EVOLUTIVE METHOD BASED ON A GENERALIZED EIGENVALUE DECOMPOSITION TO ESTIMATE TIME VARYING AUTOREGRESSIVE PARAMETERS FROM NOISY OBSERVATIONS

Hiroshi Ijima¹, Julien Petitjean² and Eric Grivel²

¹ Faculty of Education, Wakayama University, 640-8510 Wakayama, Japan
² Université de Bordeaux, IPB, UMR CNRS 5218 IMS, 33400 Talence, France

ABSTRACT

A great deal of interest has been paid to the estimation of time-varying autoregressive (TVAR) parameters. However, when the observations are disturbed by an additive white measurement noise, using standard least squares methods leads to a weight-estimation bias. In this paper, we propose to jointly estimate the TVAR parameters and the measurement-noise variance from noisy observations by means of a generalized eigenvalue decomposition. It extends to the TVAR case an off-line method that was initially proposed for AR parameter estimation from noisy observations. A comparative study is then carried out with existing methods such as the recursive errors-in-variable approach and Kalman based algorithms.

Index Terms— Time-varying autoregressive (TVAR) model, parameter estimation, parameter tracking, generalized eigenvalue decomposition, least squares.

1. INTRODUCTION

Time-varying autoregressive (TVAR) models have been used in radar processing to model the clutter [1], in biomedical applications [2], in speech enhancement or dereverberation to model time-varying channels related to moving speakers, etc.

For the last years, two kinds of methods have been proposed to estimate the TVAR parameters. On the one hand, the so-called “stochastic” family includes recursive methods such as the recursive least square (RLS) with forgetting factor and Kalman filtering [3]. In this latter case, the state vector stores the TVAR parameters to be tracked. When dealing with the “evolutive method” family (or deterministic regression based method), the parameter evolutions are assumed to be weighted linear combinations of some basis functions such as power of time, Legendre polynomials and Fourier functions. The TVAR parameters can be deduced by summing the weighted basis-functions provided that the weights are available or preliminary estimated in the least squares (LS) sense for instance. For more details, the reader may refer to various pioneering works by Grenier in the field of speech analysis, transmission and recognition like [4], [5].

However, when the TVAR process is disturbed by an additive measurement noise, using the above approaches leads to an estimation bias. To our knowledge, few papers deal with this issue. When considering stochastic methods and more particularly extended Kalman filtering and sigma point Kalman filters (SPKF) [6], the state vector stores the TVAR parameters to be tracked and the TVAR process to be estimated from noisy observations. Nevertheless, these methods require the a priori knowledge of the noise variances. Cumulants can be also used especially in the non-Gaussian case, but this may lead to poor results because their estimates may have a high variance when few data are available. In [7], two of the authors and others have recently proposed a recursive errors-in-variable based method. It combines a Newton-Raphson algorithm to iteratively estimate the variance of the additive noise and a sliding offline noise compensated Yule-Walker based method to deduce the TVAR parameters.

In this paper, we propose an alternative evolutive method to estimate the weights related to the TVAR parameters from the noisy observations by extending the method initially proposed by Davila in [8]. The relevance of the proposed method is then studied by means of a comparative study with standard Yule-Walker methods [7], standard Kalman filtering, and the EKF based approach.

2. PROBLEM STATEMENT

Let \( x(k) \) be a \( p \) th order time-varying autoregressive (TVAR) process defined by:

\[
x(k) = - \sum_{n=1}^{p} a_n (k-n)x(k-n) + u(k)
\]

where \( u(k) \) is the zero-mean driving process with variance \( \sigma_u^2 \).

This process is assumed to be disturbed by an additive zero-mean white noise \( b(k) \) with variance \( \sigma_b^2 \) uncorrelated with the driving process \( u(k) \).
The observation data are hence given by:
\[ y(k) = x(k) + b(k). \]  (2)

Let us now introduce the following vectors:
\[ \phi_u(k) = [\alpha(k) \alpha(k-1) \ldots \alpha(k-p)]^T \]
\[ (a = x, y, \text{or } b) \]
\[ \tilde{\phi}_u(k) = \begin{bmatrix} u(k) & 0 & \ldots & 0 \end{bmatrix}^T \]
and the TVAR parameter vector:
\[ \tilde{\theta}(k) = [1 \ \alpha_1(k-1) \ldots \alpha_p(k-p)]^T = [1 \ \theta^T(k)]^T. \]

Then, equ. (1) can be expressed in the matrix form:
\[ (\tilde{\phi}_u(k) - \tilde{\phi}_u(k)) \tilde{\theta}(k) = 0. \]  (3)

Premultiplying equ. (3) by \((\tilde{\phi}_u(k) - \tilde{\phi}_u(k))\) and introducing \(R_x(k) = E[\tilde{\phi}_u(k)\tilde{\phi}_u^T(k)]\), we obtain:
\[ \left( R_x(k) - \text{diag} \left[ \begin{array}{c} \sigma^2_u \ 0 \ \ldots \ 0 \end{array} \right] \right) \tilde{\theta}(k) \] 
\[ = \begin{bmatrix} 0 \ 0 \ \ldots \ 0 \end{bmatrix} \]  (4)

The TVAR parameters could be estimated provided that \(R_x(k) - \text{diag}[\sigma^2_u \ 0 \ \ldots \ 0]\) is available. However, in all cases, only the observation sample covariance matrix \(R_x(k) = E[\tilde{\phi}_u(k)\tilde{\phi}_u^T(k)]\) can be computed.

Remark: Using the conventional Yule-Walker (YW) equations directly with noisy observations would lead to a biased estimate of \(\theta(k)\); indeed, one has:
\[ \tilde{\theta}(k) = -R_x^{-1}(k) \rho(k) = -E[\tilde{\phi}_u(k)\tilde{\phi}_u^T(k)]^{-1} \rho(k) \]
where \(\rho(k) = E[\tilde{\phi}_u(k)y(k)]\). In that case, it can be shown that the bias \(b(k)\) satisfies:
\[ b(k) = \sigma^2_b R_y^{-1}(k) \tilde{\theta}(k). \]  (5)

When using an evolutive method, the TVAR parameters are expressed as follows:
\[ a_n(k) = \sum_{j=0}^m \beta_{nj} f_j(k) \]  (6)

where \(\{f_j(z)\}_{j=0,1,\ldots,m}\) are basis functions and \(\{\beta_{nj}\}_{n=1,\ldots,p \text{ and } j=0,\ldots,m}\) are constant weights.

Let us now introduce the three following column vectors of size \(p(m+1):\)
\[ \psi_a(k) = [\Lambda^T_a(k-1) \ldots \Lambda^T_a(k-p)]^T (\alpha = X, Y, \text{or } B) \]
where
\[ \Lambda_a(k) = [f_0(k) \ \ldots \ f_m(k)]^T \] 
\( (\alpha = x, y, \text{or } b) \).

Then, let us consider the weight vector:
\[ \tilde{\theta}_w = \begin{bmatrix} 1 & \beta_1 & \ldots & \beta_m & \beta_{20} & \ldots & \beta_{pm} \end{bmatrix}^T = \begin{bmatrix} 1 & \theta^T_w \end{bmatrix}^T. \]

Combining (1), (2), and (6), we obtain:
\[ \psi_y(k) = \psi_x(k) + \psi_b(k) \]  (7)
\[ x(k) = -\psi_y^T(k)\theta_w + u(k). \]  (8)

Premultiplying both sides in (8) by \(\psi_x(k)\) and taking the expectation, we have:
\[ \theta_w = -(E[\psi_x(k)\psi_y^T(k)])^{-1} E[\psi_x(k)x(k)] \]
\[ = -R_x^{-1} x. \]  (9)

This relation can be seen as the Yule-Walker equations for the TVAR-model described in [1]. As \(b(k)\) is a zero-mean white process, substituting \(\psi_y(k)\) of (7) into (9) leads to:
\[ \theta_w = -(R_y - R_p)^{-1} r_y \]  (10)
where \(r_y = E[\psi_y(k)y(k)] = r_x, \quad R_y = E[\psi_y(k)\psi_y^T(k)]\), \(R_p = E[\psi_y(k)\psi_b^T(k)]\). \(F(j) = \begin{bmatrix} f_0^2(k-j) & \ldots & f_m(k-j) f_0(k-j) & \ldots & f_0(k-j) \end{bmatrix}\)
for \(j = 1, \ldots, p\). In that case, directly using equ. (9) with noisy observations leads to the following bias \(b_w:\)
\[ b_w = R_y^{-1} R_p \theta_w \]  (12)

In this paper, our aim is to propose a method to estimate the weights \(\theta_w\) by compensating the influence of the additive noise.

**3. NOISE SUBSPACE (NS) APPROACH FOR TVAR PARAMETER ESTIMATION**

Equ. (10) can be rewritten as follows:
\[ \begin{bmatrix} f_0(k-1) E(y(k-1)y(k)) \\ f_m(k-p) E(y(k-p)y(k)) \end{bmatrix} R_y \]
\[ -\sigma^2_b \begin{bmatrix} F(1) & \ldots & 0 \\ 0 & \ldots & F(p) \end{bmatrix} \tilde{\theta}_w = 0_{p(m+1) \times 1}. \]  (13)

Given (8), for \(\tau > 0\), one can easily show that:
\[ E(x(k)x(k-\tau)) = -[f_0(k-1) E(x(k-1)x(k-\tau))] \]
\[ \cdots f_m(k-p) E(x(k-p)x(k-\tau))] \theta_w. \]  (14)

In addition, one has for \(\tau > p \geq t \geq 0:\)
\[ E(x(k-l)x(k-\tau)) = E(y(k-l)y(k-\tau)) \]  (15)

\(1\) For the sake of simplicity, the notations do not explicitly depend on the time \(k\).
Therefore for $\tau > p$, combining (14) and (15) leads to:

$$[E(y(k)y(k-\tau)) f_0(k-1) E(y(k-1)y(k-\tau)) \ldots f_m(k-p) E(y(k-p)y(k-\tau))] \theta_w = 0. \quad (16)$$

Combining (13) and (16) for $p < \tau \leq q$, we obtain the following new relationship:

$$(\hat{R}_Y - \sigma_w^2 \hat{R}_B) \theta_w = 0_{(p(m+q) \times 1)} \quad (17)$$

where $\hat{R}_B$ is a matrix satisfying:

$$\hat{R}_B = \begin{bmatrix} 0 & F(1) & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & F(p) \\ 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}_{q \times p},$$

and $\hat{R}_Y$ is the extended matrix:

$$\hat{R}_Y = \begin{bmatrix} r_Y & R_Y \\ \rho_Y & \Xi_Y \end{bmatrix},$$

with $\rho_Y = E[[y(k-p-1) \ldots y(k-q)]^T y(k)]$, and $\Xi_Y$ is a $(q-p) \times p(m+1)$ matrix:

$$\Xi_Y = E \begin{bmatrix} y(k-p-1) \\ y(k-p-2) \\ \vdots \\ y(k-q) \end{bmatrix} \begin{bmatrix} \psi^T(k-1) & \ldots & \psi^T(k-p) \end{bmatrix}.$$  

To obtain the weight vector $\theta_w$, eq. (17) has to be solved. However, $\hat{R}_Y$ and $\hat{R}_B$ are not square. As described by Davila in [8], we can first multiply both sides of (17) by $(\hat{R}_Y - \sigma_w^2 \hat{R}_B)^T$. It leads to:

$$(A_0 + \sigma_w^2 A_1 + \sigma_w^4 A_2) \theta_w = 0_{(p(m+1)+1 \times 1)} \quad (18)$$

where $A_0 = \hat{R}_Y^T \hat{R}_Y$, $A_1 = \hat{R}_Y^T \hat{R}_B + \hat{R}_B^T \hat{R}_Y$, and $A_2 = \hat{R}_B^T \hat{R}_B$. By defining $P = [A_0 \ 0 \ 0 \ 1]$ and $Q = [-A_1 \ 0 \ 1 \ 0]$, eq. (18) is equivalent to:

$$(P - \sigma_w^2 Q) \begin{bmatrix} \hat{\theta}_w \\ \hat{\theta}_w^2 \end{bmatrix} = 0_{(2(p(m+1)+1) \times 1)} \quad (19)$$

where $\hat{\theta}_w$ is a vector which is not directly concerned with estimation. Therefore, the variance and the weights can be deduced by solving the generalized eigenvalue problem of the matrices $(P, Q)$. The weights can be deduced by looking at the generalized eigenvector $[\hat{\theta}_w \ \hat{\theta}_w^2]^T$ corresponding to the generalized eigenvalue with the lowest modulus which is equal to the variance.

It should be noted that in practice, the expectation is replaced by the temporal mean over a sliding window.

Procedure for off-line TVAR estimation using subspace approach can thus be summarized as follows:

Step 1. Build $R_Y$, $\hat{R}_Y$ and $\hat{R}_B$.

Step 2. Deduce $P$ and $Q$.

Step 3. Solve the generalized eigenvalue decomposition of $P$ and $Q$.

Step 4. The weights can be deduced from the generalized eigenvector corresponding to the generalized eigenvalue with the lowest modulus.

4. SIMULATION RESULTS

In this section, the performance of the noise-subspace (NS) approach is investigated by simulations. For a simpler exposition in all simulations, the order of the simulated TVAR process is set to 2; 2048 samples are used. Basis functions are set to $f_0(k) = 1$, $f_{zm}(k) = \cos(mk\pi)$ and $f_{2zm+1}(k) = \sin(mk\pi)$.

Firstly, a comparative study with YW equations directly used with noisy observations is carried out. The poles of the
transfer functions $H(z,k) = \frac{1}{1 + \sum_{n=1}^{p} a_n(k-n)z^{-n}}$ evolve in time according to fig. 1. This leads a TVAR process, whose spectrogram is given in fig. 2. Fig. 3 points out the relevance of our approach and the error that is obtained when using YW equations. The signal-to-Noise Ratio (SNR) is equal to 7 dB. Then, our approach is compared with the REIV approach proposed in [7], the extended Kalman filtering (EKF) and the standard Kalman filtering directly used with noisy observations. The SNR is set to 5 dB. The evolution of the poles is illustrated in fig. 4. The TVAR process spectrogram is illustrated in fig. 5. Concerning our NS approach, the order of the basis is set to $m = 5$. The evolution in time of the estimated pole modulus and the pole argument obtained are reported in Fig. 6. It can be seen that NS, EIV, and EKF approaches provide estimates that almost vary along the true curves. For the NS approach, the estimated noise variance varies within $[0.77, 0.98]$ while the true variance $\sigma_n^2$ is equal to 0.82.

5. CONCLUSIONS

In this paper, a method to estimate TVAR parameters from noisy observations has been proposed. The generalized eigenvalue decomposition allows both the variance of the additive noise and the weights to be estimated. The computational costs of the NS approach and the REIV method is a little higher than the ones of the other recursive algorithms. However they do not require any a priori information on the noise variances. Furthermore, the NS approach provides the result with high accuracy. Expectation maximization algorithms are currently under study.

6. REFERENCES