Chapter VII

Design of Optimal Fuzzy Observers based on
TS Fuzzy Model
Design of Optimal Fuzzy Observers based on TS Fuzzy Model

7.1 Introduction

For a nonlinear stochastic system, the optimal state estimation from noisy measurement data has been the subject of considerable research interest for many years. Many approximate methods have been developed towards this [136,145,146] and these methods have improved the performance of the Kalman filter, which was derived for linear gaussian systems. Most important to the nonlinear estimation problem is the determination of the probability density function of the state conditioned on the available measurement data. If this a posteriori density functions were known, an estimate of the state for any performance criterion such as conditional mean, maximum a posteriori (MAP) etc., can be determined. For nonlinear stochastic systems, it is very difficult to determine the required a posteriori density. To obtain the a posteriori density for nonlinear systems, many approximate methods like the Gram-Charlier and Edgeworth expansions [147,148], Pearson-type density function by Aoki [149] are prominent. The idea of using weighted sum of Gaussian density functions for the approximation of the a posteriori density function was developed by Alpach and Sorenson [136].

In this chapter, the optimal fuzzy observer for nonlinear stochastic system is developed based on the idea of weighted Gaussian sum representation of the required a posteriori density and mathematical duality between optimal control and optimal observer problems. Ma, Sun and He [125] developed observer design for deterministic fuzzy system and separation principle between fuzzy observer and fuzzy control. Dan Simon [152] recently attempted to apply Kalman filtering for stochastic fuzzy discrete time dynamic systems. He has used time-varying premise variables, which are independent of state. Hence the problem dealt with is linear time varying rather than nonlinear. This has limited value since Kalman filter itself caters directly to linear time varying systems.
7.2 The Gaussian sum representation

The Gaussian sum representation $p_A$ of a density function $p$ associated with a vector-valued random variable, $X \in \mathbb{R}^n$ is defined as

$$p_A(X) \triangleq \sum_{i=1}^{r} \alpha_i N(X - \mu_i, P_i)$$  \hspace{1cm} (7.1)

where $N(X - \mu_i, P_i)$ are Gaussian densities with means $\mu_i$ and covariance $P_i$ for all $i = 1,2,..r$ and $\sum_{i=1}^{r} \alpha_i = 1$, $\alpha_i \geq 0$ for all $i$.

It was shown in [134,150,151] that $p_A$ converges uniformly to any density function of practical concern as the number of terms, $r$ increase and the covariance $P_i$ approach the zero matrices.

In practice, the mean values $\mu_i$ are used to establish a grid in the region of state space that contains the probability mass. The $\alpha_i$ are chosen as the normalized values $p(\mu_i)$ of the density, $p$ that is to be approximated (i.e., the $p(\mu_i)$ are normalized so that $\sum_{i=1}^{r} \alpha_i = 1$). Then, set all of the covariance matrices $P_i = b I_n$ where, $b$ is a small positive scalar. $I_n$ is identity matrix of dimension, $n$. The value of $b$ is determined so that the error in the approximation $(p - p_A)$ is minimized in some prescribed sense.

7.3 Fuzzy Model for discrete time stochastic nonlinear system

Choose the representative points, $\mu_i, i = 1,2,..r$ in the state space to establish a grid sufficiently covering the region of state space considered for the fuzzy modeling in (7.2) below. Each subsystem below corresponds to these local selected points. In the nonlinear system, whether deterministic or stochastic, the points $\mu_i$ must be suitably chosen and in the stochastic case, they can be thought of as mean values, which covers, the probability mass sufficiently in the entire state space under consideration.
Rule : i If $x_1$ is $M_{1i}$, and $x_2$ is $M_{2i}$ ...... and $x_n$ is $M_{ni}$

Then $X(k+1) = A_i(k) X(k) + B_i(k) u(k)$  

$Y_i(k) = C_i(k) X(k) ; i = 1, 2, \ldots, r$ . and $k :$ the discrete time index, $r :$ the number of if-then rules. Where, $x_1, x_2, \ldots, x_n$ are system states; $M_{1i}, M_{2i}, \ldots, M_{ni}$ are the antecedent fuzzy terms in the $i^{th}$ rule; $X(k) = [x_1(k), x_2(k) \ldots x_n (k)]^T \in \mathbb{R}^n$; $u(k) = \mathbb{R}^p$ is the system deterministic input , $Y(k) \in \mathbb{R}^q$ is the system output vector and $A_i(.), B_i(.)$ and $C_i(.)$ are respectively, $\mathbb{R}^{nxn}, \mathbb{R}^{nxp}, \mathbb{R}^{qxn}$ matrices whose elements are known constants and real valued and defined on positive real space. The overall fuzzy stochastic nonlinear system is inferred as

$$X(k+1) = \left\{ \sum_{i=1}^{r} h_i(X) \left[ A_i(k) X(k) + B_i(k) u(k) \right] \right\} + w(k)$$  \hspace{1cm} (7.3)

and the final output is

$$Y(k) = \left\{ \sum_{i=1}^{r} h_i(X) C_i(k) X(k) \right\} + v(k)$$  \hspace{1cm} (7.4)

where

$$h_i(X) = \frac{ \prod_{j=1}^{n} M_{ji}(x_j(k)) }{ \sum_{m=1}^{r} \prod_{j=1}^{n} M_{jm}(x_j(k)) } \geq 0 , \quad \sum_{i=1}^{r} h_i(X) = 1 , \quad i = 1, 2, \ldots, r.$$  \hspace{1cm} 

$w(k), v(k)$ are zero mean white-gaussian process and measurement noise vectors respectively with covariance $Q_k$ and $R_k$. $w(k), v(k)$ are uncorrelated to each other.

The system dynamics and measurement equations above are nonlinear.

For the development of optimal a posteriori density, the initial state, $X_0$ is assumed to be Gaussian distributed random variable with covariance, $P(0)$ where this Gaussian distributed $X_0$ is assumed to be exactly represented as Gaussian sum of densities as follows using the results in section:7.2

$$p(X_0) = \sum_{i=1}^{r} \alpha_i(0) N(X_0 - \mu_i(0), \ \sigma_i(0))$$  \hspace{1cm} (7.5)
7.4 The general observer problem for nonlinear system and optimal observer for fuzzy system.

The general observer problem for nonlinear system is to determine the probability density function of the state conditioned on the available measurement data. If this \textit{a posteriori} density function were known, an estimate of the state for any performance criteria can be determined.

a. The observer problem is the dual of the control problem\cite{153}. For designing the optimal observer for the above fuzzy system, it is essential to focus on the duals of all the \(r\)- subsystems. In the system in (7.3) & (7.4), the overall state is represented as the weighted sum of states of local subsystems. In the dual of this system, there are \(r\)- dual subsystems and each of them deals with respective local state in the state space rather than the global state of the overall system. This is obvious when we follow the standard methods of constructing the dual (adjoint) of a forward system \cite{128,154}. This concept is crucial in the development of optimal observer for the fuzzy system under consideration.

b. Since all the local subsystems are linear, one can design optimal local linear observers (Kalman filters) if the process and measurement noise are Gaussian and satisfy other standard requirements. The optimal local linear observer outputs can be combined appropriately using the Gaussian sum representation \cite{134,150,151} for obtaining the optimal \textit{a posteriori} density of the state conditioned on all available measurements. Thus one can identify one optimal observer for each fuzzy subsystem, which operates in specific region of the state space. A fuzzy subsystem and the corresponding fuzzy observer have one-to-one correspondence (\(i^{th}\) subsystem to \(i^{th}\) observer). At any time step, \(k\) we have \(r\)-such observers to operate in parallel to arrive at the optimum \textit{a posteriori} density of the state.

c. At any time step \(k\), the overall optimal \textit{a posteriori} density can be obtained by the convex blending of each local optimal \textit{a posteriori} Gaussian density corresponding to each local linear subsystem. The optimal convex coefficients
for blending can be obtained using the densities of the residuals of each local optimal observer.

d. All the observers corresponding to each subsystem in the overall fuzzy system have the common goal of achieving the overall optimal \textit{a posteriori} density. The time evolution of the convex weighting coefficients depends on the state at each instant. Thus the estimation process is an evolving dynamic problem i.e., the previous estimate affecting the current estimate and so on. The interaction between local observers appears vanished in the overall optimal observer, but actually they constructively co-observe for achieving the common goal.

7.5 \textbf{Duality between optimal control and optimal estimation}

In [111], the optimal fuzzy control problem was solved based on local linear models. The reasoning was based on dynamic programming formalism and linear optimal control theory. The optimal fuzzy observer problem based on these local linear models can also be analyzed with dynamic programming framework and mathematical duality between optimal control and optimal observer problems highlighted in Sections: 7.5.1 & 7.5.2 respectively.

\textbf{7.5.1 Optimal Control problem} [137,119]

Consider the discrete time stochastic system

\begin{align}
X(k) &= AX(k) + Bu(k) + w(k) \\
Y(k) &= CX(k) + v(k)
\end{align}

(7.6)

where, \( u(k) \in \mathbb{R}^p \) is the deterministic input vector, \( Y(k) \in \mathbb{R}^q \) is the output vector, \( w(k) \in \mathbb{R}^n \) is the zero mean white Gaussian process noise with \( E[w w^T] = Q_k \) and zero mean white Gaussian measurement noise \( v(k) \in \mathbb{R}^q \) with \( E[v v^T] = R_k \) and \( w(k) \) and \( v(k) \) are uncorrelated with each other for all \( k \). \( w(k) \) and \( v(k) \) are also uncorrelated with \( X(0) = X_0 \) which is Gaussian distributed with covariance \( P(0) \). Since the system in (7.6) is linear, \( X(k), Y(k) \) are also Gaussian for all \( k \). Consider the cost function to be minimized as

\begin{align}
J_c(u) &= \frac{1}{2} [X_N^T S_N X_N] + \frac{1}{2} \sum_{k=0}^{N-1} [X^T(k) F X(k) + u^T(k) G u(k)]
\end{align}

(7.7)
where \( F, G \) and \( S_N \) are weighting matrices for state, input and final state respectively.

The corresponding matrix Ricatti equation to be solved for obtaining the optimal control to minimize (7.7) is

\[
S(k) = A^T \left[ S(k+1) - S(k+1)B \left( B^T S(k+1)B + G \right)^{-1} B^T S(k+1) \right] A + F \tag{7.8}
\]

with \( S(N) = S_N \) and the optimal control to minimize (7.7) using the solution of (7.8) is

\[
u(k) = - \left[ R_k + B^T \pi(k+1)B \right]^{-1} B^T \pi(k+1)AX(k) \tag{7.9}\]

where \( \pi \) is the solution of (7.8)

### 7.5.2 Optimal Observer problem [137,123]

Consider the discrete time optimal stochastic observer of the system in (7.6)

\[
\hat{X}(k) = [A - K(k)C] \hat{X}(k-1) + \{I_n - K(k)C\} Bu(k-1) + K(k)Y(k) \tag{7.10}
\]

Define the error state as \( e_k = X(k) - \hat{X}(k) \)

Consider the cost function to be minimized as

\[
J_e = \frac{1}{2} \left[ e_0^T P_0^{-1} e_0 \right] + \frac{1}{2} \sum_{k=0}^{N-1} \left[ w^T(k) Q_k^{-1} w(k) + \{Y(k) - C \hat{X}(k)\}^T R_k^{-1} \{Y(k) - C \hat{X}(k)\} \right] \tag{7.11}
\]

subject to the constraint \( X(k+1) = AX(k) + Bu(k) + w(k) \) and to determine \( X(0) \) and \( w(k) \) for \( 0 \leq k \leq N \) such that \( J_e \) in (7.11) is minimized. This will result in minimum mean square error estimate for the state using \( X(0), w(k) \) and the constraint equation above. Thus minimizing \( J_e \) is equivalent to minimum mean square error state estimation. The corresponding Riccati equation to be solved for obtaining the optimal estimation to minimize (7.11) is

\[
P(k+1) = A[P(k) - P(k)C^T \{ CP(k)C^T + R_k \}^{-1} CP(k) ]A^T + Q_k \tag{7.12}
\]

with \( P(0) = P_0 \), the initial estimation error covariance matrix and \( P(k) \) is the covariance matrix of the state error at time, \( k \). The optimal state observer using the solution for \( P \) of (7.12) is given by (7.10), where: \( K(k) = P(k)C^T \{ CP(k)C^T + R_k \}^{-1} \).

The control problem in section:7.5.1 and the estimation problem in section:7.5.2 are
dual problems and the governing Riccati equations: equation (7.8) with $S$ and equation (7.12) with $P$ are mathematically identical if we identify input with output and weighting matrices with covariance matrices, transpose of other matrices and time reversal as follows:

<table>
<thead>
<tr>
<th>Control problem</th>
<th>$B^T$</th>
<th>$G$</th>
<th>$F$</th>
<th>$S(N)$</th>
<th>$C^T$</th>
<th>$A^T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation Problem</td>
<td>$C$</td>
<td>$R$</td>
<td>$Q$</td>
<td>$P(0)$</td>
<td>$B$</td>
<td>$A$</td>
</tr>
</tbody>
</table>

### 7.6 Optimal stochastic fuzzy observers for discrete time fuzzy system

Since all the local fuzzy observers are linear, its quadratic optimization problem (optimal estimation) is the same as the general linear quadratic (LQ) issue in sections: 7.5.1 and 7.5.2 [111,122]. Therefore solving the optimal observer problem (optimal a posteriori density) for fuzzy system can be achieved by generalizing the classical theorem in the Proposition: 7.3 in the appendix from the classical stochastic case to the fuzzy stochastic case based on the extended results in Propositions 7.1 & 7.2 of Bayesian Gaussian sum approach. We summarize this generalization result here.

**Theorem 7.1**
Consider the fuzzy system in (7.3) & (7.4). Let $A_i(k), B_i(k), C_i(k), R_k = R_k^T > 0$ and $Q_i = Q_i^T > 0$ and $P(0) = P(0)^T > 0$ be given matrices. $P(0)$ is the initial state covariance matrix. We assume that the density of initial state $p(X_0)$ can be exactly represented as the convex sum of $r$ Gaussian densities as

$$ p(X_0) = \sum_{i=1}^{r} \alpha_i(0) \, N\{X_0 - \mu_i(0), P_i(0)\}. \quad \mu_i, \ i = 1,2,..r \text{ corresponding to the } r \text{-representative points in the state space suitably selected for the fuzzy modeling in (7.3) & (7.4).} $$

If there exists on $[0, N]$ an $n \times n$ symmetric positive semi-definite solution $\pi_i(k, P_i(0))$ to the matrix Riccati difference equation,
\[ P_i(k+1) = A_i(k)\{ P_i(k) - A_i(k)C_i^T(k)\{ C_i(k)P_i(k)C_i^T(k) + R_k \}^{-1} C_i(k)P_i(k) \}A_i^T(k) + Q_k \]

then, there exists an optimal local fuzzy observer and the optimal estimate is Gaussian distributed since the local subsystems are linear. The \(i^{th}\) local optimal observer is

\[ \mu_i(k+1) = \{ A_i(k) - K_i(k+1)C_i(k)A_i(k) \} \mu_i(k) + \{ I_i - K_i(k+1)C_i(k) \} B_i(k)u(k) + K_i(k+1)Y(k+1) \]

\[ K_i(k) = \overline{\pi}_i(k)\{ C_i(k)\overline{\pi}_i(k)C_i^T(k) + R_k \}^{-1} \]

and the global a posteriori density of the state conditioned on measurements up to \(k\) is

\[ p(X_k | Y^k) = \sum_{i=1}^{r} \alpha_i(k) N\{ X(k) - \mu_i(k) , \pi_i(k) \} \tag{7.13} \]

where weights are updated as

\[ \alpha_i(k) = \frac{\alpha_i(k-1) N\{ Y(k) - C_i(k)\overline{\mu}_i(k) , \Omega_i(k) \}}{\sum_{j=1}^{r} \alpha_j(k-1) N\{ Y(k) - C_j(k)\overline{\mu}_j(k) , \Omega_j(k) \}} \]

\[ \overline{\mu}_i(k) = A_i(k-1) \mu_i(k-1) + B_i(k-1)u(k-1) \]

\[ \Omega_i(k) = [ C_i(k)\overline{\pi}_i(k)C_i^T(k) + R_k ] \quad \text{for } i = 1,2,...,r. \]

The optimal a posteriori density in (7.13) need not be Gaussian and it depends on the nonlinear dynamics and measurement equations.

\[ N\{ Y(k) - C_i(k)\overline{\mu}_i(k) , \Omega_i(k) \} \]

are Gaussian densities and \(\Omega_i(k)\) are covariance matrix of \(i^{th}\) residue vector. The weighting coefficients are updated during measurement update based on the \(r\)-residues at that instant. These residues depend on the current state and current measurement. Hence the weighting coefficients are functions of state.

**Proof:**

The proof obviously holds with Proposition 7.3 and the extended results in Propositions 7.1 and 7.2 of Bayesian sum approach in the appendix. 

**Theorem 7.2**

Consider the time invariant form of the fuzzy system in (7.3) & (7.4) with stationary process and measurement noise. Let \(A_i, B_i, C_i, Q, R\) be these constant matrices and are given. Assume \(Q = \sqrt{Q}\sqrt{Q}^T \geq 0\). If \((A_i, \sqrt{Q})\) are stabilizable and \((A_i, C_i)\) are detectable for \(i = 1,2,...,r\); then
a) There is a unique Positive definite limiting solution, $\bar{\pi}_i(\infty)$ as $(k \to \infty)$ for each of the matrix Riccati difference equation: $(i = 1,2 \ldots r)$

$$P_i(\infty) = A_i [P_i(\infty) - P_i(\infty)C_i^T \{C_i P_i(\infty)C_i^T + R_k\}^{-1} C_i P_i(\infty)] A_i^T + Q$$

b) The asymptotically stable local optimal fuzzy observers are

$$\mu_i(k+1) = (A_i - K_i(\infty)C_iA_i) \mu_i(k) + \{I_n - K_i(\infty)C_i\} B_i u(k) + K_i(\infty)Y(k+1)$$

$$K_i(\infty) = \frac{\bar{\pi}_i(\infty) C_i^T \{C_i \bar{\pi}_i(\infty)C_i^T + R\}^{-1}}{\sum_{j=1}^{r} \alpha_i(k-1) N\{Y(k) - C_i \bar{\mu}_i(k), \Omega_i(k)\}}$$

The global a posteriori density of the state conditioned on measurements up to $k$ is

$$p(X_k | Y^k) = \sum_{i=1}^{r} \alpha_i(k) N\{X(k) - \mu_i(k), \pi_i(\infty)\}$$

(7.14)

$$\pi_i(\infty) = \bar{\pi}_i(\infty) - \bar{\pi}_i(\infty)C_i^T \{C_i \bar{\pi}_i(\infty)C_i^T + R\}^{-1} C_i \bar{\pi}_i(\infty)$$

where weights are updated as

$$\alpha_i(k) = \frac{\alpha_i(k-1) N\{Y(k) - C_i \bar{\mu}_i(k), \Omega_i(k)\}}{\sum_{j=1}^{r} \alpha_j(k-1) N\{Y(k) - C_j \bar{\mu}_j(k), \Omega_j(k)\}}$$

$$\bar{\mu}_i(k) = A_i \mu_i(k-1) + B_i u(k-1)$$

$$\Omega_i(k) = [C_i \bar{\pi}_i(\infty)C_i^T + R] \text{ for } i = 1,2 \ldots r.$$  

$N\{Y(k) - C_i \bar{\mu}_i(k), \Omega_i(k)\}$ are Gaussian densities and $\Omega_i(k)$ are covariance matrix of $i$th residue vector.

The optimal a posteriori density in (7.14) need not be Gaussian and it depends on the nonlinear dynamics and measurement equations. It can be shown that the global observer given by (7.14) for the time invariant case is globally exponentially stable using the results in [155].

**Proof:**

The proof obviously holds with Propositions: 7.4 & 7.5 and the extended results in Propositions: 7.1 & 7.2 of Bayesian Gaussian sum approach in the appendix.  

**Appendix:**

**Proposition :**7.1 [133]

Given $X(k) = X_k = f_k(X_{k-1}) + w_{k-1}$

(7.15)

where $f_k(.)$ is nonlinear plant dynamics and $w_k$ is the zero mean white Gaussian process noise. Let $Z^{k-1}$ represent all measurements up to time, $k - 1$, and with the
knowledge of the density with all measurements up to $k-1$, $p(X_{k-1} | Z^{k-1})$ expressed as the sum of Gaussian terms and the knowledge of the dynamic equations of the state vector $X_k$, it is possible to calculate approximately the one-step-ahead predicted density $p(X_k | Z^{k-1})$ expressed as a sum of Gaussian terms using the dynamic equation (7.15). More precisely, consider $p(X_{k-1} | Z^{k-1})$ as the summation:

$$p(X_{k-1} | Z^{k-1}) = \sum_{i=1}^{r} \alpha_i (k-1) N\{X_{k-1} - \mu_i (k-1), P_i (k-1)\}$$

where, $\sum_{i=1}^{r} \alpha_i = 1$, $\alpha_i \geq 0$ for all $i$ and $k$. Applying the extended Kalman filter theory to yield an approximate expression for the one-step ahead predicted estimate of each Gaussian distribution $N\{X_{k-1} - \mu_i (k-1), P_i (k-1)\}$ to the gaussian distribution $N\{X_k - \bar{\mu}_i (k), \bar{P}_i (k)\}$, where

$$\bar{\mu}_i (k) = f_k (\mu_i (k-1)) \tag{7.16a}$$

$$\bar{P}_i (k) = \Phi_i (k-1) P_i (k-1) \Phi^T_i (k-1) + Q_{k-1} \tag{7.16b}$$

$$\Phi_i (k-1) = \frac{\partial f_k (X)}{\partial X} \bigg|_{X=\mu_i (k-1)} \tag{7.16c}$$

then, the one-step ahead a posteriori density $p(X_k | Z^{k-1})$ approaches the Gaussian sum

$$p(X_k | Z^{k-1}) = \sum_{i=1}^{r} \alpha_i (k-1) N\{X_k - \bar{\mu}_i (k), \bar{P}_i (k)\}$$

uniformly in $X(k-1)$ as $P_i (k-1) \to 0$ for $i = 1, 2, \ldots, r$.

**Note:** The approximate expressions for the one-step-ahead prediction above become exact if the nonlinear dynamics in (7.15) can be represented exactly by

$$X(k) = \left\{ \sum_{i=1}^{r} \alpha_i (k-1) [A_i X(k-1) + B_i u(k-1)] \right\} + w(k-1)$$

$$= \left\{ \sum_{i=1}^{r} [A_i \mu_i (k-1) + \alpha_i (k-1) B_i u(k-1)] \right\} + w(k-1)$$

where $\mu_i (k-1) = \alpha_i (k-1) X(k-1)$

as weighted sum of linear subsystems. The equations in (7.16a,b,c) also appropriately get modified.
Proposition: 7.2 [133]

Consider the measurement equation

\[ Z(k) = g_k(X_k) + v(k) \]  \hspace{1cm} (7.17)

where \( g_k(\cdot) \) is a nonlinear function of state and \( v(k) \) is the zero mean white Gaussian measurement noise. If we have the knowledge \( p(X_k | Z^{k-1}) \) as a weighted sum of Gaussian densities as in Proposition: 7.1, i.e.,

\[ p(X_k | Z^{k-1}) = \sum_{i=1}^{r} \alpha_i(k-1) \ N\{X_k - \mu_i(k), \ \bar{P}_i(k)\} \]

then the knowledge of the extended Kalman filter can be used to update the density \( p(X_k | Z^k) \), which approaches the following Gaussian

sum \( p(X_k | Z^k) = \sum_{i=1}^{r} \alpha_i(k) \ N\{X_k - \mu_i(k), \ P_i(k)\} \) uniformly in \( X(k) \) and \( Z(k) \) as

\[ P_i(k) \to 0 \quad \text{for} \quad i = 1, 2, \ldots, r. \]

\[ \mu_i(k) = \bar{\mu}_i(k) + K_i(k)[Z_k - g_k(\bar{\mu}_i(k))] \]  \hspace{1cm} (7.18a)

\[ P_i(k) = \bar{P}_i(k) - \bar{P}_i(k)C_i^T(k) [C_i(k)\bar{P}_i(k)C_i^T(k) + R_i]^{-1} C_i(k)\bar{P}_i(k) \]  \hspace{1cm} (7.18b)

\[ K_i(k) = \bar{P}_i(k)C_i^T(k) \Omega_i(k)^{-1} \]  \hspace{1cm} (7.18c)

where

\[ C_i(k) = \left. \frac{\partial g_k(X)}{\partial X} \right|_{X = \bar{\mu}_i(k)} \quad \text{and} \quad \Omega_i(k) = [C_i(k)\bar{P}_i(k)C_i^T(k) + R_i] \]

\( \Omega_i(k) \) is the Gaussian density of the \( i^{th} \) residual vector.

The weights are updated as

\[ \alpha_i(k) = \frac{\alpha_i(k-1) \ N\{Z_k - g_k(\bar{\mu}_i(k)), \ \Omega_i(k)\}}{\sum_{j=1}^{r} \alpha_j(k-1) \ N\{Z_k - g_k(\bar{\mu}_j(k)), \ \Omega_j(k)\}} \]

Note: This approximate measurement update becomes exact if the nonlinear measurement equation in (7.17) can be represented exactly by

\[ \left\{ \sum_{i=1}^{r} \alpha_i(k)C_i(k)X(k) \right\} + v(k) = \left\{ \sum_{i=1}^{r} C_i(k)\mu_i(k) \right\} + v(k) \quad \text{where} \quad \mu_i(k) = \alpha_i(k)X(k) \]

i.e., as a weighted sum of linear measurement equations. The update equations (7.18a,b,c) get appropriately modified.
**Proposition: 7.3** [137, 133, 122]
Consider the discrete time dynamic system
\[ X(k+1) = A(k)X(k) + B(k)u(k) + w(k) \]  
(7.19)
and output \[ Y(k) = C(k)X(k) + v(k) \]

\( w(k) \in \mathbb{R}^n \) and \( v(k) \in \mathbb{R}^q \) and \( w(k) \) \& \( v(k) \) are the zero mean white Gaussian process noise and measurement noise vectors with covariance matrices \( Q(k) \) and \( R(k) \) respectively and are finite and bounded and mutually uncorrelated. \( X(k) \in \mathbb{R}^n \) is the state vector, \( u \in \mathbb{R}^p \) is the deterministic input vector, and \( Y(k) \in \mathbb{R}^q \) is the output vector. \((A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{q \times n})\) is the matrix triplet. Given any fixed initial time, \( t_0 = k=0 \) and an initial nonnegative definite symmetric matrix \( P_0 \) (Covariance matrix of estimation error at \( t_0 \)) then if there exists a symmetric positive semi-definite solution \( P(k) = \pi(k, P_0) \), for the matrix Riccati equation:

\[
P(k+1) = A(k)[P(k) - P(k)C^T(k)\{C(k)P(k)C^T(k) + R_k\}^{-1} C(k)P(k)]A^T(k) + Q_k
\]
(7.20)
on the interval \( [0, N] \), then there exists the corresponding optimal observer

\[
\dot{\hat{X}}(k) = [A(k) - K(k)C(k)A(k)]\dot{\hat{X}}(k-1) + \{I_n - K(k)C(k)\}B(k)u(k-1) + K(k)Y(k)
\]
(7.21)
where, \( K(k) = \pi(k, P_0)C^T(k)\{C(k)\pi(k, P_0)C^T(k) + R_k\}^{-1} \)

The estimate \( \dot{\hat{X}}(k) \) of \( X(k) \) minimizes the mean square estimation error of the state.

**Proposition: 7.4** [123]
If the system in (7.19) is time invariant with \( A, B, C, Q \) and \( R \) are all constant matrices with \( Q \geq 0 \), \( R > 0 \) and if \((A, C)\) is detectable, then for every choice of \( P_0 \) there is a bounded limiting solution \( P \) to (7.20) as \( k \to \infty \). Further, this limiting solution \( \pi(\infty) \) is a positive semi-definite solution to the algebraic Riccati equation:

\[
P(\infty) = A[P(\infty) - P(\infty)C^T\{C P(\infty)C^T + R\}^{-1} C P(\infty)]A^T + Q
\]
(7.22)
**Proposition: 7.5** [123]

For the time-invariant case in Proposition: 7.4, with \( Q = \sqrt{Q} \sqrt{Q}^T \geq 0, R > 0 \) if \((A, \sqrt{Q})\) is stabilizable and \((A, C)\) is detectable, then

a. There is a unique positive definite limiting solution, \( P \) to the algebraic Riccati equation (7.22), which is independent of \( P_0 \).

b. The steady state error system \( e(k+1) = A(I_n - K C)e(k) + w(k) - AKv(k) \) with steady state Kalman gain \( K = PC^T(CPC^T + R)^{-1} \) is asymptotically stable.

Where, \( e(k) = X(k) - \hat{X}(k) \).

**7.7 Conclusion**

Theoretical results for the design of optimal observer for nonlinear system based on discrete time TS fuzzy model are developed. This is advance and significant.