FUZZY LINGUISTIC VARIABLE MATRIX AND PARABOLA-BASED FUZZY NORMAL DISTRIBUTION

A Method for Designing Fuzzy Linguistic terms

K.K.F. Yuen, H.C.W. Lau

Department of Industrial and Systems Engineering ,Hong Kong Polytechnic University, Hung Hom,Kowloon, Hong Kong, China,

Email:{ise.kevinyuen, mfhenry}@polyu.edu.hk

Tel:(852)-60112169,27666628

Abstract: This paper attempts to present the new approach to design sufficient number of systematic fuzzy linguistics in matrix form and map the Fuzzy Linguistic Variable Matrix, which contains linguistic terms, into numeric domain using Fuzzy Normal Distribution based on the Parabola-based Membership Function. Existing fuzzy set theory is difficult to design the systematic and sufficient fuzzy linguistics. Due to this reason, in most practice, giving insufficient fuzzy linguistics induces inaccurate calculation whilst giving excessive fuzzy linguistics induces the parameter design problems and calculation performance. This paper presents Fuzzy Linguistic Variable Matrix and Parabola-based Fuzzy Normal Distribution (FND) as preferred framework to address the problem.

Keywords: Fuzzy Set, Fuzzy Logic, Fuzzy Linguistics Variable Matrix, Parabola-based Membership Function, Fuzzy Normal Distribution, Directional Hedge Linguistics

1. INTRODUCTION

It is well known that most people use Gaussian normal distribution for the statistic model and probability model, which is widely used in many applications. However, fundamental assumption of Gaussian normal distribution is entailed by the axiom of additively where all probabilities that satisfy specific properties must add to 1. This forces the conclusion that probability of an event necessarily entails knowledge of remaining events. This articulates the challenge of modeling any uncertainty associated with an expert judgment**.**

For The fuzzy set theory, axiom of additives of where all probabilities (memberships) equal to 1 is not applied due to the fact that fuzzy set is the study of possibility instead of probability (D. Dubois and H. Prade, 1988). In addition, the motivation for selecting fuzzy set theory and fuzzy logic can be characterized by the following reasons:

1. when the measurement of the event is not given(J. C., Helton, 1997);
2. when that information is nonspecific, ambiguous, or conflicting(J. C. Helton, 1997);
3. when the information can be described by human using adverb or adjective;

For above reasons, it seems that there is a lack of models to handle the uncertainty on the basis of the classical probability. With consideration of the capability dealing with uncertainty and ambiguity for above reasons, fuzzy set theory and fuzzy logic are the preferred choices of the models. As the age of the fuzzy set theory is still young (L.A Zadeh, 1965), it has great potential to improve its theory. This project introduces the concept of Fuzzy Linguistic Variable Matrix (FLVM) and Parabola-based Fuzzy Normal Distribution (FND) for dealing with uncertainty and ambiguity.

Firstly the fundamental concept of fuzzy set theory in explained in section 2. Then the definition of FLVM and Parabola-based Membership Function (PbMF) are depicted in section 3 and section 4 respectively. The concept of FND with Fuzzy Density Function (FDF) is depicted in section 5. Finally the conclusion as well as the identification of future study is depicted in section 6.

2. FUNDAMENTAL CONCEPT

Fuzzy sets were introduced by Zadeh (1965) was specifically designed to mathematically represent uncertainty and vagueness and to provide formalized tools for dealing with the imprecision intrinsic to many problems. To illustrate the idea how this paper innovatively modifies the existing fuzzy set theory, it is necessary to review the fundamental concept of fuzzy set theory and fuzzy logic.

Definition 2.1: Let X be a nonempty set, A fuzzy set α in X is characterized by its membership function: $\mu_\alpha : X \to [0,1]$, and $\mu_\alpha(x)$ is interpreted as the degree of membership of element x in Set α for each $x \in X$ (L.A Zadeh, 1965).

Definition 2.2: A linguistic variable is characterized by a quintuple in which x is the name of variable; T(x) is the term set of x, that is, the set of names of linguistic values of x with each value being a fuzzy number defined on U; G is a syntactic rule for generating the names of values of x; and M is a semantic rule for associating with each value its meaning. (L.A Zadeh, 1978)

Most application researches apply triangular membership function. Other membership functions widely used include Cauchy, Gaussian, sigmoidal, and trapezoidal membership functions. However, the designs or/and the calculations of the membership functions are relatively complex or inefficient. To simplify them, this paper proposes FLVM and PbMF as the preferred alternatives for the fuzzy applications.

3. FUZZY LINGUISTIC VARIABLE MATRIX

This section introduces the approach to design fuzzy linguistic terms using FLVM. The human intelligence possesses the superior capability of fuzzy classification, fuzzy judgment, and fuzzy reasoning. If there is suitable linguistic schema, there will be the framework making the classification, judgment and reasoning more objective.

Definition 3.1 (Fuzzy syntactic representation)**:** The syntactic pattern of a fuzzy linguistic variable set,α, consists of three syntactic components: a direction linguistic variable set, a hedge linguistic variable set, and an atomic linguistic variable set, denoted as $\mathrm{V_d}, \mathrm{V_h}, \mathrm{V_a}$ respectively. In syntactic terms, the expression is $\alpha = (\mathrm{V_h} + \mathrm{V_d}) + \mathrm{V_a}$, where $(\mathrm{V_h} + \mathrm{V_d})$ forms syntactic terms set as a directional hedge linguistic set denoted as $\mathrm{V_{hd}}$.

Let X be the universal set, $\mathrm{V_{hd}}$ gives the column vector, $\mathrm{V_a}$ gives the row vector, and α_{ij} is the fuzzy linguistic variable (set) which its linguistic syntactic representation is determined by its row and column position.

$$X = \left[\alpha_{ij}\right] = \left[\begin{array}{c|ccc} & [\mathrm{V_a}]_1 & \cdots & [\mathrm{V_a}]_n \\ \hline [\mathrm{V_{hd}}]_1 & \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ [\mathrm{V_{hd}}]_m & \alpha_{m1} & \cdots & \alpha_{mn} \end{array}\right]$$

Definition 3.2 (Fuzzy semantic constraint): In the universal set X_{mn}, $\underset{1 \le i \le (m-1)/2}{\underset{j=n}{\forall}} \alpha_{ij} = \Phi$ and $\underset{(m+1)/2 \le i \le m}{\underset{j=1}{\forall}} \alpha_{ij} = \Phi$ where Φ is null set and is regarded as 0.

Example 1

Provided that

$$V_d = \{\text{above, below}\}$$
$$V_h = \{\text{much,quite,little,absolutely}\}$$
$$V_a = \{\text{poor,average,excellent}\}$$

V_h and V_d are used together as V_{hd} to modify the V_a. Therefore,

$$V_{hd} = \left\{ \begin{array}{l} \text{much above, quite above, little above, absolutely,} \\ \text{little below, quite below, much below} \end{array} \right\}$$

Assign V_{hd} as row matrix and V_a as column matrix, then get table 1.

Table 1: Matrix of Fuzzy Linguistic terms for a variable

	Poor	Average	Excellent
Much Above	much above poor	much above average	-
Quite Above	quite above poor	quite above average	-
Little Above	little above poor	little above average	-
Absolutely	absolutely poor	absolutely average	absolutely excellent
Little Below	-	little below average	little below excellent
Quite Below	-	quite below average	quite below excellent
Much Below	-	much below average	much below excellent

Let X be the universal set, its fuzzy subsets are represented in a matrix form on the basis of table 1 such that

$$X = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} \\ \alpha_{51} & \alpha_{52} & \alpha_{53} \\ \alpha_{61} & \alpha_{62} & \alpha_{63} \\ \alpha_{71} & \alpha_{72} & \alpha_{73} \end{vmatrix} = \begin{vmatrix} MA-P & MA-A & 0 \\ QA-P & QA-A & 0 \\ LA-P & LA-A & 0 \\ A-P & A-A & A-E \\ 0 & LB-A & LB-E \\ 0 & QB-A & QB-E \\ 0 & MB-A & MB-E \end{vmatrix}$$

4. PARABOLA-BASED MEMBERSHIP FUNCTION

"A correct and good membership function is determined by the user based on his scientific knowledge, working experience, and actual need for the particular application in question. This selection is more or less subjective, but the situation is just like in the classical probability theory and statistics where if one says 'we assume that the noise is Gaussian and while,' what he uses to start with all the rigorous mathematics is a subjective hypothesis that may not be very true, simply because the noise in question may not be exactly Gaussian and may not be perfectly white." (Chen Guanrong and Trung Tat Pham, 2001). This paper assumes the membership function shape is parabolic, and the parabolic shape can be modified with g-level method.

Definition 4.1 (symmetric fuzzy set, γ_{α_n} and d_{α_n})**:** A symmetric fuzzy set is determined by its membership function where there is only one element (or singleton), γ_{α_n}, with membership = 1 and the two end points with equal distance, d_{α_n}, to γ_{α_n}.

Definition 4.2 (PMF)**:** Let X be the universal set, and x is any element in the set X. α is the fuzzy subset. There are n subsets of α, which is a nonempty set. For each subset α_n, $\alpha_n \subseteq X$ where finite n = {1, 2, 3…i}. In other words, α_n is the subset with index n. Therefore, the Parabolic Membership Function (PMF) of α_n, $y_{\alpha_n} : \alpha_n \rightarrow [0,1]$, is defined as:

$$y_{\alpha_n}(x) = a_{\alpha_n} x^2 + b_{\alpha_n} x + c_{\alpha_n}, \text{ where } x \in \alpha_n \quad (1)$$

Theorem 1 (PMF)**:** On the basis of the definition 4.1 and 4.2**,** the parabolic membership function can be expressed as

$$y_{\alpha_n}(d_{\alpha_n}, \gamma_{\alpha_n}, x) = \frac{-1}{d_{\alpha_n}^2} x^2 + \frac{2\gamma_{\alpha_n}}{d_{\alpha_n}^2} x + \frac{d_{\alpha_n}^2 - \gamma_{\alpha_n}^2}{d_{\alpha_n}^2} \quad (2)$$

The membership function is used for the atomic linguistic variable.

Theorem 2 (fuzzy set overlap)**:** The fuzzy set overlap, δ, is defined as the cross point at a degree of the membership of two adjacent sets

($0 \leq \delta < 1$).To obtain the cross point in the parabolic membership function , d_{α_n} is defined as:

$$d_{\alpha_n} = \frac{\gamma_{\alpha_{n+1}} - \gamma_{\alpha_n}}{2\sqrt{1-\delta}} \tag{3}$$

Definition 4.3 (g-level, PbMF)**:** The shape of the PMF can be tuned as PbMF by giving the power of g_{α_n} , where $4 \geq g > 0$ suggested. The assignment of g_{α_n} is called g-level. The new membership function of PbMF, $\mu_{\alpha_n}(x)$, is defined as:

$$\mu_{\alpha_n}(x) = \left[y_{\alpha_n}(x) \right]^{g_{\alpha_n}} \tag{4}$$

The PbMF is used for the atomic linguistic set.

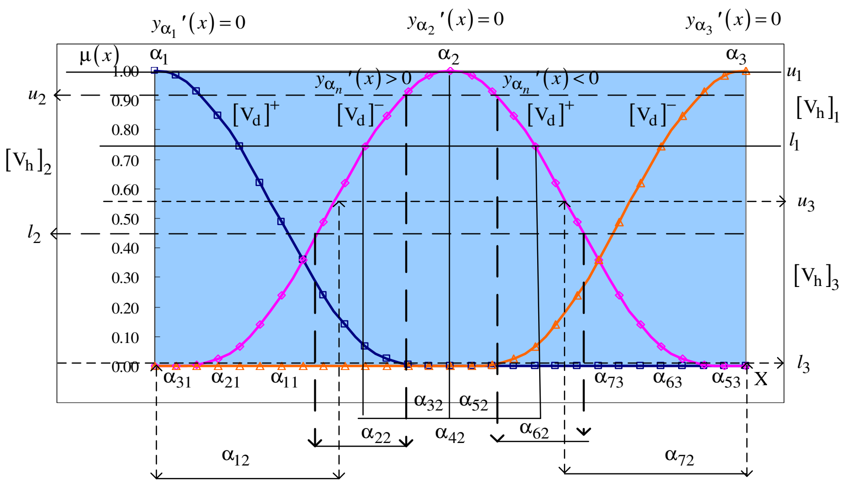

Figure 1: Fuzzy Normal Distribution

5. FUZZY NORMAL DISTRIBUTION

Fuzzy Normal Distribution (FND) is characterized by PbMF. The objective of FND is to find a suitable fuzzy number represented for a linguistic term. Figure 1 exhibits the graphical overview of FND on the basis of the PbMF, membership fuzziness, fuzzy interval, and V-partition, which are further defined as follows.

Definition 5.1 ($[\mathrm{V_a}]_j$, extension of definition 3.1)**:** an atomic set $[\mathrm{V_a}]_j$ where $1 \le j \le n$, is the super fuzzy set that contains any linguistic terms (subsets) with semantic meaning and syntactic symbol of the atomic variable itself. For all j, $\underset{1 \le i \le m}{\forall \alpha_{ij}} \in [\mathrm{V_a}]_j = \alpha_j$, i.e. $\alpha_n = \{\alpha_{1n} \quad \cdots \quad \alpha_{mn}\}$.

Definition 5.2 (L, U)**:** the lower boundary (L) and upper boundary (U) are designed by the distribution (ratio) of hedge linguistics, which is characterized by a distance function (which is used to find the crisp boundary, L\` and U\`) and the membership overlap factor (0<λ<1) (which is used for make the crisp boundary fuzzy). The distance function is defined by $dis\left([V_h]_j\right) = \dfrac{M\left([V_h]_j\right)}{\sum M\left([V_h]\right)}$ where $j \in \{1, 2, \cdots, \eta\}$, η is the maximum number of $\mathrm{V_h}$, and M() is the measure function determined by the expert judgment.

Definition 5.3 (membership fuzziness)**:** The membership [0, 1] can be fuzzified by directional hedge linguistic set $\mathrm{V_{dh}}$. The fuzzy interval of the membership is determined by L and U. Adding the membership overlap factor **(MOF),** λ, makes boundary in fuzziness. Therefore, we have

$$
Inv(\mathrm{V_{dh}}) = \begin{bmatrix} L & U \\ \hline (1-\lambda)\sum_{i=2}^{\eta} dis\left([V_h]_i\right) & 1 \\ (1-\lambda)\sum_{i=3}^{\eta} dis\left([V_h]_i\right) & \boldsymbol{min}\left((1+\lambda)\sum_{i=2}^{\eta} dis\left([V_h]_i\right), 1\right) \\ \vdots & \vdots \\ (1-\lambda)\, dis\left([V_h]_\eta\right) & \boldsymbol{min}\left((1+\lambda)\sum_{i=\eta-1}^{\eta} dis\left([V_h]_i\right), 1\right) \\ 0 & \boldsymbol{min}\left((1+\lambda)\, dis\left([V_h]_\eta\right), 1\right) \end{bmatrix} \quad (5)
$$

Example 2:

From example 1, we get $V_h = \{\text{much,quite,little,absolutely}\}$, "absolutely" can be ignored as it makes the atomic variable into a singleton after combination. Therefore, $V_h = \{\text{little,quite,much}\}$. For the measurement of V_h, we assume we have M(little) =1, M(quite)=2, M(much)=3, then

$$\text{dis}(V_h) = [\text{dis(little)} \;\; \text{dis(quite)} \;\; \text{dis(much)}] = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{bmatrix},$$

The crisp boundary is:

$$\begin{bmatrix} [V_h]_1 \\ [V_h]_2 \\ [V_h]_3 \end{bmatrix} \xrightarrow{Inv'(V_h)} \begin{bmatrix} l' & u' \\ \hline 0.8333 & 1 \\ 0.5 & 0.8333 \\ 0 & 0.5 \end{bmatrix}$$

For overlap factor (λ) = 0.1, we have the membership fuzziness interval:

$$\begin{bmatrix} L' & U' \\ \hline 0.8333 & 1 \\ 0.5 & 0.8333 \\ 0 & 0.5 \end{bmatrix} \xrightarrow[U=\boldsymbol{min}(U'(1+\lambda),1)]{L=L'(1-\lambda)} \begin{bmatrix} L & U \\ \hline 0.75 & 1 \\ 0.45 & 0.9167 \\ 0 & 0.55 \end{bmatrix}$$

Definition 5.4 (V_d, Directional Hedge)**:** Iff $y_{\alpha_n}'(x) > 0$, then V_d is negatively directional in semantic meaning. Iff $y_{\alpha_n}'(x) = 0$, V_d is static. Iff $y_{\alpha_n}'(x) < 0$, V_d is positively directional.

Definition 5.5 (L and U in V_{dh})**:** Elements of L and U in the positive direction (V_d^+) of V_h, denoted as V_h^+, are **"self-inverse-reflect"** to ones in negative direction (V_d^-) of V_h, denoted as V_h^-.

Theorem 3 (V-Partition by membership fuzziness)**:** If V_d is negatively directional in semantic meaning, then the corresponding fuzzy number x in the fuzzy boundary (L, U) **is** $x = \gamma_{\alpha_n} - d\sqrt{1-(\mu_{\alpha_n})^{1/g_{\alpha_n}}}$; If V_d is positively directional, then $x = \gamma_{\alpha_n} + d\sqrt{1-(\mu_{\alpha_n})^{1/g_{\alpha_n}}}$. If V_d is static, then $x = \gamma_{\alpha_n}$. The fuzzy set is vertically partitioned (V-Partition) by the fuzzy boundaries, which is illustrated by example 3.

From theorem 3, the fuzzy interval is assigned for each linguistic variable on the basis of membership fuzziness in the atomic fuzzy set. Definition 5.6 is to find the crisp value or fuzzy number to represent each linguistic variable.

Definition 5.6 (cen(α_{ij}))**:** The crisp value, ζ_{ij}, of a linguistic variable α_{ij} is obtained by the center function.

$$\zeta_{ij} = cen\left(\alpha_{ij}\right) = \frac{\boldsymbol{max}\left(\alpha_{ij}\right) + \boldsymbol{min}\left(\alpha_{ij}\right)}{2} \tag{6}$$

Example 3:

Continue to Example 2, we get Inv(little)=[0.75 1]; Inv(quite)=[0.45 0.9167];Inv(Much)=[0 0.55].Assume the continuous universal set is X=[1,15], the fuzzy set "average" is determined by a PMF with γ_{α_n} =8, d=7. Find the V-partitions of PMF.

By applying theorem 3, we have:

negative direction (with the linguistic word "below"):

$$\begin{array}{r} \\ LB-A \\ QB-A \\ MB-A \end{array} \begin{bmatrix} L & U \\ \hline 0.75 & 1 \\ 0.45 & 0.9167 \\ 0 & 0.55 \end{bmatrix}^{-} \xrightarrow[x_u = \gamma_{\alpha_2} - d\sqrt{1-y_{\alpha_n}(u)}]{x_l = \gamma_{\alpha_2} - d\sqrt{1-y_{\alpha_2}(l)}} \begin{bmatrix} X_L & X_U \\ \hline 4.50 & 8.00 \\ 2.81 & 5.98 \\ 1.00 & 3.30 \end{bmatrix}$$

static: As A-A is the static point or singleton, A= 8

positive direction (with the linguistic word "above"):

$$\begin{array}{r} \\ LA-A \\ QA-A \\ MA-A \end{array} \begin{bmatrix} L & U \\ \hline 1 & 0.75 \\ 0.9167 & 0.45 \\ 0.55 & 0 \end{bmatrix}^{+} \xrightarrow[x_u = \gamma_{\alpha_2} + d\sqrt{1-y_{\alpha_2}(u)}]{x_l = \gamma_{\alpha_2} + d\sqrt{1-y_{\alpha_2}(l)}} \begin{bmatrix} X_L & X_U \\ \hline 8.00 & 11.50 \\ 10.02 & 13.19 \\ 12.70 & 15.00 \end{bmatrix}$$

$$\text{Therefore, } V_{dh} = \begin{bmatrix} V_h^{+} \\ V_h^{0} \\ V_h^{-} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 12.70 & 15.00 \\ 10.02 & 13.19 \\ 8.00 & 11.50 \end{bmatrix}^{+} \\ [8.00 \quad 8.00]' \\ \begin{bmatrix} 4.50 & 8.00 \\ 2.81 & 5.98 \\ 1.00 & 3.30 \end{bmatrix}^{-} \end{bmatrix}$$

Find the fuzzy number to represent each linguistic term using cen() method, then we have table 2.

Table 2: Fuzzy numbers for the linguistic terms

		L=min	U=max	cen()
much above average	MA-A	12.70	15.00	13.85
quite above average	QB-A	10.02	13.19	11.61
little above average	LA-A	8.00	11.50	9.75
absolutely average	A-A	8.00	8.00	8.00
little below average	LB-A	4.50	8.00	6.25
quite below average	QB-A	2.81	5.98	4.40
much below average	MB-A	1.00	3.30	2.15

6. DISCUSSION AND CONCLUSION

This paper introduces the new concept of fuzzy mathematical models describing Fuzzy Linguistic Variable Matrix (FLVM) which is mapped into numeric domain by Fuzzy Normal Distribution (FND) characterized by the Parabola-based Membership Functions (PbMF) and V-partition method of membership fuzziness. Fuzzy Normal Distribution applies the fundamental assumption of the fuzzy set theory on the basis of the possibility. Similar to the assumption of Gaussian distribution, this study assumes the fuzzy distribution of atomic linguistic variable is on the basis of Parabola-based Membership Function (PbMF), which is the Parabolic Membership Function (PMF) with g-level tuning.

This model can be the preferred framework for modeling human subjective judgment which can be applied in the domain of qualitative evaluation, especially transformation the solution of the linguistics evaluation problem into the solution of an arithmetic problem.

Limitation of this approach is that the tuning method is not well defined. The future of the study will discuss the method of “fuzzy tuning for FND” using numerical analysis, which means the best practices to find out the suitable FLVM, overlap, and g-level to model the input FLVM.

Another limitation is that the new method does not merge the existing well known fuzzy logic systems. The further study discusses the fusion of the new method and existing fuzzy set theory. The main reasons include the definitions of FLVM and FND are not comparable with existing definitions of fuzzy linguistic variable, atomic variable and linguistic hedge variable.

ACKNOWLEDGEMENT

Research was supported in part through a grant of an Innovation Technology Fund (UIT/74) from the Innovation and Technology Commission HKSAR.

REFERENCE

1. Chen Guanrong, and Trung Tat Pham, “Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems”, page 6, CRC Press, 2001
2. D. Dubois and H. Prade, “Possibility Theory”, Plenum Press, New York, 1988
3. J. C Helton, “Uncertainty and Sensitivity Analysis in the Presence of Stochastic and Subjective Uncertainty.” Journal of Statistical Computation and Simulation 57: 3-76 1997.
4. L.A. Zadeh, “Fuzzy Sets”, Information and Control 8(3): 338-353 1965
5. L.A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning – I”. Inf. Sci. 8(3): 199-249 1975