On the Equivalence of OKID and Time Series Identification for Markov-Parameter Estimation

P. V. Albuquerque[*],
M. Holzel[†],
and D. S. Bernstein[‡]

April 5, 2009

Abstract

We show the equivalence of Observer/Kalman Identification (OKID) and Time Series Identification (TSI) for Markov parameters.

[*]Ph.D. Student, University of Michigan, Ann Arbor, MI
[†]Ph.D. Student, University of Michigan, Ann Arbor, MI
[‡]Professor, Aerospace Engineering Department, University of Michigan, Ann Arbor, MI, 48109

1 Introduction

Introduce paper here.

2 Markov Parameter Estimation using OKID

Observer/Kalman System Identification (OKID) involves adding an observer into an observable discrete-time system in order to help determine the Markov parameters. The method for doing so, as seen in [2], is clearly derived in this section.

2.1 Discrete-Time Linear System

Consider the linear discrete-time system

$$x(k+1) = Ax(k) + Bu(k), \tag{1}$$

$$y(k) = Cx(k) + Du(k), \tag{2}$$

where $k \geq 0$, $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^p$, and $u(k) \in \mathbb{R}^m$, and A, B, C, and D are real matrices of corresponding sizes. We assume that (A, C) is observable.

For nonnegative integers k and r we have

$$x(k+r) = A^k x(r) + \sum_{i=1}^{k} A^{i-1} Bu(k+r-i), \tag{3}$$

$$y(k+r) = CA^k x(r) + \sum_{i=1}^{k} CA^{i-1} Bu(k+r-i) + Du(k+r), \tag{4}$$

where $\sum\limits_{i=1}^{0} = 0$. Now let $s \geq r$. Then, lining up (4) for $k = 0, \ldots, s-r$ yields

$$Y_{[r:s]} = G_{[0:s-r]} X_{r,s} + H_{[0:s-r]} U_{r,s}, \tag{5}$$

where

$$\begin{aligned}
Y_{[r:s]} &\triangleq \begin{bmatrix} y(r) & y(r+1) & y(r+2) & \cdots & y(s) \end{bmatrix} \in \mathbb{R}^{p \times (s-r+1)}, \\
G_{[0:s-r]} &\triangleq \begin{bmatrix} C & CA & CA^2 & \cdots & CA^{s-r} \end{bmatrix} \in \mathbb{R}^{p \times n(s-r+1)}, \\
H_{[0:s-r]} &\triangleq \begin{bmatrix} H_0 & H_1 & H_2 & \cdots & H_{s-r} \end{bmatrix} \\
&\triangleq \begin{bmatrix} D & CB & CAB & \cdots & CA^{s-r-1}B \end{bmatrix} \in \mathbb{R}^{p \times m(s-r+1)},
\end{aligned}$$

$$X_{r,s} \triangleq \begin{bmatrix} x(r) & 0 & \cdots & 0 \\ 0 & x(r) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x(r) \end{bmatrix} \in \mathbb{R}^{n(s-r+1)\times(s-r+1)},$$

$$U_{r,s} \triangleq \begin{bmatrix} u(r) & u(r+1) & \cdots & u(s) \\ 0 & u(r) & \cdots & u(s-1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u(r) \end{bmatrix} \in \mathbb{R}^{m(s-r+1)\times(s-r+1)}.$$

2.2 Adding an Observer to the System

Adding and subtracting $Fy(k)$, where $F \in \mathbb{R}^{n\times p}$, to the right hand side of (1) yields

$$x(k+1) = Ax(k) + Bu(k) + Fy(k) - Fy(k) \tag{6}$$

$$= (A+FC)x(k) + (B+FD)u(k) - Fy(k), \tag{7}$$

which, with (2), can be written as

$$x(k+1) = \mathcal{A}x(k) + \mathcal{B}v(k), \tag{8}$$

$$y(k) = Cx(k) + \mathcal{D}v(k), \tag{9}$$

where

$$\mathcal{A} \triangleq A + FC \in \mathbb{R}^{n\times n},$$

$$\mathcal{B} \triangleq [B+FD \quad -F] \in \mathbb{R}^{n\times(m+p)},$$

$$\mathcal{D} \triangleq [D \quad 0] \in \mathbb{R}^{p\times(m+p)},$$

$$v(k) \triangleq \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \in \mathbb{R}^{m+p}.$$

For nonnegative integers k and r we have

$$x(k+r) = \mathcal{A}^k x(r) + \sum_{i=1}^{k} \mathcal{A}^{i-1}\mathcal{B}v(k+r-i), \tag{10}$$

$$y(k+r) = C\mathcal{A}^k x(r) + \sum_{i=1}^{k} C\mathcal{A}^{i-1}\mathcal{B}v(k+r-i) + \mathcal{D}v(k+r). \tag{11}$$

Now, let $s \geq r$. Then, in analogy with (5), we have

$$Y_{[r:s]} = \mathcal{G}_{[0:s-r]}X_{r,s} + \mathcal{H}_{[0:s-r]}V_{r,s}, \tag{12}$$

where

$$\begin{aligned}
\mathcal{G}_{[0,s-r]} &\triangleq \begin{bmatrix} C & C\mathcal{A} & C\mathcal{A}^2 & \cdots & C\mathcal{A}^{s-r} \end{bmatrix} \in \mathbb{R}^{p\times n(s-r+1)}, \\
\mathcal{H}_{[0:s-r]} &\triangleq \begin{bmatrix} \mathcal{H}_0 & \mathcal{H}_1 & \mathcal{H}_2 & \cdots & \mathcal{H}_{s-r} \end{bmatrix} \\
&\triangleq \begin{bmatrix} \mathcal{D} & C\mathcal{B} & C\mathcal{A}\mathcal{B} & \cdots & C\mathcal{A}^{s-r-1}\mathcal{B} \end{bmatrix} \in \mathbb{R}^{p\times (m+p)(s-r+1)}, \\
V_{r,s} &\triangleq \begin{bmatrix} v(r) & v(r+1) & \cdots & v(s) \\ 0 & v(r) & \cdots & v(s-1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & v(r) \end{bmatrix} \in \mathbb{R}^{(m+p)(s-r+1)\times(s-r+1)}.
\end{aligned}$$

Let $r \geq 0$ and $r+n < s$ so that $n < s-r$. Then,

$$\begin{aligned}
Y_{[r:s]} &= \begin{bmatrix} Y_{[r:r+n-1]} & Y_{[r+n:s]} \end{bmatrix} \in \mathbb{R}^{p\times(s-r+1)}, \\
\mathcal{G}_{[0:s-r]} &= \begin{bmatrix} \mathcal{G}_{[0:n-1]} & \mathcal{G}_{[n:s-r]} \end{bmatrix} \in \mathbb{R}^{p\times n(s-r+1)}, \\
\mathcal{H}_{[0:s-r]} &= \begin{bmatrix} \mathcal{H}_{[0:n+1]} & \mathcal{H}_{[n:s-r]} \end{bmatrix} \in \mathbb{R}^{p\times (m+p)(s-r+1)}, \\
V_{r,s} &= \left[\begin{array}{cccc|ccccc}
v(r) & v(r+1) & \cdots & v(r+n-1) & v(r+n) & v(r+n+1) & \cdots & v(s-1) & v(s) \\
0 & v(r) & \cdots & v(r+n-2) & v(r+n-1) & v(r+n) & \cdots & v(s-2) & v(s-1) \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & \cdots & 0 & v(r) & v(r+1) & v(r+2) & \cdots & v(s-n) & v(s-n+1) \\
0 & \cdots & 0 & 0 & v(r) & v(r+1) & \cdots & v(s-n-1) & v(s-n) \\
\hline
0 & \cdots & 0 & 0 & 0 & v(r) & \cdots & v(s-n) & v(s-n-1) \\
0 & \cdots & 0 & 0 & 0 & 0 & v(r) & \cdots & v(s-n-2) \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & v(r)
\end{array}\right] \\
&\triangleq \left[\begin{array}{c|cc}
V_{r,r+n-1} & & \mathcal{V}_{r+n,s} \\
0_{m+p\times n} & & \\
\hline
0_{(m+p)(s-n-r)\times n} & 0_{(m+p)(s-n-r)\times 1} & V_{r,s-n-1}
\end{array}\right] \in \mathbb{R}^{(m+p)(s-r+1)\times(s-r+1)},
\end{aligned}$$

where $\mathcal{V}_{r+n,s} \in \mathbb{R}^{(m+p)(n+1)\times(s-r-n+1)}$ is given by

$$\mathcal{V}_{r+n,s} \triangleq \begin{bmatrix} v(r+n) & \cdots & v(r+2n) \\ \vdots & \ddots & \vdots \\ v(2r+2n-s) & \cdots & v(r+n) \\ \vdots & \cdots & \vdots \\ v(r) & \cdots & v(s-n) \end{bmatrix}, \quad s < r+2n,$$

$$\mathcal{V}_{r+n,s} \triangleq \begin{bmatrix} v(r+n) & \cdots & v(r+2n) \\ \vdots & \ddots & \vdots \\ v(r) & \cdots & v(r+n) \end{bmatrix}, \quad s = r+2n,$$

$$\mathcal{V}_{r+n,s} \triangleq \begin{bmatrix} v(r+n) & \cdots & v(r+2n) & \cdots & v(s) \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ v(r) & \cdots & v(r+n) & \cdots & v(s-n) \end{bmatrix}, \quad s > r+2n,$$

Note that for $s = n+1$, and thus $r = 0$, $V_{0,0}$ denotes $v(0)$.

From (12), we have

$$Y_{[r+n:s]} = \mathcal{G}_{[n:s-r]} X_{r+n,s} + \mathcal{H}_{[0:n+1]} \mathcal{V}_{r+n,s} + \mathcal{H}_{[n:s-r]} [0_{(m+p)(s-n-r)\times 1} \quad V_{r,s-n-1}], \tag{13}$$

Note that, since $\mathcal{H}_0 = \mathcal{D} = [D \quad 0]$, $\mathcal{D}$ can be replaced by D, and thus columns $p+1, \ldots, 2p$ of $\mathcal{H}_{[0:n+1]}$ and rows $m+1, \ldots, m+p$ of $\mathcal{V}_{r+n,s}$ can be deleted. In addition, by defining

$$\mathcal{H}'_{[0:n+1]} \triangleq \begin{bmatrix} D & \mathcal{H}_1 & \mathcal{H}_2 & \cdots & \mathcal{H}_{n+1} \end{bmatrix} \in \mathbb{R}^{p\times(m+p)n+m},$$

$$\mathcal{V}'_{r+n,s} \triangleq \begin{bmatrix} u(n+r) & u(n+1+r) & \cdots & u(s) \\ v(n-1+r) & v(n+r) & \cdots & v(s-1) \\ \vdots & \vdots & \ddots & \vdots \\ v(r) & v(1+r) & \cdots & v(s-n) \end{bmatrix} \in \mathbb{R}^{(m+p)n+m\times(s-r-n+1)}.$$

Since (A, C) is observable, we choose F such that $\mathcal{A}$ is nilpotent. Hence for all $k \geq n$, it follows that $\mathcal{A}^k = 0$ so that $\mathcal{H}_k = 0$ for all $k \geq n+2$. Therefore, for $k \geq n$, (10) and (11) can be written as

$$x(k+r) = \sum_{i=1}^{\min(n,k)} \mathcal{A}^{i-1}\mathcal{B}v(k+r-i), \tag{14}$$

$$y(k+r) = \sum_{i=1}^{\min(n,k)} C\mathcal{A}^{i-1}\mathcal{B}v(k+r-i) + \mathcal{D}v(k+r). \tag{15}$$

Based on the nilpotent assumption,

$$\mathcal{G}_{[n,s-r]} = 0_{p\times n(s-r-n+1)},$$
$$\mathcal{H}_{[n:s-r]} = 0_{p\times (m+p)(s-r-n-1)}.$$

Thus (13) can be rewritten as

$$Y_{[r+n:s]} = \mathcal{H}'_{[0:n+1]}\mathcal{V}'_{r+n,s}. \tag{16}$$

If $\mathcal{V}'_{r+n,s}$ has full column rank, then the unique solution of (16) is given by

$$\mathcal{H}'_{[0:n+1]} = Y_{[r+n:s]}\mathcal{V}'^{+}_{r+n,s}, \tag{17}$$

where $(\quad)^{+}$ is the generalized inverse. When u and y are corrupted by noise, (17) provides a least squares estimate.

2.3 Deriving H_i from $\mathcal{H}_{[0:n+1]}$

Let $k \geq 1$. Then

$$\mathcal{H}_k = C\mathcal{A}^{k-1}\mathcal{B} = [\beta_k \quad \alpha_k],$$

where

$$\beta_k \triangleq C(A+FC)^{k-1}(B+FD), \tag{18}$$
$$\alpha_k \triangleq -C(A+FC)^{k-1}F. \tag{19}$$

Next, for $k \geq 3$ we have

$$\begin{aligned} H_k &= CA^{k-1}B, \qquad (20)\\ &= C\left((A+FC)-FC\right)A^{k-2}B\\ &= C\left((A+FC)A-FCA\right)A^{k-3}B\\ &= C\left((A+FC)(A+FC-FC)-FCA\right)A^{k-3}B\\ &= C\big((A+FC)^2-(A+FC)FC-FCA\big)A^{k-3}B\\ &= C\left((A+FC)^2-\sum_{i=1}^{2}(A+FC)^{i-1}FCA^{2-i}\right)A^{k-3}B. \qquad (21) \end{aligned}$$

More generally we have

$$\left.\begin{aligned}
H_k &= C\left((A+FC)^j - \sum_{i=1}^{j}(A+FC)^{i-1}FCA^{j-i}\right)A^{k-(j+1)}B \\
&= C\left((A+FC)^j A - \sum_{i=1}^{j}(A+FC)^{i-1}FCA^{(j+1)-i}\right)A^{k-(j+2)}B \\
&= C\left((A+FC)^j(A+FC-FC) - \sum_{i=1}^{j}(A+FC)^{i-1}FCA^{(j+1)-i}\right)A^{k-(j+2)}B \\
&= C\left((A+FC)^{j+1} - (A+FC)^j FC - \sum_{i=1}^{j}(A+FC)^{i-1}FCA^{(j+1)-i}\right)A^{k-(j+2)}B \\
&= C\left((A+FC)^{j+1} - \sum_{i=1}^{j+1}(A+FC)^{i-1}FCA^{j-i}\right)A^{k-(j+2)}B.
\end{aligned}\right\} \tag{22}$$

Using (22), (21) becomes

$$\begin{aligned}
H_k &= C\left((A+FC)^{k-1} - \sum_{i=1}^{k-1}(A+FC)^{i-1}FCA^{k-i-1}\right)B \\
&= C(A+FC)^{k-1}B - C\sum_{i=1}^{k-1}(A+FC)^{i-1}FCA^{k-i-1}B \\
&= \left(-\alpha_k D + C(A+FC)^{k-1}B\right) - \left(C\sum_{i=1}^{k-1}(A+FC)^{i-1}FCA^{k-i-1}B + \alpha_k D\right) \\
&= (\beta_k) - \left(\sum_{i=1}^{k-1}\left(C(A+FC)^{i-1}F\right)\left(CA^{k-i-1}B\right) + \alpha_k D\right) \\
&= \beta_k - \sum_{i=1}^{k-1}(\alpha_i)(H_{k-i}) + \alpha_k H_0
\end{aligned}$$

Thus,

$$H_k = \begin{cases} \beta_k + \sum_{i=1}^{k}\alpha_i H_{k-i}, & k = 1, \ldots, n; \\ \sum_{i=1}^{n}\alpha_i H_{k-i}, & k \geq n+1. \end{cases} \tag{23}$$

3 Alternative Derivation of OKID Using a Pseudo-FIR Model

The discrete-time state space model (1)-(2) can be represented by the infinite impulse response (IIR) model

$$y_k = \alpha_1 y_{k-1} + \ldots + \alpha_n y_{k-n} + \beta_0 u_k + \beta_1 u_{k-1} + \ldots + \beta_n u_{k-n}, \tag{24}$$

for $k \geq n$. Rearranging (24), we obtain

$$y_k = \mathcal{H}_0 v_k + \ldots + \mathcal{H}_n v_{k-n}, \tag{25}$$

where

$$\mathcal{H}_k \triangleq \left[\begin{array}{cc} \beta_k & \alpha_k \end{array} \right], \tag{26}$$

and from (24), $\alpha_0 = 0$. Furthermore, note that (25) can be written as

$$Y_{[r:s]} = \mathcal{H}_{[0:s-r]} V_{r:s}, \tag{27}$$

which has the least-squares solution

$$\hat{\mathcal{H}}_{[0:s-r]} = V_{r:s}^{+} Y_{[r:s]}, \tag{28}$$

where

$$\hat{\mathcal{H}}_k \triangleq \left[\begin{array}{cc} \hat{\beta}_k & \hat{\alpha}_k \end{array} \right], \tag{29}$$

and $\hat{\alpha}_0$ is constrained to be zero.

Next, note that if $y_0, \ldots, y_\mu$ is the impulse response of a system, then $y_k \triangleq H_k$. Impulsing (24), we obtain

$$\begin{aligned}
H_0 &= \beta_0, \\
H_1 &= \alpha_1 y_0 + \beta_1 u_0 \\
&= \alpha_1 H_0 + \beta_1, \\
H_2 &= \alpha_1 y_1 + \alpha_2 y_0 + \beta_2 u_0 \\
&= \alpha_1 H_1 + \alpha_2 H_0 + \beta_2 \\
H_3 &= \alpha_1 y_2 + \alpha_2 y_1 + \alpha_3 y_0 + \beta_3 u_0 \\
&= \alpha_1 H_2 + \alpha_2 H_1 + \alpha_3 H_0 + \beta_3 \\
&\vdots \\
H_k &= \begin{cases} \beta_k + \sum_{i=1}^{k} \alpha_k H_{k-i}, & 0 \leq k \leq n, \\ \sum_{i=1}^{k} \alpha_k H_{k-i}, & k > n. \end{cases}
\end{aligned}$$

Thus using the estimates (29), the Markov parameters can be estimated recursively by

$$\hat{H}_k = \begin{cases} \hat{\beta}_k + \sum_{i=1}^{k} \hat{\alpha}_k \hat{H}_{k-i}, & 0 \leq k \leq n, \\ \sum_{i=1}^{k} \hat{\alpha}_k \hat{H}_{k-i}, & k > n. \end{cases} \tag{30}$$

4 Conclusion

Note that (17) and (28) are reshuffled representations of the same equation. While OKID and TSI both set up the V matrix differently, each contains the same information and thus the least square approximation of $\mathcal{H}$ must be identical. This conclusion is verified in [1]. Also, OKID and TSI share a very similar method of obtaining the estimates of the Markov parameters in (23) and (30).

References

[1] S. Barnett. Inversion of partitioned matrices with patterned blocks. *International Journal of Systems Science*, 14(2):235–237, 1983.

[2] ; Wei Chen John Valasek. Observer/kalman filter identification for online system identification of aircraft. *Journal of Guidance, Control, and Dynamics*, 0731-5090(2):347–353, 2003.