Johanyák, Zs. Cs. and Berecz, A.: Survey on fuzzy model identification methods resulting in sparse rule bases, A GAMF Közleményei, Kecskemét, XXII. (2008), ISSN 0230-6182, pp. 109-114.
http://johanyak.hu

Survey on fuzzy model identification methods resulting in sparse rule bases

***Johanyák, Zsolt Csaba* [1] – *Berecz, Antónia* [2]**

Sparse rule bases are the most frequently used solutions for the problems caused by the increased complexity of the traditional covering fuzzy systems.

In this paper, we give a short survey on three automatic fuzzy model identification methods that always result in a fuzzy system with low number of rules.

1. Automatic generation of a fuzzy system

Dense rule bases are characterized by a high number of rules that grow exponentially with the number antecedent dimensions and the number of linguistic terms. In order to solve the complexity problem sparse (not covering) rule bases and reasoning methods based on rule interpolation can be applied [1].

One can produce a sparse fuzzy rule base basically in two ways. The first (e.g. [9]) starts from a completely covering rule base and reduces the number of the rules dropping out the non relevant rules or merging the similar rules.

The methods following the second way produce a rule base directly that does not cover fully the antecedent space. Usually they follow one of the following three approaches:

1. Try to identify the so-called optimal fuzzy rules [2].
2. Extend the rule base by applying the concept of Rule Base Extension [3].
3. Create the starting rules based on fuzzy clustering [4,5,6,7].

The rest of this paper is organized as follows. Section 2 presents the concept of Rule Base Extension (RBE) and the two techniques (RBE-DSS and RBE-SI) based on it. Section 3 overviews the fuzzy clustering based model identification and the method ACP.

[1] PhD, Associate professor, Institute of Information Technology, Kecskemét College, GAMF Faculty
[2] Senior lecturer, Dennis Gabor Applied University, Department of Technological and Fundamental Science

2. Model identification based on Rule Base Extension

2.1. The concept of the Rule Base Extension

The concept of Rule Base Extension (RBE) [3] suggests the creation of a fuzzy system in two steps. In the first step one defines the first two rules and initiates the rule base with them. These relations describe the typical minimal and maximal outputs. For this one looks for the two output extremes and then searches the typical data rows for them. If more data rows contain the same minimum/maximum output that one is chosen, which is closer to the border of the antecedent field.

Next one assigns fuzzy sets to the data sing trapezoidal fuzzyfication with predefined core and support width values. The reference points of the resulting linguistic terms coincide with the values of the two data rows. During the definition of the fuzzy sets some constraints [7] also have to be taken into consideration. These can cause the modification of the width values.

In order to avoid the excessive increase of the number of fuzzy sets one merges the similar (the near or equal) linguistic terms based on the following two meta rules in the antecedent dimensions:

1. If the reference points of two fuzzy sets are closer than a given $d_{i,min}$ limit value they should be merged. $d_{i,min}$ depends on the range of the partition.
2. If the average deviation of the parameters of two fuzzy sets are less than a given $dp_{i,min}$ limit value they should be merged. $dp_{i,}$ depends on the range of the partition.

The fuzzy sets are merged using the CNF union operation proposed by Kóczy. It defines the union of the sets with their convex hull. RBE applies the considerations outlined in [7] and permits a part of the two fuzzy sets of the consequent partition „to hang out” from the $\left[y_{\min}, y_{\max}\right]$ interval. The first and second rule is generated based on the starting linguistic terms.

In the second step of RBE in the course of an iterative tuning process the rule base is increased incrementally.

2.2. Parameter identification

The RBE concept extends the rule base in course of an iterative process. Each time one creates a new rule, it will describe that point of the consequent space where the deviation is the biggest between the sample data and the output calculated by the fuzzy system. The RBE-DSS and RBE-SI methods use different techniques for the determination of the shape of the new sets.

RBE-DSS (Rule Base Extension based on Default Set Shapes) uses the default set shapes that are typical for each partition. These default values are identical to those used for the generation of the first two rules.

In order to avoid the growing number of sets the above presented meta rules are also applied with a refinement. Here one examines whether the newly produced linguistic terms and the formerly created ones can be merged. The sets are merged using the CNF union. Finally one applies the constraints required by the parameterization of the fuzzy sets [7].

The benefits of RBE-DSS are its simplicity and its speed. The drawback of the method is that the insertion of the new rule into the rule base leads to a temporarily increase of the performance index (usually the relative value of the square root of the mean square error).

RBE-SI (Rule Base Extension based on Set Interpolation) tries to avoid the temporarily increase of the performance index after the generation by applying set interpolation for the creation of the new linguistic terms of the interpolated rule. To facilitate further calculation it is practical to suit the applied set interpolation technique to the rule interpolation method, i.e. one should use FRIPOC with FEAT-p, LESFRI with FEAT-LS, VEIN with VESI, etc. [7].

Similarly to the previous method the two meta rules presented above also here are applied in order to avoid unduly growth of the number of sets per partitions and the fuzzy sets considered as similar ones are merged using the CNF union. Finally one applies the constraints required by the parameterization of the fuzzy sets [7].

The advantage of the RBE-SI is that the introduction of the new rule worsens the performance index probably in a small extent at most. Its disadvantage is that the necessary calculation amount and the time need could grow.

3. Automatic System Generation Based on Fuzzy Clustering

There are two approaches in clustering based fuzzy model identification generation. The first one does a multi dimensional clustering on the whole data-set. Then it creates a rule from every cluster so that it creates the projections of the multi dimensional clusters in each antecedent and consequent dimension (e.g. Klawonn and Kruse [5]). This solution is straightforward but its disadvantage is that it is very difficult to estimate the linguistic terms from cluster projections (in some cases one can interpret them with difficulty as convex fuzzy sets).

The second approach starts from the clustering of the output of the sample data (e.g. Sugeno and Yakusawa [6], Chong, Gedeon, and Kóczy [4], Johanyák and Kovács [7,8]). Further on the ACP (Automatic fuzzy system generation based on fuzzy Clustering and Projection) [7] method is going to be presented shortly.

ACP starts with a fuzzy c-means clustering (FCM) of the output data and approximates the clusters with trapezoidal shaped fuzzy sets creating a Ruspini partition. Next for each output linguistic term it selects the belonging data rows from the sample data. For each subgroup of sample data in each output dimension an FCM clustering is done. The antecedent cluster centers will be the reference points of the antecedent fuzzy sets. For each output linguistic term several rules are generated using the corresponding antecedent fuzzy sets.

After iterating through all output linguistic terms in each input dimension the cluster centers obtained in course of different clustering are merged and Ruspini partitions are created similar to the consequent side.

In the second step of ACP a tuning algorithm is used for the identification of the parameters of the antecedent and consequent linguistic terms. It is similar to the one applied by RBE based methods. The main difference is that no new rules are created when a local optimum of the performance index is achieved. Here the algorithm is “thrown out” from the local minimum randomly generated parameter modification steps.

4. Summary

Fuzzy model identification aims the determination of the parameters, mainly rules, of a fuzzy system. Traditional fuzzy systems applying dense rule bases in most of the cases suffer on a high level of complexity owing to the increased number of rules. This problem can be solved by storing only the relevant rules (sparse rule base) and applying an approximate reasoning technique.

This paper surveyed three automatic fuzzy model identification methods. All of them generate fuzzy systems with a low complexity. They apply similar hill-climbing type tuning algorithms that differ only in the solutions they apply in the handling of the problem of local minima.

The advantage of the clustering based approach is its less time need for rule base generation because the parameter identification, which is most time consuming part of the process, starts after the generation of a relatively rich rule base. However, its disadvantage is a usually higher number of rules necessary for the achievement of the same performance index. Practical experiments showed that the number of rules can be reduced with even 40% by applying the RBE approach.

REFERENCES

[1] Kóczy, L. T. and Hirota, K.: Approximate reasoning by linear rule interpolation and general approximation, in International Journal of Approximative Reasoning, Vol. 9, 1993, pp. 197–225.

[2] Kosko, B.: Optimal fuzzy rules cover extrema, International Journal of Intelligent Systems, Vol. 10, 1995, pp. 249 - 255.

[3] Johanyák, Zs. Cs. and Kovács, Sz.: Sparse Fuzzy System Generation by Rule Base Extension, in Proceedings of the 11th IEEE International Conference of Intelligent Engineering Systems (IEEE INES 2007), Budapest, Hungary, 2007, pp. 99-104.

[4] Chong, A., Gedeon, T. D. and Kóczy, L. T.: Projection Based Method for Sparse Fuzzy System Generation, in Proceedings of the 2nd WSEAS International Conference on Scientific Computation and Soft Computing, Crete, Greece, 2002, pp. 321-325.

[5] Klawonn, F. and Kruse, R.: Constructing a fuzzy controller from data, Fuzzy Sets and Systems, Vol. 85, Issue 2, 1997, pp. 177-193.

[6] Sugeno, M. and Yasukawa, T.: A fuzzy-logic-based approach to qualitative modeling, in IEEE Transactions on Fuzzy Systems, Vol. 1, 1993, pp. 7-31.

[7] Johanyák Zsolt Csaba: Fuzzy szabály-interpolációs módszerek és mintaadatok alapján történő automatikus rendszergenerálás, Ph.D. értekezés, Miskolci Egyetem, Hatvany József Informatikai Tudományok Doktori Iskola, 2007.

[8] Johanyák, Zs. Cs. and Kovács, Sz.: The effect of different fuzzy partition parameterization strategies in gradient descent parameter identification, in Proceedings of the 4th International Symposium on Applied Computational Intelligence and Informatics (IEEE SACI 2007), Timisoara, Romania, 2007, pp. 141-146.

[9] Kóczy, L. T., Botzheim, J. and Gedeon, T. D.: Fuzzy models and interpolation. In M. Nikravesh, J. Kacprzyk, and L. A. Zadeh, editors, Forging New Frontiers: Fuzzy Pioneers I & II, volume 217 of Studies in Fuzziness and Soft Computing. Springer, 2007.

Ritka szabálybázist eredményező automatikus fuzzy modell-identifikációs módszerek áttekintése

Dr. Johanyák Zsolt Csaba, Berecz Antónia

Összefoglaló

A ritka szabálybázisok és fuzzy szabály-interpoláción alapuló következtetési módszerek alkalmazása tekinthető az egyik legelőnyösebb megoldásnak a hagyományos fedő szabálybázissal rendelkező fuzzy rendszereknél tapasztalható komplexitás növekedés elkerülésére.

Cikkünkben rövid áttekintést nyújtunk három ritka szabálybázist eredményező automatikus fuzzy modell-identifikációs módszerről. Az első két eljárás (RBE-DSS és RBE-SI) a szabálybázis kiterjesztés elvét alkalmazva a hangolási folyamat során fokozatosan bővíti szabályai körét. A harmadik módszer először fuzzy klaszterezéssel létrehoz egy kezdő szabálybázist, majd egy hangolási folyamat során törekszik a kvázi optimális paraméterek azonosítására.

Überblick über drei Fuzzy Modellidentifikationsmethoden

Dr. Johanyák, Zsolt Csaba, Berecz, Antónia

Zusammenfassung

Dünn besetzten Regelbasen und Regelinterpolation basierten Fuzzy Inferenzmethoden sind die häufigsten benutzten Lösungen für die Probleme, die durch die erhöhte Komplexität der traditionellen voll besetzten Fuzzy Systemen verursacht werden.

In diesem Beitrag geben wir einen kurzen Überblick über drei automatischen Fuzzy Modellidentifikationsmethoden, die immer Fuzzy Systemen mit geringer Anzahl der Regeln resultieren. Die erste zwei Methode (RBE-DSS und RBE-SI) ergänzen stufenweise die Wissensbasis des Systems befolgend das Regelbasis-Ausweitung Konzept. Die letzte Methode produziert eine rohe Regelbasis durch Fuzzy Clustering und danach ersucht die Identifizierung der quasi-optimalen Einflussgrößen durch einen iterativen Abstimmungsalgorithmus.