INFORMATICA, 2006, Vol. 17, No. 2, 199–206

Iterative Estimation Algorithm of Autoregressive Parameters

Kazys KAZLAUSKAS, Jaunius KAZLAUSKAS
Institute of Mathematics and Informatics
Akademijos 4, 08663 Vilnius, Lithuania
e-mail: kazlausk@ktl.mii.lt

Received: June 2005

Abstract. This paper presents an iterative autoregressive system parameter estimation algorithm in the presence of white observation noise. The algorithm is based on the parameter estimation bias correction approach. We use high order Yule–Walker equations, sequentially estimate the noise variance, and exploit these estimated variances for the bias correction. The improved performance of the proposed algorithm in the presence of white noise is demonstrated via Monte Carlo experiments.
Key words: parameter estimation, iterative approach, autoregressive system, noisy observations.

1. Introduction

Estimation of the parameters of an autoregressive (AR) system parameters from noisy observations of its output has important applications in different fields such as signal processing, speech analysis, spectral estimation, and noise cancellation (Haykin, 1996; Srinath *et al*., 1996; Kay, 1993). Estimation of the AR parameters from noisy data is an important problem in practice (Kay, 1988). The quality of estimated AR parameters can be severely deteriorated due to the presence of noise. The least squares (LS) approach will give biased estimates of the AR parameters when the measured signal is corrupted with noise (Stoica *et al*., 1987). The techniques of parameter estimation using Yule–Walker (YW) equations in the autocorrelation domain are described in (Zheng, 1999; Hasan and Khan, 2003). In such case, parameter estimation is mainly confined to the 5 dB signal to noise ratio (SNR). The technique requires estimation of noise, which is difficult to obtain, and the autocorrelation function becomes noisy at a low SNR (Kay, 1980; Sun and Yahagi, 1992). On the other hand, use of the high order Yule–Walker (HYW) equations does not require a priori estimation of noise, but it is sensitive to the penetration of noise in the positive lag data samples in the autocorrelation function and does not give acceptable results in noisy conditions (Yahagi and Hasan, 1994). A least-squares based method for noisy autoregressive signals has been developed, which needs to no prefilter noisy data (Zheng, 2000). The improved least squares method with no prefiltering (ILSNP) yields good results, as compared with the YW equations (Zheng, 1999). However, this technique cannot reach 0 dB SNR. An approach for system identification at an extremely low SNR using energy density in a discrete cosine transform domain is analyzed in (Ferdousi *et al.*, 2005). In the proposed method, the system properties like energy density in a discrete cosine transform domain have been utilized to estimate the autocorrelation function. Successive autocorrelations of this estimated function are taken for a sequential estimation of system parameters.

The paper considers the estimation algorithm of AR parameters in the presence of white observation noise. The estimation algorithm is iterative because the estimation of noise variance and parameters is sequentially repeated until some convergence criterion is satisfied. The algorithm is built upon the bias correction principle (Stoica *et al.*, 1987; Zheng, 1999; Zheng, 2000). The idea is to use the HYW equations (i.e., to use more autocorrelation samples), sequentially estimate the noise variance, and to use these variance estimates for the bias correction.

The computer simulation results are presented and compared with other methods.

2. Preliminaries

Let the signal $x(t)$ be generated by the pth order equation

$$x(t) = \sum_{i=1}^{p} a_i x(t-i) + v(t), \tag{1}$$

where $v(t)$ is white driving noise of zero mean and variance σ_v^2 (i.e., $Ev(t) = 0$, $Ev(t)v(s) = \sigma_v^2 \delta_{t,s}$), and $\{a_i\}$ are real coefficients. Let $y(k)$ denote the noise corrupted measurement of $x(t)$,

$$y(t) = x(t) + w(t), \tag{2}$$

where the noise $w(t)$ is a stationary process of zero mean and variance σ_w^2 uncorrelated with $v(t)$. N is the number of data samples.

For the noiseless case, $\{a_i, i = 1, 2, \ldots, p\}$ can be found using the Yule–Walker equations

$$r_x(k) = \sum_{l=1}^{p} a_l r_x(k-l), \quad k \geqslant 1, \tag{3}$$

where $r_x(k) = E(x(t)x(t-k))$ is the autocovariance function of $x(t)$. Clearly, any p equations are sufficient to determine the AR parameters. Generally $k = 1, 2, \ldots, p$ is chosen which results in a set of symmetric Toeplitz equations

$$\begin{pmatrix} r_x(0) & r_x(1) & \ldots & r_x(p-1) \\ r_x(1) & r_x(0) & \ldots & r_x(p-2) \\ \ldots & \ldots & \ldots & \ldots \\ r_x(p-1) & r_x(p-2) & \ldots & r_x(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_p \end{pmatrix} = \begin{pmatrix} r_x(1) \\ r_x(2) \\ \ldots \\ r_x(p) \end{pmatrix} \tag{4a}$$

or in matrix form

$$R_x a = r_x. \tag{4b}$$

When the noise is present, we cannot estimate $r_x(k)$ since only $y(t)$ is available. In such case, we have

$$\begin{pmatrix} r_y(0) & r_y(1) & \dots & r_y(p-1) \\ r_y(1) & r_y(0) & \dots & r_y(p-2) \\ \dots & \dots & \dots & \dots \\ r_y(p-1) & r_y(p-2) & \dots & r_y(0) \end{pmatrix} \begin{pmatrix} a'_1 \\ a'_2 \\ \dots \\ a'_p \end{pmatrix} = \begin{pmatrix} r_y(1) \\ r_y(2) \\ \dots \\ r_y(p) \end{pmatrix} \tag{5a}$$

or in matrix form

$$R_y a_{ls} = r_y. \tag{5b}$$

But

$$r_y(k) = \begin{cases} r_x(0) + \sigma_w^2, & \text{if } k = 0, \\ r_x(k), & \text{if } k \neq 0. \end{cases} \tag{6}$$

Then, from (6) it follows that $r_x = r_y$ and

$$R_x = R_y - \sigma_w^2 I_p, \tag{7}$$

where I_p is the $(p \times p)$ identity matrix.

Substituting (7) into (4b), we have

$$a = a_{ls} + \sigma_w^2 R_y^{-1} a, \tag{8}$$

where $a_{ls} = R_y^{-1} r_y$ is the least square estimate (Davis and Vinter, 1985).

On the other hand, since (6) is valid, we may estimate $r_x(k)$ for all $k \neq 0$ from $y(t)$, since it is assumed σ_w^2 to be unknown. Then, by choosing $k = p+1, \dots, 2p$ in (3) and using (6), the resulting equations will not involve $r_x(0)$, i.e.,

$$\begin{pmatrix} r_y(p) & r_y(p-1) & \dots & r_y(1) \\ r_y(p+1) & r_y(p) & \dots & r_y(2) \\ \dots & \dots & \dots & \dots \\ r_y(2p-1) & r_y(2p-2) & \dots & r_y(p) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_p \end{pmatrix} = \begin{pmatrix} r_y(p+1) \\ r_y(p+2) \\ \dots \\ r_y(2p) \end{pmatrix} \tag{9a}$$

or in matrix form

$$R_h a_h = r_h. \tag{9b}$$

These equations are called the high-order Yule–Walker (HYW) equations (Vergara–Dominiguez, 1990). The high-order Yule–Walker estimate

$$a_h = R_h^{-1} r_h. \tag{10}$$

3. The Iterative Estimation Algorithm

As shown in (8) (see also (Stoica *et al.*, 1987)), an asymptotic bias of the least square estimate

$$a - a_{ls} = \sigma_w^2 R_y^{-1} a \tag{11}$$

depends on the noise variance σ_w^2 and is direct proportional to the parameter vector a. From (8), we obtain

$$a = (I - \sigma_w^2 R_y^{-1})^{-1} a_{ls}. \tag{12}$$

Keeping in mind that (Bellman, 1960)

$$(I - A)^{-1} = \sum_{i=0}^{\infty} A^i, \tag{13}$$

where A is a $(p \times p)$ matrix, (12) may be written as

$$a = a_{ls} + \sigma_w^2 R_y^{-1} a_{ls} + (\sigma_w^2 R_y^{-1})^2 a_{ls} + \dots. \tag{14}$$

Define

$$\begin{aligned} a^{(0)} &= a_{ls}, \\ a^{(1)} &= a_{ls} + \sigma_w^2 R_y^{-1} a_{ls}, \\ a^{(2)} &= a_{ls} + \sigma_w^2 R_y^{-1} a_{ls} + (\sigma_w^2 R_y^{-1})^2 a_{ls}, \end{aligned} \tag{15}$$

and so on.

Next, we obtain from (15) the iterative parameter estimation algorithm

$$a^{(i)} = a_{ls} + \sigma_w^2 R_y^{-1} a^{(i-1)}, \tag{16}$$

which depends on the noise variance σ_w^2.

The high-order Yule–Walker estimate $a_h(10)$ and the estimate $a = R_x^{-1} r_x$ do not depend on the noise variance. On the other hand, the least square estimate a_{ls} (8) depends on the noise variance. So, the difference of the high-order Yule–Walker estimate and the least squares estimate is proportional to the noise variance. Substituting a_h into the left side of (8), we get

$$a_h - a_{ls} = \sigma_w^2 R_y^{-1} a. \tag{17}$$

From (17), we obtain the estimate of the noise variance in the least square sense

$$\sigma_w^2 = \frac{(R_y^{-1} a)^T (a_h - a_{ls})}{\|R_y^{-1} a\|}, \tag{18}$$

where $\|\|$ is the l_2 norm of the vectors.

We can show that our algorithm is more general as compared with (Zheng, 2000). Multiplying both sides of (17) by R_h, we obtain

$$r_h - R_h a_{ls} = \sigma_w^2 R_h R_y^{-1} a. \tag{19}$$

From (19), it follows

$$\sigma_w^2 = \frac{r_h - R_h a_{ls}}{R_h R_y^{-1} a}. \tag{20}$$

In case we use only one additional point $r_y(p+1)$ of the autocovariance function, then in (9) $R_h = [r_y(p) r_y(p-1) \ldots r_y(1)]$, $r_h = r_y(p+1)$, and (20) is equal to (24) in (Zheng, 2000).

The Iterative Algorithm:

1. Using noisy samples $y(1), \ldots, y(N)$ compute autocovariance estimates $\hat{r}_y(k)$, $k = 0, 1, \ldots, 2p$, form the autocovariance matrix estimate $\hat{R}_y$ (5), vector estimate $\hat{r}_y$ (5), autocovariance matrix estimate $\hat{R}_h$ (9), and the vector estimate $\hat{r}_h$ (9).
2. Calculate the least square estimate (14)

$$\hat{a}^{(0)} = \hat{a}_{ls} = \hat{R}_y^{-1} \hat{r}_y \tag{21}$$

   and the high-order Yule–Walker estimate (10)

$$\hat{a}_h = \hat{R}_h^{-1} \hat{r}_h. \tag{22}$$

3. Estimate the noise variance (17)

$$\hat{\sigma}_w^{2^{(i)}} = \frac{(\hat{R}_y^{-1} \hat{a}^{(i-1)})^T (\hat{a}_h - \hat{a}_{ls})}{\|\hat{R}_y^{-1} \hat{a}^{(i-1)}\|^2}. \tag{23}$$

4. Use the parameter estimation algorithm (15) to obtain

$$\hat{a}^{(i)} = \hat{a}_{ls} + \sigma_w^{2^{(i)}} \hat{R}_y^{-1} \hat{a}^{(i-1)}. \tag{24}$$

5. Terminate the calculations, if

$$\frac{\|\hat{a}^{(i)} - \hat{a}^{(i-1)}\|}{\|\hat{a}^{(i)}\|} < \varepsilon, \tag{25}$$

   where ε is a positive number; otherwise, set the iteration step $i = i + 1$ and go to 3.

4. Simulation Results

Simulation results are given to illustrate the proposed iterative estimation algorithm. We analyze the fourth order noisy AR process (1), (2), where (Zheng, 2000)

$$a^T = [1.352 \quad -1.338 \quad 0.662 \quad -0.24], \quad \sigma_v^2 = 1.0.$$

The measure of the signal to noise ratio is

$$SNR = 10\log_{10}\frac{\sigma_x^2}{\sigma_w^2}\ (\text{dB}). \tag{26}$$

Table 1

Simulation results

$\sigma_w^2 = 0.122, \quad SNR \approx 15dB, \quad N = 6000$				
True Parameter	LS	ILSNP	ILSD	Iterative
$a_1 = 1.352$	1.0779 ± 0.0163	1.3516 ± 0.0423	1.3517 ± 0.0421	1.3516 ± 0.0403
$a_2 = -1.338$	-0.8728 ± 0.0247	-1.3305 ± 0.0719	-1.3306 ± 0.0703	-1.3340 ± 0.0651
$a_3 = 0.662$	0.2303 ± 0.0246	0.6593 ± 0.0633	0.6586 ± 0.0631	0.6599 ± 0.0592
$a_4 = -0.24$	-0.04 ± 0.0163	-0.2391 ± 0.0301	-0.2372 ± 0.0300	-0.2402 ± 0.0027
$\hat{\sigma}_w^2$	–	0.1213 ± 0.0179	0.1214 ± 0.0176	0.1217 ± 0.0136
Estimation error	12.5889%	0.0017%	0.0018%	0.0005%
$\sigma_w^2 = 0.3, \quad SNR \approx 11dB, \quad N = 6000$				
True Parameter	LS	ILSNP	ILSD	Iterative
$a_1 = 1.352$	0.8874 ± 0.0161	1.3554 ± 0.0813	1.3548 ± 0.0824	1.3540 ± 0.0569
$a_2 = -1.338$	-0.594 ± 0.0223	-1.3481 ± 0.1291	-1.3471 ± 0.1304	-1.3401 ± 0.0917
$a_3 = 0.662$	0.0104 ± 0.0210	0.6738 ± 0.1164	0.6737 ± 0.1183	0.6643 ± 0.0801
$a_4 = -0.24$	0.0329 ± 0.0153	-0.2451 ± 0.0512	-0.2450 ± 0.0567	-0.2389 ± 0.0393
$\hat{\sigma}_w^2$	–	0.2931 ± 0.0346	0.2930 ± 0.0345	0.2973 ± 0.0301
Estimation error	30.8325%	0.0068%	0.0061%	$3.6242.10^{-3}$%
$\sigma_w^2 = 0.7573, \quad SNR \approx 7.0dB, \quad N = 6000$				
True Parameter	LS	ILSNP	ILSD	Iterative
$a_1 = 1.352$	0.6689 ± 0.0141	1.2521 ± 0.2563	1.2710 ± 0.2631	1.3014 ± 0.1416
$a_2 = -1.338$	-0.3648 ± 0.0167	-1.2173 ± 0.3614	-1.2210 ± 0.3691	-1.2416 ± 0.1817
$a_3 = 0.662$	-0.1371 ± 0.0163	0.5271 ± 0.2913	0.5311 ± 0.2901	0.601 ± 0.1571
$a_4 = -0.24$	0.0454 ± 0.0129	-0.2038 ± 0.0981	-0.2042 ± 0.191	-0.2201 ± 0.0671
$\hat{\sigma}_w^2$	–	0.5217 ± 0.3617	0.5131 ± 0.3501	0.7103 ± 0.2201
Estimation error	51.8758%	1.071%	0.9398%	0.38819%

The stop criterion for iterative algorithm is chosen

$$\frac{\|\hat{a}^{(i)} - \hat{a}^{(i-1)}\|}{\|\hat{a}^{(i)}\|} < \varepsilon, \tag{27}$$

where $\hat{a}^{(i)}$ is the parameter estimate at the ith iteration step; $\varepsilon = 10^{-4}$.

The parameter estimation error is

$$\frac{\|mean(\hat{a}) - a\|}{\|a\|} 100\%, \tag{28}$$

where a is the true parameter vector, *mean* $(\hat{a})$ is the mean of the parameter vector estimates in 150 Monte Carlo tests.

We consider the performance of the iterative algorithm for different SNR when noisy observations $N = 6000$ of $y(t)$ are available. Table 1 shows that the least squares method is very sensitive to the measurement noise, while ILSNP (Zheng, 1999), ILSD (Zheng, 2000) and iterative estimation algorithms yield consistent estimates of the noise variance and AR parameters. On the other hand, we see that the iterative estimation algorithm is better than the ILSNP and ILSD methods in terms of standard deviations of the estimated parameters, i.e., standard deviations of the estimates of the iterative algorithm are less as compared with the ILSNP and ILSD methods. It is due to the fact that the iterative algorithm requires more additional autocovariance points $r_y(p), r_y(p+1), \ldots, r_y(2p)$, i.e., more information about the noisy signal $y(t)$ is used for the evaluation of noise variance.

5. Conclusions

This paper introduces a new iterative estimation algorithm for system identification. The noise compensation approach presented in the paper offers a better method than the ARMA method for reducing the white observation noise effects on the parameter estimates. The bias error of the least square method introduced by noise can be reduced. An advantage of the iterative algorithm is its ability to get more accurate estimates of the AR parameters as compared with the ILSNP and ILSD methods. The use of the high order Yule–Walker equations (9) is equivalent to noise compensation in the autocovariance matrix (5). We approximate a set of the first $2p + 1$ autocovariance estimates to a noisy AR process of order p. Our approach is more general as compared with the ILSNP method (Zheng, 2000), because we use more additional samples of the autocovariance function for the noise variance estimation.

References

Bellman, R. (1960). *Introduction to Matrix Analysis*. McGraw-Hill Book Company, NJ.

Davis, M.H.A., and R.B. Vinter (1985). *Stochastic Modeling and Control*. Chapman and Hall, London.

Ferdousi, F., A.T. Connie, M. Sharmin and M.R. Khan (2005). System identification at an extremely low SNR using energy density in DCT domain. *IEEE Signal Processing Letters*, **12**(4), 289–292.

Haykin, S. (1996). *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, Prentice Hall, NJ.

Hasan, M.K., and M.R. Khan (2003). Identification of autoregressive systems at a very low SNR using cosine model based estimation of autocorrelation function. *J. Signal Processing*, **7**(1), 15–21.

Kay, S.M. (1980). Noise compensation for autoregressive spectral estimates. *IEEE Trans. Acoust., Speech, Signal Processing*, **28**(3), 292–303.

Kay, S.M. (1988). *Modern Spectral Estimation*. Englewood Cliffs, Prentice Hall, NJ.

Kay, S.M. (1993). *Fundamentals of Statistical Signal Processing*: *Estimation Theory*. Englewood Cliffs, Prentice hall, NJ.

Srinath, M.D., P.K. Rajasekaran and R. Viswanathan (1996). *Introduction to Statistical Signal Processing with Applications*. Englewood Cliffs, Prentice Hall, NJ.

Stoica, P., A. Nehorai and S.M. Kay (1987). Statistical analysis of the least squares autoregressive estimator in the presence of noise. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, **35**(9), 1273–1281.

Sun, W., and T. Yahagi (1992). A new over fitting lattice filter for ARMA parameter estimation with additive noise. *IEICE Trans. Fundament.*, **E75-A**(2), 247–254.

Vergara–Dominiquez, L. (1990). New insights into the high-order Yule–Walker equations. *IEEE Trans. on Acoustics, Speech, Signal Processing*, **38**(9), 1649–1651.

Yahagi, T., and M.K. Hasan (1994). Estimation of noise variance from noisy measurements of AR and ARMA systems: Application to blind identification of linear time-invariant systems. *IEICE Trans. Fundament.*, **E77**(5), 847–855.

Zheng, W.X. (1999). A least-squares based method for autoregressive signals in the presence of noise. *IEEE Trans. Circuits Syst. II*, **46**(1), 81–85.

Zheng, W.X. (2000). Autoregressive parameter estimation from noisy data. *IEEE Trans. Circuits Syst. II*, **47**(1), 71–75.

K. Kazlauskas received a PhD degree from Kaunas Polytechnic Institute and a doctor habilius degree from Institute of Mathematics and Informatics and Vytautas Magnus University. He is a senior researcher of the Recognition Processes Department at the Institute of Mathematics and Informatics and a professor at the Informatics Department of Vilnius Pedagogical University. His research interests include signal processing, parameter estimation, and digital system design.

J. Kazlauskas received a MSc degree from Vilnius University in 2000. Now he is a junior researcher of the Process Recognition Department at the Institute of Mathematics and Informatics. His scientific interests include digital signal processing and parameter estimation.

Iteratyvusis autoregresijos parametrų įvertinimo algoritmas

Kazys KAZLAUSKAS, Jaunius KAZLAUSKAS

Straipsnyje pasiūlytas iteratyvusis autoregresinės sistemos parametrų įvertinimo algoritmas, kai sistemos išėjimo signalas yra iškreiptas adityvaus baltojo triukšmo. Algoritmo esmė – sistemos parametrų paslinktųjų įverčių korekcijos metodas. Naudojamos aukštesniosios eilės Julo–Volkerio lygtys ir nuosekliai įvertinama triukšmo dispersija, kurios tarpiniai įverčiai panaudojami sistemos parametrų įverčių patikslinimui.