{n}$$

from (3) we have

$$(\sum_{i=1}^{n} \text{Ln } x_i) \text{Ln } \frac{a'}{a^p} = (\sum_{i=1}^{n} \text{Ln}^2 x_i)(p \cdot b - b')$$

substituting Ln $\frac{a'}{a^p}$ in this expression and taking into account that $(\sum_{i=1}^{n} \text{Ln } x_i)^2$ $-n(\sum_{i=1}^{n} \text{Ln}^2 x_i) \neq 0$, we have $b' = p \cdot b$ and $a' = a^p$. □

All of the research carried out in this section will be much used in the following sections of the paper. We are going to analyse operations of several linguistic hedges the same as Zadeh did in [10, 11], for this reason it is important to bear in mind the following consequences of the theorem above:

a) If $A' = A^2$, then $\breve{A'} = (\breve{A})^2$,

b) If $A' = A^{\frac{1}{2}}$, then $\breve{A'} = (\breve{A})^{\frac{1}{2}}$

c) If $A' = A^4$, then $\breve{A'} = (\breve{A})^4$

3 Fuzzy inference using a least square model

We have said earlier that in fuzzy logic, exact reasoning can be extended to approximate reasoning which deals with the inherent vagueness of human language. Thus, the generalized modus ponens (GMP) and the generalized modus tollens (GMT) are introduced to reach a conclusion from fuzzy premises.

Zadeh [10, 11] has introduced the operations of several linguistic hedges to convey a better understanding of human language. The operators of such hedges as *very, more o less, highly, plus* and *minus* are expressed in terms of the membership functions. Thus, taking advantage of these properties, a method for fuzzy rules of inference was proposed by T.D. Pham and S. Valliappan in [8]. This method consists of three following steps:

Method for the GMP

1.- Introduce a function $\mu_{A'}(x_i)$ in terms of $\mu_A(x)$ in order to obtain the coefficients of the relation between $\mu_{A'}(x_i)$ and $\mu_A(x_i)$. We carry out this step approximating the membership functions of the sets A and A' to ax_i^b and $a'x_i^{b'}$ type functions respectively; this is done by solving (1) for A and A'

2.- We calculate $c_1 = \frac{a'}{a^{\frac{b'}{b}}}$(obviously$0 < c_1$)and $c_2 = \frac{b'}{b}$, so that $\mu_{\breve{A'}} = c_1 \mu_{\breve{A}}^{c_2}(x_i)$.

3.- These coefficients of relation can be passed to $\mu_B(y_j)$ for the inference of $\mu_{B'}(y_j)$, that is, $\mu_{B'}(y_j) = c_1 \mu_B^{c_2}(y_j)$ for all $j = 1, \cdots, m$.

T.D. Pham and S. Valliappan in [8] did not consider the possibility that for a y_j with $j = 1, \cdots, m$, $c_1 \mu_B^{c_2}(y_j)$ could be greater than one. In this line we enunciate the following theorem.

Theorem 4. *Under the conditions of the three steps above we have:*
$0 \le c_1 \mu_B^{c_2}(y_j) \le 1$ *for all* $j = 1, \cdots, m$, *if and only if*

$$\begin{cases} 1)\, c_2 \ge 0 \quad \text{and}\; c_1 \mu_{B_{max}}^{c_2}(y_j) \le 1 \\ 2)\, c_2 < 0 \quad \text{and}\; c_1 \mu_{B_{min}}^{c_2}(y_j) \le 1. \end{cases}$$

Proof. Evident. □

Theorem 5. *Let $p \in \mathbb{R}^+ \cup \{0\}$. In the conditions in theorem 4, if $A' = A^p$, then $B' = B^p$ holds.*

Proof. It is enough to remember that $c_1 = \frac{a'}{a^{\frac{b'}{b}}}$, $c_2 = \frac{b'}{b}$ and theorem 3. □

Due to its importance, we highlight the following particular cases of the previous theorem:

a) If $A' = A^2$, then $B' = B^2$,
b) If $A' = A^{\frac{1}{2}}$, then $B' = B^{\frac{1}{2}}$
c) If $A' = A^4$, then $B' = B^4$.

4 Conclusions

In the case when this method can be applied, that is, if the conditions in theorem 4 hold, we find ourselves before a mecanism of inference that has the following two advantages:
1) It can be generally applicable when the membership functions in the premises are characterized by the fuzzy linguistic hedge operators as defined by Zadeh $[9-12]$.
2) The mathematical operations on the generalized modus ponens are simple and convenient for computer implementation.

It is important to say that the primary and fundamental *disadvantage* of this method is the following: to approximate the membership functions of fuzzy sets to ax_i^b type functions, it should be noticed that ax_i^b type functions are monotone. If the membership function $\mu_A(x_i)$ of a fuzzy set $A \in FS^*(X)$ is not monotone, such as "middle", the approximation between $\mu_A(x_i)$ and ax_i^b using the method of least squares is unreasonable. In this case, the method of fuzzy inference using a least squares model must be improved.

References

[1] J.F. Baldwin, N.C.F. Guild, Feasible algorithms for approximate reasoning using fuzzy logic, *Fuzzy Sets and Systems* **3** (1980) 225-251.

[2] T.C. Chang, K. Hasegawa and C.W. Ibbs, The effects of membership function on fuzzy reasoning, *Fuzzy Sets and Systems* **44** (1991) 169-186.

[3] Y. Ezawa and A. Kandel, Robust fuzzy inference, *Int. J. Intelligent Systems* **6** (1991) 185-197.

[4] S. Fukami, M. Mizumoto and K. Tanaka, Some considerations on fuzzy conditional inference, *Fuzzy Sets and Systems* **4** (1980) 243-273.

[5] M.B. Gorzalczany, A method of inference in approximate reasoning based on interval-valued fuzzy sets, *Fuzzy Sets and Systems* **21** (1987) 1-17.

[6] M. Mizumoto and H.-J. Zimmermann, Comparison of fuzzy reasoning methods, *Fuzzy Sets and Systems* **8** (1982) 253-283.

[7] A. Nafarieh, A new approach to inference in approximate reasoning and its application to computer vision, Ph D. Dissertation, University of Missouri Columbia, (1988).

[8] T.D. Pham and S. Valliappan, A least squares model for fuzzy rules of inference, *Fuzzy Sets and Systems* **64** (1994) 207-212.

[9] L.A. Zadeh, Fuzzy logic, *IEEE Computer* April (1988) 83-93.

[10] L.A. Zadeh, The concept of a linguistic variable, and its application to the approximate reasoning. Part III, *Information Sciences* **8** (1975) 43-80.

[11] L.A. Zadeh, The concept of a linguistic variable, and its application to the approximate reasoning. Part II, *Information Sciences* **8** (1975) 301-357.

[12] L.A. Zadeh, A fuzzy-sets theoretic interpretation of linguistic hedges, *J. Cybernetics* **2** (1972) 4-34.