Transparency Enhancement of First Order Takagi-Sugeno Systems: Promoting the Competition Between the Rules by Controlling the Overlap of Input Fuzzy Sets.

Andri Riid, Raul Isotamm and Ennu Rüstern

Department of Computer Control, Tallinn Technical University, Ehitajate tee 5, 19086, Tallinn, Estonia,
andri@dcc.ttu.ee

ABSTRACT: This paper considers transparency enhancement of 1st order Takagi-Sugeno (TS) systems. The proposed approach is based on the notion that transparency error is directly related to the overlap of antecedent fuzzy sets of the system. A new general gradient-based adaptation algorithm that maintains predetermined overlap degree on the input partition is derived and its approximation/transparency properties with various configurations are evaluated.

KEYWORDS: Fuzzy modeling, accuracy, transparency, interpretability.

1 Introduction

The fuzzy inference system proposed by Takagi and Sugeno in [1] (1st order TS system, in short) is a powerful tool for modeling complex nonlinear systems. TS modeling is a multimodel approach in which linear submodels are combined to describe the global behavior of the system. TS rules have high degrees of freedom to improve their performance that consequently makes it possible to express complicated behaviors with a small number of rules.

The excessive freedom implies that 1st order TS systems are characterized by a *lack of identifiability* [2], i.e. nonuniqueness in model structure by what different parameter vectors yield the same input/output behavior. Moreover, large perturbations of consequent parameters may have a very small effect on global approximation.

This property, which makes TS systems so effective in modeling, also implies that local models rarely admit valid interpretation as local linearizations of the modeled nonlinear system. Thus those accurate models are generally non-transparent to interpretation.

Transparency of the model allows to gain insight into the model; interpretation in terms of linearizations is useful in system analysis and local control design, for example in gain-scheduled control [3].

It is difficult to achieve both goals (accuracy and transparency) simultaneously because of the known trade-off between accuracy and transparency in fuzzy logic [4].

The system parameters in fuzzy modeling are usually obtained by using global learning strategies such as gradient descent that minimize quadratic global cost function [5]. Recently it has been shown that projection of product space fuzzy clusters and weighted least squares method can improve transparency of the model [6]. This is because local learning techniques promote competition between the rules whereas with global techniques the learning has cooperative character.

Alternatively, the competition between the rules can be promoted by reducing the overlap degree between adjacent antecedent MFs as observed in [7]. This paper develops the gradient-based method for identification that maintains the predetermined overlap degree on input fuzzy sets, thus providing the means for balancing the accuracy/transparency trade-off.

It must be taken into account, however, that the interpolation mechanism of 1st order TS systems has undesirable properties [8], because of which there is no simple solution to the problem [9].

2 System definition

We consider a multi-input/single-input first-order TS fuzzy system consisting of R rules (1), where A_{ir} denote the linguistic labels of the i^{th} input variable (i = 1…N), associated with the r^{th} rule, having one-to-one correspondence with trapezoid membership functions (MFs) μ_{ir} in the inference function (2); p_{0r}, p_{ir} denote the consequent parameters of the r^{th} rule and x_i denotes the numerical value of the i^{th} input variable.

$$\begin{aligned}&\text{IF}\, x_1 \text{ is } A_{1r} \text{ AND...AND}\, x_i \text{ is } A_{ir} \text{ AND...}\\&\text{AND } x_N \text{ is } A_{Nr} \text{ THEN } y_r = p_{0r} + ... + p_{\mathrm{i}r}x_1 + ...\mathrm{p}_{Nr}x_N,\end{aligned} \tag{1}$$

$$y = \frac{\sum_{r=1}^{R} y_r \prod_{i=1}^{N} \mu_{ir}(x_i)}{\sum_{r=1}^{R} \prod_{i=1}^{N} \mu_{ir}(x_i)} = \frac{\sum_{r=1}^{R} y_r \tau_r}{\sum_{r=1}^{R} \tau_r} \tag{2}$$

$$\mu_i^s(x_i) = \max\left(0, \min\left(\frac{x_i - b_i^s + \frac{1}{1-\alpha}\frac{b_i^s - c_i^{s-1}}{2}}{\frac{1}{1-\alpha}\frac{b_i^s - c_i^{s-1}}{2}}, \frac{c_i^s + \frac{1}{1-\alpha}\frac{b_i^{s+1} - c_i^s}{2} - x_i}{\frac{1}{1-\alpha}\frac{b_i^{s+1} - c_i^s}{2}}, 1\right)\right),\ s = 1\ldots S_i,\ i = 1\ldots N. \qquad (3)$$

The (variable-oriented) definition of antecedent fuzzy sets is given by (3), where α denotes the membership value at which the adjacent MFs μ_i^{s-1}, μ_i^s and μ_i^{s+1} intersect (Fig. 1.)

We generally choose $0 < \alpha \le 0.5$ to satisfy input transparency condition (4, see [10] for details).

$$\forall x_i \in X_i : 0 < \sum_{s=1}^{S_i} \mu_i^s(x_i) < 1\,. \qquad (4)$$

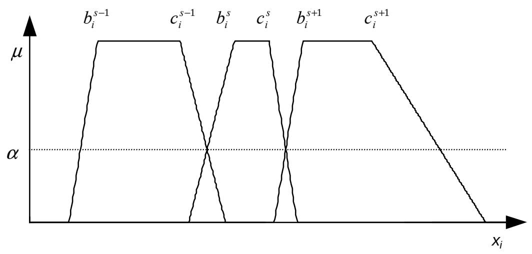

Fig.1. Fixed α-level partition

3 Adaptation method

In this section, the gradient descent based approach, extended on the basis of [11] and [12] for the determination of parameters of (1) is derived. Gradient descent method minimizes the objective function E:

$$E = \frac{1}{2}\sum_{k}^{K} \varepsilon(k) = \frac{1}{2}\sum_{k}^{K} (y(k) - \tilde{y}(k))^2\,, \qquad (5)$$

where $\tilde{y}(k)$ is the k^{th} reference value and $y(k)$ is obtained by applying the inference function (2) with the k^{th} given input pattern. The update rules based on (2) ,(3) and (5) have then the following form.

$$b_i^s(l+1) = b_i^s(l) - \frac{\partial E}{\partial b_i^s(l)} = b_i^s(l) - \frac{\eta}{2}\sum_{k=1}^{K}\frac{\partial \varepsilon(k)}{\partial b_i^s(l)} \qquad (6)$$

where η is the learning rate. To compute the derivatives in (6) and in similar expression for c_i^s we apply the chain rule (7-10).

$$\frac{\partial \varepsilon}{\partial b_i^s} = \frac{\partial \varepsilon}{\partial y}\left(\frac{\partial y}{\partial \mu_i^{s-1}}\frac{\partial \mu_i^{s-1}}{\partial b_i^s} + \frac{\partial y}{\partial \mu_i^s}\frac{\partial \mu_i^s}{\partial b_i^s} + \frac{\partial y}{\partial \mu_i^{s+1}}\frac{\partial \mu_i^{s+1}}{\partial b_i^s}\right) \qquad (7)$$

$$\frac{\partial \varepsilon}{\partial c_i^s} = \frac{\partial \varepsilon}{\partial y}\left(\frac{\partial y}{\partial \mu_i^{s-1}}\frac{\partial \mu_i^{s-1}}{\partial c_i^s} + \frac{\partial y}{\partial \mu_i^s}\frac{\partial \mu_i^s}{\partial c_i^s} + \frac{\partial y}{\partial \mu_i^{s+1}}\frac{\partial \mu_i^{s+1}}{\partial c_i^s}\right) \qquad (8)$$

Note that $r' = 1 \ldots R(\mu_i^s)$ refers to rules having A_i^s in their premise.

To identify the consequent parameters p_{0r}, p_{ir} we use global least squares method [1] as in [5].

4 Experiments

In order to analyze the adaptation/transparency properties of the new algorithm we performed identification experiments with the collections of data using a varying number of rules (R) and varying overlap degree (α).

The first data set is computed by

$$y = 0.6\sin(\pi x) + 0.3\sin(3\pi x) + 0.1\sin(5\pi x) + \rho, \qquad (11)$$

where ρ is Gaussian noise with zero mean and $\sigma^2 = 0.0019$. By using x uniformly distributed in [-1, 1], 201 samples of $y(x)$ were obtained (Fig. 2).

if $b_i^s - \frac{1}{1-\alpha}\frac{b_i^s - c_i^{s-1}}{2} < x_i < b_i^s$ (otherwise zero)

$$\frac{\partial \varepsilon}{\partial b_i^s} = \frac{(y - \tilde{y})}{\sum_{r=1}^{R}\tau_r}\left(\frac{1}{\mu_i^{s-1}}\left(\sum_{r'}^{R(\mu_i^{s-1})}\tau_{r'}(y_{r'} - y)\right)\frac{2(1-\alpha)(x_i - c_i^{s-1})}{\left(b_i^s - c_i^{s-1}\right)^2} + \frac{1}{\mu_i^s}\left(\sum_{r'}^{R(\mu_i^s)}\tau_{r'}(y_{r'} - y)\right)\frac{2(1-\alpha)(c_i^{s-1} - x_i)}{\left(b_i^s - c_i^{s-1}\right)^2}\right). \qquad (9)$$

if $c_i^s < x_i < c_i^s + \frac{1}{1-\alpha}\frac{b_i^{s+1} - c_i^s}{2}$ (otherwise zero)

$$\frac{\partial \varepsilon}{\partial c_i^j} = \frac{(y - \tilde{y})}{\sum_{r=1}^{R}\tau_r}\left(\frac{1}{\mu_i^s}\left(\sum_{r'}^{R(\mu_i^s)}\tau_{r'}(y_{r'} - y)\right)\frac{2(1-\alpha)(b_i^{s+1} - x_i)}{\left(b_i^{s+1} - c_i^s\right)^2} + \frac{1}{\mu_i^{s+1}}\left(\sum_{r'}^{R(\mu_i^{s+1})}\tau_{r'}(y_{r'} - y)\right)\frac{2(1-\alpha)\left(x_i - b_i^{s+1}\right)}{\left(b_i^{s+1} - c_i^s\right)^2}\right). \qquad (10)$$

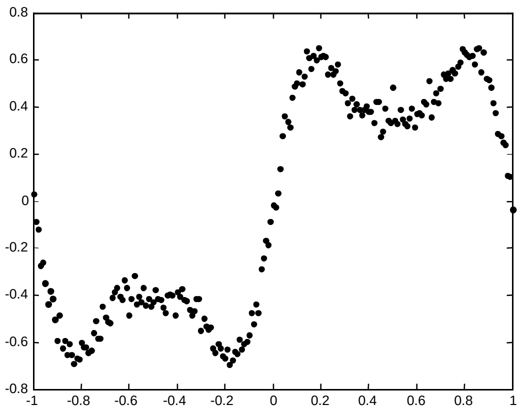

Fig. 2. Identification data (13).

Second test function is a fuzzy two-input/single-output function [10], (Fig. 3), given by noise-free 441 training samples.

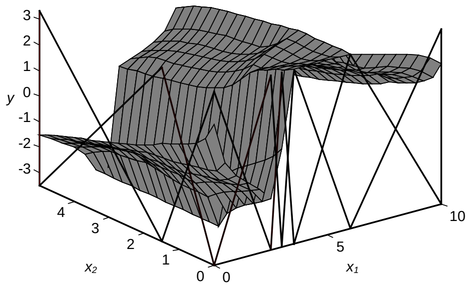

Fig. 3. Fuzzy function

With (11), 55 experiments were performed using R = 3…14 and $\alpha = \{0.1, 0.2, 0.4, 0.5, 0.7\}$. With the second test function, models with 10, 18 and 36 rules were identified, using $\alpha = \{0.2, 0.5\}$. Additional results for comparison were produced with ANFIS [5].

For each experiment approximation root mean square error (RMSE) and transparency error

$$\varepsilon_{tr} = \sqrt{\frac{\sum_{k=1}^{N}\left(\hat{y}(k) - \underset{\max \tau_r(k)}{y_r}(k)\right)^2}{N}} \qquad (12)$$

where $\underset{\max \tau_r(k)}{y_r}$ denotes the local output of the r^{th} rule having the maximum firing strength for the k^{th} input-output pair, were computed. The results are presented in Tables 1-4.

According to these results, we can speak of transparency enhancement only if $\alpha < 0.5$. Moreover, generally rather low values of α are required. Small α makes adaptation less efficient in terms of approximation error and more sensitive to the number of rules (i.e. less robust), on the other hand. Trade-off between accuracy and transparency is therefore very evident.

Another interesting property that is illustrated with obtained results is so-called transparency paradox of 1st order TS systems. Transparency error as defined by (12) implies that it directly depends on how much interpolation (or fuzziness, in other words) there is in the system. It must be stressed that the paradox does not derive from the way we define transparency error, it exists because the given interpolation schemes of 1st order TS systems do not promote transparency [8].

One extreme case is depicted in Fig. 4., where due to large α local models y_r (depicted with gray lines) clearly do not admit valid interpretation of model behavior, although the model output (bold line) follows the reference output (normal line) quite faithfully. Second extreme case is depicted in Fig. 5, where interpolation is practically non-existent, and the model has become a piecewise-linear system. We do not actually need fuzzy logic to implement the latter.

Table 1. Approximation RMSE of (11).

R \ α	0.1	0.2	0.4	0.5	0.7	ANFIS
3	0.1215	0.1180	0.1127	0.1117	0.1290	0.1135
4	0.0840	0.0665	0.0542	0.0525	0.0553	0.0492
5	0.0776	0.0652	0.0455	0.0414	0.0424	0.0385
6	0.0635	0.0599	0.0421	0.0417	0.0419	0.0423
7	0.0499	0.0508	0.0452	0.0402	0.0386	0.0397
8	0.0503	0.0493	0.0397	0.0388	0.0377	0.0373
9	0.0448	0.0450	0.0416	0.0370	0.0386	0.0384
10	0.0434	0.0415	0.0388	0.0362	0.0376	0.0383
11	0.0414	0.0403	0.0380	0.0369	0.0379	0.0359
12	0.0418	0.0404	0.0384	0.0367	0.0366	0.0359
13	0.0402	0.0394	0.0380	0.0363	0.0359	0.0357
14	0.0411	0.0367	0.0382	0.0362	0.0354	0.0341

Table 2. Transparency error of fuzzy models.

R \ α	0.1	0.2	0.4	0.5	0.7	ANFIS
3	0.0241	0.0436	0.0751	0.0810	1,8673	0.0768
4	0.0833	0.1334	0.1844	0.5580	0.2014	0.1951
5	0.0332	0.0628	0.1152	0.1239	0.1240	0.1158
6	0.0232	0.0608	0.1155	0.1336	1.1061	0.1879
7	0.0015	0.0251	0.0919	0.1320	0.1658	0.0880
8	0.0139	0.0277	0.0679	0.0839	0.1492	0.0946
9	0.0006	0.0043	0.0381	0.0761	0.1439	0.1092
10	0.0029	0.0067	0.0622	0.0913	0.1327	0.0628
11	0.0059	0.0117	0.0422	0.0793	0.1244	0.0603
12	0.0066	0.0066	0.0507	0.0851	0.1606	0.0684
13	0.0008	0.0084	0.0417	0.0647	0.1576	0.0652
14	0.0063	0.0126	0.0418	0.0941	0.1160	0.0588

Table 3. Modeling RMSE of fuzzy function.

R \ α	0.2	0.5	ANFIS
5×2	0.502556	0.153159	0.254066
6×3	0.314458	0.138636	0.122078
9×4	0.073957	0.053490	0.040321

Table 4. Transparency error of fuzzy models.

R \ α	0.2	0.5	ANFIS
5×2	0.0562	0.2385	0.4614
6×3	0.0236	0.6104	0.5216
9×4	0.0489	0.2130	0.4151

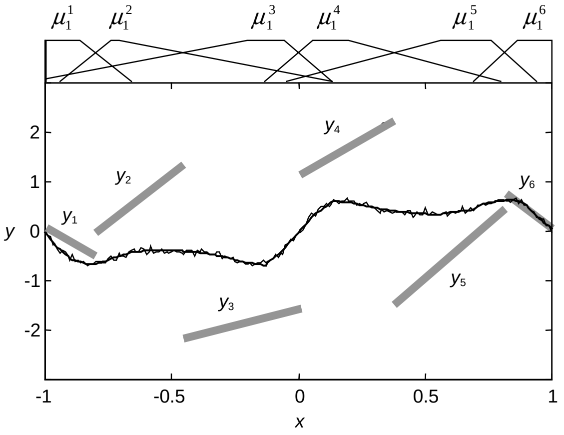

Fig. 4. Approximation of (13) with R = 6, α = 0.7.

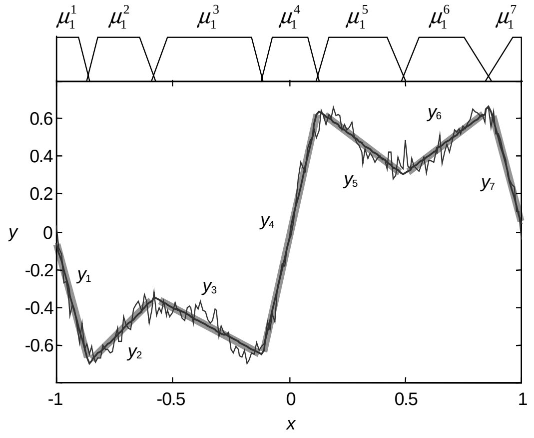

Fig. 5. Approximation of (13) with R = 7, α = 0.1.

5 Conclusions

Because of transparency paradox of 1st order TS systems some kind of compromise between transparency and accuracy must be sought. One possibility to balance those contradictory requirements is to set artificial limits on the overlap between adjacent input fuzzy sets as was done in present paper with the help of derived algorithm.

The results clearly indicate that the value of α has most immediate influence both on transparency and accuracy. With low values of α the model becomes less adaptable and more transparent and vice versa. The derived algorithm thus gives the possibility to specify appropriate α in order to reach the compromise between transparency and accuracy that is expected from the application at hand.

6 References

[1] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control". *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, No. 1, pp. 116-132, 1985

[2] T.A. Johansen, R. Shorten and R. Murray-Smith, "On the Interpretation and Identification of Dynamic Takagi-Sugeno Fuzzy Models". *IEEE Trans. Fuzzy Systems*, vol. 8, No. 3, pp. 297-313, 2000.

[3] J.-S. R. Jang and N. Gulley, "Gain Scheduling Based Fuzzy Controller Design" *Proc. Int. Joint* NAFIPS/IFIS/NASA *Conference*, San Antonio, USA, pp. 101-105, 1994.

[4] J. Casillas, O.Cordon, F. Herrera and L. Magdalena (Eds.) *Trade-off Between Accuracy and Interpretability in Fuzzy Rule-Based Modelling*. Springer-Verlag, 2002.

[5] J.-S. R. Jang. "ANFIS: Adaptive-Network-Based Fuzzy Inference System". *IEEE Trans. Syst., Man, Cybern.*, vol. 23, No. 3, pp. 665–685, 1993.

[6] J. Yen, L. Wang and G.W. Gillespie, "Improving the Interpretability of TSK Fuzzy Models by Combining Global Learning and Local Learning". *IEEE Trans. Fuzzy Systems*, vol. 6, No. 4, pp. 530-537, 1998.

[7] A. Riid and E. Rüstern, "Interpretability versus adaptability in fuzzy systems", *Proc. Estonian Acad. Sci. Eng.*, vol. 6**,** No. 2, pp. 76-95, 2000.

[8] R. Babuska, C. Fantuzzi, U. Kaymak, and H. B. Verbruggen. "Improved inference for Takagi–Sugeno models". *Proc. 5th IEEE Int. Conf. on Fuzzy Systems*, New Orleans, USA, pp. 701–706, 1996.

[9] A. Riid, R. Isotamm and E. Rüstern, "Transparency Analysis of First-Order Takagi-Sugeno Systems". *Proc. 10th Int. Conf. on System Modelling Control*, Zakopane, Poland, vol. 2, pp. 165-170, 2001.

[10] A. Riid, *Transparent Fuzzy Systems: Modeling and Control*. Ph.D thesis, Tallinn Technical University, Tallinn, 2002.

[11] R. Jager, *Fuzzy Logic in Control*, Ph.D. dissertation, Technical University of Delft, Delft, 1995.

[12] J. Abonyi, H. Andersen, L. Nagy and F. Szeifert "Inverse fuzzy-process-model based direct adaptive control". *Mathematics and Computers in Simulation* vol. 51, pp. 119-132, 1999.