Hysteresis parameter identification of hybrid dynamical systems

Masood Askari¹*, Mahmoud Moghavvemi¹ and Haider A. F. Almurib²

¹Centre for Research in Applied Electronics (CRAE), University of Malaya, 50603 Kuala Lumpur, Malaysia.
²Department of Electrical and Electronic Engineering, University of Nottingham Malaysia Campus, Jalan Borta, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.

Accepted 11 February, 2011

This paper discusses the identification of a class of hybrid dynamical systems whose discrete states have hysteresis phenomenon. The identification model is piecewise affine model along with hysteresis switching law. The first step of the proposed identification method is the estimation of the local parameter vectors for small neighbourhood of each measured data point. Then the local parameter vectors are clustered. Finally, the threshold levels of the hysteresis function which defines the discrete state are estimated.

Key words: Hybrid systems, nonlinear identification, piecewise affine systems, hysteresis.

INTRODUCTION

Hybrid systems are composed of both continuous dynamics governed by physical laws and discrete-event dynamics driven by logic and rules. Recently, much attention has been paid to hybrid systems from various viewpoints (Shorten and Narendra, 2006).

Hybrid system identification is composed of two steps. First step is identification of the sub-models governing the dynamics of the continuous states for each mode, and second step involves with identification of the discrete state or mode for each data point.

In this paper the focus is on hybrid systems which sub-models are PieceWise Affine (PWA) systems and hysteresis phenomenon appears when sub-models are switched. The proposed method consists of identification of both PWA models and hysteresis switching model.

There are many linear systems with PWA nonlinearities such as saturation or relay elements, and a general nonlinear system can be treated as a PWA system by approximating a nonlinear function by a PWA one with arbitrary accuracy. There are some applications to the identification of real systems, for example, a nonlinear electrical circuit (Ferrari-Trecate et al., 2003), a fermentation process (Fantuzzi et al., 2002), and a pick-and-place machine (Juloski et al., 2004). Several procedures are proposed for the identification PWA models (Juloski et al., 2006; Paoletti et al., 2007). For example, algebraic procedure (Bako and Vidal, 2008), bounded-error procedure (Bemporad et al., 2005) clustered-based procedure (Ferrari-Trecate et al., 2003) and Bayesian procedure (Juloski et al., 2005).

The problem of identifying PWA systems involves estimating the number of sub-models in the system, the parameters which define them, and the regions of the regression space where they are defined. If the number of sub-models and their corresponding regions are not known, the identification problem becomes complicated as it involves processes for both classifying the data and estimating the parameters and regions (Gegundez et al., 2008). Clustering techniques are powerful tools to define the regions of sub-models (Nakada et al., 2005; Nandola and Bhartiya, 2009). Although a general nonlinear system can be identified as a piecewise affine system by approximating a nonlinear function using a piecewise affine one with arbitrary accuracy, the memory feature of hysteresis elements makes switching law complicated.

Hysteresis is a nonlinear phenomenon that occurs in many mechanical and electrical systems when there are
magnetic fields, piezoelectric elements or friction in the system. Hysteresis often leads to problems in control systems because it causes tracking errors, limit cycles (Ikhouane and Gomis-Bellmunt, 2008), and undesired stick-slip motions (Li et al., 2004). The identification of hysteresis parameters of a system is quite important, because it leads to construct the inverse of the hysteresis model and inverse compensation, which is a general way to cancel out the hysteresis effect (Chan and Liu, 2007). Recursive identification and adaptive inverse control of hysteresis in smart material for example, piezoelectrics, which are used in sensors and actuators is addressed in (Tan and Baras, 2005). In (Giri et al., 2008) identification of Hammerstein systems in presence of hysteresis-backlash and hysteresis-relay nonlinearities is discussed, and the linear subsystem and the nonlinear element are estimated separately. Identification of PWA systems with hysteresis phenomenon in the Duhem model, which is widely used for friction modeling, is addressed in (JinHyoung and Bernstein, 2007).

In this paper, a new method for identification of hybrid systems, when local models can be defined by PWA sub-models and hysteresis phenomenon appears in switching law, is presented.

**PROBLEM STATEMENT**

A PWA regression problem amounts reconstructing the PWA map \( f \) from a finite set of data points \( z(k) = [y(k) \ x(k)]^T \) generated by the model

\[
y(k) = f(x(k)) + \epsilon(k)
\]  

Where \( \epsilon(k) \) are Gaussian independent identically distributed random variables. A PWA map is defined by the equation:

\[
f(x) = \begin{cases} 
\theta_1 \begin{bmatrix} x \\ 1 \end{bmatrix} & q = 1 \\
\vdots & \\
\theta_s \begin{bmatrix} x \\ 1 \end{bmatrix} & q = s
\end{cases}
\]  

Where \( \theta \in \mathbb{R}^{n+1} \) are parameter vectors and \( x \in \mathbb{R}^n \) is the vector of regressors and is defined as:

\[
x(k) = [y(k-1) \cdots y(k-n) \ u(k-1) \cdots u(k-n_b)]^T
\]  

and \( q \) corresponds the mode of the system.

The identification problem, when \( q \) can be determined by the rule:

\[
\text{If } x \subset \chi_i \text{ Then } q = i,
\]

is discussed in (Ferrari-Trecate et al., 2003), where \( \chi \subset \mathbb{R}^n \) is a bounded polyhedron and \( \{\chi_i\}_{i=1}^s \) is a polyhedral partition of \( \chi \), i.e. each set \( \chi_i \) is a (not necessarily closed) convex polyhedron s.t.

\[
\forall i \neq j \quad \chi_i \cap \chi_j = \emptyset
\]

and

\[
\bigcup_{i=1}^s \chi_i = \chi
\]

Polyhedral partitioning of \( \chi \) is not always possible. If there are some nonlinearities phenomena like hysteresis, it is not possible to partition \( \chi \) such that (5) be evaluated. This is illustrated by using the following example.

**Example 1.** Consider the single input/single output system with hysteresis phenomena depicted in Figure 1, where \( x(k) = u(k-1) \). Consider these partitions:

\[
\chi_1 = \{x| x < 2\}, \quad \chi_2 = \{x| x > 1\}.
\]

There are affine models for each \( \chi_1 \) and \( \chi_2 \), but \( \chi_1 \cap \chi_2 \neq \emptyset \). In other words, \( q \) is dependent on the previous trajectory of \( x \).
In this paper identification of piecewise affine systems in the case of overlapped partitioning of \( \mathcal{X} \) under following assumptions is discussed.

**Assumption 1.** There is prior knowledge about the behavior of the system.

**Assumption 2.** The number of sub-models \( s \) is predefined.

**Assumption 3.** Measured data is available

\[
z(k) = [y(k) \ x(k)]^T, \quad k = 1, \ldots, N \]  

and the input is persistently exciting and the sample points are not all concentrated around the boundary of the sets \( \mathcal{X}_i \). The identification problem, considered in this paper, can be stated as problem 1.

**Problem 1.** Estimate the parameter vectors \( \theta_i, i = 1, 2, \ldots, s \) and the conditions that change the system mode \( q \), from the measured data \( z(k), \ k = 1, 2, \ldots, N \) under assumptions 1, 2 and 3.

### IDENTIFICATION ALGORITHM

A general framework for the identification of a PWA system is summarized as follows.

**Phase 1:** Local regression. In this phase, the least-squares method is used to estimate the local parameter vectors for small neighborhood of each data points.

**Phase 2:** Clustering. In this phase, the local parameter vectors are clustered in \( s \) groups.

**Phase 3:** Modes estimation. In this phase, the boundary levels and hysteresis function which determine the system mode are estimated.

**Local regression**

In this step, for each data point \( (x(k), y(k)) \) \( j = 1, \ldots, N \) a local data set \( L_j \) is formed. It collects \( (x(k), y(k)) \) and the samples \( (x, y) \) corresponding to the \( c-1 \) nearest neighbors \( x \) to \( x(k) \). The parameter \( c \) should be suitably chosen by user in order to minimize the number of outliers. Local data sets collecting only data points associated to a single mode are referred to as pure local data sets. The remaining local data sets are termed mixed. Linear regression is performed on each local data set to obtain the local parameter vectors \( \theta^L_j \) and their empirical variance \( V_j \) (Ljung, 1999). The local parameter vector \( \theta^L_j \) estimated from the data in \( L_j \) is computed by least-squares estimation:

\[
\theta_j^L = (\Psi_j^T \Psi_j)^{-1} \Psi_j^T y_j, \tag{8}
\]

Where:

\[
\Psi_j = \begin{bmatrix} x(k_1) & x(k_2) & \cdots & x(k_e) \end{bmatrix}. \tag{9}
\]

The local data set centers

\[
m_j = (1/c) \sum_{(x,y) \in L_j} x \quad \text{are also computed together with the associated scatter matrix}
\]

\[
Q_j = \sum_{(x,y) \in L_j} (x-m_j)(x-m_j)^T. \tag{11}
\]

The information about the \( j \)th local model is collected in the feature vector

\[
\xi_j = \begin{bmatrix} \theta_j^L \\ m_j \end{bmatrix}. \tag{12}
\]

As for the local data sets, feature vectors are either pure or mixed. Feature vector \( \xi_j \) is interpreted as the realization of a Gaussian random variable with variance

\[
R_j = \begin{bmatrix} V_j & 0 \\ 0 & Q_j \end{bmatrix}. \tag{13}
\]

where \( V_j \) is the empirical variance of \( \theta_j^L \) (Ljung, 1999).

The problem of finding the data points associated to the same mode can be recast into the problem of finding \( s \) dense clouds of pure feature vector. However, one should be warned about the presence of mixed feature vector that do not carry any useful information on the true modes and form a pattern of isolated points in the feature vector space (Ferrari-Trecate et al., 2003).

### Clustering

The feature vectors are partitioned in \( s \) groups through clustering. For this purpose, a modified K-means algorithm which is described in (Ferrari-Trecate et al., 2003), is used. The main differences between this algorithm and the classical K-means are the use of the
matrices $R_j^{-1}$ as distance measures for assigning the vectors $\xi_j$ to $s$ clusters.

The centers $\tilde{\xi}_j$, $i=1,\ldots,s$, of the clusters are consist of the parameter vectors $\tilde{\theta}_j$ and data set centers of sub-models.

$$
\tilde{\xi}_j = \begin{bmatrix} \tilde{\theta}_j \\ \tilde{m}_j \end{bmatrix}, \quad i=1,\ldots,s
$$

$$
u(k+1) = u(k) + \Delta u
$$

**Estimation of the modes**

A new algorithm to determine the conditions that change the mode of system is proposed in this paper. After calculating the parameter vectors $\tilde{\theta}_j$, the thresholds are computed using a suitable subset of measured input-output data.

In order to detect the hysteresis characteristics of process, input-output data acquisition is repeated, while ramp signal (15) is applied to process.

$$
u(k+1) = u(k) + \Delta u
$$

where $\Delta u$ is a non-zero piecewise constant variable.

As a result of clustering, there are $s$ feature vector. Therefore, there are also $s$ parameter vectors, which estimate $s$ output for each data point. The error between estimated and measured output is denoted by $e_i(k)$.

$$
e_i(k) = y(k) - \tilde{\theta}_i \begin{bmatrix} x(k) \\ 1 \end{bmatrix}, \quad i=1,\ldots,s
$$

Sum of squared error for $p$ samples before and $p$ samples after data point at time $k$ are computed as (17) and (18).

$$
E^-(k) = \sum_{j=k-p}^{k} (e_i(j))^2
$$

$$
E^+(k) = \sum_{j=k}^{k+p} (e_i(j))^2
$$

The functions $I^-(k)$ and $I^+(k)$ defined by (19) and (20) lead us to determine the mode of the system.

$$
I^-(k) = \arg \min \{E^-(k)\}
$$

$$
I^+(k) = \arg \min \{E^+(k)\}
$$

The function $h(k)$ defined by (21) can used to identify the pure and mixed data points.

$$
h(k) = I^+(k) - I^-(k)
$$

The system mode switches at a instant between the positive and negative edge of $h(k)$. The switch time can be approximated by averaging the positive edge instant $l_p$ and negative edge instant $l_n$.

$$
l_{av} = \text{round}((l_p + l_n)/2)
$$

The values of low threshold $T_l$ and high threshold $T_h$ can be obtained from the values of data points at $k = l_{av}$. Threshold levels are expressed based on trajectory of states of the system. Since the states are observed by output and input, there is a relation between threshold levels and trajectory of input and output, which increases the complexity of algorithms to find the thresholds. The technique presented in this paper is to apply a triangular shaped input to span the whole range of admissible values of input. In this way, the input-output acquired data is used to define the values of thresholds based on outputs. Since the trajectory of output depends on input, the threshold levels implicitly depend on input while they are explicitly expressed by output. Thus, low and high threshold levels are defined as:

$$
T_l = y(k)|_{k=l_{av}}
$$

$$
T_h = y(k)|_{k=l_{av}^2}
$$

Where $l_{av1}$ and $l_{av2}$ are defined by (22).

This procedure is illustrated by the following example.

**Example 2.** Suppose that a piecewise constant input signal, which is not concentrated around the switching edge, is applied to system (25) and input-output data are acquired by simulation:

$$
\begin{cases}
y(k) = 0.99y(k-1) + 0.1u(k-1) & q = 1 \\
y(k) = 0.8y(k-1) + 1.5u(k-1) & q = 2
\end{cases}
$$

The mode $q$ has hysteresis phenomenon with two threshold levels $T_l$ and $T_h$ which are depicted in
The objective is to estimate the parameter vectors and threshold levels. The local parameter vectors for small neighbourhood of each data point is estimated and clustered in $s=2$ groups, and the following local parameter vectors are obtained:

$$\bar{\theta}_1 = [0.991 \ 0.1 \ 0]$$
$$\bar{\theta}_2 = [0.801 \ 1.499 \ 0]$$

In the next step, the threshold levels are estimated by the new method presented in this paper. Smooth subset of acquired input-output data, depicted in Figure 3, is used to identify the thresholds. The input signal is obtained by (15).

Sum of squared error for samples before and samples after each data point are obtained by (17) and (18) respectively.

The functions $I^+(k)$ and $I^-(k)$ represent the system mode for $p$ samples after and $p$ samples before data point at time $k$, and are depicted in Figures 4 and 5 respectively.

The function $h(k)$ defined by (21) is used to identify the mixed data points. Figure 6 shows the values of $h(j)$ for the data points depicted in Figure 3.

In this example, as shown in Figure 6, $h(k)$ has two pairs of edges. One of them which is depicted in Figure 7, gives us the negative edge instant $l_{n1} = 5247$ and positive edge instant $l_{p1} = 5261$. The switching instant $l_{av1} = 5254$ is obtained by (22).

$l_{av2}$ is obtained by another pair of edges, and threshold levels are obtained by (23) and (24) and finally the system
mode is estimated as:

\[
    q(k) = \begin{cases} 
        2 & y(k-1) > T_h \\
        1 & y(k-1) < T_i \\
        q(k-1) & T_i < y(k-1) < T_H
    \end{cases}
\]

Conclusion

In this paper, identification of hybrid systems, when polyhedral partitioning of data is not possible, is discussed and a new method for estimating the boundary levels is presented. This method is applicable when the system mode depends on not only the continuous states but also the discrete states in the last sample. The proposed method is illustrated by an example.

REFERENCES


