16th Triennial World Congress, Prague, Czech Republic

IDENTIFICATION OF QUASI-ARMAX MODELS OF NONLINEAR STOCHASTIC SAMPLED-DATA SYSTEMS

Bernt M. Åkesson [1] and Hannu T. Toivonen

Department of Chemical Engineering
Åbo Akademi University
FI-20500 Turku, Finland

Abstract: State-dependent parameter representations of nonlinear stochastic sampled-data systems are studied. Velocity-based linearization is used characterize sampled-data systems using nominally linear models whose parameters can be represented as functions of past outputs and inputs. For stochastic systems the approach leads to state-dependent ARMAX (quasi-ARMAX) representations. The models and their parameters are identified from input-output data using feedforward neural networks to represent the model parameters as functions of past inputs and outputs. *Copyright © 2005 IFAC*

Keywords: neural network models, nonlinear models, sampled-data systems, stochastic systems, parameter identification

1. INTRODUCTION

A widely used approach for the black-box modelling and identification of nonlinear dynamical systems is to apply various nonlinear function approximators, such as artificial neural networks or fuzzy models. A shortcoming of these models is that they do not provide much insight into the systems dynamics. For this reason various model structures, which provide such information, have been introduced. One general class of models of this type consists of models with a nominally linear structure, but with state-dependent parameters (Priestley, 1988; Hu *et al*., 2001; Young *et al*., 2001). An important class of state-dependent parameter models consists of ARX models, in which the model parameters are nonlinear functions of past system outputs and inputs. These models have been called quasi-ARX (Hu *et al*., 1998; Hu *et al*., 2001; Previdi and Lovera, 2001) or state-dependent ARX models (Priestley, 1988; Young *et al*., 2001). State-dependent parameter representations have the useful property that explicit information about the local dynamics is provided by the locally valid linear model, and in a number of situations they can be treated as linear systems whose parameters are taken as functions of scheduling variables.

For discrete-time systems, state-dependent parameter representations are usually approximative descriptions introduced for convenience. In contrast, continuous-time systems can be represented exactly by state-space models with state-dependent parameters constructed using velocity-based linearization (Leith and Leithead, 1998*b*; Leith and Leithead, 1998*a*). This fact can be applied to construct exact discrete-time state-dependent parameter representations for sampled-data systems (Toivonen, 2003). Quasi-ARX models of sampled-data systems are obtained by reconstructing the state of the state-dependent parameter representation in terms of past inputs and outputs (Toivonen, 2003).

In practice it is important to be able to deal with systems which are subject to stochastic noise. It is shown that a nonlinear sampled-data system subject

[1] Supported by the Finnish Graduate School in Chemical Engineering (GSCE)

to an additive drifting disturbance and measurement noise can be represented by a quasi-ARMAX model. By using a feedforward neural network to describe the model parameters as functions of past inputs and outputs (Hu and Hirasawa, 2002), the quasi-ARMAX model is represented with a type of recurrent network. Identification of neural network quasi-ARMAX models from input-output data is studied and illustrated with numerical examples.

2. STATE-DEPENDENT PARAMETER MODELS OF STOCHASTIC SYSTEMS

In a previous study (Toivonen, 2003), state-dependent parameter representations were derived for deterministic nonlinear sampled-data systems. In this paper, a generalization to stochastic systems is studied. Consider a nonlinear system

$$\begin{aligned} \dot{x}_P(t) &= f_P(x_P(t), u(t)) \\ y(t) &= h_P(x_P(t)) + w(t) \end{aligned} \tag{1}$$

where $x_P(t)$ denotes the state vector, $u(t)$ is the control input and $y(t)$ denotes the output. It is assumed that the nonlinear functions $f_P(\cdot,\cdot)$ and $h_P(\cdot)$ are continuous with Lipschitz continuous first derivatives. The system is subject to an additive drifting disturbance $w(t)$, which is described by a Wiener process with the incremental variance r_w.

The continuous-time input $u(t)$ to the nonlinear system is generated from a discrete-time input $u_d(k)$ by a zero-order hold and a strictly proper low-pass filter with the state-space representation (A_H, B_H, C_H). The system input is thus generated according to

$$\begin{aligned} \dot{x}_H(t) &= A_H x_H(t) + B_H u_d(k),\ t \in (kh, kh+h] \\ u(t) &= C_H x_H(t) \end{aligned} \tag{2}$$

where h denotes the sampling interval. The filter (2) generates a continuous input $u(t)$ to (1), which is differentiable for all t, except possibly at the sampling instants kh.

The sampled output is corrupted by measurement noise,

$$y_m(kh) = y(kh) + e_m(k) \tag{3}$$

which is described by a zero-mean white noise disturbance $\{e_m(k)\}$ with the variance σ_m^2.

Although a more realistic and complex disturbance model could be used, one reason for focusing on the model (3) is that it allows the construction of an exact quasi-ARMAX representation. It is therefore possible to compare identified models with the theoretically correct system description. It is also believed that the combination of a drifting disturbance and measurement noise provides a good approximation of more complex disturbances as well.

The generalized system consisting of the filter (2), the nonlinear system (1) and the disturbance model (3) can be described by a nonlinear system with a piecewise constant input,

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + Bu_d(k),\ t \in (kh, kh+h] \\ y(t) &= h(x(t)) + w(t) \\ y_m(kh) &= y(kh) + e_m(k) \end{aligned} \tag{4}$$

where $x = [x_P^T, x_H^T]^T$ is the state of the generalized system, and

$$\begin{aligned} f(x) &= \begin{bmatrix} f_P(x_P, C_H x_H) \\ A_H x_H \end{bmatrix},\ B = \begin{bmatrix} 0 \\ B_H \end{bmatrix} \\ h(x) &= h_P(x_P) \end{aligned} \tag{5}$$

Differentiation of (4) with respect to time gives a nonlinear system with jumps,

$$\begin{aligned} \ddot{x}(t) &= A(x(t))\dot{x}(t),\ t \neq kh \\ x(kh^+) &= \dot{x}(kh) + B\ \ u_d(k) \\ dy(t) &= C(x(t))dx(t) + dw(t) \end{aligned} \tag{6}$$

where the notation $x(kh^+) = \lim_{\varepsilon \downarrow 0} x(kh+\varepsilon)$ has been used, $\ \ u_d(k) = u_d(k) - u_d(k-1)$ and

$$A(x) = \frac{\ f(x)}{\ x},\ C(x) = \frac{\ h(x)}{\ x} \tag{7}$$

Integration of (6) over the sampling intervals gives (Toivonen, 2003)

$$\begin{aligned} \dot{x}(kh+h) &= F(x(kh), u_d(k))\dot{x}(kh) \\ &\quad + G(x(kh), u_d(k))\ \ u_d(k) \\ \ y(kh+h) &= H(x(kh), u_d(k))\dot{x}(kh) \\ &\quad + J(x(kh), u_d(k))\ \ u_d(k) + e_w(k+1) \end{aligned} \tag{8}$$

where $\ y(kh+h) = y(kh+h) - y(kh)$ and $e_w(k+1) = w(kh+h) - w(kh)$. Hence $\{e_w(k)\}$ is a discrete-time white noise sequence with variance $\sigma_w^2 = r_w h$. The matrices $F(\cdot,\cdot)$, $G(\cdot,\cdot)$, $H(\cdot,\cdot)$ and $J(\cdot,\cdot)$ are smooth functions given by

$$F(x(kh), u_d(k)) = \ \ (kh+h) \tag{9}$$

$$H(x(kh), u_d(k)) = \ \ _y(kh+h) \tag{10}$$

and $G(\cdot,\cdot) = F(\cdot,\cdot)B$, $J(\cdot,\cdot) = H(\cdot,\cdot)B$, where $\ \ (\cdot)$ and $\ \ _y(\cdot)$ are defined by the differential equations

$$\frac{d\ \ (t)}{dt} = A(x(t))\ \ (t),\quad \ \ (kh) = I \tag{11}$$

$$\frac{d\ \ _y(t)}{dt} = C(x(t))\ \ (t),\quad \ \ _y(kh) = 0 \tag{12}$$

where $x(t)$ is given by (4).

The parameters of (8) are functions of the system state. In order to construct a model using input-output data