REFERENCES

[1] M. O. Ahmad and J. D. Wang, "An analytical least square solution to the design problem of two-dimensional FIR filters with quadrantally symmetric or antisymmetric frequency response," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 968–979, 1989.
[2] S. C. Pei and J. J. Shyu, "Fast design of 2-D linear-phase complex FIR digital filters by analytical least squares methods," *IEEE Trans. Signal Processing*, vol. 44, pp. 3157–3161, 1996.
[3] W. H. Kwon and O. K. Kwon, "FIR filters and recursive forms for continuous time-invariant state-space models," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 352–356, 1987.
[4] W. H. Kwon, K. S. Lee, and O. K. Lee, "Optimal FIR filters for time-varying state-space models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 1011–1021, 1990.
[5] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME J. Basic Eng.*, vol. 83, pp. 95–108, 1961.
[6] A. H. Jazwinski, "Limited memory optimal filtering," *IEEE Trans. Automat. Contr.*, vol. 13, pp. 558–563, 1968.
[7] R. J. Fitzgerald, "Divergence of the Kalman filter," *IEEE Trans. Automat. Contr.*, vol. 16, pp. 736–747, 1971.
[8] A. M. Bruckstein and T. Kailath, "Recursive limited memory filtering and scattering theory," *IEEE Trans. Inform. Theory*, vol. 31, pp. 440–443, 1985.
[9] S. H. Park, W. H. Kwon, O. K. Kwon, and M. J. Kim, "Short-time Fourier analysis using optimal harmonic FIR filters," *IEEE Trans. Signal Processing*, vol. 45, pp. 1535–1542, 1997.
[10] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1787–1791, Sept. 1999.
[11] W. H. Kwon and A. E. Pearson, "A modified quadratic cost problem and feedback stabilization of a linear system," *IEEE Trans. Automat. Contr.*, vol. 22, pp. 838–842, 1977.
[12] W. H. Kwon, "Advances of predictive controls: theory and applications," in *Proc. 1st Asian Contr. Conf.*, Tokyo, Japan, Oct. 22–24, 1994.
[13] P. G. Kaminski, A. E. Bryson, and S. F. Schmidt, "Discrete square root filtering: A survey of current techniques," *IEEE Trans. Automat. Contr.*, vol. 16, pp. 727–735, 1971.
[14] B. D. O. Anderson and J. B. Moore, "New results in linear system stability," *Siam J. Contr.*, vol. 7, 1969.
[15] N. Minamide, M. Ohno, and H. Imanaka, "Recursive solutions to deadbeat observers for time-varying multivariable systems," *Int. J. Contr.*, vol. 45, 1987.
[16] K. J. Åström and B. Wittenmark, *Computer Controlled Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

A Note on Sampling and Parameter Estimation in Linear Stochastic Systems

T. E. Duncan, P. Mandl, and B. Pasik-Duncan

Abstract*— Numerical differentiation formulas that yield consistent least squares parameter estimates from sampled observations of linear, time invariant higher order systems have been introduced previously by Duncan *et al.* The formulas given by Duncan *et al.* have the same limiting system of equations as in the continuous time case. The formula presented in this note can be characterized as preserving asymptotically a partial integration rule. It leads to limiting equations for the parameter estimates that are different from the continuous case, but they again imply consistency. The numerical differentiation formulas given here can be used for an arbitrary linear system, which is not the case in the previous paper by Duncan *et al.

***Index Terms*—Estimation, linear stochastic systems, numerical differentiation for stochastic systems, sampling.**

I. INTRODUCTION

In a previous paper by the authors [2], the following parameter estimation problem in linear stochastic differential equations is considered. Let $(X(t),\, t \geq 0)$ be an n-dimensional process satisfying

$$dX^{(d-1)}(t) = \left(\sum_{i=1}^{d} f_i(\alpha)X^{(i-1)}(t) + g(\alpha)U(t)\right)dt + dW(t) \tag{1}$$

where

$$\begin{aligned} X^{(i)}(t) &= \frac{d}{dt}X^{(i-1)}(t), \qquad i = 1, 2, \cdots, d-1 \\ X^{(0)}(t) &= X(t) \end{aligned} \tag{2}$$

where $(U(t),\, t \geq 0)$ is a nonanticipative input process, $(W(t),\, t \geq 0)$ is an n-dimensional Wiener process with the local variance matrix h, $\alpha = (\alpha^1, \cdots, \alpha^p)$ is a p-dimensional unknown parameter

$$f_i(\alpha) = f_{i0} + \sum_{j=1}^{p} \alpha^j f_{ij}, \qquad g(\alpha) = g_0 + \sum_{j=1}^{p} \alpha^j g_j.$$

In these descriptions, $f_{ij},\, g_j$ for $i, j = 0, \cdots, p$ are known matrices. The true value of the unknown parameter is denoted by $\alpha_0 = (\alpha_0^1, \cdots, \alpha_0^p)$.

The least squares estimation of α_0 from the observation of

$$(X_t,\, U_t, \quad t \in [0, T])$$

is determined by minimizing the formal quadratic functional

$$\begin{aligned} \int_0^T \Bigg[& \left(X^{(d)} - \sum_{i=1}^{d} f_i(\alpha)X^{(i-1)} - g(\alpha)U\right)' \\ & \times L\left(X^{(d)} - \sum_{i=1}^{d} f_i(\alpha)X^{(i-1)} - g(\alpha)U\right) - X^{(d)'}LX^{(d)}\Bigg] dt \end{aligned} \tag{3}$$

Manuscript received January 26, 1998. Recommended by Associate Editor, J. C. Spall. This work was supported in part by the National Science Foundation under Grant DMS-9623439.

T. E. Duncan and B. Pasik-Duncan are with the Department of Mathematics, University of Kansas, Lawrence, KS 66045 USA (e-mail: duncan@math.ukans.edu).

P. Mandl is with the Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic.

Publisher Item Identifier S 0018-9286(99)08614-6.

where L is a positive semidefinite matrix. Prime denotes the transposition of vectors and matrices. The undefined term $X^{(d)'}LX^{(d)}$ is cancelled, and $X^{(d)}\,dt = dX^{(d-1)}$ in (3).

By minimizing (3), the following family of equations for the least squares estimate $\alpha^*(T) = (\alpha^{*1}(T), \cdots, \alpha^{*p}(T))$ of α_0 is obtained:

$$\begin{aligned}\sum_{k=1}^{p} \frac{1}{T}\int_0^T &\left(\sum_{i=1}^{d} f_{ij}X^{(i-1)} + g_jU\right)' \\ &\times L\left(\sum_{h=1}^{d} f_{hk}X^{(h-1)} + g_kU\right) dt\alpha^{*k}(T) \\ = \frac{1}{T}\int_0^T &\left(\sum_{i=1}^{d} f_{ij}X^{(i-1)} + g_jU\right)' \\ &\times L\left(dX^{(d-1)} - \left(\sum_{h=1}^{d} f_{h0}X^{(h-1)} + g_0U\right)dt\right) \\ &j = 1, \cdots, d. \end{aligned}\tag{4}$$

It is assumed in [2] that discrete observations of $(X(t), t \in [0, T])$ and $(U(t), t \in [0, T])$ with the uniform sampling interval $\delta > 0$ are only available yielding the observed random variables

$$\begin{aligned} X_{m,\delta} &= X(m\delta), \\ \overline{U}_{m,\delta} &= \overline{U}(m\delta), \qquad m = 0, \cdots, N+n. \end{aligned}\tag{5}$$

The notation $\overline{U}$ is used to include the case when the product $g(\alpha)U(t)$ depends only on some coordinates of U. It is assumed that all of these coordinates are observed.

To approximate (4) using only random variables in (5), a substitution for the derivatives $X^{(i)}(m\delta)$ by the forward differences

$$\begin{aligned} D^iX_{m,\delta} = \left(D^{i-1}X_{m+1,\delta} - D^{i-1}X_{m,\delta}\right)\Big/\delta, \\ i = 1, 2, \cdots, d-1 \end{aligned}\tag{6}$$

is performed and some numerical integration formulas are used to evaluate the integrals. Denoting by

$$\hat{\alpha}_{N\delta} = \left(\hat{\alpha}^1_{N\delta}, \cdots, \hat{\alpha}^p_{N\delta}\right)$$

the estimate so obtained by these substitutions it is desirable that the consistency property

$$\lim_{\delta\to 0}\lim_{N\to\infty} \hat{\alpha}_{N\delta} = \alpha_0 \tag{7}$$

is satisfied. In this expression $\lim_{N\to\infty}\hat{\alpha}_{N\delta}$ denotes the limit in probability, which under appropriate hypotheses is a nonrandom quantity.

In [4] and [6] it is noted that the forward and backward Euler approximations cannot be used, but it is shown how to modify the approximation for the highest derivative to satisfy (7). In [6] a specific approximation of the delta operator is given for the sampled data. It is shown in [1] for $n = 2$ and in [2] for $n \geq 2$ that (7) does not hold unless a correction term is introduced into the equations for $\hat{\alpha}_{N\delta}$ or unless (6) is modified. A numerical differentiation formula that determines estimates satisfying (7) is given. In this note, a different method is given for the numerical evaluation of (4) that satisfies (7). The method is based only on the random variables (5) and it employs together with (6) the backward differences

$$\begin{aligned} BX_{m,\delta} &= (X_{m-1,\delta} - X_{m,\delta})/\delta \\ B^iX_{m,\delta} &= \left(B^{i-1}X_{m-1,\delta} - B^{i-1}X_{m,\delta}\right)\Big/\delta. \end{aligned}\tag{8}$$

While the method in [2] estimated the difference from the case of continuous observations, the method presented here exploits the infinitesimal properties of the covariance function of the process. The method is more general in the sense that it applies to the case when the differentiation is subjected to white noise. This generalization has the following motivation.

It is known (see, e.g., [3]) that a stationary Gaussian process $(X(t), t \in \mathbb{R})$ with the spectral density

$$f(\lambda) = \frac{h}{|(i\lambda)^d - \alpha^d(i\lambda)^{d-1}\cdots - \alpha^1|^2}$$

can be represented as a solution of the stochastic differential equation

$$dX^{(d-1)} = \left(\alpha^1X^{(0)} + \alpha^2X^{(1)} + \cdots + \alpha^dX^{(d-1)}\right)dt + \sqrt{h}\,dW \tag{9}$$

where $(W(t), t \geq 0)$ is a standard Wiener process. Equation (9) is a particular case of (1). More generally, a Gaussian process with a rational spectral density

$$f(\lambda) = \frac{|b^d(i\lambda)^{d-1} + \cdots + b^2i\lambda + b^1|^2}{|(i\lambda)^d - \alpha^d(i\lambda)^{d-1} - \cdots - \alpha^1|^2}$$

satisfies (9) where $X^{(i)}(t)$, $i = 0, \cdots, d-1$, denote the random processes satisfying

$$dX^{(i-1)} = X^{(i)}\,dt + \beta_i\,dW, \qquad i = 1, \cdots, d-1$$

with

$$\beta_1 = b^d, \qquad \beta_i = b^{d+1-i} + \sum_{j=1}^{i-1}\beta_j\alpha^{d+1-j+i}.$$

Thus, it is important to consider the following generalization of (1), (2)

$$dX^{(i-1)}(t) = X^{(i)}(t)\,dt + dW^{(i)}(t), \qquad i = 1, \cdots, d-1 \tag{10}$$

where $((W^{(1)}(t), \cdots, W^{(d-1)}(t), W(t))', t \geq 0)$ is a $d \cdot n$-dimensional Wiener process with local variance matrix H_0.

II. Parameter Estimation

Assume that (1) and (10) are satisfied with $\alpha = \alpha_0$, and let the q-dimensional process $(U(t), t \geq 0)$ be the solution of the linear stochastic differential equation

$$dU(t) = cU(t)\,dt + dW_0(t), \qquad U(0) = U_0$$

where c is a constant matrix, and $(W_0(t), t \geq 0)$ is a q-dimensional Wiener process with local variance matrix h_0 that is independent of $(W(t), W^{(1)}(t), \cdots, W^{(d-1)}(t), t \geq 0)$.

To describe the evolution of the entire model, introduce the state vector $\mathbb{X}(t) \in \mathbb{R}^{dn+q}$ and the matrices F, H that are described in block form according to the partition of $\mathbb{X}$

$$\begin{aligned} \mathbb{X}(t) &= \begin{pmatrix} X^{(0)}(t) \\ \vdots \\ X^{(d-1)}(t) \\ U(t) \end{pmatrix} \\ F &= \begin{pmatrix} 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & I & 0 \\ f_1 & f_2 & f_3 & \cdots & f_d & g \\ 0 & 0 & 0 & \cdots & 0 & c \end{pmatrix} \\ H &= \begin{pmatrix} H_0 & 0 \\ 0 & h_0 \end{pmatrix} \end{aligned}\tag{11}$$

where I denotes the identity matrix in $\mathbb{R}^n$, and

$$f_i = f_i(\alpha_0), \qquad i = 1, \cdots, d, \qquad g = g(\alpha_0).$$

It easily follows that

$$d\mathbb{X}(t) = F\mathbb{X}(t)\,dt + d\mathbb{W}(t), \qquad X(0) = X_0$$

where $(\mathbb{W}(t), t \geq 0)$ is a $d \cdot n + q$-dimensional Wiener process with local variance matrix H.

The following assumption is made.

Assumption 1: F is a stable matrix.

This assumption implies that $\mathbb{X}(t)$ has a limiting Gaussian distribution as $t \to \infty$ with zero mean and variance matrix $\mathbb{R}$, which is the solution of the Lyapunov equation

$$F\mathbb{R} + \mathbb{R}F' + H = 0. \tag{12}$$

The partition of $\mathbb{R}$ into the blocks r_{ij} as

$$\mathbb{R} = (r_{ij}) \tag{13}$$

as introduced in (11) is used.

The matrix on the left-hand side of the system of equations (4) that acts on $\alpha^*(T)$ can be written as

$$\left(\frac{1}{T}\int_0^T \mathbb{X}' F_j' L F_k \mathbb{X}\,dt\right)$$

for $j, k = 1, \cdots, p$ where

$$F_j = (f_{1j}, \cdots, f_{dj}, g_j), \qquad j = 1, \cdots, p.$$

By Assumption 1, this family of matrices indexed by $T > 0$ converges (in quadratic mean) as $T \to \infty$ to the matrix

$$Q = (\mathrm{tr}(F_j' L F_k \mathbb{R})) \tag{14}$$

for $j, k = 1, \cdots, p$ where $\mathrm{tr}(\ )$ denotes the trace operator.

Assumption 1 and the assumption of nonsingularity of Q guarantee the consistency of the family $(\alpha^*(T), T > 0)$ as $T \to \infty$, that is,

$$\lim_{T\to\infty} \alpha^*(T) = \alpha_0 \qquad \text{a.s.}$$

The discrete approximation of (4) presented here leads in the limit to a different system of equations. The associated matrix has a form similar to (14) with the matrix $\mathbb{R}$ in (12) replaced by a matrix S which is defined subsequently. While the methods introduced in [2] estimated the difference from the case of the continuous observations, the method presented here exploits the infinitesimal properties of the covariance function.

For $\delta > 0$ define

$$R(\delta) = \lim_{t\to\infty} EX(t+\delta)X(t)' \tag{15}$$

and thus

$$R(-\delta) = R(\delta)' = \lim_{t\to\infty} EX(t-\delta)X(t)'.$$

Lemma 1: Let Assumption 1 be satisfied. For $\delta > 0$ the following equality is satisfied:

$$R(\delta) = \sum_{k=0}^{\infty} \frac{\delta^k}{k!} R_k \tag{16}$$

where

$$R_k = r_{k+1,1}, \qquad k = 0, 1, \cdots, d-1 \tag{17}$$

$$\begin{aligned} R_k = {} & f_1 R_{k-d} + f_2 R_{k+1-d} + \cdots + f_d R_{d+k-1-d} \\ & + (-1)^k g r_{d+1,k+1-d}, \qquad k = d, d+1, \cdots, 2d-1. \end{aligned} \tag{18}$$

Proof: From the theory of linear stochastic systems, it follows that for $\delta > 0$

$$R(\delta) = \left(\sum_{k=0}^{\infty} \frac{\delta^k}{k!} F^k \mathbb{R}\right)_{11} \tag{19}$$

where $(A)_{ij}$ denotes the block a_{ij} of the matrix A using the partitioning in (11). Consequently,

$$R_k = (F^k\mathbb{R})_{11}, \qquad k = 0, 1, \cdots. \tag{20}$$

Obviously, $R_0 = r_{11}$, and performing successive multiplications on $\mathbb{R}$ by F, it follows that

$$R_1 = r_{21}, \cdots, R_{d-1} = r_{d1}.$$

Furthermore, using (18), it follows that

$$\left(F^{j+1}\mathbb{R}\right)_{d1} = \left(F^{d-1}F^{j+1}\mathbb{R}\right)_{11} = R_{d+j}. \tag{21}$$

Hence,

$$F\mathbb{R} = \begin{pmatrix} R_1 & \cdots & \\ \vdots & & \\ R_d & \cdots & \\ cr_{d+1,1} & \cdots & cr_{d+1,d+1} \end{pmatrix}.$$

This equality yields (18) for $k = 0$.

By the Lyapunov equation (12)

$$cr_{d+1,j} = -r'_{j+1,d+1} = -r_{d+1,j+1}, \qquad j = 1, \cdots, d+1.$$

Consequently,

$$c^k r_{d+1,1} = (-1)^k r_{d+1,k+1}, \qquad k = 1, \cdots, d-1. \tag{22}$$

Using (21) and (22), it follows that

$$\begin{aligned} R_{d+k} &= (F(F^k\mathbb{R}))_{d1} \\ &= \left(F\begin{pmatrix} R_k & \cdots \\ \vdots & \\ c^k r_{d+1,1} & \cdots \end{pmatrix}\right)_{d1} \\ &= f_1 R_k + \cdots + f_d R_{d+k-1} + g c^k r_{d+1,1}. \end{aligned}$$

This equality and (22) yield (18) for $k = 1, \cdots, d-1$. □

Consider next the observed random variables (5), and note that it follows from (6) and (8), respectively, that

$$D^r X_{m,\delta} = \frac{1}{\delta^r} \sum_{k=0}^{r} (-1)^k \binom{r}{k} X_{m+r-k,\delta} \tag{23}$$

$$B^q X_{m,\delta} = \frac{1}{\delta^q} \sum_{k=0}^{q} (-1)^k \binom{q}{k} X_{m-q+k,\delta}. \tag{24}$$

Lemma 2: Let Assumption 1 be satisfied. For $r, s = 0, 1, \cdots,$ the following equality is satisfied:

$$\lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=s}^{N-r} D^r X_{m,\delta} B^s X'_{m,\delta} = R_{r+s}. \tag{25}$$

Proof: Using the Law of Large Numbers and (15), (19) it follows that, for $j, k = 0, 1, \cdots$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=k}^{N-j} X_{m+j,\delta} X'_{m-k,\delta} = \left(e^{(j+k)\delta F}\mathbb{R}\right)_{11}.$$

From (20), (23), and (24), it follows that

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{m=s}^{N-r} D^r X_{m,\delta} B^s X'_{m,\delta} &= \frac{1}{\delta^{r+s}}\left(\left(e^{\delta F}-I\right)^r \left(e^{\delta F}-I\right)^s \mathbb{R}\right)_{11}\\ &= (F^{r+s}\mathbb{R})_{11} + O(\delta)\\ &= R_{r+s} + O(\delta).\end{aligned}$$

From this equality, (22) follows. □

Recall that $\overline{U}(t)$ is formed by those coordinates of $U(t)$ for which $g(\alpha)U(t)$ depends. Further denote

$$\begin{aligned} r_{i\overline{u}} &= \lim_{t\to\infty} EX^{(i-1)}(t)\overline{U}(t)', \qquad i = 1, \cdots, d\\ r_{u\overline{u}} &= \lim_{t\to\infty} EU(t)\overline{U}(t)',\\ r_{\overline{u}\overline{u}} &= \lim_{t\to\infty} E\overline{U}(t)\overline{U}(t)'.\end{aligned}$$

Note that $r_{\overline{u}i} = r'_{i\overline{u}}$, $r_{\overline{u}u} = r'_{u\overline{u}}$.

Lemma 3: Let Assumption 1 be satisfied. Let $X(t)$ be r-times differentiable and $\overline{U}(t)$ be s-times differentiable (in quadratic mean), and let $r + s \geq d - 1$. Then

$$\lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1} D^i X_{m,\delta}\overline{U}'_{m,\delta} = r_{i+1,\overline{u}}, \qquad i = 0, \cdots, d-1 \tag{26}$$

$$\lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1} D^d X_{m,\delta}\overline{U}'_{m,\delta} = f_1 r_{1\overline{u}} + \cdots + f_d r_{d\overline{u}} + g r_{\overline{u}\overline{u}}. \tag{27}$$

Proof: Without loss of generality, it can be assumed that the coordinates of $U(t)$ are described in the following way:

$$U(t) = \left(\overline{U}(t), \overline{U}^{(1)}(t), \cdots, \overline{U}^{(s)}(t), U^*(t)\right) \tag{28}$$

where $\overline{U}^{(i)}(t)$ denotes the derivative of order i of $\overline{U}(t)$, and $U^*(t)$ contains the rest of the coordinates. Then the matrices c, h_0 have the form

$$\begin{aligned} c &= \begin{pmatrix} 0 & I & 0 & \cdots & 0\\ 0 & 0 & I & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ & & c^* & & \end{pmatrix}\\ h_0 &= \begin{pmatrix} 0 & 0\\ 0 & h^*\end{pmatrix}. \end{aligned} \tag{29}$$

From the partition of the matrices $\mathbb{R}$, F, etc., into block matrices a finer partition is also employed, with indices $i, j = d+1, \cdots, d+s+2$ referring to the components of (28). In particular, $r_{i\overline{u}} = r_{id+1}$, $i = 1, \cdots, d$.

Consider (26). Let $v \leq s$ be such that $i - v \geq r$. Using partial summation if $v > 0$ it follows that

$$\begin{aligned}\lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1} D^i X_{m,\delta}\overline{U}'_{m,\delta} &= \lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1} D^{i-v} X_{m,\delta} B^v \overline{U}'_{m,\delta}\\ &= (-1)^v (F^{i-v}\mathbb{R})_{1,d+1+v} = (-1)^v r_{i-v+1,d+1+v}. \end{aligned} \tag{30}$$

From (29) and (12) it follows that

$$\begin{aligned} r_{j+1,i} &= (F\mathbb{R})_{ji}\\ &= -r_{j,i+1}, \qquad i = 1, \cdots, d-1;\ j = d+1, \cdots, d+q \end{aligned}$$

so that,

$$(-1)^v r_{i-v+1,d+1+v} = r_{i+1,d+1} = r_{i+1,\overline{u}}$$

which yields (26).

Similarly, to verify (27), it follows as in (30) that

$$\begin{aligned}\lim_{\delta\to 0}\lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1} D^d X_{m,\delta}\overline{U}'_{m,\delta} &= (-1)^v (F^{d-v}\mathbb{R})_{1,d+1+v} = (-1)^v (F\mathbb{R})_{d-v,d+1+v}.\end{aligned}$$

For $v = 0$, the last term in the equality coincides with the right-hand side of (26). For $v > 0$, it follows from (12) that

$$\begin{aligned} (F\mathbb{R})_{d-v+1,d+v} &= -(F\mathbb{R})'_{d+v,d-v+1}\\ &= -r'_{d+v+1,d-v+1}\\ &= -r_{d-v+1,d+v+1}\\ &= -(F\mathbb{R})_{d-v,d+v+1}. \end{aligned}$$

Repeating this argument, it follows that

$$(-1)^v (F\mathbb{R})_{d-v,d+1+v} = (F\mathbb{R})_{d,d+1}$$

which establishes (27). □

Now the discrete observation version of (4) is introduced by letting the estimate $\hat{\alpha}_{N\delta}$ of α_0 be the solution of

$$\begin{aligned} &\sum_{k=1}^{p}\frac{1}{N}\sum_{m=0}^{N-1}\left(\sum_{i=1}^{d} f_{ij}B^{i-1}X_{m,\delta} + g_j U_{m,\delta}\right)'\\ &\quad\times L\left(\sum_{h=1}^{d} f_{hk}D^{h-1}X_{m,\delta} + g_k U_{m,\delta}\right)\hat{\alpha}^k_{N\delta}\\ &= \frac{1}{N}\sum_{m=0}^{N-1}\left(\sum_{i=1}^{d} f_{ij}B^{i-1}X_{m,\delta} + g_j U_{m,\delta}\right)'\\ &\quad\times L\left(D^d X_{m,\delta} - \sum_{h=1}^{d} f_{h0}D^{h-1}X_{m,\delta} - g_0 U_{m,\delta}\right) \end{aligned} \tag{31}$$

for $j = 1, \cdots, d$. To guarantee the consistency of $\hat{\alpha}_{N\delta}$ as $N \to \infty$ and $\delta \to 0$, a matrix $\hat{Q}$ that is analogous to Q in (14) is introduced by replacing $\mathbb{R}$ by a matrix S. Let $R_0, \cdots, R_{2d-1}, r_{d+1,1}, \cdots, r_{d+1,d}$ be given by (13), (17), and (18). Define

$$S := \begin{pmatrix} R_0 & R_1 & \cdots & R_{d-1} & r_{1,d+1}\\ R_1 & R_2 & \cdots & R_d & r_{2,d+1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ R_{d-1} & R_d & \cdots & R_{2d-1} & r_{d,d+1}\\ r_{d+1,1} & r_{d+1,2} & \cdots & r_{d+1,d} & r_{d+1,d+1}\end{pmatrix}. \tag{32}$$

The matrix $\hat{Q}$ is given in the following assumption.

Assumption 2: The matrix

$$\hat{Q} := (\mathrm{tr}(F'_j L F_k S)), \qquad j, k = 1, \cdots, p \tag{33}$$

is nonsingular.

Theorem 1: Let Assumptions 1 and 2 be satisfied, and let $\hat{\alpha}_{N\delta}$ be the solution of (31). Let $(X(t), t \geq 0)$ be r-times differentiable and $(\overline{U}(t), t \geq 0)$ be s-times differentiable (in quadratic mean), and let $r + s \geq d - 1$. Then

$$\lim_{\delta\to 0}\lim_{N\to\infty}\hat{\alpha}_{N\delta} = \alpha_0. \tag{34}$$

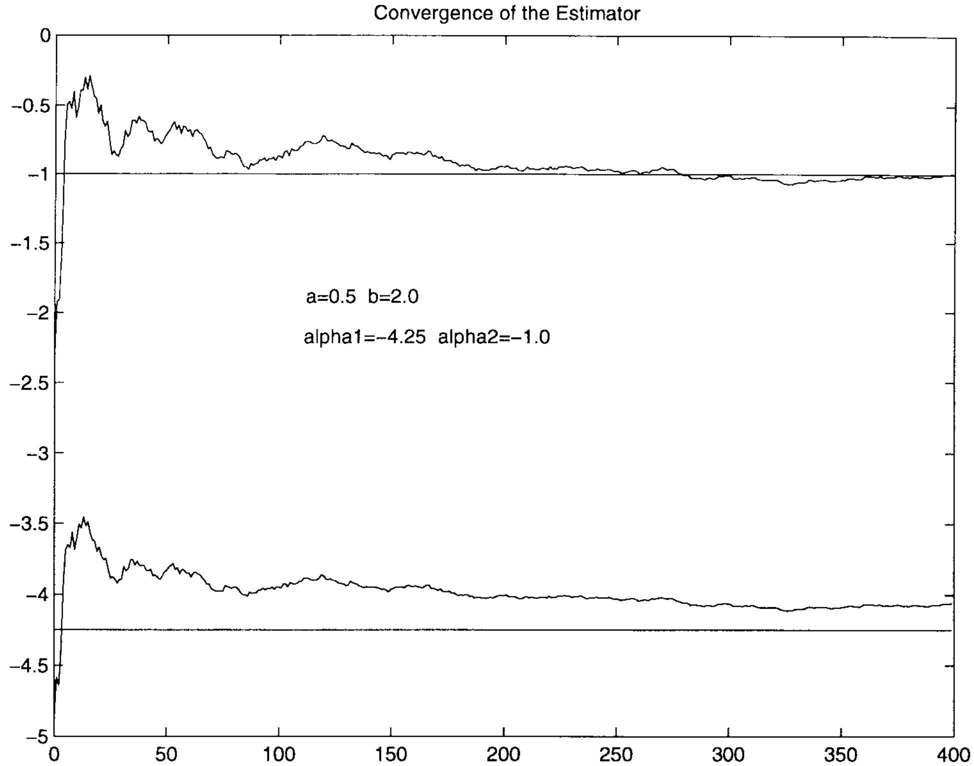

Fig. 1 Convergence of the estimator.

Proof: It is sufficient to prove that the system of equations, which is obtained by performing the two passages to the limit, that is, $N \to \infty$ and $\delta \to 0$ in (31), has the unique solution α_0.

If $N \to \infty$ and $\delta \to 0$ in (31), then the following family of linear equations for the estimate $\hat{\alpha}_{\infty,0}$ is obtained by Lemmas 2 and 3.

$$\begin{aligned}\sum_{k=1}^{p} \operatorname{tr}\Bigg(L\Bigg(&\sum_{h,i} f_{hk} R_{h+i-2} f'_{ij} + \sum_{h} f_{hk} r_{h\overline{u}} g'_j \\ &+ \sum_{i} g_k r_{\overline{u}i} f'_{ij} + g_k r_{\overline{u}\overline{u}} g'_j\Bigg)\Bigg)\hat{\alpha}^{k}_{\infty 0}\\ = \operatorname{tr}\Bigg(L\Bigg(&\sum_{i} R_{d+i-1} f'_{ij} + (f_1 r_{1\overline{u}} + \cdots + f_d r_{d\overline{u}} + g_k r_{\overline{u}\overline{u}}) g'_j \\ &- \sum_{i} f_{h0} R_{h+i-2} f'_{ij} - \sum_{i} g_k r_{\overline{u}i} f'_{ij} - \sum_{i} g_0 r_{\overline{u}i} f'_{ij}\Bigg)\Bigg),\\ &j = 1, \cdots, p. \end{aligned} \tag{35}$$

It is seen that $\hat{Q}$ is the matrix of the left hand side for the system (35) so this system of equations has a unique solution by Assumption 2.

It only remains to verify that (35) is satisfied with

$$\hat{\alpha}^{1}_{\infty 0} = \alpha^{1}_{0}, \cdots, \hat{\alpha}^{p}_{\infty 0} = \alpha^{p}_{0}. \tag{36}$$

Inserting the values in (36) into the left-hand side of (35) and recalling that $f_h = f_h(\alpha_0)$, $g = g(\alpha_0)$, the left-hand side of (35) is

$$\begin{aligned}\operatorname{tr}\Bigg(L\Bigg[&\sum_{h,i} f_h R_{h+i-1} f'_{ij} + \sum_{i} g r_{\overline{u}i} f_{ij} \\ &+ \sum_{h} f_h r_{h\overline{u}} g'_j + g r_{u\overline{u}} g'_j\Bigg]\Bigg), \qquad j = 1, \cdots, p. \end{aligned} \tag{37}$$

By Lemma 1, the first two sums in the square bracket equals

$$\sum_{i} R_{d+i-1} f'_{ij}.$$

Consequently, (37) is equal to the right-hand side of (35). □

The quadratic mean differentiability of a process is determined by the differentiability of its covariance function

Example: Let $(X(t), t \in \mathbb{R})$ be a stationary Gaussian process with the covariance function

$$R(\delta) = R_0 e^{-a|\delta|} \cos b\delta \tag{38}$$

where a, b are positive, unknown constants. The spectral density corresponding to (38) is

$$f(\lambda) = 2R_0 a \frac{|i\lambda + \sqrt{a^2+b^2}|^2}{|(i\lambda)^2 + 2ai\lambda + a^2 + b^2|^2}. \tag{39}$$

Furthermore, it follows from (38) or (39) that

$$\begin{aligned} dX &= X^{(1)}\, dt + \beta_1\, dW \\ dX^{(1)} &= \alpha_0^1 X\, dt + \alpha_0^2 X^{(1)}\, dt + \beta_2\, dW \end{aligned}$$

where

$$\alpha_0^1 = -(a^2+b^2), \qquad \alpha_0^2 = -2a \tag{40}$$

$$\beta_1 = \sqrt{2R_0 a}, \qquad \beta_2 = \sqrt{2R_0 a}(\sqrt{a^2+b^2} - 2a). \tag{41}$$

The parameter $\alpha_0 = (\alpha_0^1, \alpha_0^2)$ is estimated by (4) and (35) and β_1^2 is estimated by the quadratic variation of $(X(t), t \geq 0)$. Estimates of a, b and R_0 are obtained from (40) and (41). Expanding $R(\delta)$ in terms of δ it follows that

$$R(\delta) = R_0\big(1 - a|\delta| + \tfrac{1}{2}(a^2-b^2)\delta^2 + \tfrac{1}{6}(3ab^2 - a^3)|\delta|^3 + \cdots\big).$$

Letting $L = I$ in (3), it is expressed as

$$\int_0^T \left[\left(X^{(2)} - \alpha^1 X - \alpha^2 X^{(1)}\right)^2 - \left(X^{(2)}\right)^2\right] dt$$

so (4) is

$$\frac{1}{T}\int_0^T (X)^2\,dt\,\alpha^{*1}(T)+\frac{1}{T}\int_0^T XX^{(1)}\,dt\,\alpha^{*2}(T) = \frac{1}{T}\int_0^T X\,dX^{(1)} \tag{42}$$

$$\frac{1}{T}\int_0^T XX^{(1)}\,dt\,\alpha^{*1}(T)+\frac{1}{T}\int_0^T \left(X^{(1)}\right)^2 dt\,\alpha^{*2}(T) = \frac{1}{T}\int_0^T X^{(1)}\,dX^{(1)}. \tag{43}$$

In the limit as $T \to \infty$, there are the two equations

$$R_0\alpha^{*1}(\infty) - aR_0\alpha^{*2}(\infty) = R_0\alpha_0^1 - aR_0\alpha_0^2 \tag{44}$$

$$aR_0\alpha^{*1}(\infty) + r_{22}\alpha^{*2}(\infty) = \alpha_0^1 aR_0 + \alpha_0^2 r_{22}. \tag{45}$$

For the discretized version (31) of (42), (43), the limit as $N \to \infty$ and $\delta \to 0$ is different from (44), (45), specifically

$$R_0\hat{\alpha}^1(\infty) - aR_0\hat{\alpha}^2(\infty) = R_0(a^2-b^2) \tag{46}$$

$$-aR_0\hat{\alpha}^1(\infty) + R_0(a^2-b^2)\hat{\alpha}^2(\infty) = R_0(3ab^2-a^3). \tag{47}$$

The solution is $\hat{\alpha}_{\infty 0} = \alpha_0$.

The quadratic variation can be used to estimate β_1^2 by

$$\hat{\beta}_1^2 = \frac{1}{N\delta}\sum_m (X_{m+1,\delta} - X_{m,\delta})^2.$$

A numerical example is described graphically (see Fig. 1) where $a = 0.5$, $b = 2.0$, and $R_0 = 1$ so that $\alpha_1 = -4.25$ and $\alpha_2 = -1.0$. The convergence of family of estimates is relatively fast. If a longer time interval is used, then the family of estimates for α_1 are closer to the true value.

Acknowledgment

The authors thank Y. Yan for performing the computations for the numerical example. The authors also thank the referees for useful comments that improved the note.

References

[1] T. E. Duncan, P. Mandl, and B. Pasik-Duncan, "On statistical sampling for system testing," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 118–122, 1994.

[2] ——, "Numerical differentiation and parameter estimation in higher order stochastic systems," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 522–532, 1996.

[3] R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes, Vol. 2: Applications*. New York: Springer-Verlag, 1978.

[4] T. Söderström, H. Fan, S. Begi, and B. Carlsson, "Can a least-squares fit be feasible for modeling continuous-time autoregressive processes from discrete-time data?," in *Proc. 34th IEEE Conf. Decision Contr.*, New Orleans, 1995, pp. 1795–1800.

[5] T. Söderström, H. Fan, B. Carlsson, and S. Begi, "Least squares parameter estimation of continuous-time ARX models from discrete-time data," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 659–673, 1997.

[6] T. Söderström, H. Fan, B. Carlsson, and M. Mossberg, "Some approaches on how to use the delta operator when identifying continuous-time processes," in *Proc. 36th IEEE Conf. Decision Contr.*, San Diego, CA, 1997, pp. 890–895.

Stochastic Control of Discrete Systems: A Separation Principle for Wiener and Polynomial Systems

M. J. Grimble

***Abstract*—A new separation principle is established for systems represented in discrete frequency-domain Wiener or polynomial forms. The LQG or H_2 optimal controller can be realized using an observer based structure estimating noise free output variables that are fed back through a dynamic gain control block. Surprisingly, there are also two separation principle theorems, depending upon the order in which the *ideal output optimal control* and the *optimal observer* problems are solved.**

***Index Terms*—Optimal control, polynomial systems, Wiener theory.**

I. Introduction

The separation principle of stochastic optimal control theory has often been utilized for systems represented in state equation form. However, no such results have been established for systems represented in transfer-function or polynomial matrix form. The frequency domain approach to optimal control and estimation was initiated by Wiener [1], but two seminal contributions later established the main tools for synthesis. These contributions were undertaken in the same period by Youla *et al.* [2] and by Kucera [3].

The separation principle that is well known in state-space LQG synthesis was not used in the frequency-domain solutions, although Kucera [4] provided independent solutions of the LQ state feedback control and the *Kalman filtering* problems. Thus, in this case, if the polynomial models are related back to a system described in state equation form, it is possible to use the polynomial solutions to calculate the constant control and filter gains. The state-space separation principle results can then be invoked to obtain the LQG output feedback controller. The *separation principle* was not, however, established in the polynomial setting. Moreover, there was no attempt to generalize the results to the case where the control law feedback included a reduced set of variables, such as plant output estimates. The objective of the analysis that follows is to use frequency domain models and analysis, to establish a new separation principle result for systems represented in frequency domain matrix fraction form.

II. Polynomial System Description

The linear time-invariant discrete-time multivariable, finite-dimensional system of interest is illustrated in Fig. 1. The noise free system output sequence is denoted by $\{y(t)\}$, where $y(t) \in R^r$, and the observations signal is denoted by $\{z(t)\}$. The white driving noise signals $\{\xi(t)\}$ and $\{v(t)\}$ represent the disturbance, and measurement noise signals, respectively. These signals are statistically independent and the covariance matrices $(R_f > 0)$:

$$\operatorname{cov}[\xi(t),\,\xi(\tau)] = I_q\delta_{t\tau} \quad \text{and} \quad \operatorname{cov}[v(t),\,v(\tau)] = R_f\delta_{t\tau}. \tag{1}$$

Manuscript received January 22, 1999. Recommended by Associate Editor, G. Gu. This work was supported by the Engineering and Physical Sciences Research.

The author is with the Industrial Control Centre, University of Strathclyde, Glasgow G1 1QE, U.K. (e-mail: m.grimble@eee.strath.ac.uk).

Publisher Item Identifier S 0018-9286(99)08615-8.