Copyright © 2002 IFAC
15th Triennial World Congress, Barcelona, Spain

VARIANCE ASPECTS OF L_2-MODEL REDUCTION WHEN UNDERMODELING – THE OUTPUT ERROR CASE

Fredrik Tjärnström

Department of Electrical Engineering,
Linköpings universitet, SE-581 83 Linköping, Sweden.
Email: fredrikt@isy.liu.se

Abstract: In this contribution, variance properties of L_2 model reduction are studied. That is, given an estimated model of high order we study the resulting variance of an L_2 reduced approximation. The main result of the paper is showing that estimating a low order output error (OE) model via L_2 model reduction of a high order model gives a *smaller* variance compared to estimating a low order model directly from data in the case of undermodeling. This has previously been shown to hold for FIR (Finite Impulse Response) models, but is in this paper extended to general linear OE models.

Keywords: System identification, Model reduction, Modeling errors, Covariance, System order reduction, Output error identification

1. INTRODUCTION

Model reduction has been subject to considerable interest, both from a modeler's perspective and from a control designer's. There exists quite a wide range of alternatives when performing model reduction. This paper studies one of these methods, namely L_2 model reduction. L_2 model reduction is a topic which has been studied by a large number of people; early references are Wilson (1970) and Aplevich (1973) and two more recent once are Spanos *et al.* (1992) and Ferrante *et al.* (1999). Most papers in the area of model reduction are focused on developing algorithms for different types of reduction methods or analyzing numerical properties of existing algorithms. This paper, on the other hand, deals with the variance aspects of the L_2 model reduction. We analyze the covariance of low order models obtained through L_2 model reduction of high order ones and compares the result with direct estimation of a low order model.

One motivation for this study is that low order models with error bounds are good alternatives for control designers, since they lead to less complicated controllers. Ninness and Goodwin (1995) study this from a system identification perspective, i.e., estimating low order models with tight error bounds, where the model and its error bound is used for control design.

In the system identification literature, there have been a couple of ideas on how to utilize model reduction like methods to achieve good low order models. In Hsia (1977), three rather similar two-step approaches to identification are given. Zhu and Backx (1993) propose a way of estimating models by starting from a high order ARX model, which is then reduced to an OE model using a reduction method close to L_2 model reduction. In these two publications no analysis of the variance of the low order model is given. The number of contributions with a clear variance perspective on model reduction seems to be rather limited. Two of the first papers in this directions are Porat and Friedlander (1985) and Porat (1986). Here ARMA estimation based on sample covariances is studied. The approach taken has strong connections with the model reduction approach of this contribution. Only a limited number of references found, deal with the case when an input signal is present in the identification setup. These are Söderström *et al.* (1991), Wahlberg (1989), Wahlberg (1987), and our previous contributions in this area (Tjärnström and Ljung, 2000; Tjärnström, 2000; Tjärnström and Ljung, 2001). In Tjärnström

and Ljung (2000) it was (among other things) shown that estimating a low order (undermodeled) FIR model by L_2 reduction of a high order model gives strictly lower variance compared to estimating the low order model directly from data. This contribution extends the previous result to hold for general linear output error models, which is a much stronger result.

The outline of the paper is as follows. The next section covers the basic identification setup and notation. The tools used to calculate the variance of L_2 reduced models are summarized in Section 3. Section 4 presents and proves the main result. An illustrating example is presented in Section 5, and conclusions are given in Section 6.

2. PREDICTION ERROR METHODS

Throughout the paper, we denote the input signal by $u(t)$, the output signal by $y(t)$, and N is the total number of measured data. We assume that $y(t)$ is generated according to

$$y(t) = G_0(q)u(t) + e(t) \tag{1}$$

where $G_0(q)$ is a linear time-invariant system, usually referred to as the "true system", q is the discrete-time shift operator, i.e., $qu(t) = u(t+1)$. Furthermore, we assume that the additive noise, $e(t)$, is a zero-mean, white noise sequence, independent of the input having a variance equal to λ.

Output error (OE) models are fitted to the measured data. These models are parameterized by a d-dimensional real-valued parameter vector θ, i.e.,

$$y(t) = G(q,\theta)u(t) + e(t), \tag{2}$$

where

$$G(q,\theta) = \frac{B(q,\theta)}{F(q,\theta)}, \tag{3}$$

$$B(q,\theta) = b_1 q^{-n_k} + \cdots + b_{n_b} q^{-n_k-n_b+1}, \tag{4}$$

$$F(q,\theta) = 1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f}, \tag{5}$$

$$\theta = \begin{pmatrix} b_1 & \dots & b_{n_b} & f_1 & \dots & f_{n_f} \end{pmatrix}^T. \tag{6}$$

If $F = 1$, the model is said to be of finite impulse response (FIR) type. The true system is said to belong to the model class if, for some $\theta = \theta_0$, it holds that $G_0(q) = G(q,\theta_0)$.

The loss function is defined as the sample mean of the squared sum of the prediction errors (in this case the output errors)

$$V_N(\theta) = \frac{1}{2N}\sum_{t=1}^{N} \varepsilon^2(t,\theta), \tag{7}$$

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = y(t) - G(q,\theta)u(t). \tag{8}$$

The estimate of θ is taken as the minimizer of (7)

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta) \tag{9}$$

i.e., we use prediction error methods (PEM). The basic result is then (Ljung, 1999, Chapter 8) that under weak conditions

$$\hat{\theta}_N \to \theta^* = \arg\min_{\theta} \frac{1}{2}\bar{\mathrm{E}}\varepsilon^2(t,\theta), \text{ as } N \to \infty. \tag{10}$$

Here $\bar{\mathrm{E}} f(t) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \mathrm{E}\, f(t)$. That is, $\hat{\theta}_N$ converges to the best model provided by the model class. (If the minimizer is not unique, the convergence will be to some value in the set of minimizers.)

By undermodeling a system, we mean that there exists *no* $\theta = \theta^*$ such that $G(q,\theta^*) = G_0(q)$. Therefore, in case of undermodeling, the prediction errors are correlated with the input via

$$\begin{aligned} \varepsilon(t,\theta^*) &= (G_0(q) - G(q,\theta^*))u(t) + e(t) \\ &= \tilde{G}(q,\theta^*)u(t) + e(t). \end{aligned} \tag{11}$$

The limiting loss function (as $N \to \infty$) is denoted

$$\bar{V}(\theta) = \frac{1}{2}\bar{\mathrm{E}}\varepsilon^2(t,\theta). \tag{12}$$

Moreover, denote the first and second order derivatives of $\bar{V}(\theta)$, evaluated at θ^* by

$$\bar{V}^{'}(\theta^*), \quad \bar{V}^{''}(\theta^*) \tag{13}$$

and similarly for the derivative of $V_N(\theta)$. The expression for the distribution of the estimate is based on the central limit theorem, see (Ljung, 1999, Chapter 9)

$$\sqrt{N}(\hat{\theta}_N - \theta^*) \in AsN(0, P_\theta), \tag{14}$$

where

$$P_\theta = \lambda \left[\bar{V}^{''}(\theta^*)\right]^{-1} Q \left[\bar{V}^{''}(\theta^*)\right]^{-1} \tag{15}$$

$$Q = \lim_{N\to\infty} N \cdot \mathrm{E}\{\left[V_N^{'}(\theta^*)\right]\left[V_N^{'}(\theta^*)\right]^T\}. \tag{16}$$

To evaluate these derivatives, the negative gradient of the prediction errors needs to be defined

$$\Psi(t,\theta) = -\frac{d}{d\theta}\varepsilon(t,\theta), \tag{17}$$

and also

$$\Psi^{'}(t,\theta) = -\frac{d}{d\theta}\Psi(t,\theta). \tag{18}$$

By simple calculations (cf. Ljung (1999, Chapter 9)) we get

$$\begin{aligned} Q = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N}\sum_{s=1}^{N} &\mathrm{E}\Psi(t,\theta^*)\varepsilon(t,\theta^*) \\ &\times \varepsilon(s,\theta^*)\Psi^T(s,\theta^*) \end{aligned} \tag{19}$$

and

$$\bar{V}^{''}(\theta^*) = \bar{\mathrm{E}}\Psi(t,\theta^*)\Psi^T(t,\theta^*) - \bar{\mathrm{E}}\Psi^{'}(t,\theta^*)\varepsilon(t,\theta^*). \tag{20}$$

The equations (15) - (20) constitute the basis for the calculation of the covariance of low order models. When the true system actually lies in the model class, the covariance expression (15) simplifies to

$$P_\theta = \lambda\left[\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\theta_0)\right]^{-1}. \tag{21}$$

The calculation of the distributions for other statistics, such as $G(e^{i\omega},\hat{\theta}_N)$, is based on a linear approximation

of the mapping from the parameter distribution given by (14) to the statistic of interest. This mapping is usually referred to as Gauss' approximation formula. It states that if $\hat{\theta}_N$ is sufficiently close to $\theta^* = \mathrm{E}\,\hat{\theta}_N$, the following approximation is valid

$$\begin{aligned}\operatorname{Cov} f(\hat{\theta}) &\approx \left[f'(\theta^*)\right] P_\theta \left[f'(\theta^*)\right]^T \\ &\approx \left[f'(\hat{\theta}_N)\right] P_\theta \left[f'(\hat{\theta}_N)\right]^T.\end{aligned} \tag{22}$$

3. MODEL REDUCTION

This section contains a brief summary of some model reduction results presented in Tjärnström and Ljung (2000) and Tjärnström (2000). The covariance expression for a reduced model given below is the foundation for the derivation of the main result in the next section.

Let two model classes $\mathcal{M}_1$ and $\mathcal{M}_2$ be given. These are parameterized by θ and η, respectively. Assume that a model $\mathcal{M}_1(\theta)$ is estimated and that the parameter estimate, $\hat{\theta}$, and its covariance matrix, $\operatorname{Cov}\hat{\theta}$, is available. The model $\mathcal{M}_1(\hat{\theta})$ is subjected to model reduction to $\mathcal{M}_2(\eta)$, described by a loss function J, and the parameters $\hat{\theta}$ and η

$$\hat{\eta} = \arg\min_{\eta} J(\eta, \hat{\theta}). \tag{23}$$

It is also assumed that J is twice continuously differentiable. Let η^* be defined from

$$\eta^* = \lim_{N\to\infty} \hat{\eta}, \tag{24}$$

and θ^* from (10). Then the low order model $\hat{\eta}$ has an asymptotic covariance given by

$$\begin{aligned}\operatorname{Cov}\hat{\eta}_N \approx\ & [J''_{\eta\eta}(\eta^*,\theta^*)]^{-1}[J''_{\eta\theta}(\eta^*,\theta^*)] \\ & \times \operatorname{Cov}\hat{\theta}\,[J''_{\theta\eta}(\eta^*,\theta^*)][J''_{\eta\eta}(\eta^*,\theta^*)]^{-1}.\end{aligned} \tag{25}$$

This expression is derived via Gauss' approximation formula (22) in Tjärnström and Ljung (2000).

When considering L_2 model reduction, the loss function equals

$$J(\eta,\theta) = \frac{1}{4\pi}\int_0^{2\pi} |G(e^{i\omega},\theta) - G(e^{i\omega},\eta)|^2 \Phi_u(\omega) d\omega, \tag{26}$$

where $\Phi_u(\omega)$ is the spectrum of the input. By the use of Parseval's formula, (26) can be rewritten according to

$$\begin{aligned}J(\eta,\theta) =& \frac{1}{2}\bar{\mathrm{E}}\left[(G(q,\theta) - G(q,\eta))u(t)\right]^2 \\ =& \frac{1}{2}\bar{\mathrm{E}}\left(\tilde{G}(q,\eta,\theta)u(t)\right)^2 = \frac{1}{2}\bar{\mathrm{E}}\gamma^2(t,\eta,\theta)\end{aligned} \tag{27}$$

Here $\tilde{G}(q,\eta,\theta) = G(q,\theta) - G(q,\eta)$ is the model error and $\gamma(t,\eta,\theta) = \tilde{G}(q,\eta,\theta)u(t)$ are the the residuals from the reduction step. The loss function (27) is studied closer in the next section.

In Tjärnström and Ljung (2000) and Tjärnström (2000), the expression (25) is used to analyze the variance properties of L_2 model reduction. The basic results are:

- When estimating finite impulse response models of low order (undermodeling), it is strictly better to estimate the low order model by model reduction of a high order model compared to estimating the low order model directly from data.
- When estimating output error models (of correct order), nothing is lost in terms of variance, when getting the estimate via reduction of a high order model compared to estimating the low order model directly from data.

It is the first of these results that is generalized to hold for general linear undermodeled output error models in this contribution.

The tool (25) is also used in Tjärnström and Ljung (2001) to study a two-step least-squares procedure for estimating the system dynamics. Here a covariance expression is derived for the final estimate and optimization issues around this are discussed.

4. MAIN RESULT

Before presenting the main result of the paper, the difference between the *direct* estimate and the *reduced* estimate must be made clear. The direct estimate $G(q,\hat{\eta}_d)$ of the system is obtained by directly minimizing the loss function $V_N(\eta)$ defined in (9) from measured data. On the other hand, the reduced estimate $G(q,\hat{\eta}_r)$ is obtained by first estimating a high order model $G(q,\hat{\theta})$ that minimizes $V_N(\theta)$ in (9) and then reducing this estimate by L_2 model reduction defined by $J(\eta,\hat{\theta})$ in (26).

Note: When comparing the variance expressions for the direct and the reduced estimate, the expectation is taken over *both* e and u. Therefore it is essential to use the true input spectrum Φ_u as weighting in the L_2 reduction, and not an estimate of it like the absolute square of the FFT of the input sequence used. If the latter is used, the results changes. For finite impulse response models, it is easy to see that the reduced and the direct estimates are identical in this case, see (Tjärnström and Ljung, 2000; Tjärnström, 2000).

Theorem 1. Assume that data is generated according to the assumptions in Section 2. Let a high order OE model, $G(q,\theta)$, be parameterized such that there exists a unique $\theta = \theta_0$ giving $G(q,\theta_0) = G_0(q)$, and let η parametrize a low order OE model, $G(q,\eta)$. Then by first estimating $G(q,\theta)$ by minimizing (9) and then reducing $G(q,\hat{\theta})$ by L_2 model reduction according to (26) we get a model $G(q,\hat{\eta}_r)$ with $\operatorname{Cov}\hat{\eta}_r = P_r$. Also assume that a low order model $G(q,\eta_d)$ with the same parameterization as $G(q,\eta_r)$ is estimated directly by

minimizing (9). This model is denoted by $G(q, \hat{\eta}_d)$ having $\mathrm{Cov}\,\hat{\eta}_d = P_d$. Then it holds that

$$P_r \leq P_d.$$

Note that since the existence of a unique θ_0 that minimizes $\bar{V}(\theta)$ is assumed, it follows that $\theta^* = \theta_0$ in (10). It is important to note that $\eta_d^* = \eta_r^*$. The reason for this is that both estimates asymptotically minimize the same criterion, i.e., $\bar{V}(\eta)$ given by (12). In the following η_d^* and η_r^* are therefore replaced by η^*.

4.1 *Direct estimate*

To prove Theorem 1, some expressions involving the covariance matrices for the two estimates are needed. We start by looking at the covariance expression (15) for the low order model $\hat{\eta}_d$.

In the special case of undermodeling for OE models we are considering, it is possible to simplify the expression for Q in (19). Using (11) we get

$$\begin{aligned} Q = &\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N}\sum_{s=1}^{N} \mathrm{E}\, \Psi(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \\ &\qquad \times \tilde{G}(q,\eta^*)u(s)\Psi^T(s,\eta^*) \\ &+ \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N}\sum_{s=1}^{N} \mathrm{E}\, \Psi(t,\eta^*)e(t) \\ &\qquad \times e(s)\Psi^T(s,\eta^*) \\ &+ \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N}\sum_{s=1}^{N} 2\,\mathrm{E}\, \Psi(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \\ &\qquad \times e(s)\Psi^T(s,\eta^*) \end{aligned} \tag{28}$$

Since e is white and independent of u the last term vanishes and the second term can be simplified

$$\begin{aligned} Q = &\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N}\sum_{s=1}^{N} \mathrm{E}\Psi(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \\ &\qquad \times \tilde{G}(q,\eta^*)u(s)\Psi^T(s,\eta^*) \\ &+ \lim_{N\to\infty} \frac{\lambda}{N} \sum_{t=1}^{N} \mathrm{E}\Psi(t,\eta^*)\Psi^T(t,\eta^*) \end{aligned} \tag{29}$$

Continuing to work with the first term in (29) (from now on denoted by S) we find that

$$\begin{aligned} S = &\lim_{N\to\infty} \frac{1}{N}\, \mathrm{E} \left(\sum_{t=1}^{N} \Psi(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right) \\ &\qquad \times \left(\sum_{t=1}^{N} \Psi(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right)^T \geq 0 \end{aligned} \tag{30}$$

From the definition of $\bar{\mathrm{E}}$ we also have that the second term equals

$$\lambda \bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*). \tag{31}$$

That is, (29), can be written as

$$Q = \lambda \bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) + S, \tag{32}$$

where S is positive semidefinite. Finally we simplify the expression (20) by utilizing that e and u are independent

$$\begin{aligned} \bar{V}''(\eta^*) = &\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \\ &- \bar{\mathrm{E}} \Psi'(t,\eta^*) \left(\tilde{G}(q,\eta^*)u(t) + e(t) \right) \\ = &\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \\ &- \bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t). \end{aligned} \tag{33}$$

Expressions (15), (32), and (33) together give the covariance of $\hat{\eta}_d$, i.e.,

$$\begin{aligned} P_d = &\left[\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \right. \\ &\left. -\bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right]^{-1} \\ &\times \left[\lambda \bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) + S \right] \\ &\times \left[\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \right. \\ &\left. -\bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right]^{-1}. \end{aligned} \tag{34}$$

4.2 *Reduced estimate*

To calculate the variance of the reduced order model we need to derive the expressions for $J''_{\eta\eta}(\eta^*,\theta_0)$ and $J''_{\eta\hat{\theta}}(\eta^*,\theta_0)$ (see (25)). Taking derivatives of (27) gives

$$J'_\eta(\eta,\theta) = -\bar{\mathrm{E}} \Psi(t,\eta)\gamma(t,\eta,\theta) \tag{35}$$

$$\begin{aligned} J''_{\eta\eta}(\eta,\theta) = &\bar{\mathrm{E}} \Psi(t,\eta)\Psi^T(t,\eta) \\ &- \bar{\mathrm{E}} \frac{d}{d\eta} \Psi(t,\eta)\gamma(t,\eta,\theta) \end{aligned} \tag{36}$$

$$J''_{\eta\theta}(\eta,\theta) = -\bar{\mathrm{E}} \Psi(t,\eta)\Psi^T(t,\theta) \tag{37}$$

Note that since the true system belongs to the model class parameterized by θ, then $G(q,\theta_0) = G_0(q)$ and it follows that

$$\begin{aligned} \gamma(t,\eta,\theta_0) = &(G(q,\theta_0) - G(q,\eta))u(t) \\ = &(G_0(q) - G(q,\eta))u(t) \\ = &\tilde{G}(q,\eta)u(t) \end{aligned} \tag{38}$$

This gives

$$\begin{aligned} J''_{\eta\eta}(\eta^*,\theta_0) = &\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \\ &- \bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \end{aligned} \tag{39}$$

$$J''_{\eta\theta}(\eta^*,\theta_0) = -\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\theta_0) \tag{40}$$

Expressions (21), (25), (39), and (40) together give the covariance of $\hat{\eta}_d$, i.e.,

$$\begin{aligned} P_r = &\lambda \left[\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \right. \\ &\left. -\bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right]^{-1} \\ &\times \left[\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\theta_0) \right] \\ &\times \left[\bar{\mathrm{E}} \Psi(t,\theta_0)\Psi^T(t,\theta_0) \right]^{-1} \\ &\times \left[\bar{\mathrm{E}} \Psi(t,\theta_0)\Psi^T(t,\eta^*) \right] \\ &\times \left[\bar{\mathrm{E}} \Psi(t,\eta^*)\Psi^T(t,\eta^*) \right. \\ &\left. -\bar{\mathrm{E}} \Psi'(t,\eta^*)\tilde{G}(q,\eta^*)u(t) \right]^{-1}. \end{aligned} \tag{41}$$

4.3 *Proof of the main result*

After simplifying the expressions for the covariance matrices (34) and (41), the main result can now be proved. Showing that $P_r \leq P_d$ is equivalent to showing that

$$\begin{aligned}\lambda &\left[\bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\theta_0)\right]\ \left[\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\theta_0)\right]^{-1}\\ &\times\left[\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\eta^*)\right]\\ &\leq \lambda\bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\eta^*)+S\end{aligned} \tag{42}$$

Moreover, since S is positive semidefinite it is sufficient to show that

$$\begin{aligned}&\left[\bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\theta_0)\right]\ \left[\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\theta_0)\right]^{-1}\\ &\times\left[\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\eta^*)\right]\leq\bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\eta^*)\end{aligned} \tag{43}$$

In order to complete the proof of the main result we need a result of the Schur complement (Zhang, 1999, Theorem 6.13).

Theorem 2. Let $A,\ B$, and C be real matrices, where A is an $n\times n$ positive definite matrix and B is an $n\times m$ matrix. Then for any positive semidefinite C $(m\times m)$

$$\begin{pmatrix}A & B\\ B^T & C\end{pmatrix}\geq 0 \quad\Leftrightarrow\quad C\geq B^TA^{-1}B.$$

Since every covariance matrix is *at least* positive semidefinite it follows that

$$\begin{aligned}0&\leq\bar{\mathrm{E}}\begin{pmatrix}\Psi(t,\theta_0)\\ \Psi(t,\eta^*)\end{pmatrix}\begin{pmatrix}\Psi^T(t,\theta_0)\ \ \Psi^T(t,\eta^*)\end{pmatrix}\\ &=\begin{pmatrix}\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\theta_0) & \bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\eta^*)\\ \bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\theta_0) & \bar{\mathrm{E}}\Psi(t,\eta^*)\Psi^T(t,\eta^*)\end{pmatrix}\end{aligned} \tag{44}$$

By noting that $\bar{\mathrm{E}}\Psi(t,\theta_0)\Psi^T(t,\theta_0)$ is positive definite due to the assumption that θ_0 is unique, Theorem 2 is applicable. Direct application of this shows that (44) is equivalent to (43), and the main result follows, i.e., $P_r\leq P_d$.

For $P_r=P_d$ to hold, two things have to be fulfilled. First of all $S=0$ is needed. Secondly, equality in (43) has to hold. These things are obviously fulfilled if the model error, $\tilde{G}$, is zero, and no other configuration fulfilling both of these conditions is obvious. It is worth noting that the difference between P_r and P_d grows with S, and S grows with the bias error squared according to (30).

5. ILLUSTRATING EXAMPLE

This example is included to illustrate two things. First, to show that the covariance matrix of the L_2 reduced estimate really is smaller than the covariance matrix for the direct estimate. Second, to illustrate that the difference in size is linked to the size of the model error.

Let the true system be given by

$$y(t)=\frac{B(q)}{F(q)}u(t)+e(t),$$

where

$$\begin{aligned}B(q)=&2q^{-1}-q^{-2}\\ F(q)=&1-0.7q^{-1}+0.52q^{-2}\\ &-0.092q^{-3}-0.1904q^{-4}.\end{aligned}$$

The input, u, and noise, e, are jointly independent, zero-mean, white noise processes, both with variance 1. The system is simulated with $N=250$ data and low order models of OE type are calculated. To illustrate how the size of the model error affects the result, 3 levels of approximation are calculated. This is represented by estimating models of orders $\{n_b=1,\ n_f=1,\ n_k=1\}$, $\{n_b=2,\ n_f=2,\ n_k=1\}$, and $\{n_b=3,\ n_f=3,\ n_k=1\}$, respectively. The approximations obtained are shown in the Bode plot in Figure 1. As seen from the Bode plot, both the 2nd and 3rd order approximations give reasonably sized model errors, while the first order approximation is very bad. From data, both the direct and the reduced

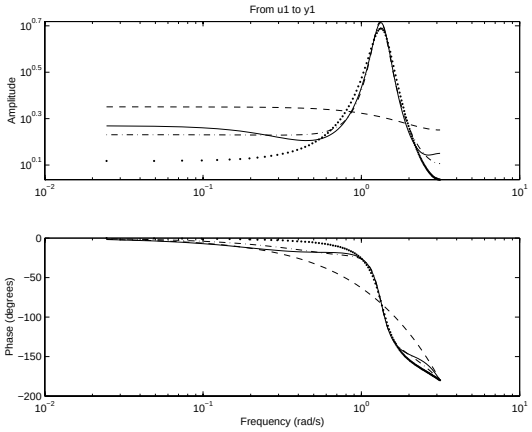

Fig. 1. True system (solid), 3rd order approximation (dash-dotted), 2nd order approximation (dotted), 1st order approximation (dashed).

estimates are calculated. The reduced order models are calculated from an estimated model of true order, i.e., $\{n_b=2,\ n_f=4,\ nk=1\}$. This is repeated 250 times and from the obtained estimates, Monte Carlo based estimates of the covariance matrices are calculated. From each (of the six) covariance matrices the eigenvalues are determined to represent the size of the covariance matrices. The results are presented in Table 1. From the table it is immediate that there is a difference in size of the covariance matrix between the two different methods. The difference can be large as in the OE(1,1,1) case (where the bias is large) or rather small as in the OE(2,2,1) and the OE(3,3,1) case (where the bias is smaller).

Table 1. Monte Carlo estimates of the eigenvalues of the covariance matrices for OE models of order 1, 2 and 3. Both the direct and the reduced estimates are included. "d" is short fort direct estimate and "r" is short for reduced estimate.

Order	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
d 1, 1, 1	2.643	0.216	-	-	-	-
r 1, 1, 1	0.971	0.076	-	-	-	-
d 2, 2, 1	1.866	0.918	0.088	0.040	-	-
r 2, 2, 1	1.696	0.868	0.077	0.037	-	-
d 3, 3, 1	241.0	3.266	0.889	0.381	0.037	0.023
r 3, 3, 1	176.8	2.193	0.841	0.339	0.033	0.022

6. CONCLUSIONS

In this contribution we have shown that it is *never* better to estimate a low order OE model by direct estimation, compared to estimating the model using a two step procedure. This procedure consists of first estimating a high order OE model which is subjected to L_2 model reduction to produce a low order model. It has also been pointed out that the difference grows with the bias error.

This result clearly has an impact on the way low order modeling should be addressed. By starting with a high order model which is then reduced to low order, bias errors can be tracked, and at the same time variance is kept small. Applications of this are obvious, e.g., in robust control where one would like to use low order models with tight error bounds to get simple regulators with good performance.

REFERENCES

Aplevich, J. D. (1973). Approximations of discrete linear systems. *International Journal of Control* **17**(3), 565–575.

Ferrante, A., W. Krajewski, A. Lepschy and U. Viaro (1999). Convergent algorithm for l_2 model reduction. *Automatica* **35**, 75–79.

Hsia, T. C. (1977). *Identification: Least Squares Methods*. Lexington Books. Lexington, Mass.

Ljung, L. (1999). *System Identification: Theory for the User*. 2nd ed.. Prentice-Hall. Upper Saddle River, NJ.

Ninness, B. and G. Goodwin (1995). Estimation of model quality. *Automatica* **31**(12), 1771–1797.

Porat, B. (1986). On the estimation of the parameters of vector gaussian processes from sample covariances. In: *Proceedings of the 25th Conference on Decision and Control*. Athens, Greece. pp. 2002–2005.

Porat, B. and B. Friedlander (1985). Asymptotic accuracy of ARMA parameter estimation methods based on sample covariances. In: *Preprints 7th IFAC Symposium on Identification and System Parameter Estimation*. York, U.K.. pp. 963–968.

Söderström, T., P. Stoica and B. Friedlander (1991). An indirect prediction error method for system identification. *Automatica* **27**, 183–188.

Spanos, J. T., M. H. Milman and D. L. Mingori (1992). A new algorithm for L_2 optimal model reduction. *Automatica* **28**(5), 897–909.

Tjärnström, F. (2000). Quality Estimation of Approximate Models. Licentiate thesis LIU-TEK-LIC-2000:06. Department of Electrical Engineering, Linköpings universitet. SE-581 83 Linköping, Sweden.

Tjärnström, F. and L. Ljung (2000). L_2-model reduction and variance reduction. In: *Proc IFAC Symposium, SYSID 2000*. Santa Barbara. pp. 929–935.

Tjärnström, F. and L. Ljung (2001). Variance properties of a two-step ARX estimation procedure. In: *Proceedings of the European Control Conference*. Porto, Portugal. pp. 1840–1845.

Wahlberg, B. (1987). On the Identification and Approximation of Linear Systems. PhD thesis 163. Department of Electrical Engineering, Linköping University.

Wahlberg, B. (1989). Model reduction of high order estimated models: The asymptotic ML approach. *Int. J. Control* **49**(1), 169–192.

Wilson, D. A. (1970). Optimum solution of model reduction problem. *Proceedings IEE* **117**(6), 1161–1165.

Zhang, F. (1999). *Matrix Theory - Basic Results and Techniques*. Springer-Verlag, New York.

Zhu, Y. and T. Backx (1993). *Identification of Multivariable Industrial Processes for Simulation, Diagnosis and Control*. Springer-Verlag.