

DESIGNING A KALMAN FILTER WHEN NO NOISE COVARIANCE INFORMATION IS AVAILABLE

Robert Bos ^{*,1} Xavier Bombois ^{*}
Paul M. J. Van den Hof ^{*}

** Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, email: r.bos@dcsc.tudelft.nl*

Abstract: A problem when designing Kalman filters using first principles models is often that these models lack a description of the noises that affect the states and the measurements. In these cases, the Kalman filter has to be estimated from data. For this purpose many algorithms have been presented in the literature. All methods in the literature assume that the system under consideration has an observability matrix that has no small singular values. In this paper it will be shown that small singular values can lead to poor performance of estimated Kalman filters. Also a method will be introduced for estimating the Kalman filter in the case that the system has small singular values. This method is able to construct a good filter, even if the first principles model is badly observable. *Copyright*© 2005 IFAC.

Keywords: State estimation, Kalman filter design, noise modelling

1. INTRODUCTION

Unmeasurable states of a process can be estimated using measurements related to the state and a state filter such as the Kalman filter. The standard design methodology for such a filter requires not only a full description of the relation between the states and measurements, but also a full description of the noise affecting the states and the measurements. A model relating the measurements to the states can be obtained using first principles modelling. At this point it is important to note that first principles models of physical plants are generally not fully controllable and observable, due to the limited amount of measurements (i.e. the limited dimension of the measurement vector with respect to the dimension of the state vector). Moreover, although the

deterministic behavior can often be found using first principles modelling, this form of modelling only very rarely admits an accurate description of the stochastic behavior. Based on these considerations, we will here address the problem of designing a reliable Kalman filter when no noise properties are available and when some singular values of the observability matrix are small or even zero. For simplicity, we will consider linear systems of the form

$$x(k+1) = Ax(k) + Bu(k) + w(k) \quad (1)$$

$$y(k) = Cx(k) + v(k), \quad (2)$$

with $x(k) \in \mathbb{R}^{n \times 1}$ the state of the system at time index k , $u(k) \in \mathbb{R}^{m \times 1}$ the known input of the system, $w(k) \in \mathbb{R}^{n \times 1}$ the unknown process noise, $y(k) \in \mathbb{R}^{p \times 1}$ the measurement vector and $v(k) \in \mathbb{R}^{p \times 1}$ the measurement noise. The noises $w(k)$ and $v(k)$ are assumed to be gaussian with:

¹ The work of Robert Bos is supported by Dutch research organization TNO-TPD.

$$\mathbb{E} \begin{bmatrix} w(k) \\ v(k) \end{bmatrix} = 0 \quad \mathbb{E} \begin{bmatrix} w(k) \\ v(k) \end{bmatrix} \begin{bmatrix} w(l) \\ v(l) \end{bmatrix}^T = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{k,l}.$$

In this situation the optimal estimator for the state vector $x(k)$ is the Kalman filter (Anderson and Moore, 1979). The optimal predictor or state estimator can be written in the Kalman innovations form:

$$\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + Ke(k) \quad (3)$$

$$e(k) = y(k) - C\hat{x}(k), \quad (4)$$

with $K \in \mathbb{R}^{n \times p}$ the Kalman gain and $e(k) \in \mathbb{R}^{p \times 1}$ the so-called innovation sequence. The optimal Kalman gain is generally computed as a complex function of (A,B,C) and (Q,R,S). As already mentioned, we consider the case where the matrices Q, R, S are not available.

In the literature, many methods can be found that are able to construct a Kalman filter using test data if the covariance matrices Q, R, S are unknown. These methods can be separated into two main classes. The first class of methods parameterizes the matrices Q, R, S , then these parameters are estimated using test data $u(1) \dots u(N), y(1) \dots y(N)$. Once the matrices Q, R, S have been estimated, the Kalman filter can be constructed in the usual manner. The second class of methods parameterize the Kalman gain K . Then the Kalman gain is directly estimated for data. The advantage of directly parameterizing K is that if the system is fully observable, that the K matrix will be identifiable, but a Q, R, S representation may not be.

Once a user has chosen to either estimate K or Q, R, S , there are also different methods for estimating the required matrix or matrices from test data. Most estimation methods can be used to estimate Q, R, S as well as K . The estimation methods can be roughly classified into four categories, Bayesian estimation (Magill, 1965), Maximum Likelihood (ML) estimation (Zhou and Luecke, 1994), Prediction error identification (Ljung, 1999) and correlation or covariance matching. The first three estimation methods are general purpose parameter estimation techniques and thus we shall not discuss them in detail here.

The fourth class of estimators are the covariance matching methods. In these methods, required matrices are solved by considering the autocovariance function of the innovations sequence $e(k)$ (see (4)). For an optimal filter it can be shown that this sequence should be a white noise (Anderson and Moore, 1979). This property is used in (Mehra, 1970) and (Carew and Bélanger, 1973) to iteratively construct the optimal choices for either Q, R, S or K . Instead of considering the covariance of $e(k)$, it is also

possible to use the covariance of the stochastic part of $y(k)$. Then it is possible to directly identify the Kalman gain matrix K without K having to be parameterized. This method is described in (Mehra, 1973). The advantage of this method over all other methods listed earlier is that for the construction of the Kalman gain K , only linear operations are necessary. As a consequence, we will consider this method in order to design our Kalman filter.

In all methods, it is assumed that the system described by (1)-(2) is completely observable and also that the corresponding observability matrix has no small singular values. This can be a problem, because for many physical systems, this is not necessarily the case.

In this paper we consider a covariance method similar the covariance method introduced in (Mehra, 1973). The first contribution is to show that an ill conditioned observability matrix for the system (1)-(2) can lead to poor results of the Kalman Filter derived using the covariance-based approach. Then we will also show that this problem can be solved by using the normal covariance based approach for only a part of the state directions. For the remaining state directions, a more robust estimator is used. Also we will show that even if a system is completely unobservable in certain directions of the state-space, it is still possible to construct an useful Kalman filter.

This paper is organized as follows: section 2 starts with a summary of the covariance based Kalman filter design method. Then, in the second part of section 2, we demonstrate how this Kalman derivation can be used to construct an approximately optimal filter using known matrices A, B, C and a series of test measurements. Section 3 then discusses the sensitivity of this method to stochastic influences, and also shows how this sensitivity can be reduced. Section 4 presents a simulation example, in which the results of the previous sections are applied. The paper ends with a summary of the main conclusions.

2. COVARIANCE BASED DESIGN OF A KALMAN FILTER

Our methodology for designing a state filter is based on a covariance based design procedure of the Kalman filter. In the first part of this section a short overview of the covariance based design procedure is given. The second part of this section describes how this method can be used to obtain a state filter using only knowledge of the matrices A, B, C , and test data consisting of a series of inputs $u(k)$ and corresponding outputs $y(k)$.

2.1 Summary of the covariance based Kalman filter design

As opposed to the standard method for the design of a Kalman filter which requires knowledge of the matrices (Q,R,S), the covariance based method does not require this information. Instead, the covariance of the measurement signal y_k is used to construct the Kalman filter. The resulting filter is equivalent to the filter designed using the common design method.

The model (1)-(2) can be split into two parts, a purely deterministic part, and a purely stochastic part. The deterministic part is given by:

$$x_d(k+1) = Ax_d(k) + Bu(k) \quad (5)$$

$$y_d(k) = Cx_d(k). \quad (6)$$

The stochastic part of the systems is given by:

$$x_s(k+1) = Ax_s(k) + w_k \quad (7)$$

$$y_s(k) = Cx_s(k) + v_k. \quad (8)$$

The deterministic and stochastic systems are related to the complete system by:

$$x(k) = x_d(k) + x_s(k) \quad (9)$$

$$y(k) = y_d(k) + y_s(k). \quad (10)$$

If we assume that the matrix A is stable and the process is in steady-state, the covariance function of the stochastic part of the measurements $y_s(k)$ can be written as:

$$R_{y_s}(i) = \mathbb{E}[y_s(k)y_s(k+i)^T] = \begin{cases} CA^{i-1}M & i > 0 \\ C^T\Phi C + R_e(0) & i = 0 \\ M^T(A^T)^{(-i-1)}C & i < 0 \end{cases}$$

In the equation above, M is the cross-covariance between the stochastic part of the state $x_s(k+1)$ and stochastic part of the output $y_s(k)$:

$$M = \mathbb{E}[x_s(k+1)y_s(k)^T], \quad (11)$$

$R_e(k)$ is the covariance of the innovation sequence $e(k)$ (see (4)):

$$R_e(k) = \mathbb{E}[e(k)e(k)^T] \quad (12)$$

and $\Phi(k)$ is the covariance of the stochastic part of the state:

$$\Phi(k) = \mathbb{E}[x_s(k)x_s(k)^T]. \quad (13)$$

Using the definitions in this section, the optimal steady state Kalman filter (for prediction) is given by (3)-(4), with

$$K(k) = [M - A\Sigma(k)C^T]R_e(k)^{-1} \quad \text{with (14)}$$

$$R_e(k) = R_{y_s}(0) - C\Sigma(k)C^T,$$

$$\Sigma(k+1) = A\Sigma(k)A^T + K(k)R_e(k)K(k)^T,$$

$$\Sigma(0) = 0.$$

In these last equations, $\Sigma(k)$ can be interpreted as the covariance of the estimated state $\hat{x}_s(k)$:

$$\Sigma(k) = \mathbb{E}[\hat{x}_s(k)\hat{x}_s(k)^T]. \quad (15)$$

As can be seen from (14) B plays no role when computing the Kalman gain K . To simplify the notation in the remainder of this report without loss of generality it is assumed that $B = 0$.

2.2 Estimation of M , $R_{y_s}(0)$

In order to be able to use the Kalman filter with the gain computed using (14), we will need to estimate the unknown quantities M and $R_{y_s}(0)$. In this section it is demonstrated how these can be estimated, using N points of test-data generated by (1)-(2). Note that (once the process is in steady state) these quantities M and $R_{y_s}(0)$ are constants for all possible input signals.

We start by showing how M can be estimated. The estimator for M is easily derived by realizing that the following relation holds:

$$\mathcal{O}M = [R_{y_s}^T(1) \ \dots \ R_{y_s}^T(N-1)]^T \quad \text{with (16)}$$

$$\mathcal{O} = [C^T \ (CA)^T \ \dots \ (CA^{N-2})^T]^T. \quad (17)$$

In the previous equation, the covariance function $R_{y_s}(i)$ is unknown, but it can be estimated using the available test data. The covariance function $R_{y_s}(i)$ would be easy to estimate if the signal $y_s(k)$ would be available. If A is stable, the signal $y_s(k)$ can be approximately recovered using:

$$y_s(k) \approx \tilde{y}_s(k) = y(k) - \sum_{i=0}^{k-1} CA^{k-1-i}Bu(i) \quad \forall k \gg 1. \quad (18)$$

This approximate signal $\tilde{y}_s(k)$ can then be used to estimate the covariance function $R_{y_s}(i)$. In the remainder we shall denote the estimated covariance function by $\hat{R}_{y_s}(i)$.

Denote the left inverse of $\mathcal{O} \in \mathbb{R}^{(N-1)p \times n}$ as \mathcal{O}^\dagger , i.e. $\mathcal{O}^\dagger = (\mathcal{O}^T\mathcal{O})^{-1}\mathcal{O}^T$. Using the estimated covariance function $\hat{R}_{y_s}(i)$, a natural estimator for M is:

$$\hat{M} = \mathcal{O}^\dagger \left[\hat{R}_{y_s}(1) \ \dots \ \hat{R}_{y_s}^T(N-1) \right]^T. \quad (19)$$

Using the estimator for M and the estimate $\hat{R}_{y_s}(0)$ of $R_{y_s}(0)$, an approximate Kalman filter

can be constructed using (14), after replacing \hat{M} for M and $\hat{R}_{y_s}(0)$ for $R_{y_s}(0)$.

3. IMPROVED ESTIMATION

If the technique described in the previous section is directly used to obtain an approximate Kalman filter, the accuracy of the estimates of the obtained Kalman filter can be very poor if the matrix \mathcal{O} has one or several very small singular values. In this section we will show how (given that \mathcal{O} has one or several small singular values) we can still obtain good state estimates. In this section we will thus use the following assumption:

Assumption 1. Denote the decreasing singular values of \mathcal{O} as $\sigma_1 \geq \dots \geq \sigma_l \geq \dots \geq \sigma_n$. We will assume that $\sigma_i \ll 1 \quad \forall \quad i > l$.

We will start by explaining why the estimated states can be very poor under this assumption. The poor quality of estimated states can be easily explained if we consider the estimation error for \hat{M} . Denote the SVD of \mathcal{O} as:

$$\mathcal{O} = U \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_n & \\ & & & 0 \end{bmatrix} V^T \quad (20)$$

in which $U \in \mathbb{R}^{(N-1)p \times (N-1)p}$, $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices. Using this notation, the estimation error for \hat{M} can be written as:

$$\hat{M} - M = [V_{1:l} \quad V_{l+1:n}] \times \begin{bmatrix} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & & \sigma_n^{-1} \end{bmatrix} \begin{bmatrix} U_{1:l}^T \\ U_{l+1:n}^T \\ U_{n+1:(N-1)p}^T \end{bmatrix} \Delta R, \quad (21)$$

with

$$\Delta R = \begin{bmatrix} \hat{R}_{y_s}(1) - R_{y_s}(1) \\ \vdots \\ \hat{R}_{y_s}(N-1) - R_{y_s}(N-1) \end{bmatrix}.$$

and $V_{i:j}$ and $U_{i:j}$ matrices consisting of the columns i through j of V and U , respectively. From (21) we see that an error ΔR in the direction of U_i corresponds to an error $\hat{M} - M$ in the direction V_i after multiplication with σ_i^{-1} . Due to Assumption 1 small errors ΔR in the directions of $U_{l+1:n}$ induce large errors $\hat{M} - M$ in the directions of $V_{l+1:n}$. This relatively large error for $\hat{M} - M$ can cause severe state estimation errors if an approximate Kalman filter is obtained by substituting \hat{M} in (14).

In the remainder of this section we shall show how a good estimate of the state can be obtained under

Assumption 1. When deriving this estimator, it is convenient to first apply the following similarity transform:

$$\tilde{x}(k) = V^T x(k) = \begin{bmatrix} V_{1:l}^T \\ V_{l+1:n}^T \end{bmatrix} x(k) = \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \end{bmatrix} \quad (22)$$

The state $\tilde{x}_1(k)$ is thus the component of the original state in the directions of $V_{1:l}$, and $\tilde{x}_2(k)$ corresponds to the component of $x(k)$ in the directions $V_{l+1:n}$. After having derived an estimate $\hat{\tilde{x}}(k)$ the corresponding estimate $\hat{x}(k)$ can be simply computed via $\hat{x}(k) = V \hat{\tilde{x}}(k)$. Due to the similarity transform, the various state related matrices also need to be transformed:

$$\tilde{A} = V^T A V = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \quad (23)$$

$$\tilde{C} = C V = \begin{bmatrix} \tilde{C}_1 & \tilde{C}_2 \end{bmatrix} \quad (24)$$

$$\tilde{M} = V^T M = \begin{bmatrix} \tilde{M}_1 \\ \tilde{M}_2 \end{bmatrix} \quad (25)$$

$$\tilde{\Sigma}(k) = V^T \Sigma(k) V = \begin{bmatrix} \tilde{\Sigma}_{11}(k) & \tilde{\Sigma}_{12}(k) \\ \tilde{\Sigma}_{21}(k) & \tilde{\Sigma}_{22}(k) \end{bmatrix} \quad (26)$$

$$\tilde{K}(k) = V^T K(k). \quad (27)$$

Note that after the similarity transform, \tilde{M}_1 corresponds to the directions in which M can be accurately estimated using (19) and \tilde{M}_2 corresponds to the directions in which M is difficult to estimate.

The estimator for $\tilde{x}(k)$ will be derived in two steps, first we shall create an estimator for $\tilde{x}_1(k)$ then an estimator for $\tilde{x}_2(k)$ will be discussed.

3.1 Estimator for $\tilde{x}_1(k)$

Our estimator for the state component $\tilde{x}_1(k)$ is based on the following proposition:

Proposition 2. Define \tilde{M}_l as:

$$\tilde{M}_l = [\tilde{M}_1^T \quad 0]^T.$$

Suppose that $\sigma_i = 0$ for $i > l$ and assume we construct a Kalman filter using (14) using \tilde{M}_l instead of \tilde{M} . Denote the vectors and matrices $\tilde{\Sigma}(k)$, $\tilde{R}_e(k)$, $\hat{\tilde{x}}(k)$ that are computed using \tilde{M}_l as $\tilde{\Sigma}_l(k)$, $\tilde{R}_{e,l}(k)$, $\hat{\tilde{x}}_l(k)$. Then it holds that:

- $R_{e,l}(k) = R_e(k)$
- $\tilde{\Sigma}_{11,l}(k) = \tilde{\Sigma}_{11}(k)$
- $\hat{\tilde{x}}_{1,l}(k) = \hat{\tilde{x}}_1(k)$.

Proof. See (Bos *et al.*, 2005). \square

This proposition implies that the estimation result $\hat{\tilde{x}}_1(k)$ does not change if, instead of using the exact

Kalman filter, we would use a Kalman filter obtained by replacing \tilde{M} by \tilde{M}_l in (14). Combining the result with the fact that we can accurately estimate $\hat{\tilde{M}}_1$ with $\hat{\tilde{M}}_1 = V_{1,l}^T \hat{M}$, a natural estimator for $\tilde{x}_1(k)$ is thus the approximate Kalman Filter computed with (14) where \tilde{M} is replaced by $[\hat{\tilde{M}}_1 \ 0]^T$.

Note that even though Proposition 2 holds for $\sigma_i = 0 \ \forall i > l$, we assume that it still approximately holds under Assumption 1.

3.2 Estimator for $\tilde{x}_2(k)$

Now that we have an estimator for $\tilde{x}_1(k)$ we still need to construct an estimator for the remaining states $\tilde{x}_2(k)$. The optimal estimate for $\tilde{x}_2(k)$ corresponding to the optimal Kalman filter is:

$$\hat{\tilde{x}}_2(k+1) = \tilde{A}_{21} (\hat{\tilde{x}}_1(k) + \mathbb{E}[(\tilde{x}_1(k) - \hat{\tilde{x}}_1(k))|e(k)]) + \tilde{A}_{22} (\hat{\tilde{x}}_2(k) + \mathbb{E}[(\tilde{x}_2(k) - \hat{\tilde{x}}_2(k))|e(k)]).$$

Both conditional expectations can be expressed as complex functions of \tilde{M} , $\tilde{\Sigma}(k)$, \tilde{C} , $\tilde{R}_e(k)$ and $e(k)$ using the following proposition:

Proposition 3. Suppose that $\sigma_i = 0$ for $i > l$, it holds that:

$$\begin{aligned} \tilde{A}_{21} \mathbb{E}[(\tilde{x}_1(k) - \hat{\tilde{x}}_1(k))|e(k)] = \\ \tilde{A}_{21} (\tilde{A}_{11}^{-1} \tilde{M}_1 - \tilde{\Sigma}_{11}(k)) \tilde{C}_1^T R_e^{-1}(k) e(k). \end{aligned} \quad (28)$$

$$\begin{aligned} \tilde{A}_{22} \mathbb{E}[(\tilde{x}_2(k) - \hat{\tilde{x}}_2(k))|e(k)] = \\ (\tilde{M}_2 - \tilde{A}_{21} \tilde{A}_{11}^{-1} \tilde{M}_1 - \tilde{A}_{22} \tilde{\Sigma}_{21}(k) \tilde{C}_1^T) R_e^{-1}(k) e(k). \end{aligned}$$

Proof. See (Bos *et al.*, 2005). \square

In this proposition we see that the conditional expectation for the error in $\tilde{x}_1(k)$ can be approximately computed, since the expression contains only elements that are either known, or can be accurately estimated.

Computing the expectation $\mathbb{E}[(\tilde{x}_2(k) - \hat{\tilde{x}}_2(k))|e(k)]$ requires an accurate estimate of $\hat{\tilde{M}}_2$, which is unavailable. This means we cannot reliably compute the conditional expectation for the error in $\tilde{x}_2(k)$. Therefore we chose not to use this conditional expectation and thus the best reliably obtainable estimator for \tilde{x}_2 becomes:

$$\begin{aligned} \hat{\tilde{x}}_2(k+1) = \tilde{A}_{21} (\hat{\tilde{x}}_1 + \mathbb{E}[(\tilde{x}_1(k) - \hat{\tilde{x}}_1(k))|e(k)]) \\ + \tilde{A}_{22} \hat{\tilde{x}}_2. \end{aligned} \quad (29)$$

Combining (28) and (29), a natural estimator for $\tilde{x}_2(k+1)$ is obtained by replacing \tilde{M}_1 with the estimate $\hat{\tilde{M}}_1$.

$$\begin{aligned} \hat{\tilde{x}}_2(k+1) = \tilde{A}_{21} \hat{\tilde{x}}_1(k) + \tilde{A}_{22} \hat{\tilde{x}}_2(k) \\ + \tilde{A}_{21} (\tilde{A}_{11}^{-1} \hat{\tilde{M}}_1 - \tilde{\Sigma}_{11,l}(k)) \tilde{C}_1^T R_{e,l}^{-1}(k) e(k) \end{aligned} \quad (30)$$

Summarizing, an approximate Kalman filter can be estimated using the procedure described in section 2.2. However the resulting filter can produce poor estimates if some of the singular values of \mathcal{O} are small. In this particular case, a better and more robust filter can be constructed. This filter is constructed by first applying a similarity transform $\tilde{x}(k) = V^T x(k)$. Then the elements corresponding to the first l columns of V ($= \tilde{x}_1(k)$) are estimated using an approximate Kalman filter estimate that is constructed by inserting $[\hat{\tilde{M}}_1^T \ 0]^T$ for M in (14). Using the results of this filter, $\tilde{x}_2(k)$ can be estimated using (30).

4. SIMULATION EXAMPLE

In this simulation example 500 measurements y were generated using a randomly chosen system. This system can be written in the form (1)-(2), with:

$$\begin{aligned} A = \begin{bmatrix} -0.1821 & -0.3703 & 0.2848 & -0.2656 \\ -0.5478 & -0.0379 & -0.4672 & -0.2570 \\ -0.2200 & -0.2308 & 0.0153 & 0.0206 \\ -0.1037 & -0.3984 & 0.3201 & 0.6234 \end{bmatrix} \\ B = [1 \ 1 \ 2 \ 1]^T \quad C = [1 \ 0 \ 0 \ 0] \\ Q = I, \quad R = 1, \quad S = [0 \ 0 \ 0 \ 0]^T. \end{aligned}$$

The goal of the simulation example is to test, assuming that the matrices A , B , C , are known, and the set of test measurements y is available, if it possible to construct a Kalman filter using the procedure outlined in the previous sections. The accuracy of the filter designed using this procedure will be compared with the accuracy of the optimal filter, that can be constructed since in the simulations the true values of Q , R and S are also available.

The first step in constructing an approximate Kalman filter is to estimate the auto-covariance function of the stochastic part of the measurements. After subtracting the deterministic part of the system using (18), the last 250 measurements in the sequence y_s were used to estimate the autocovariance function of the stochastic part of measurements, $\hat{R}_{y_s}(k)$ using an $AR(10)$ model.

After estimating the auto-covariance function $\hat{R}_{y_s}(k)$, the estimated matrix \hat{M} can be constructed using (19). Finally the approximate Kalman filter can be designed by substituting the estimated values $\hat{R}_{y_s}(0)$ and \hat{M} for $R_{y_s}(0)$ and M in (14) as was described in section 2.2.

In order to evaluate the performance of the approximate Kalman filter design method, the ac-

Table 1. Mean squared one-step-ahead prediction errors \mathcal{E} of both the optimal Kalman filter and approximations obtained using methods outlined in sections 2 and 3. Results were averaged over 30 runs. Values in brackets denotes the standard deviation over the 30 runs.

Filter	\mathcal{E}
Optimal filter	5.49 (0.030)
Direct approx. filter (sect. 2)	561 (906)
Improved filter $l = 1$ (sect. 3)	5.60 (0.24)
Improved filter $l = 2$ (sect. 3)	5.56 (0.20)
Improved filter $l = 3$ (sect. 3)	6.09 (0.89)

accuracy of the resulting filters are compared with the optimal Kalman filter. The comparison was conducted by using the true system to generate additional series of 1000 measurements. The error measure that is used to compare the performance is the averaged one-step-ahead prediction error, which will be denoted by \mathcal{E} : $\mathcal{E} = \frac{1}{N} \sum_{k=0}^{N-1} (x(k) - \hat{x}(k))^2$.

The entire simulation procedure (data generation, filter design, performance evaluation) was repeated 30 times, the averaged errors can be found in Table 1.

As can be seen in Table 1, the average accuracy of the resulting approximate filter is poor, and the standard deviation of the prediction error is very large, indicating that the accuracy strongly varies. This effect of varying accuracy may be explained by the theory of section 3 if \mathcal{O} has small singular values. To verify this the singular values were computed. The singular values for \mathcal{O} is this example are:

$$\sigma_1 = 1.22 \quad \sigma_2 = 0.77 \quad \sigma_3 = 0.26 \quad \sigma_4 = 0.0047$$

The poor results of the approximate Kalman filter are likely caused by the smallest value, as illustrated in (21).

After the direct method of section 2.2, the improved filter of section 3 was also tested. In order to apply the results of section 3, we first need to apply the similarity transform $\tilde{x}(k) = V^T x(k)$ and choose the number of elements in $\tilde{x}_1(k)$ and $\tilde{x}_2(k)$. Normally a single partition would be chosen using the computed singular values but in this example we chose to try all options. We shall denote the number of elements in $\tilde{x}_1(k)$ as l . The results using the improved estimator can also be found in Table 1.

Results in Table 1 show a vast improvement if the improved filter of section 3 is used (using $l = 3$) instead of the direct approximation of section 2.2. Apparently the error in $\hat{M} - M$ corresponding to the smallest singular value was the main reason for the severe prediction error of the filter of section 2.2. This large error is avoided in the improved estimator.

Using $l = 2$ instead of $l = 3$ again increases the accuracy of the resulting filter. This indicates that

the prediction error caused by the error in $\hat{M} - M$ corresponding to the small value of σ_3 was larger than the decrease in possibly obtainable accuracy due to the suboptimal estimate for $\tilde{x}_2(k)$.

Finally using $l = 1$ the error increases slightly compared to the filter designed using $l = 2$. This was to be expected since σ_2 is not that small, which means that we ignore information contained in M that we could have used to improve the prediction accuracy of the state.

5. CONCLUSIONS

Designing a state filter for first principles models is difficult since first principles modelling often does not provide the required information on the noise covariances. This paper introduces a method for constructing a Kalman filter without resorting to manual tuning or nonlinear optimization. The approximate Kalman filter is constructed using the estimated autocovariance of the outputs $y(k)$. The sensitivity of the design method to estimation errors in the estimated autocovariance function is also discussed and methods to reduce this sensitivity are given. The proposed method was illustrated in a simulation example.

REFERENCES

- Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering*. Prentice-Hall. Englewood Cliffs, N.J.
- Bos, R., X. Bombois and P.M.J Van den Hof (2005). Designing a Kalman filter for poorly observable systems without noise covariance information. Technical Report 05-003. Delft Center for Systems and Control, Delft University of Technology.
- Carew, B. and P.R. Bélanger (1973). Identification of Optimum Filter Steady-State Gain for Systems with Unknown Noise Covariances. *IEEE Trans. on Automatic Control* **6**(18), 582–587.
- Ljung, L. (1999). *System Identification-Theory for the User*. second ed.. Prentice Hall. Upper Saddle River.
- Magill, D.T. (1965). Optimal Adaptive Estimation of Samples Stochastic Processes. *IEEE Trans. On Automatic Control* **10**(4), 434–439.
- Mehra, R.K. (1970). On the Identification of Variances and Adaptive Kalman Filtering. *IEEE Trans. On Automatic Control* **15**(2), 175–184.
- Mehra, R.K. (1973). Approaches to Adaptive Filtering. *IEEE Trans. On Automatic Control* **17**(5), 693–698.
- Zhou, J. and R.H. Luecke (1994). Estimation of the Covariances of the Process Noise and Measurement Noise for a Linear Discrete Dynamic System. *Computers chem. Engng.* **19**(2), 187–195.