Technical Report
SCCH–TR–0009

Numerical Optimization of Fuzzy Systems

Josef Haslinger
Software Competence Center Hagenberg
e-mail *josef.haslinger@scch.at*

Ulrich Bodenhofer
Software Competence Center Hagenberg
e-mail *ulrich.bodenhofer@scch.at*

Abstract — The tuning of the one-dimensional Sugeno controller with respect to different membership functions (B-splines, cubic, trapezoidal) is considered. Taking into account that the linguistic interpretability of the optimization results is of vital importance, we end up with a constrained nonlinear least squares problem. Following the lines of spline approximation/smoothing theory, we give conditions when the observation matrix has full rank (Schoenberg-Whitney condition, regularization) and the reduction of the original optimization problem to a smaller one is possible.

Kompetenzzentren-Programm

1 Introduction

Fundamentally, the idea of fuzzy sets and systems, dated back to Zadeh (1965), is to provide a mathematical model that can present and process vague, imprecise and uncertain knowledge. It has been modeled on human thinking and the ability of humans to perform approximate reasoning, so that precise and yet significant statements can be made on the behaviour of a complex system.

It is out of doubt that fuzzy logic control has become the major application of fuzzy systems. Successful applications of fuzzy logic control include automatic train operation systems, elevator control, temperature control, power plant control, fuzzy refrigerators, washing machines, etc.

There are three essential components of fuzzy logic controllers [2]:

1. The rules, i.e. a verbal description of the relationships usually of a form as the following (n the number of rules):

   if x is A_i then u is B_i $\qquad (i = 1, \ldots, n)$

2. The fuzzy sets (membership functions), i.e. the semantics of the vague expressions used in the rules. More precisely (cf. [1]): Given a universe of discourse X a fuzzy subset A of X is characterized by its membership function

   $$\mu_A : X \to [0, 1] \tag{1}$$

   where for $x \in X$ the number $\mu_A(x)$ is interpreted as the degree of membership of x in the fuzzy set A.

3. An inference machine, i.e. a mathematical methodology for processing a given input through the rule base. The general inference process proceeds in three (or four) steps.

   (a) Under *Fuzzification*, the membership functions defined on the input variables are applied to their actual values, to determine the degree of truth for each rule premise.

   (b) Under *Inference*, the truth value for the premise of each rule is computed, and applied to the conclusion part of each rule. This results in one fuzzy subset to be assigned to each output variable for each rule. Usually only MIN or PRODUCT are used as inference rules as special cases of a triangular norm (t-norm) [1].

   (c) Under *Composition*, all of the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable. Again, usually MAX or SUM are used.

   (d) Finally is the (optional) *defuzzification*, which is used when it is useful to convert the fuzzy output set to a crisp number. Two of the more common defuzzification methods are the CENTROID (center of gravity) and the MAXIMUM method.

In the following, we assume that a reasonable inference scheme (Mamdani or Sugeno controller, see below) is given. There are still two components left which have to be specified in order to design a fuzzy controller - the rules and the fuzzy sets. As the designing of fuzzy controllers should no longer remain an ad hoc trial and error exercise, recent effort has been concentrated on developing new techniques which may be able to design the membership functions and rule base automatically. Genetic Algorithm has played a special role in fuzzy control design as well as methods treating fuzzy systems as artificial neural networks to adjust membership functions using back propagation. For references see [10]. Also optimization algorithms, such as the method of steepest descent have been applied in tuning small and medium sized controllers.

However, all these methods suffer from several drawbacks. Most importantly, it seems that the spirit of fuzzy systems could be lost, resulting in systems that may be not understandable, nor modifiable by human beings [10].

- The resulting membership functions are hardly understandable, impossible connecting them to linguistic terms, which can be easily performed in normal fuzzy sets.
- The resulting rule base has less meaning, partly because the membership functions themselves have less meaning. In this case, injection of human knowledge into the control rule base is almost impossible.
- Smooth control is crucial to many practical applications. The resulting control rule base and control surface often has some abrupt changes, leading to difficulties in implementing such a controller.
- The algorithms often take a long time to find an optimal solution or even fail to converge.

To the best knowledge of the authors, there has been made no attempts to apply highly sophisticated optimization algorithms taking care of the linguistic interpretability of the results to the design of fuzzy controllers. Hence this paper is devoted to the theoretical and experimental study how the parameters describing fuzzy systems can be tuned with advanced numerical algorithms for constrained optimization.

In the following section we specialize to Sugeno controllers (one-dimensional case). In Section 2 the Sugeno controller is introduced and three different classes of membership functions are considered. In Section 3 the tuning of the Sugeno controller is investigated, which is equivalent to solving a (reduced) nonlinear least squares problem. Also the case of incomplete data samples incorporating regulariztion methods is considered here. Finally Section 4 gives an outlook on future work.

2 Sugeno controllers - The single input single output case

Next we look at a fuzzy controller from the point of view of mappings which assign to each crisp observation a crisp value (vector) in the output space, i.e. for any given fuzzy controller there is a function $F : X \to Y$ associating to each input x its corresponding output y. Then, in principle, the design of a fuzzy controller is an interpolation process. Given a set of measured input and output data, representing expert knowledge, the aim of the designing and optimization is to find a configuration of the fuzzy controller such that the corresponding outputs match the sample outputs as well as possible.

We restrict ourselves to the case of a Sugeno controller, where only the input functions are fuzzy sets, but the output are crisp values (fuzzy singletons).

Definition 1. Let X be an input space, let $A_1, A_2, \ldots, A_n$ be normalized fuzzy subsets of X with $\sum \mu_{A_i}(x) > 0$ for all $x \in X$, and $f_1, f_2, \ldots, f_n$ be functions from X to $\mathbb{R}^m$, and consider the rule base $(i = 1, 2, \ldots, n)$

$$\text{if } x \text{ is } A_i \text{ then } u = f_i(x).$$

Then the Sugeno controller defines the following input-output function $F : X \to \mathbb{R}^m$

$$F(x) = \frac{\sum \mu_{A_i}(x)\, f_i(x)}{\sum \mu_{A_i}(x)} \tag{2}$$

In the following we consider the special case, that for $i = 1, 2, \ldots, n$ the functions f_i are constant, that is $f_i(x) = \alpha_i$.

2.1 Sugeno controller with special classes of input membership functions

For the systematic design of optimal Sugeno controllers, we will concentrate on three different kinds of input membership functions: B-spline models, cubic and trapezoidal membership functions. For each input membership class, we proceed under the following conditions:

- the granularity of the input space (i.e. the number of input membership functions) is given a priori.
- Product is used as fuzzy inference rule
- Summation is used as composition scheme.
- Centroid is the defuzzification method.

In a first step, we restrict ourselves to the one-dimensional case (single input-single output controller).

2.1.1 Sugeno controller with B-spline membership functions

B-spline membership functions for Sugeno controllers were proposed in [11]. They showed that the computation of such a fuzzy controller is equivalent to that of a general B-spline hyper-surface under the assumptions mentioned above.

Assume that x is an input variable of a Sugeno controller that is defined on the universe of discourse $[a, b]$. Given a sequence of ordered knots $t = \{t_i\}$ where

$$t_1 = \ldots = t_k = a < t_{k+1} \leq \ldots \leq t_n < b = t_{n+1} = \ldots = t_{n+k}$$

the j-th normalized B-spline basis function $B_{j,k,t}$ of order k for the knot sequence t is recursively defined as

$$B_{j,1,t}(x) := \begin{cases} 1 & \text{if } t_j \leq x < t_{j+1}, \\ 0 & \text{otherwise} \end{cases}$$

$$B_{j,k,t}(x) := \omega_{j,k}(x)\, B_{j,k-1,t}(x) + (1 - \omega_{j+1,k}(x)\, B_{j+1,k-1,t}(x) \qquad \text{for } k > 1$$

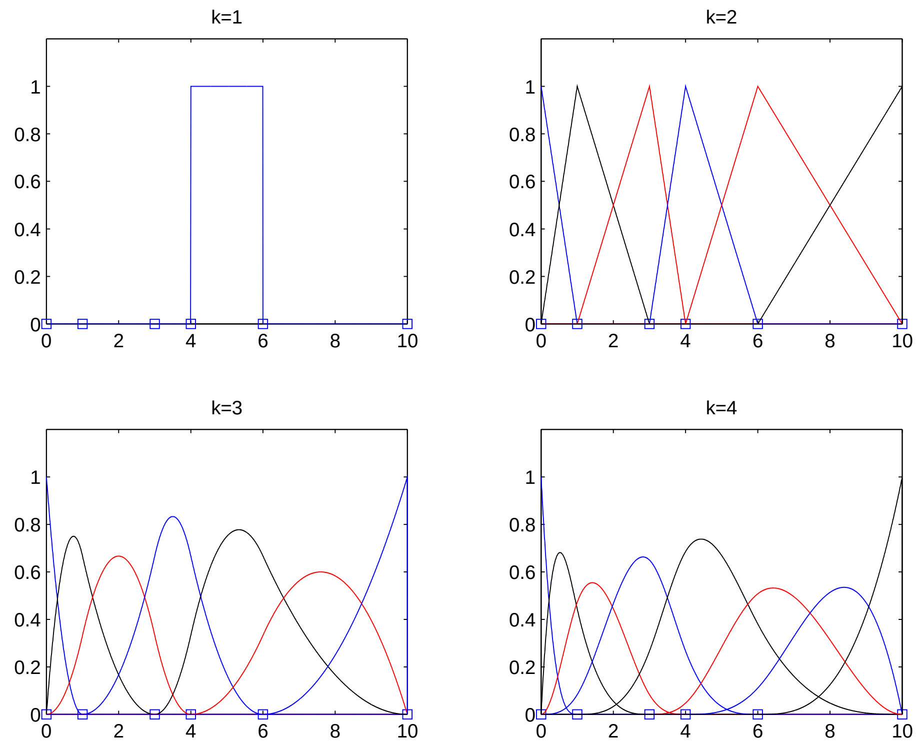

Figure 1: B-spline basis functions of order 1–4 for a non-uniform knot sequence

where

$$\omega_{j,k}(x) := \begin{cases} \frac{x-t_j}{t_{j+k-1}-t_j} & \text{if } t_j < t_{j+k-1}, \\ 0 & \text{otherwise} \end{cases}$$

The complete knots consist of two parts, the interior knots that lie within the universe of discourse, and extended knots that are generated at both ends of the universe for a unified definition of B-splines (marginal linguistic terms [11]).

Some properties of B-splines:

- Positivity: $B_{j,k,t}(x) \geq 0 \quad$ for all $\quad x \in [a,b]$
- Local support: $B_{j,k,t}(x) = 0 \quad$ if $\quad x \notin [t_j, t_{j+k}]$
- C^{k-2} continuity: if the knots $t_k, \dots t_{n+1}$ are pairwise different from each other, then $B_{j,k,t}(x)$ is $(k-2)$ times continuously differentiable.
- Partition of unity:

$$\sum_j B_{j,k,t}(x) = 1 \tag{3}$$

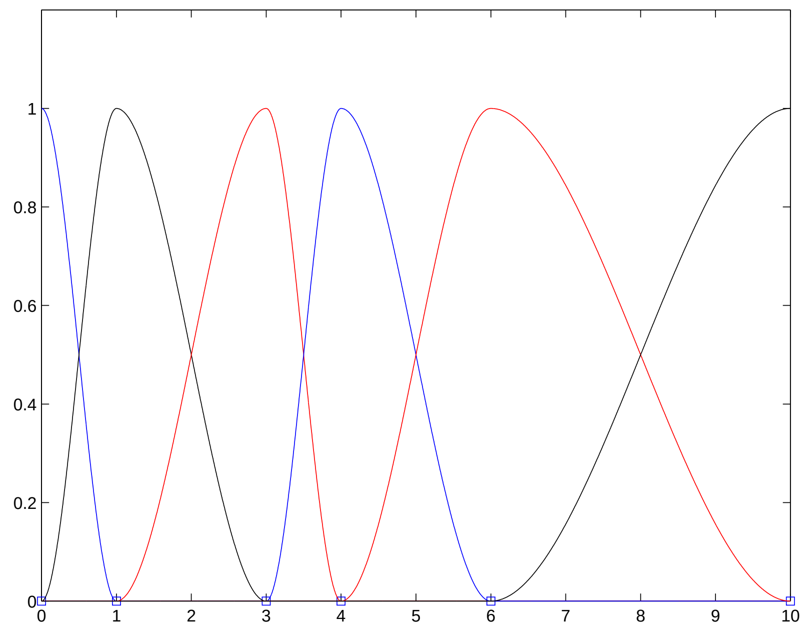

Figure 2: Cubic membership functions

2.1.2 Sugeno controller with cubic membership functions

B-spline basis function suffer from the drawback that the are not necessarily normalized membership function, i.e. the largest membership grade is not necessarily one. Additionally for higher order B-splines, the linguistic interpretation of membership degree is rather complicated. Hence, we consider the following simpler class of membership functions. Let the knot sequence $t = \{t_i\}$ where

$$t_1 = t_2 = a < t_3 < \ldots < t_n < b = t_{n+1} = t_{n+2}$$

be a partition of the universe of an input variable defined over $[a, b]$, corresponding to n linguistic terms. The construction principle is just that the membership function is a polynomial of a given (odd) degree with positive values over two neighboring intervals, zero at the endpoints of the union of the two intervals and one at the inner knot point and additional assumptions on the derivatives for higher order polynomials . We take into account cubic membership functions, although linear, quintic, ... membership functions could be constructed in the same way . Then the membership functions $f_{j,t}(j \in \{1, \ldots, n\})$ (Fig. 2) are defined as

$$f_{j,t}(x) := \begin{cases} -2(\frac{x-t_j}{t_{j+1}-t_j})^3 + 3(\frac{x-t_j}{t_{j+1}-t_j})^2 & \text{if } t_j \leq x < t_{j+1}, \\ 2(\frac{x-t_{j+1}}{t_{j+2}-t_{j+1}})^3 - 3(\frac{x-t_{j+1}}{t_{j+2}-t_{j+1}})^2 + 1 & \text{if } t_{j+1} \leq x < t_{j+2}, \\ 0 & \text{otherwise} \end{cases}$$

Properties similar to that of B-spline membership functions hold, especially that of positivity, local support $[t_j, t_{j+2}]$ and partition of unity.

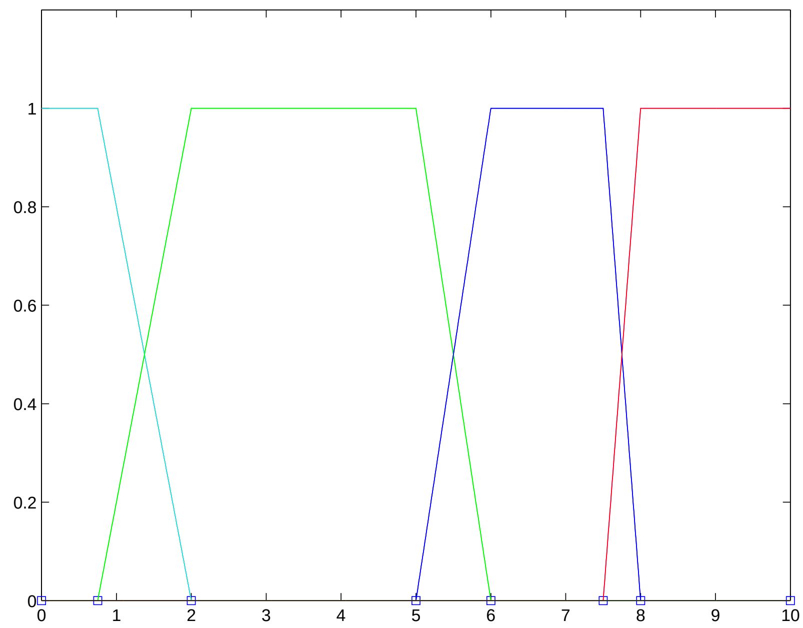

Figure 3: Trapezoidal membership functions

2.1.3 Sugeno controller with trapezoidal membership functions

Next we consider the classical trapezoidal membership functions. Let the knot sequence $t = \{t_i\}$ where

$$a = t_1 < t_2 < ... < t_{2n-1} < t_{2n} = b$$

be a partition of the universe of an input variable defined over $[a, b]$, corresponding to n linguistic terms. Then the mathematical formulation of the trapezoidal membership functions $f_{j,t} \quad (j \in \{2, \ldots, n-1\})$ is as follows:

$$f_{j,t}(x) := \begin{cases} \frac{x-t_{2j-2}}{t_{2j-1}-t_{2j-2}}) & \text{if } x \in (t_{2j-2}, t_{2j-1}] \\ 1 & \text{if } x \in [t_{2j-1}, t_{2j}] \\ \frac{-x+t_{2j}}{t_{2j+1}-t_{2j}}) & \text{if } x \in (t_{2j}, t_{2j+1}] \\ 0 & \text{otherwise} \end{cases}$$

$f_{1,t}$ and $f_{n,t}$ defined analogously. Fig. 3 shows a typical example. Again similar properties to the B-spline case hold.

3 Tuning of Sugeno controllers — The single input single output case

In this section we consider the optimization of Sugeno controllers with membership functions described in the section above. Because of the property of the partition of unity (3), the input-output function of the Sugeno controller (2) takes the following special form:

$$F(x) = \sum_{j=1}^{n} B_{j,k,t}(x)\alpha_j \tag{4}$$

where $\alpha_j (j = \{1, \dots, n\})$ are the consequences of the rules or in terms of spline approximation, the spline coefficients. Completely analogue formulas hold for cubic and trapezoidal membership functions.

As (4) indicates, the optimization of a Sugeno controller with B-spline input membership functions is equivalent to the approximation/interpolation of measured data within the space of B-splines, where both the spline coefficients and the optimal position of knots has to be determined, corresponding to the optimal shape of input membership functions and fuzzy output singletons. In the following, we summarize several results of spline approximation/smoothing with free knots ([8], [7], [3]).

Let $\{x_i, y_i\}, (i = 1, \dots, m)$ be given data with strongly increasing abscissae $x_i \in [a, b]$ and (noisy) measurements $\{y_i\}$ of output values. We want to approximate these data by a function from the n-dimensional spline space consisting of all polynomial splines of order $k \geq 1$. The parameters of the spline F (coefficients and knot positions) have to be chosen in such a way that the natural measure of the goodness of the fit (l_2 norm) provided by F

$$\varphi := \frac{1}{2} \sum_{i=1}^{m} [y_i - F(x_i)]^2 \tag{5}$$

is minimal (nonlinear least squares problem).

Let us begin with some general remarks about free knot spline approximation:

- Simple examples show that the problem of best approximation by splines with free knots has not always a solution in the set of splines with simple knots, whereas it can be proved that there always exists a solution in the set of splines with multiple knots.
- Lethargy theorem [5]: φ is a non-convex function of the knots; poor convergence properties of approximation algorithms at the boundary.
- [6] refers to the potentially high number of local extrema, which may occur in approximation by splines with variable knots.
- When having only discrete data points (in contrast to approximation of functions), missing data points within knot intervals lead to non-uniqueness of spline approximation with fixed knots (cf. Schoenberg-Whitney regularity assumption).

With the observation matrix

$$B(t) := \begin{pmatrix} B_{1,k,t}(x_1) & \cdots & B_{n,k,t}(x_1) \\ \vdots & & \vdots \\ B_{1,k,t}(x_m) & \cdots & B_{n,k,t}(x_m) \end{pmatrix} \tag{6}$$

the minimum least squares functional is given by

$$\varphi = \frac{1}{2}\|y - B(t)\alpha\|^2 \tag{7}$$

where y and α represent column vectors of measured output data and spline coefficients. For fixed knot sequence t, the least squares problem (7) has a unique solution if, and only if, B has full rank n. The following characterization of the full rank property is a consequence of the Schoenberg-Whitney Theorem [3]:

Theorem 1 (Schoenberg-Whitney). *The observation matrix B of format m x n has full rank if, and only if, there exists a strictly increasing subsequence $\{x_{i_j}\}$ with*

$$B_{j,k,t}(x_{i_j}) \neq 0 \tag{8}$$

for all $j \in \{1, \ldots, n\}$. A necessary (and in case of $t_{j+k} > t_j$ also sufficient) condition for (8) to hold is the Schoenberg-Whitney-condition

$$t_j < x_{i_j} < t_{j+k} \tag{9}$$

Condition (9) is often referred to as interlacing property.

In case of a rank-deficient observation matrix B (i.e., r :=rank(B) $< n$), the least squares problem (7) is no longer uniquely solvable. The set of solutions consists of the linear manifold

$$x^\dagger + N(B).$$

$x^\dagger$ denotes the unique least squares solution of minimal (Euclidean) norm, given by $x^\dagger = B^\dagger y$ ($B^\dagger$ the Moore-Penrose inverse or pseudo inverse), $N(B)$ denotes the nullspace of B with dimension $n - r$. Furthermore it is well-known, that the Moore-Penrose inverse is a non-continuous, unbounded operator with regard to noisy data $\{B, y\}$, i.e., an ill-posed problem. In general ill-posed problems are solved by regularization. Famous regularization methods are truncated singular value decomposition and Tikhonov regularization, in which case we consider the problem of approximating the least squares solution, which does not minimizes the natural l_2-norm but a different (semi)norm in order to enforce special features of the regularized approximation [4], i.e.:

$$\|y - B(t)\alpha\|^2 + \gamma_1 \|L\, B(t)\alpha\|^2_{L_2} \tag{10}$$

Important choices include those where L is a differential operator. Especially in classical approximation theory instead of the spline approximation problem, a spline smoothing problem is often considered, where the smoothing term

$$\rho(F) := \frac{1}{2}\int_a^b [F^{(r)}(x)]^2 dx \tag{11}$$

characterizes the smoothness of the spline (r denotes the order of differentiation). The functional to be minimized ($\mu > 0$)

$$\phi := \frac{1}{2}\sum_{i=1}^{m}[y_i - F(x_i)]^2 + \mu\frac{1}{2}\int_a^b [F^{(r)}(x)]^2 dx \tag{12}$$

is called Schoenberg functional.

In the case of $t_j < t_{j+k-r}$ the $r-th$ derivative of a spline F exists and is a spline of order $k-r$ to the same knot sequence. The coefficients of this spline are related linearly to the coefficients α_j, but non-linearly to the knot sequence t_j ([7]). Analogously to the observation matrix, the smoothing term can be represented by a matrix $S_r(t) \in \mathbb{R}^{n-r,n}$, which has full rank $n-r$.

Theorem 2 (full rank property of the regularized system matrix). *If the regularity conditions $m \geq r$ and $\mu > 0$ hold, then the regularized observation matrix*

$$B_\mu(t) := \begin{pmatrix} B(t) \\ \sqrt{\mu}\, S_r(t) \end{pmatrix}$$

has full rank n.

It follows immediately that for fixed knot sequence t and any $\mu > 0$, $m \geq r$ the minimization of the Schoenberg functional has a unique solution independent on whether the Schoenberg-Whitney condition is fulfilled or not. Furthermore the solution $\alpha = \alpha(t)$ is continuously differentiable as a function of t. In the approximation problem (7), the rank of the observation matrix depends on the position of the free knots with respect to the data $\{x_i\}$. Therefore to guarantee unique and continuously differentiable solvability the Schoenberg-Whitney condition (9) has to be satisfied on open neighborhoods of all t encountered in the optimization process. In the case of a rank deficient $B(t)$ the solution can be made unique by choosing the minimum-norm solution $B(t)^\dagger y$, but then the continuity is lost in a neighborhood of t. Moreover, then the reduction of the problem described below is not justified. Both the approximation and the smoothing problems are nonlinear least squares problems where a subset of the variables occurs linearly. For any fixed t, by replacing the variable α by its optimal value $\alpha(t) = B(t)^\dagger y$, the reduced problem

$$\|[I - B(t)B(t)^\dagger]y\|^2 \rightarrow \min_t \tag{13}$$

is obtained.

Since the knot sequence t has to satisfy the order relation $t_{j+1} - t_j \geq 0 \quad (j = 1, \ldots, n-1)$, the free knots occurring in (13) can not be chosen arbitrarily but have to fulfill the constraints

$$t_{j+1} - t_j \geq 0$$

These constraints do not prevent the knots from coalescing and, hence, have to be modified for numerical purposes:

$$t_{j+1} - t_j \geq \varepsilon \tag{14}$$

with a sufficiently small separation parameter $\varepsilon > 0$. The equivalent matrix formulation is given by

$$Ct \geq h \tag{15}$$

with a certain matrix C and a vector h (cf. [7]).

To summarize, we end up with one of the following constrained nonlinear least squares problem:

Full approximation problem:

$$\min\{\varphi(t,\alpha) := \frac{1}{2}\|y - B(t)\alpha\|^2 \,:\, Ct \geq h, t \in \mathbb{R}^{n-k}, \alpha \in \mathbb{R}^n\} \tag{16a}$$

Full smoothing problem:

$$\min\{\phi(t,\alpha) := \frac{1}{2}\left\|\begin{pmatrix} y \\ 0 \end{pmatrix} - \begin{pmatrix} B(t) \\ \sqrt{\mu}\, S_r(t) \end{pmatrix}\alpha\right\|^2 \,:\, Ct \geq h, t \in \mathbb{R}^{n-k}, \alpha \in \mathbb{R}^n\} \tag{16b}$$

Reduced approximation problem:

$$\min\{\varphi_r(t) := \frac{1}{2}\|[I - B(t)B(t)^\dagger]y\|^2 \,:\, Ct \geq h, t \in \mathbb{R}^{n-k}\} \tag{16c}$$

Reduced smoothing problem:

$$\min\{\phi_r(t) := \frac{1}{2}\left\|\left[\begin{pmatrix} I_m \\ I_{n-r} \end{pmatrix} - \begin{pmatrix} B(t) \\ \sqrt{\mu}\, S_r(t) \end{pmatrix}\begin{pmatrix} B(t) \\ \sqrt{\mu}\, S_r(t) \end{pmatrix}^\dagger\right]\begin{pmatrix} y \\ 0 \end{pmatrix}\right\|^2 \,:\, Ct \geq h, t \in \mathbb{R}^{n-k}\} \tag{16d}$$

We have the following correspondence between full and reduced problem ([9]):

Theorem 3. *1. Suppose that B has full rank at a feasible knot sequence t^*, (c.f. Schoenberg-Whitney condition). If t^* is a local minimizer (or a stationary point) of the reduced problem (16c) then (t^*, α^*) with*

$$\alpha^* = B(t^*)^\dagger y$$

is also a local minimizer (or a stationary point) of the full problem (16a).

2. *Suppose that $m \geq r$ and $\mu > 0$ holds. Then the reduced problem (16d) is always solvable, and t^* is a local minimizer (or a stationary point) of the reduced problem (16d) if and only if (t^*, α^*) with*

$$\alpha^* = \begin{pmatrix} B(t^*) \\ \sqrt{\mu} S_r(t^*) \end{pmatrix}^\dagger \begin{pmatrix} y \\ 0 \end{pmatrix}$$

is a local minimizer (or a stationary point) of the full smoothing problem (16b).

The solution of these constrained optimization problems by a Gauss-Newton like method (including the Kaufmann model, exploiting the band structure of the system matrix) is described in [7], [8].

So far we have concentrated on B-spline membership functions. Analogue results hold for cubic and (smoothed) trapezoidal membership functions. Only the Schoenberg-Whitney condition (9) has to be slightly modified:

- cubic membership functions:

$$t_j < x_{i_j} < t_{j+2} \qquad (j = 1, \ldots, n) \tag{17}$$

- trapezoidal membership functions:

$$\begin{aligned} t_1 &< x_{i_1} < t_3 \\ t_{2j-2} &< x_{i_j} < t_{2j+1} \qquad\qquad (j = 2, \ldots, n-1) \\ t_{2n-2} &< x_{i_n} < t_{2n} \end{aligned}$$

4 Future work

Future work will concentrate onto several points:

1. **Tuning of Sugeno controllers - The multiple input single output case:** extension of the input-output function F of a Sugeno controller to the multi-dimensional case is straightforward:

$$F(x_1, x_2, \ldots, x_d) = \\ \alpha_{j_1, j_2, \ldots, j_d} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_d=1}^{n_d} B_{j_1,k_1,t_1}(x_1) \cdot B_{j_2,k_2,t_2}(x_2) \cdot \ldots \cdot B_{j_d,k_d,t_d}(x_d) \tag{18}$$

   F represents a $d-$dimensional tensor product spline. If the data is given on a regular grid (e.g. a rectangular grid in the 2D case), then the $d-$dimensional tuning problem splits up into d one-dimensional problems. For irregular data, it is hard to define a Schoenberg-Whitney like condition; practical examples show that the observation matrix B is very often rank-deficient. Hence regularization is strongly recommended or even a must.

2. **Numerical experiments:** including non-differentiable optimization algorithms.

3. **Mamdani controller:** Analogously to the Sugeno controller, the input-output function can be given explicitly for certain types or Mamdani controllers (triangular or trapezoidal membership functions, certain methods of defuzzification). However this relationship is much more complicated (often a rational function).

References

[1] P. Bauer, E. P. Klement, B. Moser, and A. Leikermoser. Modeling of control functions by fuzzy controllers. In H. T. Nguyen, M. Sugeno, R. M. Tong, and R. R. Yager, editors, *Theoretical Aspects of Fuzzy Control*, chapter 5, pages 91–116. John Wiley & Sons, New York, 1995.

[2] U. Bodenhofer. *Genetic Algorithms: Theory and Applications*. Joh. Kepler University, Linz, 1999. Lecture Notes.

[3] C. de Boor. *A practical guide to splines*. Springer-Verlag, New York, Heidelberg, Berlin, 1978.

[4] H.W.Engl, M.Hanke, and A.Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.

[5] D.L.B. Jupp. The lethargy theorem - a property of approximation by γ-polynomials. *J. Approx. Theory*, 14:204–217, 1975.

[6] D.L.B. Jupp. Approximation to data by splines with free knots. *SIAM J. Numer. Anal*, 15(2):328–343, 1978.

[7] T. Schütze. *Diskrete Quadratmiitelapproximation durch Splines mit freien Knoten*. PhD thesis, Technical University Dresden, Department of Mathematics, 1997.

[8] T. Schütze and H. Schwetlick. Constrained approximation by splines with free knots. *BIT*, 37(1):105–137, 1997.

[9] H. Schwetlick and T. Schütze. Least squares approximation by splines with free knots. *BIT*, 35(3):361–384, 1995.

[10] G. V. Tan and X. Hu. On designing fuzzy controllers using genetic algorithms. In *Proc. FUZZ-IEEE'96*, volume II, pages 905–911, 1996.

[11] J. Zhang and A. Knoll. Constructing fuzzy controllers with B-spline models. In *Proc. FUZZ-IEEE'96*, volume I, pages 416–421, 1996.