Fuzzy membership function optimization for system identification using an extended Kalman filter

Srikiran Kosanam
Department of Electrical Engineering
Cleveland State University
2121 Euclid Avenue, SH 332
Cleveland, OH 44115
s.kosanam@csuohio.edu

Dan Simon
Department of Electrical Engineering
Cleveland State University
2121 Euclid Avenue, SH 332
Cleveland, OH 44115
d.j.simon@csuohio.edu

Abstract: The generation of membership functions for fuzzy systems is a challenging problem. In this paper, we use an extended Kalman filter to optimize the membership functions for system modeling, or system identification. We describe the algorithm and then show the result as sub-optimal novel method of system identification. The ideas described in this paper are illustrated for system identification of a nonlinear dynamic system of a permanent magnet synchronous motor. The other interesting observation made is that the proposed system acts as a noise-reducing filter. We demonstrate that the extended Kalman filter can be an effective tool for identifying the parameters of a fuzzy system model.

I. Introduction

The performance of a fuzzy system depends on both the rule base and its membership functions. In most cases, membership function and rule base determination is based on heuristics, making system identification using fuzzy logic error-prone. Given a rule base, the membership functions can be optimized in order to obtain the best possible performance from the fuzzy system. This paper does not address the generation of rule bases, but does suggest a method for generation of membership functions. Many methods (both derivative-based and derivative free) have been proposed for the optimization of fuzzy membership functions. See [1] for a list of references. This paper extends the results of [1] and demonstrates that the Kalman filter can be effectively used to optimize the membership functions of a Mamdani type fuzzy system model.

This paper demonstrates how the extended Kalman filter can be applied to system identification using fuzzy logic. We demonstrate its performance on fuzzy model identification and show its efficacy on a real life nonlinear dynamic system of a permanent magnet stepper motor. It is shown that the Kalman filter identifies the system model effectively with a slight computational effort. The interesting observation that was made with this system identification is that it is acting as an effective low pass filter inherently, thus minimizing the effect of high frequency instrumentation noise in most real life scenarios.

II. Optimization

In this paper we will assume that the membership functions are triangular. We denote the centroid, lower half-width, and upper half width of the i^{th} membership function of the j^{th} input by c_{ij}, a_{ij} and b_{ij} respectively. The degree of membership of a crisp input x in the i^{th} category of the j^{th} input is therefore given by

$$f_{ij}(x)=\begin{cases}1+(x-c_{ij})/a_{ij} & \text{if } -a_{ij}\le(x-c_{ij})\le 0\\ 1-(x-c_{ij})/b_{ij} & \text{if } 0\le(x-c_{ij})\le b_{ij}\\ 0 & \text{otherwise}\end{cases} \tag{1}$$

In this paper we will limit our discussion to single output systems for ease of notation and illustration. Similarly the output membership functions are characterized by membership function parameters of p_{ik}, q_{ik}, r_{ik} for centroid, lower half-width and upper half-width respectively with k being the category of the i^{th} output. Conceptually the results of this paper can be easily extended to multiple output systems, but the notation becomes quite cumbersome.

The fuzzy output is mapped into a crisp numerical value using centroid defuzzification [2]. Centroid defuzzification can be expressed as

$$\text{crisp output} = \frac{\sum_{j=1}^{n} m(\gamma_j)\gamma_j J_j}{\sum_{j=1}^{n} m(\gamma_j) J_j} \tag{2}$$

where γ_j and J_j are the centroid and area of the j^{th} output fuzzy membership function, and n is the number of fuzzy output sets. For the special case of two fuzzy inputs, the fuzzy output function $m(\gamma)$ is computed as

$$m(\gamma) = \text{fuzzy output function} = \sum_{i,k} m_{ik}(\gamma) \tag{3}$$

where $m_{ik}(\gamma)$ is defined as the consequent fuzzy output function when input 1 is in class i and input 2 is in class k.

$$m_{ik}(\gamma) = w_{ik} m_{oik}(\gamma) \tag{4}$$

$m_{oik}(\gamma)$ is the fuzzy function of the consequent that is activated when input 1 is in class i and input 2 is in class k, and w_{ik} is the activation level of that consequent.

$$w_{ik} = \min[f_{i1}(input1), f_{k2}(input2)] \quad (5)$$

If the fuzzy membership functions are triangular as assumed in this paper, derivative based methods can be used to optimize the centroids and the widths of the input and output membership functions. Consider an error function given by

$$E = \frac{1}{2N}\sum_{q=1}^{N} g_q^2 (y_q - \hat{y}_q)^2 \quad (6)$$

where N is the number of training samples, g_q is a user-defined problem-specific weighting function, y_q is the target value of the fuzzy system and $\hat{y}_q$ is the output of the fuzzy system. We can optimize E by using the partial derivatives of E with respect to the centroids and half-widths of the input and output fuzzy membership functions. We can obtain expressions for these derivatives using (1) and following. Then, using the differentiation chain rule on (6), we can obtain expressions for the derivative of the error function with respect to the half-widths and centroids. We can then use those derivatives in an optimization scheme to minimize the error function with respect to the fuzzy membership function parameters. This idea was first suggested in [3] and was extended in [4, 5]. In this paper we will consider optimization with respect to the input fuzzy membership functions and the output fuzzy membership functions c_{ij} , b_{ij} , a_{ij} , p_{ik}, q_{ik} and r_{ik}.

III. THE EXTENDED KALMAN FILTER

Derivations of the extended Kalman filter are widely available in the literature [6]. In this section we briefly outline the algorithm and show how it can be applied to fuzzy membership function optimization. Consider a discrete time system of the form

$$\begin{aligned} x_{n+1} &= f(x_n) + w_n \\ d_n &= h(x_n) + v_n \end{aligned} \quad (7)$$

where the vector x_n is the state of the system at time n, w_n is the process noise, d_n is the observation vector, v_n is the observation noise, and $f(.)$ and $h(.)$ are nonlinear vector functions of the state. Assume that the sequences v_n and w_n are zero mean and the intial state x_0, v_n and w_n are gaussian and independent from each other with

$$\mathrm{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0 \quad (8)$$

$$\mathrm{E}(w_n w_l^T) = Q\delta_{nl} \quad (9)$$

$$\mathrm{E}(v_n v_l^T) = R\delta_{nl} \quad (10)$$

where $E(.)$ is the expectation operator and δ_{nl} is the Kronecker delta. The problem addressed by the extended Kalman filter is to find an estimate

$$\hat{x}_{n+1} \text{ of } x_{n+1} \text{ given } \mathrm{d}_{\mathrm{j}}(j = 0,..n).$$

If the nonlinearities in (7) are sufficiently smooth, the system can be approximated as

$$\begin{aligned} x_{n+1} &= F_n x_n + w_n + f(\hat{x}_n) - F_n \hat{x}_n \\ d_n &= H_n^T x_n + v_n + h(\hat{x}_n) - H_n^T \hat{x}_n \end{aligned}$$

(11)

where

$$\begin{aligned} F_n &= \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}_n} \\ H_n^T &= \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}_n} \end{aligned} \quad (12)$$

It can then be shown that the desired estimate $\hat{x}_n$ can be obtained by the recursion

$$\begin{aligned} K_n &= P_n H_n (R_n + H_n^T P_n H_n)^{-1} \\ \hat{x}_n &= f(\hat{x}_{n-1}) + K_n[d_n - h(\hat{x}_{n-1})] \\ P_{n+1} &= F_n(P_n - K_n H_n^T P_n)F_n^T + Q_n \end{aligned} \quad (13)$$

In order to reduce the computational effort of the Kalman filter, a pseudo-steady assumption can be made in (13) that

$$H_n^T \approx H_0^T = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}_0} \quad (14)$$

So the calculation of the partial derivative matrix can be performed only once. This assumption is only valid if the partial derivative of the system output $h(.)$ with respect to the state estimate $\hat{x}_n$ does not change much from iteration to iteration. But the extended Kalman filter is only an approximation in the first place because of the linearization discussed above, so if computational effort is any consideration at all, the pseudo-steady-state assumption certainly deserves consideration.

IV. APPLICATION TO FUZZY MODEL IDENTIFICATION

Inspired by the successful use of Kalman filter for training neural networks [7] and for defuzzification strategies [8], we can apply a similar technique for the training of fuzzy systems for model identification. In general we can view the optimization of fuzzy membership functions as weighted least squares minimization problem, where the error vector is the difference between the fuzzy system outputs and the target values for those outputs. Consider a fuzzy system that has L outputs. We use d to denote the target vector for the fuzzy system outputs, and $h(k)$ to denote the actual outputs at the k^{th} iteration of the optimization algorithm.

$$\begin{aligned} d &= [d_1 \ldots d_L]^T \\ h(k) &= [h_1(k) \ldots h_L(k)]^T \end{aligned} \quad (15)$$

In order to cast the membership function optimization in a form suitable for Kalman filtering, we let the membership function parameters constitute the state of a nonlinear system, and we let the output of the fuzzy system constitute the output of the nonlinear system to which the Kalman filter is applied.

In this paper we will consider optimizing membership function parameters of inputs as well as outputs.

We will consider a two input, one output fuzzy system model. The theory developed is not limited to this case but we chose this for notational convenience. Consider a fuzzy system which has μ fuzzy sets for the first input, ν fuzzy sets for second input and κ fuzzy sets for the output. As above we define the centroid and lower and upper half-widths of the i^{th} fuzzy membership function of the j^{th} input *by* c_{ij}, a_{ij}, *and* b_{ij} respectively and we denote the centroid and half-widths of outputs by p_k, q_k and r_k. (Since we have only one output the first subscript is omitted). The state of the fuzzy system can be represented as

$$\begin{aligned} x = [&a_{11}\, b_{11}\, c_{11} \ldots a_{\mu 1}\, b_{\mu 1}\, c_{\mu 1} \\ &a_{12}\, b_{12}\, c_{12} \ldots a_{\nu 2}\, b_{\nu 2}\, c_{\nu 2} \\ &p_1\, q_1\, r_1 \ldots p_\kappa\, q_\kappa\, r_\kappa]^T \end{aligned} \tag{16}$$

The vector x thus consists of all of the fuzzy membership function parameters arranged in a linear array. The nonlinear system model to which the Kalman filter can be applied is

$$\begin{aligned} x_{n+1} &= x_n \\ d_n &= h(x_n) \end{aligned} \tag{17}$$

where $h(x_n)$ is the fuzzy system's nonlinear mapping between the membership function parameters and the single output of the fuzzy system. In order to execute a Kalman filter algorithm, we need to add some artificial process noise and measurement noise to the system model. So (17) can be re-written as

$$\begin{aligned} x_{n+1} &= x_n + w_n \\ d_n &= h(x_n) + v_n \end{aligned} \tag{18}$$

where v_n and w_n are artificially added noise processes. Now we can apply the Kalman recursion. In previous section, *f(.)* is the identity mapping, d_n is the target output of the fuzzy system, and $h(\hat{x}_n)$ is the actual output of the fuzzy system given the current membership function parameters. H_n is the partial derivative of the fuzzy output with respect to the membership function parameters, and F_n is the identity matrix. The Q_n and R_n matrices are tuning parameters, which are covariance matrices of the artificial noise processes.

V. Experimental and Simulation Results

In this section, we describe and illustrate the use of the Kalman filter for identifying the membership function parameters for a fuzzy model. In order to optimally identify the membership function parameters for the fuzzy model defining a permanent magnet synchronous motor, we propose to estimate them using Kalman filter. This is an important and challenging problem, which will be critical in many applications like, closed loop motor control. The main idea here is to use an estimator to estimate the motor winding currents and then use the reference to tune the membership function parameters using a Kalman filter algorithm. The estimator structure that we use to obtain the estimates of the motor current *y* is given by

$$\begin{aligned} \hat{y}_k^- &= \hat{y}_{k-1}^+ + T\dot{\hat{y}}_{k-1} \\ \hat{y}_k^+ &= \hat{y}_k^- + g(z_k, \hat{y}_k^-) \end{aligned} \tag{19}$$

where $\hat{y}_k^-$ denotes the estimate of y at time k before the measurement at time k is processed and $\hat{y}_k^+$ denotes the estimate at time k after the measurement is processed. T is the update period of the estimator and z_k is the noisy measurement of the winding current. The estimate of the rate of change of the winding current is computed using the method of undermined coefficients [2] as

$$\hat{y}_k^{'} = \frac{1}{T}\left[-\frac{1}{3}\hat{y}_{k-3}^+ + \frac{3}{2}\hat{y}_{k-2}^+ - 3\hat{y}_{k-1}^+ + \frac{11}{6}\hat{y}_k^+\right] \tag{20}$$

The fuzzy correction mapping *g(.)* has two arguments:

$$\begin{aligned} (\text{input } 1)_k &= z_k - \hat{y}_k^- \\ (\text{input } 2)_k &= (\text{input } 1)_k - (\text{input } 1)_{k-1} \end{aligned} \tag{21}$$

So the correction mapping depends on the estimation error and the rate at which it has changed. The fuzzy rule base for the mapping *g(.)* is chosen as per [4]. The rule base has seven membership parameters for both the inputs as wells as the output. So the estimator has total of 21 membership functions with 3 parameters associated with each of them. So the estimator has 63 parameters to identify. These 63 parameters arranged in a vector comprise the state of the Kalman filter.

In order to implement the membership function optimization discussed in this paper, we collected motor winding currents with a digital oscilloscope at a rate of one sample every 200 µs. We then used created a training waveform for the data by taking a 51-point moving average. This is to use as a reference as the actual measurement has high frequency instrumentation noise. The actual measurement and the reference measurement are shown in Figure 1 and 2 respectively.

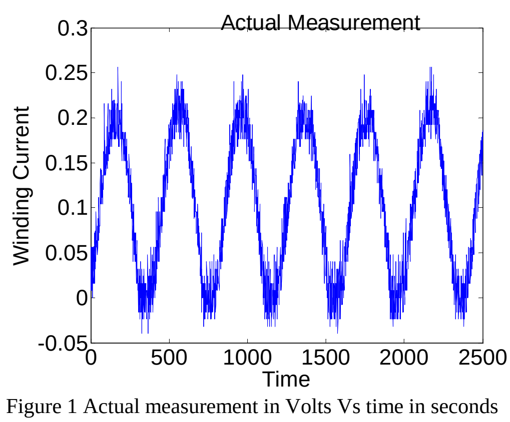

Figure 1 Actual measurement in Volts Vs time in seconds

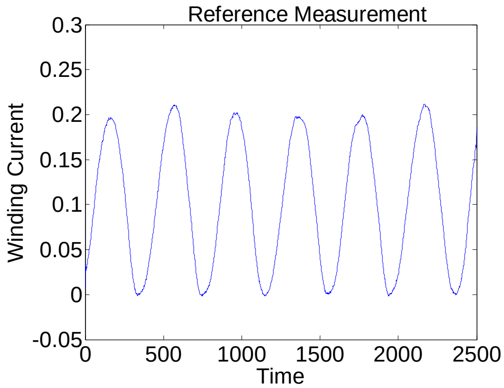

Figure 2 Reference measurement for training Vs time in seconds

During the training the error has reduced considerably about 81%. The change in the error is shown in Figure 3.

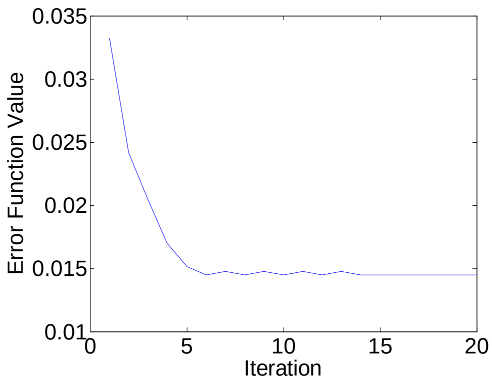

Figure 3 Training progress

The result of the system identification is shown in Figure 4. It can be clearly seen that the measurement obtained after the trained fuzzy membership parameters are used to identify the model, the result matched the reference model better than the raw measurement. The important observation that can be seen is that the high frequency measurement noise that is present in the raw measurement is reduced considerably. This proves that the identifier has the inherent property of disturbance rejection associated with it. The other observation that was made is that the initial membership function parameters used were evenly spaced triangular membership functions but the trained membership functions were not evenly spaced. This justifies the point that the heuristic methods for identifying membership functions for fuzzy system modeling may not be optimal.

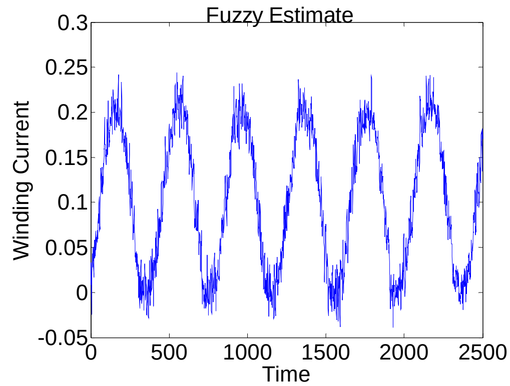

Figure 4 Measurement after training

VI. Conclusions

We have shown that optimization of the fuzzy parameters can be viewed as a system identification problem. An extended Kalman filter with a search algorithm can be an effective solution to the system identification problem. The results shown are preliminary proving that this can be a way of looking at the problem. The results could be further improved with effective tuning of parameters of Kalman filter. Further work could be to include the sum normalization method [9] for effective modeling by the fuzzy systems. We have shown that the theoretical strength of the Kalman filter could lead to be a fruitful application for the system identification problem of fuzzy systems. Now that it has been proved that system identification can be achieved using fuzzy logic and extended Kalman filtering, it would be interesting to see the application of other filtering approaches like robust Kalman filter [10], H-infinity or unscented filter [6], etc. The other area of interest would be able to prove the convergence of the system identification problem using Kalman filtering.

References

[1] D.Simon, "Training Fuzzy systems with the Extended Kalman Filter," IEEE Trans. on Fuzzy systems, vol. 132, pp. 189-199, December 2002.
[2] B.Kosko, Neural Networks and Fuzzy Systems, Prentice Hall: Englewood Cliffs, NJ, 1992.
[3] F. Guely and P. Siarry, "Gradient Descent Method for Optimizing Various Fuzzy Rule Bases," IEEE Int. Conf. on Fuzzy Systems, pp. 1241-1246, March 1993.
[4] D. Simon and H. El-Sherief, "Fuzzy Logic for Digital Phase-Locked Lopp Filter Design," IEEE Trans. on Fuzzy Systems, (3)2, pp. 211-218, May 1995.
[5] D. Simon, "Design and Rule Base Reduction of a Fuzzy Filter for the Estimation of Motor Currents," Int. Journal of Approximate Reasoning, vol. 25, pp. 145-167, October 2000.
[6] D. Simon, Optimal State Estimation, John Wiley & Sons, 2006.
[7] G. Puskorius and L. Feldkamp, "Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks," IEEE Trans. Neural Networks, pp. 279-297, 1994.
[8] T. Jiang, Y. Li, "Generalized defuzzification strategies and their parameter learning procedures," IEEE Trans. Fuzzy Systems, pp. 64-71, 1996.
[9] D. Simon, "Sum Normal Optimization of Fuzzy Membership Functions," Intl Journal of Uncertainity, Fuzziness, and Knowledge-Based Systems, vol. 10 pp. 363-384, August 2002.
[10] S. Kosanam and D. Simon, "Kalman Filtering with Uncertain Noise Covariances," Intelligent Systems and Control, pp. 375-379, August 2004.