2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Fast estimation of continuous-time ARX parameters from unevenly sampled data

Magnus Mossberg and Erik K. Larsson

***Abstract*— The problem of estimating the parameters in a continuous-time ARX (CARX) process from unevenly sampled data is studied. A solution where the differentiation operator is replaced by a difference operator is suggested. In the paper, results are given for how a difference operator is found so that consistent parameter estimates are obtained. The proposed method is considerably faster than conventional methods, such as the maximum likelihood method (MLM).**

I. Problem description

Consider the CARX model structure

$$A(p)y(t) = B(p)u(t) + e(t), \tag{1}$$

where $A(p) = p^n + a_1 p^{n-1} + \ldots + a_n$ and $B(p) = b_1 p^{n-1} + \ldots + b_n$, with p denoting the differentiation operator, and where $\mathrm{E}\{e(t)e(s)\} = \sigma_e^2 \delta(t-s)$. A CARX process can be very useful in at least three cases; when a control design is made in continuous-time, when it is important to have a model where the parameters have a physical meaning, and when data are sampled unevenly. The last case is perhaps the most important. In such a case, it might be difficult to find a discrete-time system description involving the conventional shift operator. Even if such a description can be found, it will be time-varying and the computational demands for the MLM will in general increase drastically [1]. Instead, an attractive alternative is to choose a continuous-time model in which the differentiation operator can be approximated with some difference operator. This alternative also provides good numerical properties for short sampling intervals.

The paper is devoted to the problem of estimating the CARX parameters

$$\boldsymbol{\theta}_0 = [a_1, \ldots, a_n, b_1, \ldots, b_n]^T, \tag{2}$$

where $\boldsymbol{\theta}_0$ denotes the true parameter vector, from irregularly sampled data $\{u(t_k)\}_{k=1}^N$ and $\{y(t_k)\}_{k=1}^N$ at the moments of observations $\{t_k\}_{k=1}^N$. The proposed solution is a direct method (DM) where the differentiation operator p^j in (1) is replaced by a difference operator D_k^j, to be specified later.

In order to certify a well-posed problem, some important assumptions must be introduced. First, denote the distance between two sampling instances by h_k, i.e., $h_k = t_{k+1} - t_k$, and assume that $\underline{h} \leq h_k \leq \bar{h}$, $\forall k$, where $\underline{h} > 0$ and $\bar{h}$ is sufficiently bounded. Moreover, assume that the involved processes are ergodic, that the input signal $u(t)$ is sufficiently differentiable, persistently exciting and quasi-stationary, and that the process operates in open loop. More details regarding these assumptions are found in [2].

M. Mossberg is with the Dept. of Electrical Eng., Karlstad Univ., SE-651 88 Karlstad, Sweden. E-mail: Magnus.Mossberg@kau.se
E. K. Larsson is with the Dept. of Systems and Control, Uppsala Univ., SE-751 05 Uppsala, Sweden. E-mail: ekl@it.uu.se

II. Proposed solution

Substitute the derivatives in (1) by approximations at time t_k as

$$p^j f(t_k) = D_k^j f(t_k) + \mathcal{O}(\bar{h}), \quad j = 0, \ldots, n, \tag{3}$$

where $f(t)$ is a smooth enough function. The discrete-time linear regression

$$w(t_k) = \boldsymbol{\varphi}^T(t_k)\boldsymbol{\theta} + \varepsilon(t_k), \tag{4}$$

$$w(t_k) = D_k^n y(t_k), \tag{5}$$

$$\boldsymbol{\varphi}^T(t_k) = [-D_k^{n-1} y(t_k), \ldots, -y(t_k), D_k^{n-1} u(t_k), \ldots, u(t_k)], \tag{6}$$

where $\boldsymbol{\theta}$ denotes the model parameters and $\varepsilon(t_k)$ is an equation error, can be constructed. The least squares method then gives the estimate

$$\hat{\boldsymbol{\theta}}_N = \left(\sum_{k=1}^N \boldsymbol{\varphi}(t_k)\boldsymbol{\varphi}^T(t_k) \right)^{-1} \left(\sum_{k=1}^N \boldsymbol{\varphi}(t_k) w(t_k) \right). \tag{7}$$

The following two important questions must be answered:

1) How are difference operators D_k^j, $j = 0, \ldots, n$, that fulfill (3) found?
2) Will the estimate (7) be consistent, i.e., will

$$\lim_{\bar{h} \to 0} \lim_{N \to \infty} \hat{\boldsymbol{\theta}}_N = \boldsymbol{\theta}_0 \tag{8}$$

hold for such difference operators?

The first question is answered by the following lemma:

Lemma 1: The difference operators that fulfill (3) can be generated by the recursion formula

$$D_k^j f(t_k) = \frac{1}{\tilde{h}_k(j)} \big(D_{k+1}^{j-1} f(t_{k+1}) - D_k^{j-1} f(t_k) \big), \tag{9}$$

$j = 1, \ldots, n$, where $\tilde{h}_k(j) = \frac{1}{j}\lambda_k(j)$, with

$$\lambda_k(j) = t_{k+j} - t_k = \begin{cases} 0, & j = 0, \\ \sum_{s=0}^{j-1} h_{k+s}, & j \geq 1. \end{cases} \tag{10}$$

Proof: See [1]. □

The second question is answered by the following lemma:

Lemma 2: An estimate

$$\hat{\boldsymbol{\theta}} \triangleq \lim_{N \to \infty} \hat{\boldsymbol{\theta}}_N = \boldsymbol{\theta}_0 + \boldsymbol{\mathcal{O}}(\bar{h}), \tag{11}$$

is obtained if the condition (3) is fulfilled and if

$$w(t_k) = D^n_{k+n-1} y(t_{k+n-1}), \qquad (12)$$

i.e., if a shift is introduced in the data when forming the nth order derivative approximation.

Proof: See [2]. □

The proposed estimation approach is summarized as follows:

1) Compute the approximate signal derivatives using Lemma 1.
2) Shift the data used when forming the nth order derivative approximation according to (12).
3) Construct the discrete-time linear regression (4)–(6).
4) Compute the least squares estimate (7).

III. Numerical studies

In order to study the proposed estimation approach numerically, data are generated from the second order CARX process $(p^2 + a_1 p + a_2)y(t) = (b_1 p + b_2)u(t) + e(t)$, where $a_1 = a_2 = 2$, $b_1 = 3$, $b_2 = 1$, and where the intensity σ_e^2 of the continuous-time white noise $e(t)$ is equal to one. The input signal $u(t)$ is chosen as the continuous-time AR process $(p^3 + c_1 p^2 + c_2 p + c_3)u(t) = v(t)$, where $c_1 = c_2 = c_3 = 2$, and where $v(t)$ is a continuous-time white noise, independent on $e(t)$, of intensity $\sigma_v^2 = 1$.

More exactly, a uniformly distributed sampling strategy is considered. Here, $t_n = nT + \sum_{k=1}^{n} \delta_k$, $n = 1, \ldots, N$, where δ_k is uniformly distributed between $-\rho_0$ and ρ_0; δ_k is independent of $e(t)$ for all t and k, and δ_k is independent of δ_j for all $j \neq k$. The choice $\rho_0 = T/5$ is made.

The DM proposed in the paper is considered in a Monte Carlo study with 50 realizations, in which the parameters a_1, a_2, b_1 and b_2 are estimated from $N = 1000$ data points, generated according to the sampling strategy described above.

The MLM [3] is also considered for estimating the parameters. It is well-known that estimates given by the MLM are the parameters that minimize the loss function

$$V_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{k=1}^{N} \Big\{ \frac{\varepsilon^2(t_k, \boldsymbol{\theta})}{r(k, \boldsymbol{\theta})} + \log\{r(k, \boldsymbol{\theta})\} \Big\}, \qquad (13)$$

where $\varepsilon(t_k, \boldsymbol{\theta}) = y(t_k) - \hat{y}(t_k | t_{k-1})$ is the prediction error, and where $r(k, \boldsymbol{\theta})$ is the variance of $\varepsilon(t_k, \boldsymbol{\theta})$. Here, the prediction errors are given by a Kalman filter. In the simulations, it is assumed that the input signal is constant between the sampling instants for the MLM.

The mean values and the empirical standard deviations for the estimates of a_1, a_2, b_1 and b_2 as functions of T, given by the DM proposed in the paper and the MLM are shown in Fig. 1 and Fig. 2, respectively. The biases are larger for the DM than for the MLM, especially for larger values of T, whereas the standard deviations are slightly smaller for the DM. However, the average computational times are considerably shorter for the DM, as seen in Table I.

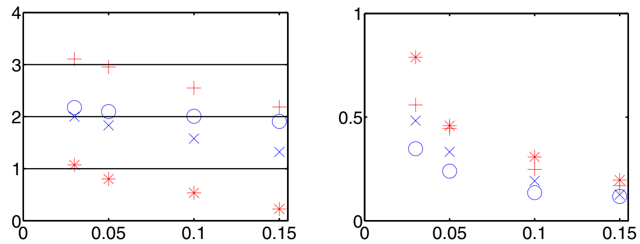

Fig. 1. The mean values (left) and the empirical standard deviations (right) for the estimates of a_1 (○), a_2 (×), b_1 (+) and b_2 (∗) as functions of T, for the DM proposed in the paper with uniformly distributed sampled data. The true parameter values (left) are indicated with horizontal lines.

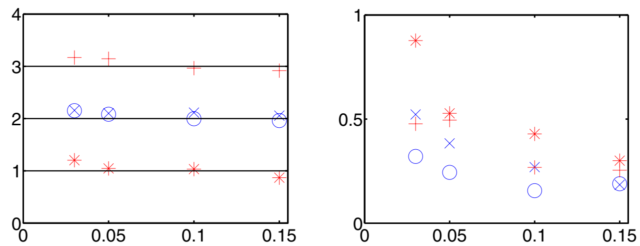

Fig. 2. The mean values (left) and the empirical standard deviations (right) for the estimates of a_1 (○), a_2 (×), b_1 (+) and b_2 (∗) as functions of T, for the MLM with uniformly distributed sampled data. The true parameter values (left) are indicated with horizontal lines.

TABLE I

The average computational times, in seconds, for the DM and the MLM for different T.

T	DM	MLM
0.03	0.099	16
0.05	0.084	25
0.10	0.073	29
0.15	0.074	30

IV. Conclusions

A direct approach for estimating the parameters in CARX processes from irregularly sampled data was suggested. The method is based on that the differentiation operator is replaced by a difference operator. Results were given that states how the difference operator should be found in order to get parameter estimates with a bias of $\mathcal{O}(\bar{h})$, where $\bar{h}$ is the upper bound for the distance between two sampling instants. Numerical studies showed that the method is very fast compared to the MLM. The method is an interesting choice when data are sampled quickly, and for giving initial values of high quality for the MLM.

References

[1] E. K. Larsson and T. Söderström, "Identification of continuous-time AR processes from unevenly sampled data," *Automatica*, vol. 38, no. 4, pp. 709–718, 2002.

[2] E. K. Larsson, M. Mossberg, and T. Söderström, "Identification of continuous-time ARX models from unevenly sampled data," Karlstad University, Karlstad, Sweden, Tech. Rep. EE-2004-04, 2004.

[3] L. Ljung, *System Identification*, 2nd ed. Upper Saddle River, NJ: Prentice–Hall, 1999.