**Proceedings of the 17th World Congress**
**The International Federation of Automatic Control**
**Seoul, Korea, July 6-11, 2008**

# Recursive Identification of EIV ARMA Processes ⋆

**Han-Fu Chen** *

* *Key Laboratory of Systems and Control, Institute of Systems Science, AMSS, Chinese Academy of Sciences, Beijing 100080, China (Tel: 86-10-62579540; e-mail: hfchen@iss.ac.cn).*

**Abstract:** Easily computable recursive algorithms are proposed for estimating coefficients of $A(z)$, $C(z)$, and the covariance matrix $R_w$ of $w_k$ for the multivariate ARMA process $A(z)y_k = C(z)w_k$ on the basis of the noise-corrupted observations $\eta_k \triangleq y_k + \epsilon_k$. It is shown that the estimates converge to the true ones under reasonable conditions. An illustrative example is provided, and the simulation results are shown to be consistent with the theoretical analysis.

Keywords: Multivariate ARMA, EIV, recursive identification, convergence

## 1. INTRODUCTION

There is a vast amount of publications on identification of ARMA processes, e.g., Box and Jenkins (1970), Brockwell and Davis (2001), Choi (1992), Hannan (1975), Stoica (1983), Stoica, Söderström, and Friedlander (1985), and Stoica, McKelvey, and Mari (2000) among others. The problem may be dealt with by using various methods such as Yule-Walker equation, extended least squares, maximum likelihood, and many others. The problem may also be solved from the viewpoint of spectral factorization (Chen (2007b), Claerbout (1966), Lai and Ying (1992), and Wilson (1969)). However, all works mentioned here concern the case where the observation data are free of noise.

In this paper we consider the case where the observation is corrupted by noise, i.e., we estimate the parameters in an errors-in-variables (EIV) ARMA process. A survey of EIV methods is provided by Söderström (2007). A recursive and strongly consistent estimate is given in Chen (2007a) for a class of EIV systems.

The multivariate ARMA process considered in the paper is as follows:

$$A(z)y_k = C(z)w_k, \quad y_k \in \mathbb{R}^m, \tag{1}$$

where the matrix polynomials in backward-shift operator $z$

$$A(z) \triangleq I + A_1 z + \cdots + A_p z^p \tag{2}$$

and

$$C(z) \triangleq I + C_1 z + \cdots + C_q z^q, \quad zy_k = y_{k-1} \tag{3}$$

are with unknown coefficients

$$\theta_A^T \triangleq [A_1, \ldots, A_p] \qquad \text{and} \qquad \theta_C^T \triangleq [C_1, \ldots, C_q]$$

but with known orders $(p, q)$.

Assume the process $\{y_k\}$ is observed with additive noise $\{\epsilon_k\}$

$$\eta_k \triangleq y_k + \epsilon_k. \tag{4}$$

The purpose of the paper is to give recursive estimates for $\theta_A$, $\theta_C$, and the covariance matrix $R_w$ of $w_k$ on the basis of observations $\{\eta_k\}$.

The conditions used in the paper are as follows:

A1. $\det A(z) \neq 0$, $\forall |z| \leq 1$, $\det C(z) \neq 0$, $\forall z : |z| < 1$.

A2. $A(z)$ and $C(z)$ have no common left-factor. The matrix $C_q$ is nonsingular, and the matrix $A_p$ is also nonsingular in the case $p > q$.

A3. $\{w_k\}$ are mutually independent and identically distributed (iid) with $Ew_k = 0$ and $Ew_k w_k^T \triangleq R_w > 0$.

A4. $\{\epsilon_k\}$ is iid and independent of $\{w_k\}$ with $E\epsilon_k = 0$ and known $E\epsilon_k\epsilon_k^T \triangleq R_\epsilon \geq 0$.

Based on the observation data $\{\eta_k\}$ the recursive algorithms proposed in the paper for estimating $\theta_A$, $\theta_C$, and $R_w$ are given with the help of stochastic approximation with expanding truncations (Chen (2002)). The estimates are proved convergent as time tends to infinity.

The rest of the paper is arranged as follows. In Section 2 the observation process $\{\eta_k\}$ is represented as an ARMA process with some desired properties. In Section 3 the coefficients $\theta_A$ contained in $A(z)$ are recursively estimated, while estimates for $\theta_C$ and $R_w$ are given in Section 4. A numerical example verifying convergence of the algorithms is given in Section 5. Some concluding remarks are given in the last section. The results in Section 4 are based on a convergence theorem, which is presented as Theorem A in Appendix.

## 2. REPRESENTATION OF OBSERVATION PROCESS

From (1)(4) it follows that the observation process $\{\eta_k\}$ is generated by the following equation:

$$A(z)\eta_k = C(z)w_k + A(z)\epsilon_k. \tag{5}$$

⋆ The work is supported by NSFC (Projects 60221301, 60334040, 60474004) and the Key Laboratory of Systems and Control, CAS.

By A3 and A4 it is clear that $\eta_k$ is stationary and ergodic. We now show that $\eta_k$ can be expressed as a stable ARMA process of minimum phase with some desired properties.

**Lemma 1.** Assume A2-A4 hold. Then the observation process $\{\eta_k\}$ can be expressed as an ARMA process

$$A(z)\eta_k = D(z)\xi_k \text{ with } D(z) = I + D_1 z + \cdots + D_r z^r,\quad r = \max(p,q) \qquad (6)$$

satisfying the following conditions:

i) $A(z)$ and $D(z)$ have no common left-factor;

ii) $[A_p \vdots D_r]$ is of row-full-rank;

iii) $\det D(z) \neq 0\ \forall z : |z| \leq 1$; and

iv) $E\xi_k = 0,\ E\xi_k\xi_k^T \triangleq R_\xi > 0\ \forall k,\ E\xi_k\xi_s^T = 0$ if $k \neq s$.

*Proof.* By the innovation representation (Anderson and Moore (1979)) the stationary process $\zeta_k \triangleq [C(z)\vdots A(z)]\Delta_k$ with $\Delta_k \triangleq [w_k^T, \epsilon_k^T]^T$ can be represented as

$$\zeta_k = H(z)\xi_k$$

with $E\xi_k = 0,\ E\xi_k\xi_k^T = R_\xi > 0\ \forall k,\ E\xi_k\xi_s^T = 0$ if $k \neq s$, where $H(z)$ is an $m \times m-$matrix of rational functions with $H(0) = I$, and both $H(z)$ and $H^{-1}(z)$ are stable (without root in $\{z : |z| \leq 1\}$).

Denote by $\mathcal{F}_k$ the $\sigma-$algebra generated by $\{\Delta_j,\ j \leq k\}$. Then, $\zeta_k \in \mathcal{F}_k$, and hence $\xi_k \in \mathcal{F}_k$. By stability of $H(z)$, $\zeta_k$ can be represented as

$$\zeta_k = \sum_{i=0}^{\infty} D_i \xi_{k-i},\quad D_0 = I. \qquad (7)$$

Since $\zeta_k \triangleq [C(z)\vdots A(z)]\Delta_k$, $\zeta_k$ is independent of $\Delta_{k-i}\ \forall i \geq r+1$ and hence independent of $\mathcal{F}_{k-i}\ \forall i \geq r+1$. Consequently, $E\zeta_k\xi_{k-i}^T = 0\ \forall i \geq r+1$, and hence

$$D(z)\xi_k = [C(z)\vdots A(z)]\Delta_k, \qquad (8)$$

where $D(z)$ is a monic polynomial of order $r$. By the definition of $\zeta_k$ and (5) we derive (6). Noticing that $H(z)$ is stable and $H(z) = D(z)$, we conclude that $D(z)$ is stable. It remains to show the properties i)-ii).

i) Assume the converse that there is $G(z)$ with $\det G(z) \neq$ constant such that $A(z) = G(z)F(z)$, and $D(z) = G(z)L(z)$, where the orders of $F(z)$ and $L(z)$ are less than $r$.

Writing the spectral functions of both sides of (8), we have

$$D(z)R_\xi D^T(z^{-1}) = A(z)R_\epsilon A^T(z^{-1}) + C(z)R_w C^T(z^{-1}), \qquad (9)$$

or

$$G(z)[L(z)\vdots F(z)]\begin{bmatrix} R_\xi & 0 \\ 0 & -R_\epsilon \end{bmatrix}\begin{bmatrix} L^T(z^{-1}) \\ F^T(z^{-1}) \end{bmatrix} G^T(z^{-1}) = C(z)R_w C^T(z^{-1}),$$

which implies that $C(z)$ also contains a left-factor $G(z)$. But, this is impossible, because $A(z)$ and $C(z)$ have no common left-factor by A2. The obtained contradiction proves i).

ii) Setting $A_i \triangleq 0,\quad C_j \triangleq 0$ for $i > p,\quad j > q$, respectively, we show that $[A_r \vdots D_r]$ is of row-full-rank.

Assume the converse: there is a non-zero $m-$vector $\mu$ such that $\mu^T[A_r \vdots D_r] = 0$.

Comparing the coefficients of $z^r$ in both sides of (9) we find that

$$D_r R_\xi = C_r R_w + A_r R_\epsilon. \qquad (10)$$

Multiplying (10) from left by $\mu^T$ and taking notice of the converse assumption, we derive $\mu^T C_r R_w = 0$, and hence $\mu^T C_r = 0$ by non-singularity of $R_w$. Thus, we have $\mu^T[A_r \vdots C_r] = 0$. However, $[A_r \vdots C_r]$ is of row-full-rank by A2, and hence $\mu$ must be zero. The obtained contradiction shows that $[A_r \vdots D_r]$ is of row-full-rank.

If $p \geq q$, then $r = p$ and we have shown $[A_p \vdots D_r]$ is of row-full-rank. If $p < q$, then $A_r = 0$ and the row-full-rank of $[A_r \vdots D_r] = [0 \vdots D_r]$ is equivalent to the non-singularity of $D_r$. This implies that $[A_p \vdots D_r]$ is of row-full-rank. $\square$

## 3. ESTIMATION OF $A(Z)$

In the case where $A(z)$ is stable, the process $\{\eta_k\}$ is stationary and ergodic.

Set

$$E\eta_k\eta_{k-i}^T \triangleq R_i. \qquad (11)$$

Write

$$\varphi_k^T = [\eta_k^T, \ldots, \eta_{k-p+1}^T],\quad \psi_k^T = [\eta_k^T, \ldots, \eta_{k-\mu+1}^T],\quad \mu = pm,$$

$$E\varphi_k\psi_{k-r}^T \triangleq \Gamma = \begin{bmatrix} R_r & R_{r+1} & \cdots & R_{r+\mu-1} \\ R_{r-1} & R_r & \cdots & R_{r+\mu-2} \\ \vdots & \vdots & & \vdots \\ R_{r-p+1} & R_{r-p+2} & \cdots & R_{r+\mu-p} \end{bmatrix},\quad r = \max(p,q), \qquad (12)$$

and

$$E\eta_k\psi_{k-1-r}^T \triangleq W = [R_{r+1} \cdots R_{r+\mu}]. \qquad (13)$$

From (5) we have

$$\eta_k = -\theta_A^T\varphi_{k-1} + C(z)w_k + A(z)\epsilon_k. \qquad (14)$$

By noticing $E(C(z)w_k)\psi_{k-1-r}^T = 0$ and $E(A(z)\epsilon_k)\psi_{k-1-r}^T = 0$, multiplying (14) by $\psi_{k-1-r}^T$ from the right, and taking expectation lead to the following generalized Yule-Walker equation

$$W = -\theta_A^T\Gamma,$$

which is clearly equivalent to

$$\Gamma\Gamma^T\theta_A + \Gamma W^T = 0. \qquad (15)$$

It is worth noting that the equation (15) is linear with respect to $\theta_A$.

Let $M_k$ be a sequence of positive real numbers increasingly diverging to infinity.

The matrix coefficient $\theta_A$ is recursively estimated by the following algorithms:

$$\Gamma_k = \Gamma_{k-1} - \frac{1}{k}(\Gamma_{k-1} - \varphi_{k-1}\psi_{k-r-1}^T) \quad (16)$$

$$W_k = W_{k-1} - \frac{1}{k}(W_{k-1} - \eta_k\psi_{k-r-1}^T) \quad (17)$$

$$\theta_k = \big(\theta_{k-1} - \frac{1}{k}\Gamma_k(\Gamma_k^T\theta_{k-1} + W_k^T)\big) \cdot I_{[\|\theta_{k-1}-\frac{1}{k}\Gamma_k(\Gamma_k^T\theta_{k-1}+W_k^T)\|\le M_{\lambda_k}]} \quad (18)$$

$$\lambda_k = \sum_{i=1}^{k-1} I_{[\|\theta_{i-1}-\frac{1}{i}\Gamma_i(\Gamma_i^T\theta_{i-1}+W_i^T)\|> M_{\lambda_i}]}, \qquad \lambda_0 = 0 \quad (19)$$

with arbitrary initial values $\Gamma_0$, $W_0$, $\theta_0$.

**Theorem 1.** Assume A2-A4 hold and $A(z)$ is stable. Then

$$\theta_k^T \xrightarrow[k\to\infty]{} [A_1, \dots, A_p] \qquad \text{a.s.},$$

where $\theta_k$ is given by (16)-(19).

*Proof.* Based on Lemma 1, by Theorem 1 in Stoica (1983) we conclude that $\Gamma\Gamma^T > 0$. Then the assertion of the theorem follows from the convergence theorem of stochastic approximation (Chen(2002)) for the linear regression function $f(\theta) \triangleq \Gamma\Gamma^T\theta + \Gamma W^T$. $\square$

**Remark 1.** In the literature there are many papers on AR parameter estimates, e.g., Stoica, Söderström, and Friedlander (1985), but estimates given there mostly are nonrecursive.

## 4. ESTIMATION OF $C(Z)$ AND $R_W$

In the last section we have estimated $A(z)$ in (1). It remains to estimate $C(z)$ and $R_w$.

In Stoica, McKelvey, and Mari (2000) it was pointed out that "the parameter estimation of moving-average (MA) signals from second-order statistics was deemed for a long time to be a difficult nonlinear problem for which no computationally convenient and reliable solution was possible". In what follows a recursive and easily computable algorithm is proposed to solve this problem. Under reasonable conditions the algorithm leads to the strongly consistent estimates.

Let

$$\rho_k \triangleq C(z)w_k (= A(z)y_k) \text{ and } E\rho_k\rho_{k-i}^T \triangleq S_i^\rho, \quad (20)$$

$$S^\rho(z) \triangleq \sum_{j=-r}^{r} S_j^\rho z^j, \quad S^\rho \triangleq [S_0^\rho, \cdots, S_r^\rho]^T \quad (21)$$

$$\zeta_k \triangleq D(z)\xi_k (= A(z)\eta_k) \text{ and } E\zeta_k\zeta_{k-i}^T \triangleq S_i^\zeta, \quad (22)$$

$$S^\zeta(z) \triangleq \sum_{j=-r}^{r} S_j^\zeta z^j, \quad S^\zeta \triangleq [S_0^\zeta, \cdots, S_r^\zeta]^T. \quad (23)$$

Write $\theta_k^T$ as

$$\theta_k^T = [A_{1k}, \dots, A_{pk}] \text{ with } A_{ik} \in \mathbb{R}^{m\times m}, i = 1, \dots, p. \quad (24)$$

Set

$$\hat\zeta_k \triangleq \eta_k + A_{1k}\eta_{k-1} + \cdots + A_{pk}\eta_{k-p}, \quad (25)$$

which serves as an estimate for $\zeta_k (= A(z)\eta_k)$.

Recursively define $S_k^\zeta \triangleq [S_{0k}^\zeta, \cdots, S_{rk}^\zeta]^T$ as follows:

$$S_{ik+1}^\zeta = S_{ik}^\zeta - \frac{1}{k+1}\Big(S_{ik}^\zeta - \hat\zeta_{k+1}\hat\zeta_{k+1-i}^T\Big), \quad S_{i0}^\zeta = 0, \quad i = 0, 1, \dots, r. \quad (26)$$

Further, define

$$S_k^\rho(z) \triangleq S_k^\zeta(z) - A_k(z)R_\epsilon A_k^T(z^{-1}), \quad (27)$$

where $S_k^\zeta(z) \triangleq \sum_{j=-r}^{r} S_{jk}^\zeta z^j$ with $S_{-jk}^\zeta \triangleq S_{jk}^{\zeta T}$, and $A_k(z) \triangleq I + A_{1k}z + \cdots + A_{rk}z^r$.

**Lemma 2.** Assume A2-A4 hold and $A(z)$ is stable. Then $S_k^\rho(z)$ a.s. converges to $S^\rho(z)$, the spectral function of $\rho_k$, as $k \to \infty$.

*Proof.* Noticing

$$\hat\zeta_k = \zeta_k + (A_{1k} - A_1)\eta_{k-1} + \cdots + (A_{pk} - A_p)\eta_{k-p},$$

and

$$\begin{aligned}
&\frac{1}{n}\Big\|\sum_{k\ge i,j}^{n}(A_{ik} - A_i)\eta_{k-i}\eta_{k-j}^T\Big\| \\
&\le \Big(\frac{1}{n}\sum_{k\ge i,j}^{n}\|A_{ik} - A_i\|\|\eta_{k-i}\|^2\Big)^{\frac{1}{2}} \\
&\cdot \Big(\frac{1}{n}\Big\|\sum_{k\ge i,j}^{n}\|A_{ik} - A_i\|\|\eta_{k-j}\|^2\Big)^{\frac{1}{2}} \xrightarrow[n\to\infty]{} 0 \quad a.s.
\end{aligned}$$

by Theorem 1 and ergodicity of $\{\eta_k\}$, $i, j = 1, \dots, p$, we find

$$\Big\|\frac{1}{k}\sum_{j=1}^{k}\zeta_j\zeta_{j-i}^T - \frac{1}{k}\sum_{j=1}^{k}\hat\zeta_j\hat\zeta_{j-i}^T\Big\| \xrightarrow[k\to\infty]{} 0 \quad \text{a.s.} \quad (28)$$

This incorporating with (26) implies

$$\lim_{k\to\infty} S_{jk}^\zeta = S_j^\zeta, \quad \text{and} \quad \lim_{k\to\infty} S_k^\zeta(z) = S^\zeta(z). \quad (29)$$

Noticing $S^\zeta(z) = D(z)R_\xi D^T(z^{-1})$ by (22)(23), $S^\rho(z) = C(z)R_wC^T(z^{-1})$ by (20)(21), and $A_k(z)R_\epsilon A_k^T(z^{-1}) \xrightarrow[k\to\infty]{} A(z)R_\epsilon A^T(z^{-1})$ by Theorem 1, from (9)(27)(29) we conclude the assertion of the lemma. $\square$

Define

$$U(X) \triangleq \begin{bmatrix} I & U_1(X) & \cdots & & U_r(X) \\ 0 & I & U_1(X) & \cdots & U_{r-1}(X) \\ \vdots & & \ddots & \ddots & \vdots \\ & & & I & U_1(X) \\ 0 & \cdots & & 0 & I \end{bmatrix},$$

$$\Phi(X) \triangleq \begin{bmatrix} I & 0 & \cdots & & 0 \\ 0 & X(0) & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \\ & & & & 0 \\ 0 & \cdots & & 0 & X(0) \end{bmatrix}, \quad (30)$$

$$U_0(X) \triangleq I, \quad U_1(X) = -X(1),$$
$$U_l(X) = -X(1)U_{l-1}(X) - X(2)U_{l-2}(X) - \cdots - X(l), \quad l = 2, \dots, r.$$

It is straightforward to verify that $X^\rho \triangleq [R_w, C_1, \cdots, C_r]^T$ satisfies the following algebraic equation

$$\Phi(X)X = \overline{U(X)S^\rho}, \tag{31}$$

where $S^\rho$ is given by (21), and the matrix $\overline{U(X)S^\rho}$ is obtained from $U(X)S^\rho$ by replacing its first $m \times m$-matrix $\sum_{i=0}^{r} U_i(X)S^{\rho T}(i)$ with

$$[(\sum_{i=0}^{r} U_i(X)S^{\rho T}(i))(\sum_{i=0}^{r} U_i(X)S^{\rho T}(i))^T]^{\frac{1}{2}}$$

and keeping the rest unchanged, i.e.,

$$\overline{U(X)S^\rho} = \begin{bmatrix} [(\sum_{i=0}^{r} U_i(X)S^{\rho T}(i))(\sum_{i=0}^{r} U_i(X)S^{\rho T}(i))^T]^{\frac{1}{2}} \\ \sum_{i=1}^{r} U_{i-1}(X)S^{\rho T}(i) \\ \vdots \\ S^{\rho T}(r) \end{bmatrix}.$$

Let $\nu_k$ be a sequence of positive real numbers increasingly diverging to infinity. The unknown matrix $X^\rho$ is estimated by the following algorithm:

$$X^\rho_{k+1} = \begin{cases} X^\rho_k - \frac{1}{k+1}\big(\Phi(X^\rho_k)X^\rho_k - \overline{U(X^\rho_k)S^\rho_k}\big), & \text{if} \\ \quad \|X^\rho_k - \frac{1}{k+1}\big(\Phi(X^\rho_k)X^\rho_k - \overline{U(X^\rho_k)S^\rho_k}\big)\| \le M_{\nu_k}, \\ X^\rho_0, & \text{otherwise}, \end{cases} \tag{32}$$

$$\nu_k = \sum_{j=0}^{k-1} I_{[\|X^\rho_j - \frac{1}{j+1}\big(\Phi(X^\rho_j)X^\rho_j - \overline{U(X^\rho_j)S^\rho_j}\big)\| > M_{\nu_j}]}, \tag{33}$$

$$\nu_0 = 0,$$

where $\Phi(X)$ and $U(X)$ are given by (30), $X^\rho_0 = [\nu I, 0, \cdots, 0]$ with $\nu \ge 1$, $S^\rho_k = [S^\rho_{0k}, \ldots, S^\rho_{rk}]^T$ is defined from (27) with $S^\rho_k(j)$ being the coefficient of $z^j$ in $S^\zeta_k(z) - A_k(z)R_\epsilon A^T_k(z^{-1})$ and $S^\rho_k(-j) \triangleq S^{\rho T}_k(j)$, and $\overline{U(X^\rho_k)S^\rho_k}$ is formed in the same way as $\overline{U(X)S^\rho}$ for (31) but with $X$ and $S^\rho$ replaced by $X^\rho_k$ and $S^\rho_k$, respectively.

**Theorem 2.** Assume A1-A4 hold. Then $\{X^\rho_k\}$ produced by (25)–(27), (32), and (33) converges to $J \triangleq \{X \in \mathbb{R}^{(m(r+1)\times m)} : \Phi(X)X = \overline{U(X)S^\rho}\}$: $X^\rho_k \xrightarrow[k\to\infty]{} X'^\rho \triangleq [X'^\rho(0), \cdots, X'^\rho(q)]^T$ a.s., and $X'^\rho \in J$. In other words, $\{X^\rho_k\}$ converges to a solution to (31). Further, $X'^\rho(0) = R_w$, $X'^\rho(i) = C_i$, $i = 1, \cdots, q$, whenever $\det(I + X'^\rho(1)z + \cdots + X'^\rho(q)z^q) \ne 0\ \forall z : |z| < 1$.

*Proof.* By Lemma 2 the assertions of the theorem follow from Theorem A in the Appendix.

$\square$

## 5. NUMERICAL EXAMPLE

Consider the following EIV ARMA process:

$$y_k + A_1 y_{k-1} = w_k + C_1 w_{k+1},$$
$$\eta_k = y_k + \epsilon_k,$$

where

$$A_1 = \begin{bmatrix} 0.5 & 1 \\ 0 & -0.5 \end{bmatrix}, C_1 = \begin{bmatrix} 0.6 & 1 \\ 0 & 0.3 \end{bmatrix}, \tag{34}$$

$w_k \in \mathcal{N}(0, R_w)$, $\epsilon_k \in \mathcal{N}(0, R_\epsilon)$, $R_w = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $R_\epsilon = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\{w_k\}$ and $\{\epsilon_k\}$ are sequences of iid random vectors, and they are mutually independent.

The spectral function $S^\rho(z)$ of $\rho_k$ is expressed by $S^\rho(z) = C(z)R_w C^T(z^{-1})$. Therefore, by (34)

$$S^\rho(z) = \begin{bmatrix} 1+0.6z & z \\ 0 & 1+0.3z \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1+0.6z^{-1} & 0 \\ z^{-1} & 1+0.3z^{-1} \end{bmatrix} \tag{35}$$

However, it is straightforward to check that

$$\begin{bmatrix} 1+0.6z & z \\ 0 & 1+0.3z \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1+0.6z^{-1} & 0 \\ z^{-1} & 1+0.3z^{-1} \end{bmatrix} = \begin{bmatrix} 1+\frac{5}{3}z & z \\ 0 & 1+0.3z \end{bmatrix} \begin{bmatrix} 0.72 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1+\frac{5}{3}z^{-1} & 0 \\ z^{-1} & 1+0.3z^{-1} \end{bmatrix}. \tag{36}$$

It is seen that the factorization (35) corresponds to the stable $C(z)$ with $C_1$ given in (34), while the factorization given at the right-hand side of (36) corresponds to an unstable $C(z)$ with

$$C_1 = \begin{bmatrix} \frac{5}{3} & 1 \\ 0 & 0.3 \end{bmatrix}, \qquad R_w = \begin{bmatrix} 0.72 & 0 \\ 0 & 1 \end{bmatrix}.$$

According to (16)-(19) for estimating $\theta_A$ and (25)-(27),(32)(33) with $\nu = 1$ in $X^\rho_0$ for estimating $\theta_C$ and $R_w$, more than 20 samples have been computed. As expected, the estimates for the elements $a_{11}$, $a_{12}$, $a_{21}$, $a_{22}$ of $A_1$ converge to the true ones for all samples. It turns out that all estimates for elements of $C_1$ and $R_w$ converge to the stable factorization.

For an arbitrarily chosen sample Fig.1, Fig.2 and Fig.3 respectively demonstrate the estimates (denoted by dotted lines) for $A_1$, $C_1$ and $R_w$ (denoted by solid lines), while Fig.4 shows the square errors $\frac{1}{k}\sum_{i=1}^{k}\|\theta - \theta_i\|^2$ of the estimates, where $\theta$ denotes $A_1$, $C_1$, and $R_w$, respectively, while $\theta_i$ denotes the corresponding estimate at time $i$.

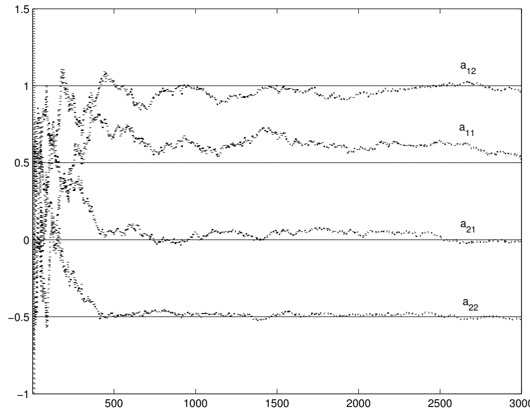


Fig.1 Estimates for $A_1$

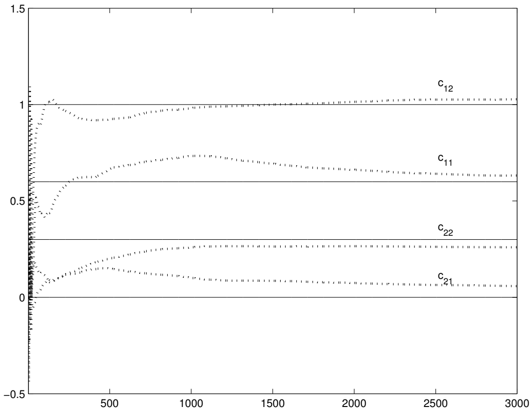


Fig.2 Estimates for $C_1$

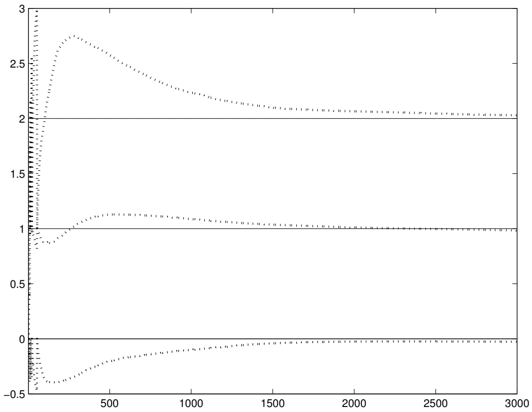


Fig.3 Estimates for $R_w$

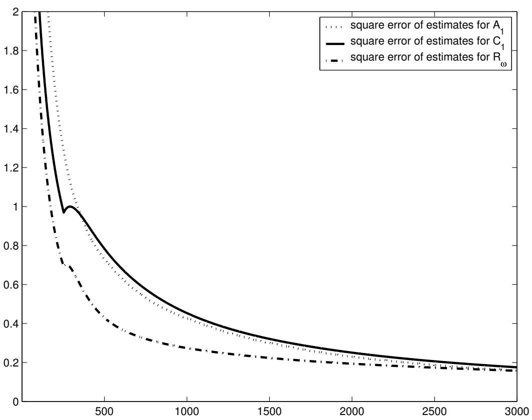


Fig.4 Square errors of the estimates

## 6. CONCLUDING REMARKS

In the special case where the observation noise $\epsilon_k \equiv 0$, the problem is reduced to the widely-studied identification of ARMA processes. However, for the multivariate ARMA process there is no recursive method for identifying its coefficients in time series analysis until the recent work Chen (2007b), while for ELS (extended least-squares) used in systems and control for recursively identifying an ARMA process, the restrictive SPR (strictly positive realness) condition is normally required.

In this paper the observation is allowed to be noisy, and the estimation problem is more complicated in comparison with Chen(2007b). Nevertheless, the recursive estimates are given for the unknowns of an EIV ARMA process, and they are proved to be convergent.

The numerical simulation is fully consistent with the theoretical analysis.

In A4 it is required that the covariance matrix $R_w$ of the observation noise is available. How to remove this undesirable condition belongs to further research.

**Acknowledgement** The author would like to thank Mr. Wenxiao Zhao for his skilled computer simulation.

## Appendix

Consider the stationary MA process $z_k = C(z)w_k$ with $Ez_k z_{k-i}^T \triangleq S(i)$, where $C(z) = \sum_{i=0}^{q} C_i z^i$ with $C_0 = I$, $Ew_k = 0$, and $Ew_k w_k^T = R_w$.

Let $\{S_k\}$ be a matrix sequence converging to $S \triangleq [S(0), S(1), \cdots, S(q)]^T : S_k \xrightarrow[k\to\infty]{} S$ *a.s.*, and let $\{M_k\}$ be a sequence of positive real numbers increasingly diverging to infinity.

The following algorithm is used to estimate $X^* \triangleq [R_w, C_1, \cdots, C_q]^T$ :

$$X_{k+1} = \begin{cases} X_k - \dfrac{1}{k+1}\big(\Phi(X_k)X_k - \overline{U(X_k)S_k}\big), & \text{if} \\ \quad \|X_k - \dfrac{1}{k+1}\big(\Phi(X_k)X_k - \overline{U(X_k)S_k}\big)\| \le M_{\sigma_k} \\ X_0, & \text{otherwise}, \end{cases} \tag{37}$$

$$\sigma_k = \sum_{j=0}^{k-1} I_{[\|X_j - \frac{1}{j+1}\left(\Phi(X_j)X_j - \overline{U(X_j)S_j}\right)\| > M_{\sigma_j}]}, \quad \sigma_0 = 0, \tag{38}$$

where $X_0 \triangleq [\nu I, 0, \cdots, 0]$ with $\nu \ge 1$, $\Phi(X)$, $U(X)$, and $\overline{U(X_k)S_k}$ are given by (30)(31) with $X$, $S_\zeta$ replaced by $X_k$ and $S_k$, respectively.

The following theorem is proved in Chen (2007b).

**Theorem A.** Assume $\{\epsilon_k\}$ is iid and $C_q$ is nonsingular. Then $\{X_k\}$ produced by (37)(38) converges to an $X' \in G \triangleq \{X \in \mathbb{R}^{(m(q+1)\times m)} : \Phi(X)X = \overline{U(X)S}\}$ a.s., and $X' = X^*$ whenever $\det C(z) \neq 0$, $\forall z : |z| < 1$, and $\det(I + X'(1)z + \cdots + X'(q)z^q) \neq 0$ $\forall z : |z| < 1$, where $X'(i)$ are sub-matrices of $X'^T = [X'(0), X'(1), \cdots, X'(q)]$.

## 7. REFERENCES

Anderson, B. D. O., and Moore, J. B. (1979). Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall.

Box, G. E. P., and Jenkins, G. (1970). Time Series Analysis, Forecasting and Control, Holden-Day, San-Francisco.

Brockwell, P. J. , and Davis, R. A. (2001). Time Series, Theory and Methods, Second Edition, Springer Verlag, Heidelberg.

Chen, H. F. (2002). Stochastic Approximation and Its Applications, Kluwer, Dordrecht.

Chen, H. F. (2007a). Recursive identification for multivariate errors-in-variables systems, Automatica, Vol. 43, No. 7, 1234-1242.

Chen, H. F. (2007b). New approach to recursive identification for ARMA systems, Technical Report of LSC, AMSS, Chinese Academy of Sciences.

Choi, B. S. (1992). ARMA Model Identification, Springer Verlag, New York.

Claerbout, J. F. (1966). Spectral factorization of multiple time series, Biometrica, Vol. 53, No. 1, 264-267.

Hannan, E. J. (1975). The estimation of ARMA models, The Annals of Statistics, Vol. 3, No. 4, 975–981.

Lai, T. L., and Ying, Z. (1992). Recursive selections of estimating algorithms and adaptive spectral factorization, IEEE Trans. Autom. Control, Vol. 37, No. 2, 240-243.

Söderström, T. (2007). Errors-in-variables methods in system identification, Automatica, Vol. 43, No. 6, 939-958.

Stoica, P. (1983). Generalized Yule-Walker equations and testing the order of multivariate time series, Int. J. Control, Vol. 37, No. 5, 1159–1166.

Stoica, P., Söderström, T., and Friedlander, B. (1985). Optimal instrumental variable estimates of the AR parameters of an ARMA process, IEEE Trans. Autom. Control, Vol. 30, No. 11, 1066-1074.

Stoica, P., McKelvey, T., and Mari, J. (2000). MA estimation in polynomial time, IEEE Trans. Signal Processing, Vol. 48, No. 7, 1999-2012.

Wilson, G. (1969). Factorization of the covariance generating function of a pure moving average process, SIAM J. Numer. Analysis, Vol. 6, No. 1, 1-7.