H\(_\infty\) Estimation for Optimization of Rational-Powered Membership Functions

Mortaza Aliasghary\(^1\), Nasser Arghavan \(^2\)
Control Engineering Department, Faculty of Electronic Engineering
Istanbul Technical University, Istanbul, Turkey
\(^1\)aliasghary@itu.edu.tr, \(^2\)arghavanin@itu.edu.tr

Abstract—Determination of membership functions of a fuzzy logic system to achieve the best performance is of great importance. Constraining the membership functions to a specific shape which is parameterized by a few variables and then applying a parameter optimization method can be a solution. The \(H\infty\) filter has been applied to optimize only parameters of triangular membership functions. In this study, the \(H\infty\) filter is extended to optimize the input rational-powered membership functions and output singletons. The corresponding equations and derivatives are given and finally the simulation results are shown and compared with Gradient Descent and extended Kalman filtering methods to discuss the efficiency of this approach.

I. INTRODUCTION

During the past few decades, fuzzy logic has been applied to many engineering problems. Many successful applications of fuzzy logic system (FLS) have been reported in literature [1-4]. However, the overall performance of a fuzzy logic system is highly dependent on the parameters of the membership functions, rule base of inference mechanism and input/output scaling factors. These criteria are usually chosen due to theoretical and practical knowledge of an expert or simply by trial and error method. However, in critical applications, having designed the basic FLS, an optimization algorithm is applied to tune the above mentioned factors.

A great deal of work has been done on optimization of a FLS. Some studies have been devoted to proper selection of rule base in inference mechanism [5, 6], while some others propose various methods to choose proper scaling factors [5, 7, 8]. But most researches have focused on shaping the membership functions since fuzzy systems show high sensitivity on the parameters of membership functions. These methods can be broadly divided into two groups: those that explicitly use the derivatives of fuzzy system’s performance index with respect to parameters of membership functions, and those that do not use these derivatives.

Derivative-free approaches show more robustness and tend to converge to global minima and are applicable to wide range of objective functions and membership functions as well. The main drawback that they suffer is slow convergence in comparison with derivative-based approaches. Some of the derivative-free methods are genetic algorithm [9], neural network [10], evolutionary programming [11], geometric methods [12], etc.

The advantage of derivative-based approach is fast convergence however it tends to converge to local minimum. Other drawback can be related as its limitation to specific type of membership functions, inference method and objective functions due to analytical derivative. Gradient descent [13, 14], least squares [15], simplex method [16], Kalman filtering [17, 18] and \(H\infty\) filtering [19] are some examples of this approach.

\(H\infty\) filtering, categorized in derivative-based approaches, was firstly introduced in [19], was applied to triangular MFs. In this paper, we extend this approach to optimize input rational-powered (RP) membership functions, which are extended form of triangular types. Moreover, due to simplicity, output singleton membership functions are chosen in this study.

This paper is organized in six sections. Section II briefly presents the fuzzy system and used membership functions. The optimization methods, \(H\infty\) filtering and Gradient Descent, are discussed in section III. In section IV, the derivatives of fuzzy system with respect to membership functions are derived. Simulation results of the proposed method and comparisons with gradient descent and extended Kalman are given in section V. And finally section VI contains summary and conclusion.

II. FUZZY SYSTEM AND ITS COMPONENTS

In the considered fuzzy system, rational-powered membership functions in the premise part and singletons for the consequent part are used. The rule of fuzzy system can be written as

\[m(\hat{y}) = \sum_{k=1}^{M} \tilde{m}_k(\hat{y})\]  

If \(z_1\) is \(A_1\) and \(z_2\) is \(B_2\) Then output \(m(\hat{y})\) is \(\Gamma_u\)  

where \(z_1\) and \(z_2\) are inputs, \(A_1\) and \(B_2\) fuzzy sets and \(m(\hat{y})\) is the output placed at \(\Gamma_u\). Contribution of this rule in fuzzy output is

\[\tilde{m}_k(\hat{y}) = \omega_k m_k(\hat{y}) = \omega_k \Gamma_u\]  

where \(\omega_k\) is the activation level of the \(k^{th}\) rule:

\[\omega_k = \min \left[f_{A_1}(z_1), f_{B_2}(z_2)\right]\]  

If the system has a total of \(M\) rules, the aggregated fuzzy output is

\[m(\hat{y}) = \sum_{k=1}^{M} \tilde{m}_k(\hat{y})\]
The crisp output due to weighted average defuzzification becomes

\[ \hat{y} = \frac{m(\hat{y})}{\sum_{k=1}^{M} \omega_k m_k(\hat{y})} = \frac{\sum_{k=1}^{M} \omega_k m_k(\hat{y})}{\sum_{k=1}^{M} \omega_k} \]  

(5)

In this study, the rational-powered membership function [20] is used. A typical shape of RP membership functions is shown in Fig. 1. The mathematical expression of the membership functions are expressed as

\[
 f_{ij}(z_j) = \begin{cases} 
 (1 + \frac{z_j - c_{ij}}{b_{ij}})^{a_{ij}} & \text{if } c_{ij} - b_{ij} \leq z_j \leq c_{ij} \\
 (1 - \frac{z_j - c_{ij}}{b_{ij}})^{a_{ij}} & \text{if } c_{ij} \leq z_j \leq c_{ij} + b_{ij}^+ \\
 0 & \text{otherwise}
\end{cases} 
\]  

(6)

where \( c_{ij} \) is the modal point, \( b_{ij}^+ \) and \( a_{ij}^+ \) are the right half-width and its power, \( b_{ij}^- \) and \( a_{ij}^- \) are the left half-width and its power, respectively. Note that \( a_{ij} > 0 \) and \( a_{ij}^+ \) and \( a_{ij}^- \) is equal to 1, RP return to triangular membership functions.

III. OPTIMIZATION METHODS

In order to optimize the parameters of fuzzy system, gradient descent and \( H_{\infty} \) filter are applied to the premise and the consequent parts, respectively. Considering the output of the fuzzy system (\( \hat{y}_n \)) and knowing its desired value (\( y_n \)), the cost function is defined as

\[ E = \frac{1}{2N} \sum_{n=1}^{N} E_n^2 \]  

(7)

\[ E_n = \hat{y}_n - y_n \]  

(8)

where \( N \) is the number of training samples.

A. Gradient Descent-based Optimization

The gradient descent method can be applied to optimize the parameters of RP membership function as follows:

\[ c_{ij}(k + 1) = c_{ij}(k) - \eta \frac{\partial E(k)}{\partial c_{ij}} \]  

(9)

\[ a_{ij}^{+/−}(k + 1) = a_{ij}^{+/−}(k) - \eta \frac{\partial E(k)}{\partial a_{ij}^{+/−}} \]  

(10)

\[ b_{ij}^{+/−}(k + 1) = b_{ij}^{+/−}(k) - \eta \frac{\partial E(k)}{\partial b_{ij}^{+/−}} \]  

(11)

where \( \eta \) is the gradient descent step size.

B. \( H_{\infty} \) filtering-based Optimization

In this subsection, a derivation of \( H_{\infty} \) filter derived in [19, 21] is briefly explained. Derivative of an \( m \)-element vector \( a \) with respect to a \( p \)-element vector \( b \) is as following

\[ \frac{\partial a}{\partial b} = \begin{bmatrix} \frac{\partial a_1}{\partial b_1} & \cdots & \frac{\partial a_1}{\partial b_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_m}{\partial b_1} & \cdots & \frac{\partial a_m}{\partial b_p} \end{bmatrix} \]  

(12)

A nonlinear time-invariant finite dimensional discrete time system is described as

\[ x_{n+1} = f(x_n) + Bw_n + \delta_n \]  

(13)

\[ d_n = h(x_n) + v_n \]

where the vector \( x_n \) is the state of the system at time \( n \), \( w_n \) and \( v_n \) are white noise, \( \delta_n \) is an arbitrary noise sequence, \( d_n \) is the observation vector, and \( f(\cdot) \) and \( h(\cdot) \) are nonlinear vector functions of the state. \( H_{\infty} \) filter estimates \( \hat{x}_{n+1} \) of \( x_{n+1} \) given \( \{d_0, ..., d_n\} \) while \( w_n \) and \( v_n \) are independent unity variance noise processes, but the noise sequence \( \delta_n \) is arbitrary. We define the augmented noise vector and the estimation error as follows:

\[ e_n = [w_n^T \ v_n^T] \]  

(14)

\[ \tilde{x}_n = x_n - \hat{x}_n \]  

(15)

The infinity norm of the transfer function from \( e \) to \( \tilde{x}_n \) must be smaller than a use-defined quantity \( \gamma \) when we apply \( H_{\infty} \) filter to estimate \( \hat{x}_n \).

\[ \|G_{\infty}\|_\infty < \gamma \]  

(16)
Using $H_\infty$ estimator, it is possible to estimate $\hat{x}_n$ as following [19]:

$$F_n = \frac{\partial f(x)}{\partial x} \bigg|_{x=\hat{x}_n} \quad (17)$$

$$H_n = \frac{\partial h(x)}{\partial x} \bigg|_{x=\hat{x}_n} \quad (18)$$

$$Q_0 = E(x_0x_0^T) \quad (19)$$

$$Q_n(I - H^T H P_n) = (I - Q_n/y^2) P_n \quad (20)$$

$$Q_{n+1} = F P_n F^T + BB^T \quad (21)$$

$$K_n = FP_nH^T \quad (22)$$

$$\hat{x}_{n+1} = F\hat{x}_n + K_n(d_n - H\hat{x}_n) \quad (23)$$

where $Q_n$ and $P_n$ are nonsingular sequences of matrices, $K_n$ is the $H_\infty$ gain. If the system is linear, the covariance of the estimation error is bounded by $Q_n$.

$$E[(x_n - \hat{x}_n)(x_n - \hat{x}_n)^T] \leq Q_n \quad (24)$$

In the case of nonlinear systems, because of undefined transfer function, the following inequalities must hold at each time step $n$.

$$I - Q_n/y^2 > 0 \quad (25)$$

$$I + HQ_nH^T > 0$$

Optimization of fuzzy membership functions can be viewed as a weighted least-squares minimization problem. In order to make the problem suitable for $H_\infty$ filtering, membership function parameters constitute the state of a nonlinear system, and the output of the fuzzy system constitutes the output of the nonlinear system to which the $H_\infty$ filter is applied. If the fuzzy system has $\mu_1$ and $\mu_2$ fuzzy sets for the first and second inputs respectively and $k$ singletons as output, then state vector becomes

$$x = [c_{i1} b_{i1}^+ b_{i1}^- a_{i1}^+ a_{i1}^- \cdots c_{i2} b_{i2}^+ b_{i2}^- a_{i2}^+ a_{i2}^- \cdots \Gamma_k]^T \quad (26)$$

which has $5(\mu_1 + \mu_2) + k$ elements. The following shows the nonlinear system model to which the $H_\infty$ filter is applied:

$$x_{n+1} = x_n + R \omega_n + \delta_n \quad (27)$$

$$d_n = h(x_n) + v_n$$

where $h(x_n)$ is the fuzzy system’s nonlinear mapping from the membership function parameters to the single fuzzy system output, and $\omega_n$, $\delta_n$, and $v_n$ are artificial noise processes. The determination of the tuning parameters in $H_\infty$ filtering approach $(B, y)$ is challenging like covariance matrices of Kalman filter [19].

### IV. DERIVATIVE FORMULAS FOR OPTIMIZATION

In this section, the partial derivative of function $E$ in (7) with respect to state parameters and hence matrix $H$ in (18) is calculated as

$$H = \left[ \frac{\partial E}{\partial c_{ij}} \frac{\partial E}{\partial b_{i1}^+} \frac{\partial E}{\partial b_{i1}^-} \frac{\partial E}{\partial a_{i1}^+} \frac{\partial E}{\partial a_{i1}^-} \cdots \frac{\partial E}{\partial \Gamma_j} \right] \quad (28)$$

This matrix is used in the recursion formulas to estimate the next state of the fuzzy system. In the following subsections, the partials of the cost function with respect to the parameters of membership functions are determined by using chain rule [18]. To reduce computational time, we define $r_{ij} = 1$ for active membership functions and if the membership functions are not fired $r_{ij} = 0$.

- **Derivative with respect to input modal points**

$$\frac{\partial E}{\partial c_{ij}} = \frac{1}{N} \sum_{n=1}^{N} E_n \frac{\partial \hat{y}_n}{\partial c_{ij}} \quad (29)$$

$$\frac{\partial \hat{y}_n}{\partial c_{ij}} = \sum_{k=1}^{M} \frac{\partial \hat{y}_n}{\partial \omega_k} \frac{\partial \omega_k}{\partial c_{ij}} \quad (30)$$

$$\frac{\partial \hat{y}_n}{\partial \omega_k} = \frac{m_k(y) \sum_{i=1}^{M} \omega_i - \sum_{i=1}^{M} \omega_i m_i(y)}{(\sum_{i=1}^{M} \omega_i)^2} \quad (31)$$

$$\frac{\partial \omega_k}{\partial c_{ij}} = \frac{\partial f_{ij}(z_j)}{\partial c_{ij}} r_{ijk} \quad (32)$$

$$\frac{\partial f_{ij}(z_j)}{\partial c_{ij}} = \begin{cases} 
\frac{a_{ij}^-}{c_{ij} - b_{ij}^- - z_j} f_{ij}(z_j) & \text{if } c_{ij} - b_{ij}^- < z_j < c_{ij} \\
0 & \text{otherwise}
\end{cases} \quad (33)$$

- **Derivative with respect to upper/lower half-widths**

$$\frac{\partial E}{\partial b_{ij}^\pm} = \frac{1}{N} \sum_{n=1}^{N} E_n \frac{\partial \hat{y}_n}{\partial b_{ij}^\pm} \quad (34)$$
\[
\frac{\partial y_n}{\partial b_{ij}^+} = \sum_{k=1}^{M} \frac{\partial y_n}{\partial \omega_k} \frac{\partial \omega_k}{\partial b_{ij}^+}
\]
\[
\frac{\partial \omega_k}{\partial b_{ij}^+} = \frac{\partial f_{ij}(z_j)}{\partial b_{ij}^+} r_{ijk}
\]

\[
\frac{\partial f_{ij}(z_j)}{\partial b_{ij}^+} = \left\{ \begin{array}{ll}
\frac{a_{ij}^+}{b_{ij}^+} - \frac{a_{ij}^-}{b_{ij}^-} f_{ij}(z_j) & \text{if } c_{ij} - b_{ij} < z_j < c_{ij} \\
0 & \text{otherwise}
\end{array} \right.
\]

\[
\frac{\partial f_{ij}(z_j)}{\partial a_{ij}^-} = \left\{ \begin{array}{ll}
\left( f_{ij}(z_j) \ln f_{ij}(z_j) \right) & \text{if } c_{ij} < z_j < c_{ij} + b_{ij}^+ \\
0 & \text{otherwise}
\end{array} \right.
\]

\[
\frac{\partial f_{ij}(z_j)}{\partial a_{ij}^-} = \left\{ \begin{array}{ll}
\frac{1}{a_{ij}^-} f_{ij}(z_j) \ln f_{ij}(z_j) & \text{if } c_{ij} - b_{ij} < z_j < c_{ij} \\
0 & \text{otherwise}
\end{array} \right.
\]

V. SIMULATION

In order to show the efficiency of the proposed optimization method, we consider the problem of approximating a two-dimensional function given by [18]

\[
f(x, y) = \cos(\pi x) + \cos(\pi y)
\]

where \(0 \leq x, y \leq 2\) with steps of 0.1, i.e. 21 points for each and thus \(N = 21 \times 21 = 441\) samples for the ideal output. A two-input one-output fuzzy system is used to approximate this function. The domain of inputs is divided into five membership functions and five singletons are considered for the output. The rule base is shown in Table 1.

<table>
<thead>
<tr>
<th>(y)</th>
<th>VS</th>
<th>S</th>
<th>M</th>
<th>L</th>
<th>VL</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS</td>
<td>PL</td>
<td>PS</td>
<td>Z</td>
<td>PS</td>
<td>PL</td>
</tr>
<tr>
<td>S</td>
<td>PS</td>
<td>Z</td>
<td>NS</td>
<td>Z</td>
<td>PS</td>
</tr>
<tr>
<td>M</td>
<td>Z</td>
<td>NS</td>
<td>NL</td>
<td>NS</td>
<td>Z</td>
</tr>
<tr>
<td>L</td>
<td>PS</td>
<td>Z</td>
<td>NS</td>
<td>Z</td>
<td>PS</td>
</tr>
<tr>
<td>VL</td>
<td>PL</td>
<td>PS</td>
<td>Z</td>
<td>PS</td>
<td>PL</td>
</tr>
</tbody>
</table>

Fig. 2 is the ideal surface which is plotted by substituting \(x\) and \(y\) samples in (46), while Fig. 3 shows the approximated surface by the non-optimized fuzzy system having initial symmetric input membership functions with \(\alpha^+ = \alpha^- = 1.5\). Since \(x\) and \(y\) are samples in the same interval, membership functions for both of them are chosen identical as well. Fig. 4 depicts the approximated surface by applying \(r\) filtering optimization to all the parameters; i.e. singleton outputs and modal points, upper/lower half-widths and their powers for input membership functions. The tuning parameters are taken as \(\gamma = 300\) and \(B = 10\).
VI. CONCLUSION

In this study, the $H_{\infty}$ filtering approach for membership function optimization was applied to fuzzy system with input rational-powered and output singleton membership functions. The fuzzy logic system with the proposed optimization method was used to model a two-dimensional nonlinear function and then compared with Gradient descent and extended Kalman filtering optimization methods to discuss the efficiency of this approach.

REFERENCES


