Proceedings of the 37th IEEE
Conference on Decision & Control
Tampa, Florida USA • December 1998

# Least Squares Identification Using $\mu$-Markov Parameterizations

Tobin H. Van Pelt
Department of Aerospace Engineering
The University of Michigan
vanpelt@engin.umich.edu

Dennis S. Bernstein [1]
Department of Aerospace Engineering
The University of Michigan
dsbaero@umich.edu

**Abstract**

In this paper we introduce $\mu$-Markov parameterizations for use in least squares estimation. These parameterizations explicitly contain the system impulse response parameters, or Markov parameters, and, under very general noise models, the least squares estimates of the Markov parameters are consistent regardless of model order choice when the input is white noise. A numerical example is given to illustrate this result.

## 1 Introduction

Although least squares techniques are commonly used for identification, it is well known [1, 2] that for most types of noise models, least squares estimates of parameters in a full parameterization are not consistent, and thus biased model estimates result. Moreover, accuracy of these techniques is highly dependent on knowledge of the true system order.

An alternate approach is to use a finite impulse response (FIR) model structure to approximate the infinite impulse response (IIR) system when performing least squares identification. FIR model parameterizations contain only impulse response, or Markov, parameters, and it is interesting to note that the estimates of these parameters are consistent [1, 2] even when the model structure is not correct. The significance of this result arises from the fact that the estimates of the Markov parameters can then be used to construct a model of the system using realization theory [3]. Therefore it is natural to seek a more general model parameterization that contains an arbitrary number of Markov parameters, but retains the appropriate IIR model structure. The objective in using such a model parameterization is to obtain consistent estimates of the Markov parameters with faster convergence than with FIR models.

## 2 $\mu$-Markov Parameterizations

Consider the discrete-time single-input single-output transfer function $G(q)$ with the fully parameterized $p$th-order parameterization

$$G(q) \sim \frac{b_0 q^p + b_1 q^{p-1} + \cdots + b_{p-1} q + b_p}{q^p + a_1 q^{p-1} + \cdots + a_{p-1} q + a_p}, \tag{1}$$

where $q$ is the forward shift operator.

**Proposition 1.** Let $G(q)$ be a transfer function of relative degree $r$ and let $h_0, h_1, \ldots$ denote the Markov parameters of $G(q)$. Consider the parameterization (1) of order $p$ and let $\mu$ be a positive integer such that $\mu \leq p-r$. Then $b_i = h_i$, $i = 0, \ldots, \mu + r$, if and only if $a_i = 0$, $i = 1, \ldots, \mu$.

**Definition 1.** Let $G(q)$ be a transfer function with McMillan degree $n$ and consider the non-minimal parameterization of $G(q)$ as $G(q) \sim \frac{B(q)}{A(q)}$, where $A(q) = q^p + a_1 q^{p-1} + \cdots + a_p$. Furthermore, let $\mu$ be a positive integer such that $\mu \leq p - n$. Then $\frac{B(q)}{A(q)}$ is a *$\mu$-Markov parameterization* of $G(q)$ if $a_i = 0$, $i = 1, \ldots, \mu$.

We note that all $\mu$-Markov parameterizations are non-minimal by definition. Moreover, when $p = n + \mu$ it can be shown that for each positive integer $\mu$ there exists a unique $\mu$-Markov parameterization.

## 3 $\mu$-Markov Least Squares Identification

In this section we consider least squares identification using $\mu$-Markov parameterizations. Consider the output $y(k)$ generated by the linear time-invariant system

$$y(k) = G(q)u(k) + H(q)w(k), \tag{2}$$

where $G(q)$ and $H(q)$ are stable, proper transfer functions, $u(k)$ is the input signal, and $w(k)$ is the disturbance signal. Next consider the $p$-th order $\mu$-Markov parameterization of $G(q)$ and write (2) in regression form

$$y(k) = \phi_\mu^{\mathrm{T}}(k)\theta_0 + v(k), \tag{3}$$

where

$$\phi_\mu(k) \triangleq \begin{bmatrix} -y(k-\mu-1) \\ \vdots \\ -y(k-p) \\ u(k) \\ \vdots \\ u(k-p) \end{bmatrix}, \quad \theta_0 \triangleq \begin{bmatrix} \alpha \\ h_0 \\ \vdots \\ h_\mu \\ \beta \end{bmatrix}, \tag{4}$$

[1]This research was supported in part by the Air Force Office of Scientific Research under grant F49620-97-1-0406 and F49620-98-1-0037 and the University of Michigan Office of the Vice President for Research.

0-7803-4394-8/98 $10.00 © 1998 IEEE

and where $\alpha = [a_{\mu+1} \cdots a_p]$ and $\beta = [b_{\mu+1} \cdots b_p]$. Next assume that $G(q)$ and $H(q)$ are unknown, that $w(k)$ is a zero-mean stochastic sequence, and that measurements of the inputs and outputs of the system are available. We then define the *prediction* of $y(k)$ as

$$\hat{y}(k) \triangleq \mathbb{E}[y(k)] = \phi_\mu^{\mathrm{T}}(k)\theta + \mathbb{E}[v(k)] = \phi_\mu^{\mathrm{T}}(k)\theta. \quad (5)$$

We note that when $H(q)$ is unknown (5) is equivalent to the $k$-step ahead prediction of $y(k)$ for $k = \mu$ [1].

Next define the *prediction error*

$$\varepsilon(k) \triangleq y(k) - \hat{y}(k) = y(k) - \phi^{\mathrm{T}}(k)\theta, \quad (6)$$

and the least squares cost function

$$J_N(\theta) \triangleq \frac{1}{2}\sum_{k=\mu+p}^{N} \varepsilon^2(k) = \frac{1}{2}\sum_{k=\mu+p}^{N} [y(k) - \phi^{\mathrm{T}}(k)\theta]^2. \quad (7)$$

The parameter estimate $\hat{\theta}_N$ that minimizes $J_N(\theta)$ using data from $k = \mu + p$ through $k = N$, where $N \geq 2p + \mu + 1$ and $(\Phi_N^{\mathrm{T}}\Phi_N)^{-1}$ is invertible, is given by

$$\hat{\theta}_N = (\Phi_N^{\mathrm{T}}\Phi_N)^{-1}\Phi_N^{\mathrm{T}}Y_N \quad (8)$$

$$= \left[\sum_{k=\mu+p}^{N} \phi_\mu(k)\phi_\mu^{\mathrm{T}}(k)\right]^{-1}\left[\sum_{k=\mu+p}^{N} \phi_\mu(k)y(k)\right]. \quad (9)$$

If $\hat{\theta}_N \to \theta_0$ as $N \to \infty$, then the estimates of the parameters are *consistent*. The following theorem gives conditions for consistency of $\mu$-Markov parameterized models.

**Theorem 1.** Consider the $n$-th order system (2), and let $u(k)$ and $w(k)$ be realizations of independent, zero-mean, white noise stochastic processes. Furthermore, let $h_0, h_1, \ldots$ be the Markov parameters of $G(q)$ and consider the least squares estimate (9) using the $\mu$-Markov parameterization of order $p$ in (4). Then the estimate of the Markov parameters satisfy $\hat{h}_{iN} \to h_i$ as $N \to \infty$ for $i = 0, \ldots, \mu$ in both mean square and with probability one.

We note that Theorem 1 does not place any restrictions on the choice of $p$ in the $\mu$-Markov parameterization, and thus the least squares estimates of the Markov parameters are consistent *regardless of model order choice.*

## 4 Simulation

Data was obtained from $k = 0$ to $k = 600$ using the sampled-data system in (2), where

$$G(q) = \frac{0.04833q+0.04813}{q^2-1.641q+0.9881}, \quad H(q) = \frac{0.9063}{q-0.8187},$$

and where the input $u(k)$ and disturbance $w(k)$ were realizations of a zero-mean, white noise Gaussian random sequence with variances $\sigma_u^2 = 1$ and $\sigma_w^2 = 0.1$, respectively. Figures 1 and 2 show the estimates of the Markov parameters of $G(q)$ obtained by extracting the Markov parameters (+) from the least squares estimate of a $\mu$-Markov parameterization with $\mu = 40$ and $p = 1$ and $p = 2$. For comparison the true Markov parameters (○) and the Markov parameters (×) obtained from the impulse response of the least squares estimate of the full parameterization ($\mu = 0$) are also shown.

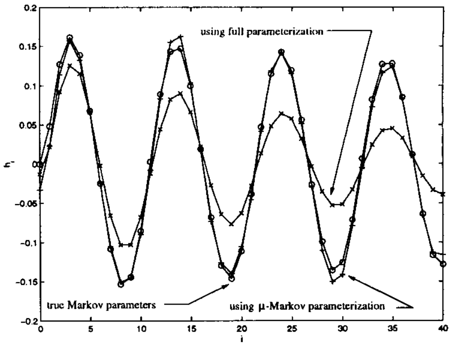


**Figure 1:** Markov parameter estimates with correct order ($p = 2$).

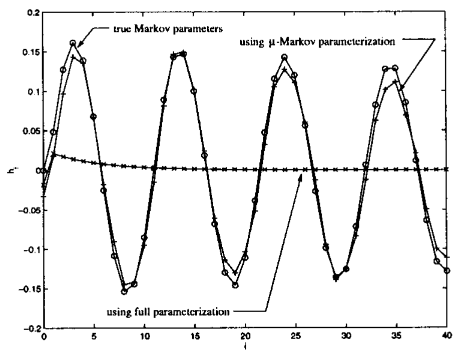


**Figure 2:** Markov parameter estimates with incorrect order ($p = 1$).

## 5 Conclusions

In this paper we have introduced $\mu$-Markov parameterizations which can be viewed as an IIR extension of FIR models in the sense that they explicitly incorporate Markov parameters in the parameterization. It was shown that these parameterizations yield consistent least squares estimates of Markov parameters without correct knowledge of the system order and for a general class of noise models.

## References


[1] K. J. Astrom and B. Wittenmark. *Adaptive Control.* Addison Wesley, second edition, 1995.

[2] T. Soderstrom and P. Stoica. *System Identification.* International Series in Systems and Control Engineering. Prentice Hall, 1989.

[3] J. Juang. *Applied System Identification.* Prentice–Hall, Englewood Cliffs, NJ, 1994.