PARAMETER ESTIMATION IN A CLASS OF NONLINEAR STATE - SPACE MODELS

Mihai Enescu and Visa Koivunen

Signal Processing Laboratory, SMARAD CoE
Helsinki University of Technology
P.O. Box 3000, FIN-02015 HUT, Finland

ABSTRACT

Many applications rely on a state-space description of the input-output dynamics of the system. In this paper we address the problem of parameter estimation in a class of nonlinear state-space models. The state transition coefficient and the process noise variance are estimated from the received observations in a recursive manner. Some properties of the characteristic function (in this case for normally distributed variables) are exploited. Two applications are presented, estimation of the carrier frequency offset in OFDM and propagation path parameter estimation in channel sounding. Illustrative examples show the reliable performance of the introduced parameter estimation technique.

1. INTRODUCTION

State-space models have been extensively used in various applications. When the linearity of the system and the Gaussianity of the noise are assumed, this type of models lead to optimal filtering using the Kalman Filter. In many applications such as communications and radar, the system model may be nonlinear. In such case the Extended Kalman Filter (EKF) can be used.

The optimality of the Kalman Filter (KF) is partly based on the fact that all the parameters describing the state space model are known except for the state which has to be estimated. When state-space models are used in various applications, a typical assumption is that the state transition scalar (matrix) and the noise statistics are known. However, this may not be the case in practice. In real applications one may have to accurately estimate these parameters in order to obtain a close to optimal performance of the KF. The problem of recursively estimating the state transition matrix and the noise statistics was considered in [1] for a linear model.

In this paper, the goal is to determine the state transition scalar and the noise statistics in a class of non-linear models. This type of models are encountered in various communications applications. For example, in OFDM transmission [2], the carrier frequency offset may be modeled using a nonlinear state-space formulation [4]. Another potential application is in channel sounding and propagation parameter estimation [5]. The main challenge is in modeling real-world measured data where no prior information on state transition coefficients and statistics of process noise is available. In both applications, i.e. OFDM and channel sounding, when applying EKF to estimate the state, the parameters describing the state-space model are commonly selected by trial and error. These parameters have a crucial role in the performance of the estimator, as has been observed in [5]. Hence, it is highly desirable to estimate both the state transition coefficient and the noise statistics from the data.

The rest of the paper is organized as follows. In the next section we briefly present the system model and in section three we describe how the state scalar and noise statistics are estimated from the received data. In section four we focus our attention on estimating the noise statistics. Illustrative examples are presented in section five.

2. MODEL AND ASSUMPTIONS

Let us start by introducing the scalar state-space model used in this paper:

$$\theta_{k+1} = a\theta_k + v_k \quad (1)$$
$$z_k = h(\theta_k) + w_k, \quad (2)$$

where $h(\theta_k) = e^{j\omega\theta_k}$. This specific nonlinearity is needed in our applications that are later described in the paper. The noise sequences v and w are zero mean, mutually uncorrelated and also uncorrelated with the state. We assume that v is real noise and w is circular white noise, hence $E[w_t w_s^*] = \sigma_w^2 \delta_{t,s}$ and $E[w_t w_s] = 0$ for any t and s. We also assume that $|a| < 1$. It can be easily verified that if v_k is zero mean, the state θ_k is also zero mean. The task is to estimate a and the variances of v_k and w_k, from the received data z_k.

Let us start by considering equation (1). Multiplying from the right by v_k and taking the expectation we obtain:

$$E[\theta_{k+1}v_k] = aE[\theta_k v_k] + E[v_k^2] \quad (3)$$

We can write $E[\theta_k v_k] = E[\theta_k]E[v_k] = 0$ since θ_k depends only on past values (up to $k-1$) of v_k. We obtain that:

$$E\left[\theta_{k+1}v_k\right] = \sigma_v^2. \quad (4)$$

It can be also shown that $\sigma_v^2 = \sigma_\theta^2(1 - a^2)$. If θ is a Gaussian distributed random parameter obeying $\mathcal{N}(0, \sigma_\theta^2)$, then we have the following property of the characteristic function:

$$\Phi_\theta(\omega) = E\left[e^{j\omega\theta}\right] = e^{-\frac{1}{2}\sigma_\theta^2\omega^2}. \quad (5)$$

For other possible distributions of θ, one can use the corresponding characteristic function [3].

3. ESTIMATION

In this section we introduce the method for estimating the state transition scalar and the noise statistics. For this purpose, considering the equations (1)-(2), let us start by investigating $E\left[z_k z_k^*\right]$:

$$E\left[z_k z_k^*\right] = 1 + E\left[w_k w_k^*\right].$$

As a result:

$$\boxed{\hat{\sigma}_w^2 = \frac{1}{k-2}\sum_{n=2}^{k} z_n z_n^* - 1.} \quad (6)$$

We continue by inspecting $E\left[z_k z_k\right]$:

$$E\left[z_k z_k\right] = E\left[e^{2j\omega\theta_k}\right] + E\left[w_k w_k\right].$$

The noise is assumed to be circular and white, hence $E\left[w_k w_k\right] = 0$. Using the definition of the characteristic function we obtain:

$$\frac{1}{k-2}\sum_{n=2}^{k} z_n z_n = e^{-2\omega^2\hat{\sigma}_\theta^2} \quad (7)$$

from which we get:

$$\boxed{\hat{\sigma}_\theta^2 = -\frac{1}{2\omega^2}\ln\left[\frac{1}{k-2}\sum_{n=2}^{k} z_n z_n\right].} \quad (8)$$

Let us now investigate $E\left[z_k z_{k-1}^*\right]$:

$$\begin{aligned} E\left[z_k z_{k-1}^*\right] &= E\left[\left(e^{j\omega\theta_k} + w_k\right)\left(e^{j\omega\theta_{k-1}} + w_{k-1}\right)^*\right] \\ &= E\left[e^{j\omega(\theta_k - \theta_{k-1})}\right] \\ &= E\left[e^{j\omega(a-1)\theta_{k-1}}\right] E\left[e^{j\omega v_{k-1}}\right]. \end{aligned} \quad (9)$$

This means that the previous result can be written as:

$$E\left[z_k z_{k-1}^*\right] = \Phi_{\theta_{k-1}}[\omega(a-1)]\Phi_{v_{k-1}}[\omega] \quad (10)$$

which can be further written as:

$$\frac{1}{k-2}\sum_{n=2}^{k} z_n z_{n-1}^* = e^{-\frac{1}{2}\hat{\sigma}_\theta^2[\omega(a-1)]^2} e^{-\frac{1}{2}\hat{\sigma}_v^2[\omega]^2}. \quad (11)$$

Using the result that $\sigma_v^2 = \sigma_\theta^2(1 - a^2)$, the previous relation can be rewritten as:

$$-\frac{1}{2}\hat{\sigma}_\theta^2\omega^2(\hat{a} - 1)^2 - \frac{1}{2}\hat{\sigma}_\theta^2\omega^2(1 - \hat{a}^2) = \ln\left[\frac{1}{k-2}\sum_{n=2}^{k} z_n z_{n-1}^*\right] \quad (12)$$

$$-2\hat{\sigma}_\theta^2\pi^2\left[\hat{a}^2 - 2\hat{a} + 1 + 1 - \hat{a}^2\right] = \ln\left[\frac{1}{k-2}\sum_{n=2}^{k} z_n z_{n-1}^*\right]. \quad (13)$$

Finally, the state transition scalar is given by:

$$\boxed{\hat{a} = 1 + \frac{\ln\left[\frac{1}{k-2}\sum_{n=2}^{k} z_n z_{n-1}^*\right]}{\hat{\sigma}_\theta^2\omega^2}.} \quad (14)$$

4. APPLICATIONS

In the previous derivations, for simplicity, we have used a scalar model. The results can be easily extended to a vector state-space model. In this section we apply the method to two potential applications, OFDM and channel sounding.

4.1. OFDM

In OFDM, the carrier frequency offset (CFO) may be modeled in a state-space form, leading to a nonlinear model [4]. The CFO, denoted here by θ is assumed to obey the following dynamics over time:

$$\theta_k = a\theta_{k-1} + v_k, \quad (15)$$

where $|a| < 1$ for stability, v_k is the Gaussian state noise with variance σ_v^2. The quantity $\theta \in [0, 1)$ is referred to as the normalized frequency offset with respect to the intercarrier spacing.

The observation equation, which gives us the input-output relation of the system is:

$$\mathbf{z}_k = \mathcal{G}\left(\theta_k\right) + \mathbf{w}_k, \quad (16)$$

where $\mathcal{G} : [0, 1) \to \mathbf{C}_N$ defined as $\mathcal{G}(\theta_k) = \mathbf{C}_\theta \widetilde{\mathbf{H}}_k \tilde{\mathbf{x}}_k$ is a non-linear function of the state variable θ. $\widetilde{\mathbf{H}}_k$ is the $N \times N$ frequency-domain circulant channel matrix, with the (i,j)th entry given by $h_{(i,j) \mod N}(k)$. The diagonal matrix $\mathbf{C}_\theta$ of size $N \times N$ introduces the frequency offset and is defined as $\mathbf{C}_\theta = \text{diag}\left\{\exp\left(j\frac{2\pi n\theta}{N}\right)\right\}$ with $n = L, \ldots, N + L - 1$, where L is the length of the cyclic prefix.

Equations (15)-(16) form the nonlinear state space model on which the estimation of the state, θ, is performed.

4.2. Channel Sounding

In channel sounding, one of the key tasks is to estimate parameters describing the propagation model (for example the direction of departure at the mobile station, the time delay of arrival or the complex paths weights) [5]. Using state-space models improves considerably the propagation path parameter tracking and thus, enhances the accuracy of channel analysis. Parameter tracking can considerably increase the path parameter pairing over time. This can be used to answer more precisely the questions of path lifetime in different measurement scenarios. It also allows detailed investigation of the parameter statistics of scattering clusters. In this application the state-space model is given by:

$$\theta_{p,k+1} = a\theta_{p,k} + v_{p,k} \tag{17}$$

$$\mathbf{z}_k = \mathbf{g}(\theta_{p,k}) + \mathbf{w}_k, \tag{18}$$

where $\theta_{p,k+1}$ is the parameter of interest (e.g. mean path delays, angles of arrival, direction of departure). This parameter is mapped to the measurements by a nonlinear function that is parameter dependent, for example the delay of the path p is mapped to the measurements by:

$$\mathbf{g}\left(\theta_{p,k}\right) = \frac{1}{\sqrt{N_r}}\left[e^{-j\theta_{p,k}\frac{N_r-1}{2}} \cdots 1 \cdots e^{+j\theta_{p,k}\frac{N_r-1}{2}}\right]^T \in \mathbb{C}^{M\times 1} \tag{19}$$

Estimation of the channel parameters using the model (17)-(18) has been performed in [5]. However, the state transition coefficient and state noise variance were selected by trial and error. Moreover, it was observed that the estimator in highly sensitive to the values of these parameters. This is due to the fact that the technique is operating on real world measured data, where the dynamics change over a long period of time (one measurement can reach up to several minutes). Hence, there is a need to estimate and possibly track these parameters from the received data.

Using the rationale presented in Section 3, it can be shown that an estimate of the transition scalar is given by:

$$\hat{a} = 1 + \frac{1}{\sigma_\theta^2||(\mathbf{m}\odot\mathbf{m})||^2}(\mathbf{m}\odot\mathbf{m})^T \ln\left[\mathbf{q}\right] \tag{20}$$

where

$$\mathbf{q} = diag\left\{\frac{1}{k-2}\sum_{n=2}^{k}\mathbf{z}_n\mathbf{z}_{n-1}^*\right\}, \tag{21}$$

$diag$ operation is selecting the diagonal elements of a matrix, the result being a vector, and $\odot$ is denoting the Hadamard product. We have used also the notation $\mathbf{m} = \left[\frac{N_r-1}{2}\cdots\frac{N_r+1}{2}\right]^T$.

5. EXAMPLES

In our first example, all the parameters involved in the model are known. This type of simulation is useful since we can verify the convergence of estimated parameters to the true ones. Let us start by generating 1000 samples according to equations (1)-(2). The state noise v is real white Gaussian noise, the observation noise w is circular complex white Gaussian noise. The two noise sequences are mutually uncorrelated and also uncorrelated with the state. The state transition coefficient is $a = 0.9$. In Figure 1 we have depicted with the solid line the estimated observation noise variance Figure 1-(a) and the estimated state variance, Figure 1-(b). Dashed line represents the true values. We have used equation (6) for the measurement noise variance estimation and equation (8) for the state noise variance estimation. We have estimated the state transition scalar using equation (14), the result is depicted in figure 2.

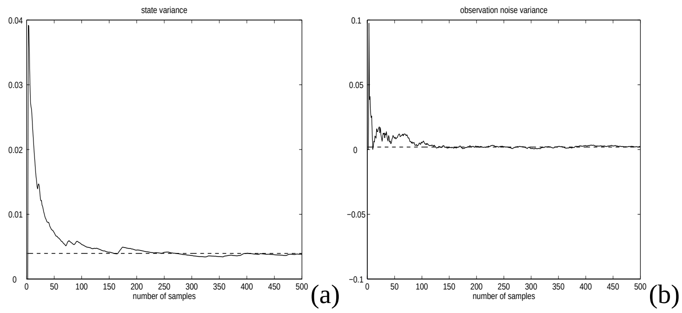

Fig. 1. Tracking result for: (a) process noise variance, true value is $\sigma_v^2 = 0.004$, (b) measurement noise variance, true values is $\sigma_w^2 = 0.0019$. Dash line represents true values, continuous line represents estimated values.

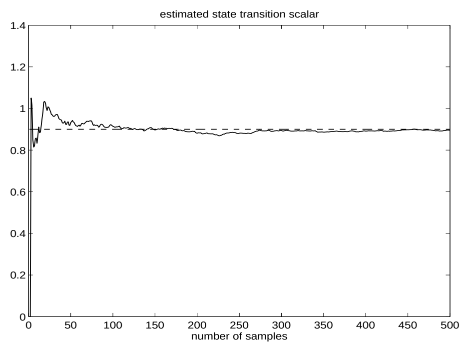

Fig. 2. Estimated value of the state transition coefficient, true value is $a = 0.9$. Dash line represents true value, continuous line represents estimated value.

The second example is related to the channel sounding application. In our study we have considered the normalized Direction Of Arrival (DoA) azimuth as the parameter of interest. The results obtained by the proposed technique and ML estimator [6] are presented. The received signal is modeled using (18). Real sounder data is used [6]. Hence the true value is not known. For this purpose, the state has been considered known (it has been already estimated from the real sounder measurements using a ML technique [6]).

It is important to note that the parameter estimation is applied in advance to the data to acquire the state-space parameters. The estimated state transition coefficient is presented in figure 3. In this case the state-transition parameter is not know. One should use a trial and error approach in order to find good approximations for the state-space model parameters, as has been used in [5]. Moreover, this example shows the estimation over about 10.5 seconds. Typically the channel sounding data recordings can span over several minutes, time in which the parameters change due to the movement of the mobile station and the dynamic environment. This makes the trial and error selection of the parameters a tedious task, if not impossible.

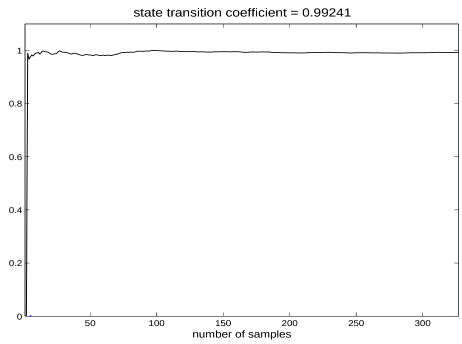

Fig. 3. Estimated value of the state transition coefficient.

In this example the measurement noise is known, hence we can depict a converge to the true value (figure 4). The estimated state using estimated state-space model parameters is depicted in Figure 5, along with the true state.

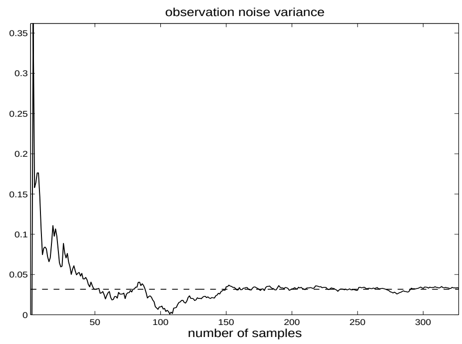

Fig. 4. Measurement noise variance. Estimated (continuous line) versus true value (dash line).

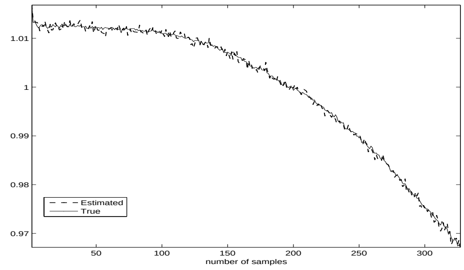

Fig. 5. Estimated normalized DoA azimuth using estimated parameters (continuous line) versus normalized DoA azimuth estimated with the ML method (dash line).

6. CONCLUSIONS

In this paper we have introduced a parameter estimation method applicable for a class of non-linear state-space models. The state transition coefficient and the process noise variance are estimated from the received observations in a recursive manner, exploiting some properties of the characteristic function. The reliable performance of the technique has been demonstrated in illustrative examples.

7. REFERENCES

[1] M. Enescu, V. Koivunen, "On the estimation of state transition matrix and noise statistics in state-space models", Vehicular Technology Conference, 2002. Proceedings. IEEE VTC 2002-Fall. Vol. 4, pp.: 2192 -2196, 2002.

[2] R. van Nee, R. Prasad, "OFDM for Wireless Multimedia Communications", Artech House universal personal communications library, 2000.

[3] A. Papoulis, "Probability, Random Variables, and Stochastic Processes", third edition, McGraw-Hill, 1991.

[4] T. Roman, M. Enescu, V. Koivunen, "Joint time-domain tracking of channel and frequency offset for OFDM systems", IEEE Workshop on Signal Processing Advances in Wireless Communications, pp.: 605 - 609, 2003.

[5] A. Richter, M. Enescu, V. Koivunen, "A state space approach to propagation path parameter estimation and tracking", IEEE Workshop on Signal Processing Advances in Wireless Communications, 2005.

[6] A. Richter, M. Landmann, and R. S. Thomä, "Maximum likelihood channel parameter estimation from multidimensional channel sounding measurements", in Proc. IEEE Vehicular Techn. Conf., Cheju, South Korea, Apr. 2003.