

The Solution to the Problem of Curse of Dimensionality in Fuzzy Modelling

Author: Shin-Jye Lee¹; Ching-Hsun Tseng²

Affiliation: Institute of Management of Technology, National Chiao Tung University, Taiwan¹;

Department of Computer Science, University of Manchester, United Kingdom²

E-mail: camhero@gmail.com¹; hank131415go61@gmail.com²

DOI: 10.26821/IJSHRE.9.6.2021.9619

ABSTRACT

As for the solid fuzzy modeling, well-understanding interpretability, highly accuracy of the approximation, less effort, and resources saving are extremely significant features to design a high-performance fuzzy inference model in the coming future. Though the current novel learning algorithms can effectively optimize the system performance, it still can't entirely break the problem of curse of dimensionality (COD). Basically, it is easy to cause the problem of curse of dimensionality as processing high-dimensional data by sophisticated fuzzy learning algorithms, so it is necessary to consider the problem of curse of dimensionality as designing solid fuzzy learning algorithms.

Keywords: Curse of Dimensionality, Fuzzy Modelling

1. INTRODUCTION

As for the solid fuzzy modeling, well-understanding interpretability, highly accuracy of the approximation, less effort, and resources saving are extremely significant features to design a high-performance fuzzy inference model in the coming future. Though the current novel learning algorithms can effectively optimize the system performance, it still can't entirely break the problem of curse of dimensionality.

Basically, it is easy to cause the problem of curse of dimensionality as processing high-dimensional data by sophisticated fuzzy learning algorithms, so it is necessary to consider the problem of curse of dimensionality as designing solid fuzzy learning algorithms. Curse of Dimensionality is a term coined by Richard Bellman [1] applied to the problem caused by the rapid increase in volume associated with adding extra dimensions to a (mathematical) space, and it is a significant obstacle in machine learning problems that involve learning from few data samples in a high-dimensional feature space. Simply, the curse of dimensionality means the size of a data-set grows exponentially with its dimension N. For example, there are ten inputs in a fuzzy system, and each input with two membership functions. Therefore, the partitioning will lead to rules, and it's indeed large to any learning algorithm. Also, the partition result from curse of dimensionality can be calculated as follows:

$$\prod_{i=1}^n X_i \quad (1)$$

where n is the number of input fuzzy sets of the input variable X, is the input variable in the fuzzy system.

2. FUZZY MODELLING IN HIGH-DIMENSIONAL FEATURE SPACE

In a triangle membership function, the total number

of fuzzy rules can be defined as follows:

$$\Delta = f(X, \dots, Z) = \sum_{i=1}^l \dots \sum_{k=1}^n A_i(x) \dots A_k(z) \Delta_{i\dots k} \quad (2)$$

where Δ means the fuzzy rule, X, \dots, Z present the input variables of membership functions from X to Z , $\Delta_{i\dots k}$ presents each tuple or data set with respect to the relative database.

In a two-input-one output fuzzy system, the total number of fuzzy rules can be defined as follows:

$$\Delta = f(X, Y) = \sum_{i=1}^l \sum_{j=1}^n A_i(x) A_j(y) \Delta_{ij} \quad (3)$$

Based on (3), the partition of the two-input-one-output fuzzy system can be illustrated as Fig.

1.

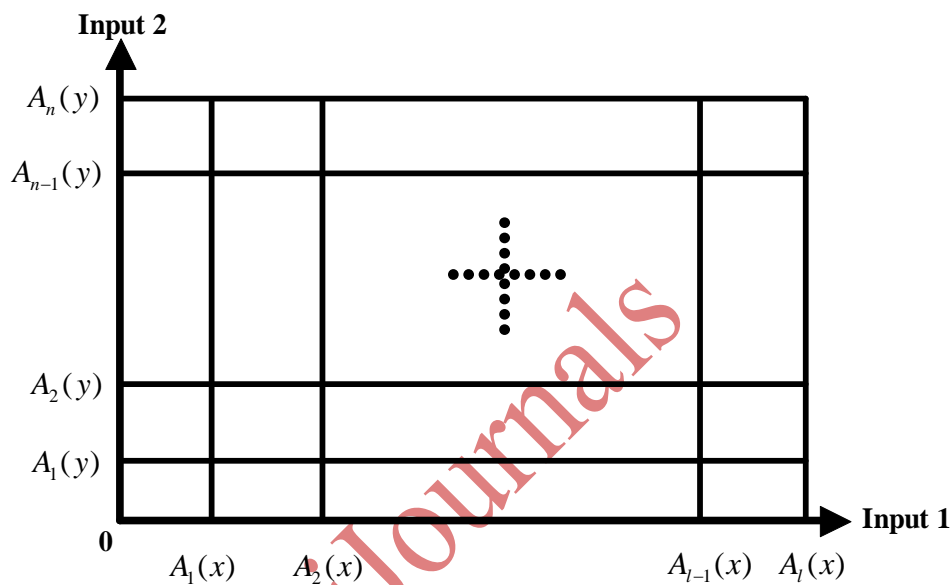


Fig. 7.1. The partition of a two-input-one-output fuzzy system

3. THE SOLUTION TO THE PROBLEM OF CURSE DIMENSIONALITY IN FUZZY MODELLING

The problem of curse of dimensionality costs much time to any learning algorithm and spend many resources in the entire system when converges. In order to break the curse of dimensionality, how to refine and simply partitioning or clustering is a significant issue on solving the problem of curse of dimensionality. It's also a good way to break the curse of dimensionality by reducing the regular partition into the irregular partition. Meanwhile, the irregular

partition of a two-input-one-output fuzzy system can be defined as follows:

$$\Delta = f(X, Y) = \sum_{k=1}^n \Psi_k(\theta) \Delta_{ij} \quad (4)$$

$$\theta = \hat{A}_i(x) \hat{A}_j(y) \quad (5)$$

where Δ means the fuzzy rule, X and Y present the input variables of fuzzy system, Ψ_k presents the irregular partition of the fuzzy system, θ

presents the integration of $A_i(x)$ and $A_j(y)$ by a certain method, Δ_{ij} presents each tuple or data set with respect to the relative database.

Generally, based on (4), the regular partition of membership function could be illustrated as Fig. 2.

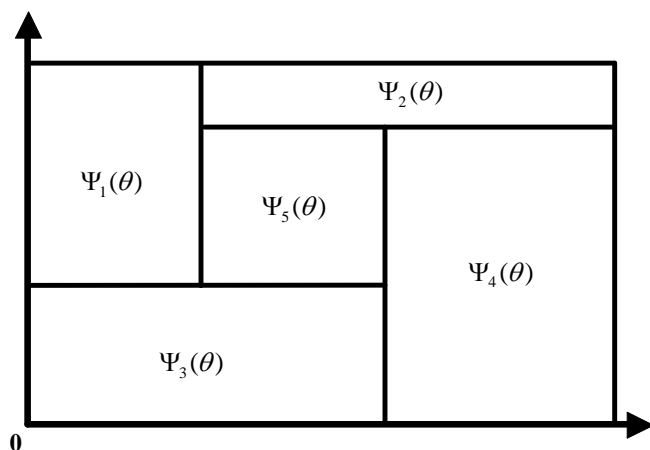


Fig. 2. The irregular partition of a two-input-one-output fuzzy system

According to the comparison of Fig. 1 and Fig. 2, Fig. 2 breaks the problem of curse of dimensionality, because the partition of Fig. 2 is less than that of Fig. 1.

Based on (2), the regular membership function can be calculated as follows:

$$\mu_{A_i}(x) = \left\{ \begin{array}{l} \frac{x - a_i}{b_i - a_i}, a_i \leq x \leq b_i, \\ \frac{c_i - x}{c_i - b_i}, b_i \leq x \leq c_i, \\ 0, otherwise. \end{array} \right\} \quad (6)$$

Based on (4), the irregular membership function can

be calculated as follows:

$$\mu_{A_i}(x) = \left\{ \begin{array}{l} \frac{x - \hat{x}_i}{\sigma}, x \in \sigma \\ 0, otherwise \end{array} \right\} \quad (7)$$

where \hat{x}_i presents a specific value defined by certain clustering of the irregular membership function, σ presents the range of the irregular partitioning.

4. CONCLUSION

How to find the best solution or the optimal clustering to break the curse of dimensionality is a very significant research in the coming future. For the time being, it could be put focus on doing research about optimizing irregular partition of fuzzy clustering as the short-term goal of the future work. Though the irregular partition of membership function can possibly lift the problem of curse of dimensionality, meanwhile but it decreases the unique feature of the fuzzy system – Interpretability. Therefore, in addition to optimize regular partition of fuzzy clustering, how to keep performing the interpretability of the fuzzy system at the same time will be the mid-term goal of the future work. So far, there are so many theories and methods working on Computer Science. Therefore, to discover useful theories and methods and try to integrate which to make the best compromise between these conditions is really a significant topic, and which will be set as a long-term goal of the future work.

5. REFERENCE

[1] Bellman, R., E., (1961), Adaptive Control Processes, Princeton University Press, Princeton, NJ.