On the Asymptotic Estimates of Least Squares Identification

Söderström, Torsten

1973

Document Version:
Publisher's PDF, also known as Version of record

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

ON THE ASYMPTOTIC ESTIMATES OF LEAST SQUARES IDENTIFICATION

T. SÖDERSTRÖM

Report 7327(C) November 1973
Lund Institute of Technology
Division of Automatic Control

ON THE ASYMPTOTIC ESTIMATES OF LEAST SQUARES IDENTIFICATION

T. Söderström

Abstract

Least squares identification of a first-order system is considered. The identified model is assumed to be of a general order. Expressions for the parameter estimates and some other quantities are derived using asymptotic theory.

TABLE OF CONTENTS

I. INTRODUCTION

The least squares identification method has several advantages. It is very easy to apply, but in a general case the resulting parameter estimates are biased. In order to overcome this fact, a model of high order can be tried see e.g. Åström - Eykhoff (1971). In their paper this idea is called the repeated least square method. The purpose of this report is to analyze the effect of the least squares method applied to a model of an arbitrary order. Asymptotic theory will be used in all calculations (assuming infinitely many samples).

The considered situation will be the following. Assume that the data, i.e. the input signal u(t) and the output signal y(t), are related through

$$A(q^{-1})\, y(t) = B(q^{-1})\, u(t) + C(q^{-1})\, e(t) \tag{1.1}$$

where e(t) is white noise of variance λ, q^{-1} the backward shift operator and

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_{n_o} q^{-n_o}$$

$$B(q^{-1}) = \quad b_1 q^{-1} + \dots + b_{n_o} q^{-n_o}$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \dots + c_{n_o} q^{-n_o}$$

The model of the process is assumed to be

$$\hat{A}(q^{-1})\, y(t) = \hat{B}(q^{-1})\, u(t) + \varepsilon(t) \tag{1.2}$$

where $\{\varepsilon(t)\}$ is the residuals and

$$\hat{A}(q^{-1}) = 1 + \hat{a}_1 q^{-1} + \dots + \hat{a}_n q^{-n}$$

$$\hat{B}(q^{-1}) = \quad \hat{b}_1 q^{-1} + \dots + \hat{b}_n q^{-n}$$

It is well known, that $C(q^{-1}) = 1$ implies consistent estimates. The idea behind the repeated least square method can be expressed in the following way. Consider a polynomial $L(q^{-1})$ of order $n - n_o$ satisfying approximately

$$L(q^{-1})\, C(q^{-1}) = 1$$

If this is true exactly, (1.1) can be written as

$$A(q^{-1})\, L(q^{-1})\, y(t) = B(q^{-1})\, L(q^{-1})\, u(t) + e(t)$$

and the estimates will be given by

$$\hat{A}(q^{-1}) = A(q^{-1})\, L(q^{-1})$$

$$\hat{B}(q^{-1}) = B(q^{-1})\, L(q^{-1}) \tag{1.3}$$

The estimates described by (1.3) will give the true transfer function of the process. It can be proved by a slight modification of the results in Söderström (1972) that, provided the signal to noise ratio is large, the real parameter estimates satisfy

$$\hat{A}(q^{-1}) \approx A(q^{-1})\ \tilde{L}(q^{-1})$$

$$\hat{B}(q^{-1}) \approx B(q^{-1})\ \tilde{L}(q^{-1})$$

The coefficients of the operator $\tilde{L}(q^{-1}) = 1 + \ell_1 q^{-1} + \ldots + \ell_{n-n_o} q^{-(n-n_o)}$ gives the minimum point of

$$E[\tilde{L}(q^{-1})\ C(q^{-1})\ e(t)]^2$$

There are several ways to analyze the model (1.2) and compare it with the true system (1.1). One way is to compare the static gains

$$\frac{\Sigma b_i}{1+\Sigma a_i} \qquad\qquad \frac{\Sigma \hat{b}_i}{1+\Sigma \hat{a}_i}$$

Another is to evaluate the loss function

$$E\varepsilon^2(t) = E\left[\frac{\hat{A}(q^{-1})\ B(q^{-1}) - A(q^{-1})\ \hat{B}(q^{-1})}{A(q^{-1})}\ u(t)\right]^2 + $$

$$+\ E\left[\frac{\hat{A}(q^{-1})\ C(q^{-1})}{A(q^{-1})}\ e(t)\right]^2 \qquad (1.4)$$

Clearly

$$E\varepsilon^2(t) \geq \lambda^2$$

Then, if the loss function is close to λ^2, the model can be considered as good.

A third way is to compute the minimal variance strategy from the model (1.2), apply it to the true system (1.1) and evaluate the variance of the output of the closed loop system.

Computer program for the second and third way of evaluating the model can be found in Söderström (1973a) where it is assumed that the input signal is white noise, independent of e(t).

In the calculations in this report it is generally assumed that

i) $n_o = 1$
ii) u(t) is white noise of variance σ, and independent of $\{e(t)\}$.

In the next section the parameter estimates are computed for a general model order n. In the third section the ways of model evaluation discussed above are penetrated.

In order to decrease the number of minussigns in the calculations introduce

$\alpha = -a, \gamma = -c$

II. COMPUTATION OF THE PARAMETER ESTIMATES

The least squares estimates $\hat{a}_1, \ldots \hat{a}_n, \hat{b}_1 \ldots \hat{b}_n$ are given by, see e.g. Åström (1968),

$$E\begin{bmatrix} y(t-1) \\ \vdots \\ y(t-n) \\ -u(t-1) \\ \vdots \\ -u(t-n) \end{bmatrix} [y(t-1) \ldots y(t-n) \; - u(t-1) \ldots -u(t-n)] \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_n \\ \hat{b}_1 \\ \vdots \\ \hat{b}_n \end{bmatrix} = Ey(t) \begin{bmatrix} -y(t-1) \\ \vdots \\ -y(t-n) \\ u(t-1) \\ \vdots \\ u(t-n) \end{bmatrix}$$

This system of linear equations can be written as

$$\left[\begin{array}{cccc|cccc}
r_o+\rho & \cdot & & \alpha^{n-2}r_1+\alpha^{n-1}\rho & 0 & -b\sigma^2 & \cdot & -\alpha^{n-2}b\sigma^2 \\
r_1+\alpha\rho & \cdot & & & & \cdot & & \\
\cdot & \cdot & & & & 0 & \cdot & \\
\cdot & & \cdot & & & & & \cdot \\
\alpha^{n-2}r_1+\alpha^{n-1}\rho & & & r_o+\rho & & & & 0 \\
\hline
0 & & & & \sigma^2 & & & \\
-b\sigma^2 & \cdot & & & & \cdot & & 0 \\
\cdot & & \cdot & 0 & & 0 & \cdot & \\
\cdot & & & & & & & \cdot \\
-\alpha^{n-2}b\sigma^2 & & \cdot & 0 & & & & \sigma^2
\end{array}\right]
\begin{bmatrix} \hat{a}_1 \\ \cdot \\ \cdot \\ \hat{a}_n \\ \hat{b}_1 \\ \cdot \\ \cdot \\ \hat{b}_n \end{bmatrix}
=
\begin{bmatrix} -r_1-\alpha\rho \\ \cdot \\ \cdot \\ -\alpha^{n-1}r_1-\alpha^n\rho \\ b\sigma^2 \\ \cdot \\ \cdot \\ \alpha^{n-1}b\sigma^2 \end{bmatrix} \qquad (2.1)$$

where

$$\rho = \frac{b^2\sigma^2}{1-a^2} = E\left[\frac{bq^{-1}}{1+aq^{-1}}\,u(t)\right]^2$$

$$r_o = \frac{\lambda^2(1+c^2-2ac)}{1-a^2} = E\left[\frac{1+cq^{-1}}{1+aq^{-1}}\,e(t)\right]^2$$

$$r_1 = \frac{\lambda^2(c-a)(1-ac)}{1-a^2} = E\left[\frac{1+cq^{-1}}{1+aq^{-1}}\,e(t)\right]\left[\frac{1+cq^{-1}}{1+aq^{-1}}\,e(t-1)\right]$$

In this section the solution of (2.1) for a general n will be derived. The following variables will be useful

$$\beta_o = 1 \tag{2.2}$$

$$\beta_i = \alpha^i + \hat{a}_1\alpha^{i-1} + \dots + \hat{a}_i \qquad 1 \le i \le n$$

The <u>last</u> n equations of (2.1) give

$$\hat{b}_i = \beta_{i-1}b \qquad 1 \le i \le n \tag{2.3}$$

The estimates $\{a_i\}$ are found by (2.2)

$$\hat{a}_i = \beta_i - \alpha\beta_{i-1} \qquad 1 \le i \le n \tag{2.4}$$

Thus it is sufficient to compute β_i, $i = 1 \dots n$.

Insertion of (2.3) and (2.4) in the <u>first</u> n equations of (2.1) will give

$$\begin{bmatrix} r_o+\rho & \cdot & \cdot & \alpha^{n-2}r_1+\alpha^{n-1}\rho \\ r_1+\alpha\rho & \cdot & & \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ \alpha^{n-2}+\alpha^{n-1}\rho & & & r_o+\rho \end{bmatrix} \left\{ \begin{bmatrix} 1 & 0 & & 0 \\ -\alpha & 1 & & 0 \\ & \cdot & \cdot & \\ 0 & & \cdot & \cdot \\ & & -\alpha & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix} + \begin{bmatrix} -\alpha \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \right\}$$

$$-b^2\sigma^2 \begin{bmatrix} 0 & 1 & \alpha & & \alpha^{n-2} \\ & \cdot & & & \cdot \\ & & \cdot & & \cdot \\ & & & \cdot & \cdot \\ & 0 & & & 1 \\ & & & & 0 \end{bmatrix} \cdot \left\{ \begin{bmatrix} 0 & & & \\ 1 & \cdot & & 0 \\ & \cdot & \cdot & \\ 0 & & \cdot & \cdot \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \right\}$$

$$= \begin{bmatrix} 1 \\ \alpha \\ \cdot \\ \cdot \\ \alpha^{n-1} \end{bmatrix} (-r_1-\alpha\rho) \qquad (2.5)$$

This system of equations can be simplified as follows

$$\left\{ \begin{bmatrix} r_o-\alpha r_1 & r_1(1-\alpha^2) & \alpha r_1(1-\alpha^2) & \cdot\ \cdot & \alpha^{n-2}r_1 \\ r_1-\alpha r_o & \cdot & & & \\ 0 & & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & & & \cdot & r_o-\alpha r_1 \\ 0 & & & & r_1-\alpha r_o \quad r_o \end{bmatrix} + \right.$$

$$\left. +\rho \begin{bmatrix} 1-\alpha^2 & \alpha(1-\alpha^2) & & \alpha^{n-1} \\ & 1-\alpha^2 & \cdot & \\ 0 & & \cdot & \\ & & 1-\alpha^2 & \\ & & & 1 \end{bmatrix} - (1-\alpha^2)\rho \begin{bmatrix} 1 & \alpha & \alpha^{n-2} & 0 \\ & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ & & \cdot & \cdot \\ & & 1 & \\ & & & 0 \end{bmatrix} \right\}$$

$$\cdot \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix} = (-r_1-\alpha\rho) \begin{bmatrix} 1 \\ \alpha \\ \cdot \\ \cdot \\ \alpha^{n-1} \end{bmatrix} + \alpha \begin{bmatrix} r_o+\rho \\ r_1+\alpha\rho \\ \cdot \\ \cdot \\ \alpha^{n-2}r_1+\alpha^{n-1}\rho \end{bmatrix}$$

or rewritten

$$\left\{\begin{bmatrix} r_o-\alpha r_1 & r_1(1-\alpha^2) & & \alpha^{n-2}r_1 \\ r_1-\alpha r_o & \cdot & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ 0 & & \cdot & r_o-\alpha r_1 \\ & & r_1-\alpha r_o & r_o \end{bmatrix} +\rho \begin{bmatrix} & \alpha^{n-1} \\ & \cdot \\ 0 & \\ & \cdot \\ & 1 \end{bmatrix}\right\} \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} (-r_1+\alpha r_o)$$

Premultiplying this system with

$$\begin{bmatrix} 1 & -\alpha & & \\ & 1 & \cdot & 0 \\ & & \cdot & \cdot \\ 0 & & \cdot & -\alpha \\ & & & 1 \end{bmatrix}$$

the following equivalent system will be obtained

$$\begin{bmatrix} r_o(1+\alpha^2)-2\alpha r_1 & r_1-\alpha r_o & & & 0 \\ r_1-\alpha r_o & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot \\ 0 & & \cdot & r_o(1+\alpha^2)-2\alpha r_1 & r_1-\alpha r_o \\ & & & r_1-\alpha r_o & r_o+\rho \end{bmatrix} \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} (-r_1+\alpha r_o) \quad (2.6)$$

The matrix elements can easily be evaluated. The result is

$$r_o(1+\alpha^2) - 2\alpha r_1 = \lambda^2(1+c^2)$$

$$r_1 - \alpha r_o = \lambda^2 c$$

Then the following difference equation must be fulfilled

$$c\beta_{i-1} + (1+c^2)\beta_i + c\beta_{i+1} = 0 \qquad 2 \leq i \leq n-1 \qquad (2.7)$$

The solution of (2.7) can be written as

$$\beta_i = K_1\gamma^i + K_2\gamma^{-i} \qquad 1 \leq i \leq n \qquad (2.8)$$

The constants K_1 and K_2 are to be determined from the first and the last equation of (2.6). The first one gives

$$\lambda^2(1+c^2)(K_1\gamma+K_2\gamma^{-1}) + \lambda^2 c(K_1\gamma^2+K_2\gamma^{-2}) = -\lambda^2 c$$

or after simplifying calculations

$$K_1 + K_2 = 1$$

Thus (2.8) can be substituted with an expression valid also for β_o, namely

$$\beta_i = K\gamma^i + (1-K)\gamma^{-i} \qquad 0 \leq i \leq n \qquad (2.9)$$

The last equation of (2.6) gives after simple calculations

$$K = \frac{r_o-\lambda^2\gamma^2+\rho}{r_o-\lambda^2\gamma^2+\rho - \gamma^{2n}(r_o-\lambda^2+\rho)}$$

Introduce a new variable δ through

$$\delta = \frac{(r_o-\lambda^2\gamma^2+\rho) - (r_o-\lambda^2+\rho)}{r_o-\lambda^2+\rho} = \frac{\lambda^2(1-\gamma^2)\ (1-a^2)}{\lambda^2(1+c^2-2ac-1+a^2)+b^2\sigma^2}$$

$$= \frac{\lambda^2(1-a^2)(1-c^2)}{\lambda^2(c-a)^2 + b^2\sigma^2} \qquad (2.10)$$

With use of the variable δ, which always is positive, K can be simplified to

$$K = \frac{1 + \delta}{1 + \delta - \gamma^{2n}} \qquad (2.11)$$

Finally, inserting (2.9) and (2.11) in (2.3) and (2.4) the estimates are found to be

$$\hat{a}_i = \frac{(1+\delta)(-c)^{i-1}(a-c) - (-c)^{2n-i}(1-ac)}{1 + \delta - c^{2n}}$$

$$\hat{b}_i = b\,\frac{(1+\delta)(-c)^{i-1} - (-c)^{2n-i+1}}{1 + \delta - c^{2n}} \qquad 1 \le i \le n \qquad (2.12)$$

III. COMPUTATIONS OF FUNCTIONS OF THE PARAMETER ESTIMATES

In the parts 3.1 - 3.4 different quantities evaluating the model are computed. In part 3.5 some general comments to the results are given.

3.1 The static gain

The static gain of the model is

$$S = \frac{\sum_{i=1}^{n} \hat{b}_i}{1+\sum_{i=1}^{n} \hat{a}_i} \qquad (3.1)$$

while the value of the system is $S_o = \frac{b}{1-\alpha}$

With use of (2.12) it is straightforward to calculate S as follows

$$S = b \frac{(1+\delta)\frac{1-\gamma^n}{1-\gamma} - \gamma^{n+1}\frac{1-\gamma^n}{1-\gamma}}{1+\delta-\gamma^{2n}+(1+\delta)(\gamma-\alpha)\frac{1-\gamma^n}{1-\gamma} - (1-\alpha\gamma)\gamma^n \frac{1-\gamma^n}{1-\gamma}} =$$

$$= S_o(1-\alpha) \frac{(1-\gamma^n)(\delta+1-\gamma^{n+1})}{\delta[1-\gamma+(\gamma-\alpha)(1-\gamma^n)]+(1-\gamma^n)[(1+\gamma^n)(1-\gamma)+(\gamma-\alpha)-\gamma^n(1-\alpha\gamma)]} =$$

$$= S_o \frac{(1-\alpha)(1-\gamma^n)(\delta+1-\gamma^{n+1})}{\delta[1-\alpha-(\gamma-\alpha)\gamma^n]+(1-\gamma^n)(1-\alpha)(1-\gamma^{n+1})} =$$

$$= S_o - S_o \frac{(-c)^n\delta(1+c)}{(1+a)[1-(-c)^n][1-(-c)^{n+1}]+\delta[1+a-(-c)^{n+1}-a(-c)^n]} \qquad (3.2)$$

From (3.2) it is especially seen that S is close to S_o if either δ is small or n large.

3.2 The identification loss function

The identification loss function is given by

$$V_{LS} = E\,[\hat{A}(q^{-1})\,y(t) - \hat{B}(q^{-1})\,u(t)]^2 =$$

$$= E\left[\frac{\hat{A}(q^{-1})\,bq^{-1} - (1+aq^{-1})\,\hat{B}(q^{-1})}{1+aq^{-1}}u(t)\right]^2 + E\left[\frac{\hat{A}(q^{-1})(1+cq^{-1})}{1+aq^{-1}}e(t)\right]^2 \tag{3.3}$$

First, the coefficients of the polynomial

$$G(q^{-1}) = \sum_{i=1}^{n+1} g_i q^{-i} = \hat{A}(q^{-1})\,bq^{-1} - (1+aq^{-1})\,\hat{B}(q^{-1})$$

are computed. Since $\hat{b}_1 = b_1, g_1 = 0$. With use of (2.3) and 2.4)

$$g_i = \hat{a}_{i-1}\,b - \hat{b}_i - a\hat{b}_{i-1} = 0 \text{ for } 2 \le i \le n$$

Finally

$$g_{n+1} = \hat{a}_n b - a\hat{b}_n = b\beta_n = b\,\frac{\delta\gamma^n}{1+\delta-\gamma^{2n}} \tag{3.4}$$

Define the polynomial $H(q^{-1})$ through

$$H(q^{-1}) = 1 + \sum_{i=1}^{n+1} h_i q^{-i} = \hat{A}(q^{-1})\,(1+cq^{-1})$$

Thus

$$h_1 = \hat{a}_1 - \gamma$$

$$h_i = \hat{a}_i - \hat{a}_{i-1}\gamma \qquad 2 \le i \le n$$

$$h_{n+1} = -\hat{a}_n\gamma$$

The loss function (3.3) can be expressed as

$$V_{LS} = \frac{b^2 \sigma^2}{1-a^2} g_{n+1}{}^2 +$$

$$+ \frac{\lambda^2}{1-a^2} \begin{bmatrix} 1 & h_1 \cdot \cdot h_{n+1} \end{bmatrix} \begin{bmatrix} 1 & \alpha & & \alpha^{n+1} \\ \alpha & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ \alpha^{n+1} & & & 1 \end{bmatrix} \begin{bmatrix} 1 \\ h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_{n+1} \end{bmatrix} \tag{3.5}$$

Now, equation (2.4) implies

$$\begin{bmatrix} 1 \\ h_1 \\ \\ \\ \\ h_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ -\alpha & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ 0 & & \cdot & 1 \\ & & & -\alpha \end{bmatrix} \begin{bmatrix} \beta_o \\ \beta_1 - \gamma\beta_o \\ \cdot \\ \cdot \\ \beta_n - \gamma\beta_{n-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ (\alpha-\gamma)\beta_n \end{bmatrix}$$

which gives

$$V_{LS} = \frac{b^2 \sigma^2}{1-a^2} g_{n+1}{}^2 + \lambda^2 \left[\beta_o^2 + \sum_{i=1}^{n} (\beta_i - \gamma\beta_{i-1})^2 + \frac{(\alpha-\gamma)^2}{1-\alpha^2} \beta_n^2 \right] \tag{3.6}$$

With use of (2.9) and (2.11)

$$\sum_{i=1}^{n} [\beta_i - \gamma\beta_{i-1}]^2 = \sum_{i=1}^{n} [(1-K)\gamma^{-i}(1-\gamma^2)]^2 =$$

$$= \sum_{i=1}^{n} \left[- \frac{\gamma^{2n-i}}{1+\delta-\gamma^{2n}} (1-\gamma^2) \right]^2 = \frac{(1-\gamma^2)^2 \gamma^{2n}}{(1+\delta-\gamma^{2n})^2} \frac{1-\gamma^{2n}}{1-\gamma^2}$$

Inserting $\beta_o = 1$ and $g_{n+1} = \beta_n b$ from (3.4) the expression (3.6) is simplified as follows

$$V_{LS} = \frac{b^2\sigma^2\lambda^2(\alpha-\gamma)^2}{1-a^2}\left[\frac{\delta\gamma^n}{1+\delta-\gamma^{2n}}\right]^2 + \lambda^2\left[1+\frac{\gamma^{2n}(1-\gamma^2)(1-\gamma^{2n})}{(1+\delta-\gamma^{2n})^2}\right]$$

$$= \lambda^2 + \frac{\lambda^2(1-c^2)}{\delta}\,\frac{\delta^2c^{2n}}{(1+\delta-\gamma^{2n})^2} + \frac{\lambda^2(1-c^2)c^{2n}(1-c^{2n})}{(1+\delta-c^{2n})^2} =$$

$$= \lambda^2 + \frac{\lambda^2(1-c^2)c^{2n}}{(1+\delta-c^{2n})^2}\left[\delta + 1 - c^{2n}\right]$$

or finally

$$V_{LS} = \lambda^2\left[1 + c^{2n}\,\frac{1-c^2}{1+\delta-c^{2n}}\right] \tag{3.7}$$

3.3 The control loss function

The control loss function is given by

$$V_{MVS} = E\left[\frac{\hat{B}(q^{-1})\ (1+cq^{-1})}{(1+aq^{-1})\ \hat{B}(q^{-1}) - \hat{A}(q^{-1})\ bq^{-1} + bq^{-1}}\ e(t)\right]^2 \tag{3.8}$$

Define a new noise through

$$v(t) = \frac{bq^{-1}}{bq^{-1} + (1+aq^{-1})\ \hat{B}(q^{-1}) - \hat{A}(q^{-1})\ bq^{-1}}\ e(t)$$

$$= \frac{1}{1-\beta_n q^{-n}}\ e(t)$$

The covariance function of v(t) clearly satisfies

$$r_v(o) = \frac{1}{1-\beta_n}\lambda^2$$

$$r_v(i) = 0 \qquad 1 \le i \le n-1$$

$$r_v(n) = \frac{\beta_n}{1-\beta_n^2}\lambda^2$$

Introduce the polynomial $J(q^{-1})$ through

$$J(q^{-1}) = \sum_{1}^{n+1} j_i q^{-i} = \hat{B}(q^{-1})(1+cq^{-1})$$

Then

$$j_1 = \hat{b}_1$$

$$j_i = \hat{b}_i - \gamma\hat{b}_{i-1} \qquad 2 \le i \le n$$

$$j_{n+1} = -\hat{b}_n\gamma$$

The loss function can now be written as

$$V_{MVS} = \frac{1}{b^2} r_v(o)\left[\sum_{i=1}^{n+1} j_i^2\right] + \frac{1}{b^2} 2r_v(n)\, j_1 j_{n+1} =$$

$$= \frac{\lambda^2}{b^2(1-\beta_n^2)}\left[\hat{b}_1^2 + \sum_{i=2}^{n}(\hat{b}_i-\gamma\hat{b}_{i-1})^2 + \hat{b}_n^2\gamma^2 - 2\beta_n\hat{b}_1\hat{b}_n\gamma\right] =$$

$$= \frac{\lambda^2}{1-\beta_n^2}\left[1 + \sum_{i=2}^{n}(\beta_{i-1}-\gamma\beta_{i-2})^2 + \beta_{n-1}{}^2\gamma^2 - 2\beta_n\beta_{n-1}\gamma\right] =$$

$$= \lambda^2 + \lambda^2\frac{1}{1-\beta_n^2}\sum_{i=2}^{n+1}(\beta_{i-1}-\gamma\beta_{i-2})^2 =$$

$$= \lambda^2 + \lambda^2\frac{(1+\delta-\gamma^{2n})^2}{(1+\delta-\gamma^{2n})^2-\delta^2\gamma^{2n}}\;\frac{(1-\gamma^2)^2}{(1+\delta-\gamma^{2n})^2}\,\gamma^{2n}\,\frac{1-\gamma^{2n}}{1-\gamma^2}$$

or finally

$$V_{MVS} = \lambda^2 \left[1+c^{2n} \frac{1-c^2}{(1+\delta)^2-c^{2n}}\right] \quad (3.9)$$

3.4 Test of common factors

In Söderström (1973b) a systematic way of testing common factors has been proposed. It leads to the minimization of

$$F(\hat{a}_1, \ldots \hat{\ell}_k) = \frac{E\left[\hat{A}(q^{-1})\hat{L}(q^{-1})y(t)-\hat{B}(q^{-1})\hat{L}(q^{-1})u(t)\right]^2 - V_{LS}}{V_{LS}} N \quad (3.10)$$

with respect to $\hat{a}_1 \ldots \hat{a}_{n-n}\ \hat{b}_1 \ldots \hat{b}_{n-n}\ \hat{\ell}_1 \ldots \hat{\ell}_n$. If the minimum value of F is small, it is assumed that the polynomial obtained from the LS identification can be replaced by $\hat{A}(q^{-1})\hat{L}(q^{-1})$ and $\hat{B}(q^{-1})\hat{L}(q^{-1})$. In order to give some indication of the properties of the LS estimates in this respect, the function F will be considered. It is probably impossible to obtain explicit expression for the minimum point. However, it is easy to treat the situation of a restricted minimization. It will be assumed that k = n-1 and the restrictions $\hat{A}(q^{-1}) = A(q^{-1})$, $\hat{B}(q^{-1}) = B(q^{-1})$ are introduced. Then F is minimized with respect to $\hat{\ell}_1 \ldots \hat{\ell}_{n-1}$.

Thus the following quantity has to be examined

$$V^*_{LS} = \min_{\hat{\ell}\ldots\hat{\ell}_{n-1}} E\left[A(q^{-1})\hat{L}(q^{-1})\left\{\frac{B(q^{-1})}{A(q^{-1})}u(t) + \frac{C(q^{-1})}{A(q^{-1})}e(t)\right\} - B(q^{-1})\hat{L}(q^{-1})u(t)\right]^2 =$$

$$= \min_{\hat{\ell}\ldots\hat{\ell}_{n-1}} E\left[\hat{L}(q^{-1})C(q^{-1})e(t)\right]^2 \quad (3.11)$$

However, this minimization is just LS identification of the system $y(t) = C(q^{-1})e(t)$ with a model of order n-1. Thus V^*_{LS} can be obtained from V_{LS} through the changes

$n \rightarrow n-1$

$a \rightarrow 0$

$bo \rightarrow 0$

Thus, for this case,

$$\delta = \frac{1-c^2}{c^2}$$

and from (3.7)

$$V^*_{LS} = \lambda^2 \left[1 + c^{2n-2} \frac{1-c^2}{1+\frac{1-c^2}{c^2} - c^{2n-2}} \right] =$$

$$= \lambda^2 \left[1 + c^{2n} \frac{1-c^2}{1-c^{2n}} \right] \qquad (3.12)$$

This result gives after trivial calculations the following value of the function F

$$F = \delta c^{2n} \frac{1-c^2}{1-c^{2n}} \frac{1}{1+\delta-c^{2n+2}} . N \qquad (3.13)$$

3.5 Some general comments

For an ideal model the quantities considered in this section would give the following

$$
\begin{aligned}
S &= S_o \\
V_{LS} &= \lambda^2 \\
V_{MVS} &= \lambda^2 \\
F &= 0
\end{aligned}
\tag{3.14}
$$

It is seen from the derived expressions that (3.14) is approximately true if either

1) δ is small (say a large signal to noise ratio)
2) $|c|$ is small compared with 1
3) n is large

It should not be needed to stress that the validity of the calculations is limited. One limitation is the use of asymptotic theory. If n is large the number of data perhaps must be very large in order to give estimates fulfilling the calculated expressions.

REFERENCES

Åström, K. J. (1968):
Lectures on the Identification Problem - the Least Squares Method. Report 6806, Division of Automatic Control, Lund Institute of Technology.

Åström, K. J. and Eykhoff, P. (1971):
System Identification - A Survey. Automatica 7, 199 - 210.

Söderström, T. (1972):
On the Convergence Properties of the Generalized Least Squares Identification Method. Report 7228, Division of Automatic Control, Lund Institute of Technology.

Söderström, T. (1973a):
Programs for Evaluation of Identified Models of Simulated Data. Report 7315(c), Division of Automatic Control, Lund Institute of Technology.

Söderström, T. (1973b):
Test of Common Factors of Identified Models. Application to the Generalized Least Squares Method. Report 7328, Division of Automatic Control, Lund Institute of Technology.