THE RECURSIVE MAXIMUM LIKELIHOOD ALGORITHM FOR NON-STATIONARY SIGNALS

Christian Debes and Abdelhak M. Zoubir

Signal Processing Group
Darmstadt University of Technology
Merckstr. 25, D-64283 Darmstadt, Germany
e-mail: {cdebes, zoubir}@ieee.org

ABSTRACT

In this paper we address the problem of parametric spectral estimation for non-stationary signals. An extension of the recursive maximum likelihood (RML) algorithm which iteratively tracks the time-varying signature of the process parameters is proposed. In particular we deal with the problem of estimating the parameters of time-varying autoregressive (TVAR) processes. Computer simulations are conducted that demonstrate the performance of the new method for parameter estimation as well as for time-varying spectral estimation.

Index Terms— Spectral analysis, Time-frequency analysis, Time-varying filters, Time series

1. INTRODUCTION

Estimating the parameters of autoregressive (AR) processes is a fundamental issue in time series analysis [1] and signal processing [2]. Estimates of the process parameters can be used to form a parametric spectral estimate as well as predict future samples of the process. Applications include radar [3], geophysics [4], biomedicine [5], image processing [6] and speech signal processing [7] to name a few.

However, in all these applications one is often faced with signals whose spectral characteristics naturally change with time or space. In order to take non-stationarity into account, e.g. to get an estimate of the time-varying spectrum or to obtain an improved predictor, one has to allow for the process parameters to vary with time. This situation is depicted in Figure 1. The TVAR process $x(n)$ can be described by the output of a linear filter with time-varying transfer function $H(e^{j\omega}, n) = 1/A(e^{j\omega}, n)$ driven by stationary white Gaussian noise $e(n)$ with $E[e(n)] = 0$ and autocovariance function $c_{ee}(m) = \sigma_e^2 \delta(m)$, where $\delta(m)$ is Kronecker's delta function.

The polynomial $A(e^{j\omega}, n)$ is given by

$$A(e^{j\omega}, n) = 1 + \sum_{k=1}^{p} a_k(n) e^{-j\omega k}$$

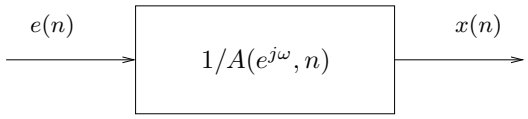


Fig. 1. Time-varying autoregressive process

with $a_k(n)$ being the k-th AR parameter at time instance n, p denoting the order of the autoregressive process.

Given N samples of the observed process $x(n)$ it is our aim to estimate the TVAR parameters $a_k(n)$ for $k = 1, ..., p$ and $n = 0, ..., N-1$. These can be used to form the parametric spectral estimate

$$\hat{C}_{XX}(e^{j\omega}, n) = \frac{\hat{\sigma}_e^2}{|\hat{A}(e^{j\omega}, n)|^2}$$

or to construct a predictor

$$\hat{x}(n+1) = -\sum_{k=1}^{p} \hat{a}_k(n) x(n-i+1)$$

The remainder of the paper is organised as follows: Section 2 gives an overview of the recursive maximum likelihood (RML) algorithm [8] including a short demonstration of its ability to track time-varying systems. In Section 3 an extension to the RML algorithm is proposed which iteratively tracks the time-varying signature of the process parameters. In Section 4 simulation results are shown which demonstrate the performance of the proposed method and section 5 provides conclusions and an outlook.

2. THE RML ALGORITHM

In this section we will shortly review the Recursive Maximum Likelihood (RML) algorithm which in the past has successfully been used for ARMA parameter estimation [9]. Given an ARMA process of order (p, q)

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + \sum_{k=1}^{q} b_k e(n-k) + e(n) \quad (1)$$

with $n = 0, ..., N-1$. Let θ denote the parameter vector

$$\theta = [a_1, ..., a_p, b_1, ..., b_q]^T$$

where $[\cdot]^T$ denotes matrix transpose. Further, let ϕ_n and ψ_n denote a data vector and its filtered version:

$$\begin{aligned} \phi_n &= [-x(n-1), ..., -x(n-p), r(n-1), ..., r(n-q)]^T \\ \psi_n &= [-\tilde{x}(n-1), ..., -\tilde{x}(n-p), \tilde{r}(n-1), ..., \tilde{r}(n-q)]^T \end{aligned}$$

with $r(n)$ being the residual error at time instance n. With each new data point $x(n)$ the following set of equations will be solved:

- **Prediction error:** $\epsilon(n) = x(n) - \phi_n^T \hat{\theta}_{n-1}$
- **Error covariance matrix:**

$$P_n = \frac{1}{\lambda}\left[\frac{P_{n-1} - P_{n-1}\psi_n\psi_n^T P_{n-1}}{\lambda + \psi_n^T P_{n-1}\psi_n}\right]$$

- **Parameter update:** $\hat{\theta}_n = \hat{\theta}_{n-1} + P_n\psi_n\epsilon_n$
- **Residual error:** $r(n) = x(n) - \psi_n^T\hat{\theta}_n$

The filtered data vectors $\tilde{x}(n)$ and $\tilde{r}(n)$ can be obtained by applying the filter with frequency response $\hat{B}(e^{j\omega}) = 1 + \sum_{k=1}^{q}\hat{b}_k e^{-j\omega k}$ to $x(n)$ and $r(n)$. λ is the so called forgetting factor, typically a constant close to unity.

When using the algorithm in recursive mode one automatically obtains an estimate of the time-varying signature of the process parameters, denoted by $\hat{\theta}_n$, $n = 0, ..., N-1$. This is visualised in Figure 2. For this example we chose an TVAR(1) process with $N = 100$ samples and $a_1(n) = 1.0 - 0.2 \cdot n/N$ being the time-varying signature of the AR parameter.

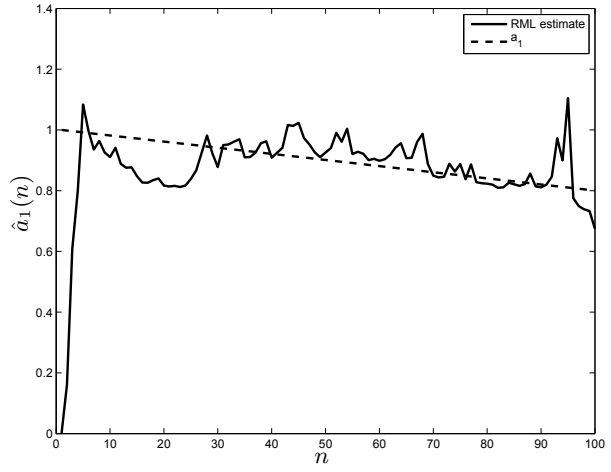


Fig. 2. RML estimate of a TVAR(1) process

3. PROPOSED METHOD

It was shown in the previous section that the (unmodified) RML algorithm has the ability to track time-varying model parameters. However the large variance of the estimates is a major drawback and may be unacceptable in many applications. Furthermore much of the data is wasted due to the transient phase of the algorithm which is needed to lock onto the TVAR signature. This problem becomes especially crucial for small data records.

In the following, we present a new method to iteratively estimate the TVAR parameters: Assume that the time variation of the model parameters can accurately be expressed as a linear combination of a finite number of orthonormal basis functions, i.e.

$$a_k(n) = \sum_{j=0}^{L} f_j(n)\alpha_{j,k}, \qquad 1 \le k \le p$$

with $\{f_j(n)\}_{j=0}^{L}$ being orthonormal basis functions, e.g. Legendre or Fourier. Then the expansion coefficients $\alpha_{j,k}$ can be estimated by means of a least squares regression of an initial guess $\hat{\theta}_n$, $n = 0, ..., N-1$ obtained by using the RML algorithm as described in Section 2. This leads to a revised estimate $\hat{\Theta}_n^0$ where the superscript 0 stands for the 0-th iteration:

$$\hat{\Theta}_n^0 = \begin{pmatrix} \hat{a}_1(1) & \hat{a}_1(2)\cdots & \hat{a}_1(N-1) \\ \vdots & \ddots & \vdots \\ \hat{a}_p(1) & \hat{a}_p(2)\cdots & \hat{a}_p(N-1) \end{pmatrix}$$

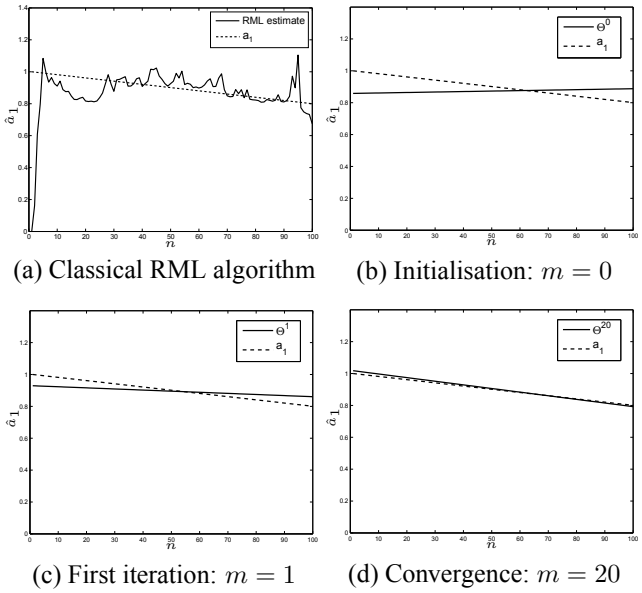
The revised estimate shall now be included in the next iteration of the RML algorithm. For this purpose we propose the following framework:

1. Obtain an initial estimate of the TVAR signature $\hat{\theta}_n^0$ by using the classical RML algorithm
2. Estimate a revised signature $\hat{\Theta}_n^0$ by expressing $\hat{a}_k(n)$ as a linear combination of orthonormal basis functions
3. For $m = 1, ..., M$
   - $\epsilon_n = x_n - \phi_n^T\hat{\theta}_{n-1}^m$
   - $P_n = \frac{1}{\lambda(m)}\left[\frac{P_{n-1} - P_{n-1}\psi_n\psi_n^T P_{n-1}}{\lambda(m) + \psi_n^T P_{n-1}\psi_n}\right]$
   - $\hat{\theta}_n = (1 - c(m))\hat{\theta}_{n-1}^m + c(m)\hat{\Theta}_n^{m-1} + P_n\psi_n\epsilon_n$
   - $r_n = x_n - \psi_n^T\hat{\theta}_n$
   - Estimate $\hat{\Theta}_n^m$ by expressing $\hat{a}_k(n)$ as a linear combination of orthonormal basis functions

When comparing each iteration (Step 3) of the new method with the original RML algorithm we note the following modifications:

- The parameter update step has been changed: When obtaining the estimate $\hat{\theta}_n$ at time instance n we do not rely solely on the estimate of the previous time instance, $\hat{\theta}_{n-1}$, but we also incorporate the information of the previous iteration, namely the revised estimate $\hat{\Theta}_n^{m-1}$. These two estimates are weighted by $(1 - c(m))$ and $c(m)$ respectively. With an increasing number of iterations $c(m)$ changes from 0 to 1, e.g. $c(m) = 1/M \cdot m$. Thus in the first iterations by chosing a small value for $c(m)$ we mainly trust $\hat{\theta}_{n-1}$ (initialisation phase). The more iterations we perform the more we can trust the revised estimate $\hat{\Theta}_n^{m-1}$ by chosing a larger $c(m)$ (modification phase) until finally $c(m)$ approaches 1 and we mainly trust $\hat{\Theta}_n^{m-1}$ (convergence phase).
- When calculating the error covariance matrix P_n the forgetting factor $\lambda(m)$ is also varying from iteration to iteration. At the beginning a relatively small forgetting factor is prefered, such that the time-varying signature of the parameters can roughly be tracked (at the cost of a large variance). With an increasing number of iterations the time-varying signature is already incorporated by $\hat{\Theta}_n^{m-1}$ and λ can be increased towards 1 in order to reduce the variance of the estimate.

In Figures 3(a)-(d) the estimation process of the proposed method is shown. We again used the simple example of a linear TVAR(1) process with $a_1(n) = 1.0 - 0.2 \cdot n/N$ with $N = 100$ samples to demonstrate the concept. The true time-varying signature is plotted as a dotted curve. Figure 3(a) shows the estimation result of the classical RML algorithm which is used for initialisation. We now perform a linear regression on the RML estimate which is depicted in Figure 3(b). It can be seen that the 0-th revised estimate $\hat{\Theta}_n^0$ is rather poor as it is mainly dominated by the transient phase of the RML algorithm. We now use $\hat{\Theta}_n^0$ as a basis for the next iteration and perform again a linear regression on the (modified) RML estimate which is depicted in Figure 3(c). Finally Figure 3(d) shows the estimated time-varying signature after 20 iterations. It can be seen that we obtain an estimate which closely follows the true $a_1(n)$ signature and significantly improves the parameter estimates from Figure 3(a) and (b).



(a) Classical RML algorithm (b) Initialisation: $m = 0$

(c) First iteration: $m = 1$ (d) Convergence: $m = 20$

Fig. 3. Estimating the TVAR signature using the proposed method

4. SIMULATION RESULTS

In order to demonstrate the performance of the proposed method for parameter estimation as well as for the estimation of time-varying spectra we chose the following setup: $x(n)$ is modeled as an TVAR(2) process with

$$a_1(n) = 0.8 - \frac{n}{N}; \quad a_2(n) = 0.55 - 1.2\left(\frac{n - N/2}{N - 1}\right)^2$$

σ_e^2 is chosen to be 1. A total of 1000 Monte Carlo simulations have been conducted with varying sample length. A polynomial order of $L = 2$ has been chosen and $M = 20$ iterations of the proposed method were performed with $c(m) = (m/M)^2$. and $\lambda(m) = 0.9 + 0.1 \cdot (m/M)^2$ which gave good performance. Convergence of the estimate was achieved in all 1000 Monte Carlo simulations.

The proposed method is compared to the unmodified RML algorithm with a subsequent polynomial fitting over $\hat{\theta}_n$. The mean absolute deviation of the second estimated TVAR parameter, $\hat{a}_2(n)$, versus the number of samples is shown in Figure 4. Similar results are obtained for $\hat{a}_1(n)$. It can clearly be seen that the proposed method outperforms the classical RML algorithm. Especially when a small sample size is chosen the iterative procedure allows accurate tracking of the time-varying signature and overcomes the disadvantages of the classical RML algorithm (high variance, transient phase). When chosing a large sample size both methods perform similarly which is due to the fact that the transient phase of the classical RML algorithm can be neglected and the high variance is compensated by the polynomial fit over a large set of data points.

The effect of the proposed method on the estimate of the time-varying spectrum is shown in Figure 5 for a sample size of $N = 256$. The true spectrum of the process is shown in Figure 5(a). The result of the classical RML algorithm shown in Figure 5(b) is rather poor which is mainly due to the large variance of the parameter estimates. A polynomial fitting and a first run of the proposed method already shows a vast improvement which can be seen in Figure 5(c). Further runs of the proposed method lead to convergence and a good approximation of the true time-varying spectrum as shown in Figure

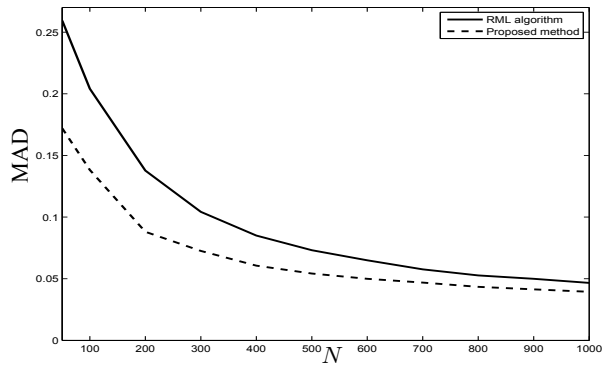


Fig. 4. Mean absolute deviation of the TVAR parameter estimates

5(d). Figures 5(e) and (f) display the spectrum obtained by Grenier's method [10] and the spectrogram for comparison. In addition to the spectra the location of the maximum (which in the true spectrum is at $\omega = 0.58\pi$ and $n = 120$) is shown. It can be seen that the proposed method is able to detect the point of maximal energy accurately whereas the spectrum obtained by Grenier's method is somewhat distorted and does not allow for accurate detection of its maximum.

5. CONCLUSIONS

An extension to the recursive maximum likelihood algorithm for non-stationary signals has been presented. It allows to iteratively track the time-varying signature of the model parameters and shows significantly better performance compared to the classical RML algorithm, especially for small sample sizes. Future work will be done in extending the proposed method to the case of autoregressive moving-average, as well as autoregressive plus noise processes. Further, the use of the proposed method for the estimation of the instantaneous frequency of chirp signals in noise will be studied.

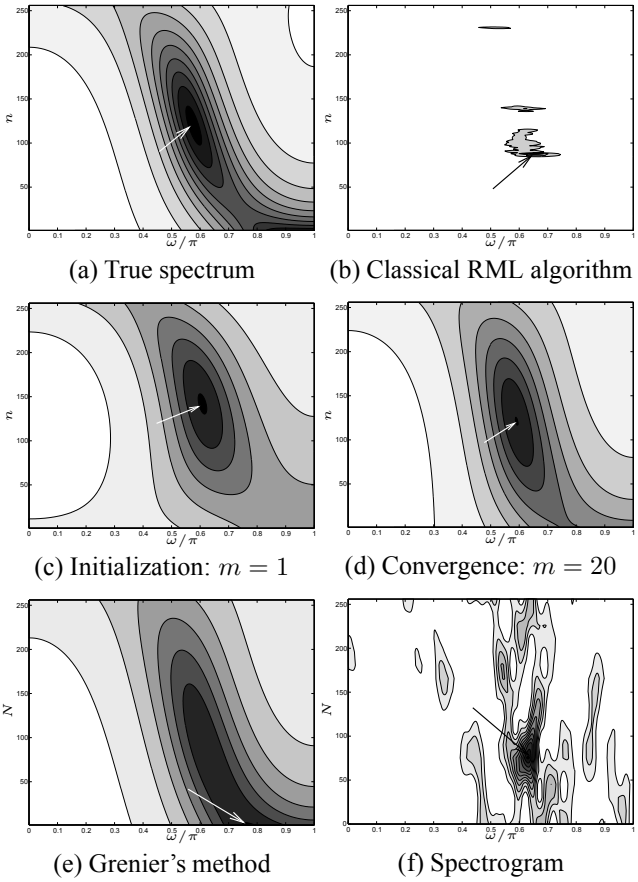


Fig. 5. Estimating the time-varying spectrum using the proposed method. Arrows indicate the maximum.

6. REFERENCES

[1] D. R. Brillinger, *Time Series: Data Analysis and Theory*, Holt, R & W, 1975.

[2] S.M. Kay and S.L. Marple, "Spectrum analysis - a modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981.

[3] S. Kesler and S. Haykin, "The maximum entropy method applied to the spectral analysis of radar clutter," *IEEE Transactions on Information Theory*, vol. 24, no. 2, pp. 269–272, 1978.

[4] L.J. Griffiths and R. Prieto-Diaz, "Spectral analysis of natural seismic events using autoregressive techniques," *IEEE Transactions on Geoscience Electronics*, vol. 15, no. 1, pp. 13–25, 1977.

[5] W. Gersch, "Spectral analysis of EEG's by autoregressive decomposition of time series," *Math. Biosc.*, vol. 7, pp. 205–222, 1970.

[6] J. Cadzow and K. Ogina, "Two-dimensional spectral estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, pp. 396–401, 1981.

[7] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975.

[8] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, 1983.

[9] B. Friedlander, "A recursive maximum likelihood algorithm for ARMA spectral estimation," *IEEE Trans. on Information Theory*, vol. 28, no. 4, pp. 639–646, 1982.

[10] Y. Grenier, "Time-dependent arma modeling of nonstationary signals," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 899–911, 1983.