STOCHASTIC MULTIVARIABLE SELF–TUNING TRACKER FOR NON–GAUSSIAN SYSTEMS

VOJISLAV FILIPOVIC

Regionalni Centar za Talente (RCT)
PO Box 126, 15 300 Loznica, Serbia and Montenegro
e-mail: vfilip@eunet.yu

This paper considers the properties of a minimum variance self-tuning tracker for MIMO systems described by ARMAX models. It is assumed that the stochastic noise has a non-Gaussian distribution. Such an assumption introduces into a recursive algorithm a nonlinear transformation of the prediction error. The system under consideration is minimum phase with different dimensions for input and output vectors. In the paper the concept of Kronecker's product is used, which allows us to represent unknown parameters in the form of vectors. For parameter estimation a stochastic approximation algorithm is employed. Using the concept of the stochastic Lyapunov function, global stability and optimality of the feedback system are established.

Keywords: ARMAX model, self-tuning tracker, non-Gaussian noise, robust statistics, global stability, optimality

1. Introduction

Adaptive control is a very important topic in control theory and practice. A vast amount of literature already exists on parameter estimation and adaptive control of stochastic systems (Åström and Wittenmark, 1989; Caines, 1988; Chen and Guo, 1991; Duflo, 1997; Goodwin and Sin, 1984; Kumar and Varaija, 1986). In those references it is assumed that stochastic disturbance has a Gaussian distribution. In some cases a dominant aspect in a control problem is the unmodeled dynamics, and then it is very important to assume the robustness of the control scheme (Ioannou and Sun, 1996; Landau *et al.*, 1998; Sastry and Bodson, 1989).

The problem of stochastic adaptive control of linear ARMAX systems has received considerable attention in the literature. In (Goodwin *et al.*, 1981), self-optimality and global stability for minimum variance regulation and tracking were proved. Self-optimality means that the time average value of the square of the tracking error is minimal. In (Becker *et al.*, 1985), for a stochastic gradient algorithm, the self-tuning property for the regulation problem was shown. This means that the adaptive control law converges to an optimal control law. The same results were obtained for the tracking problem in (Kumar and Praly, 1987). The results of (Lin *et al.*, 1985) show that the self-tuning regulation with the minimum variance cost criterion is asymptotically optimal. That does not occur for other cost criteria (for example, the quadratic cost criterion). An exception is the class of systems with large delays.

In the above papers it is shown that, in the case of the minimum variance problem, the closed-loop identifiability problem does not prevent self-tuning because every possible limit of parameter estimates leads to an optimal control law. Moreover, in (Becker *et al.*, 1985) it is shown that the parameter estimate converges to some random multiple of the true parameter. For consistent parameter estimation it is necessary to introduce an additional signal: continually disturbed control (Caines, 1988), a diminishing excitation signal (Chen and Guo, 1991), or an occasional excitation (Lai and Wei, 1986). The problem of the robustness of the minimum-variance controller is considered in (Praly *et al.*, 1989). It is shown that an adaptive controller for linear stochastic systems is optimal for all ideal plants and remains stable with respect to violations of the positive real condition and with respect to perturbations of the system in the graph topology from all ideal plants. In the case of multiplicative and additive system perturbations, the problem of adaptive control was considered in (Radenkovic and Michel, 1992). The underlying idea for the above problem is the construction of a suitable Lyapunov function for different periods of adaptation.

In this paper we will consider a minimum-variance controller when the disturbance is non-Gaussian. The non-Gaussian assumption introduces a nonlinear transformation of the tracking error in the estimation algorithm. A special case of such a situation is when one has *a-priori* information about the class of distributions to which the actual real disturbance belongs. In such a situation the theory of min-max estimation can be applied and the corresponding algorithm is known as a robust algorithm (Filipovic and Kovacevic, 1994). Robustness here is with respect to a change in the disturbance distribution. In (Filipovic, 1999), a robust ELS algorithm was considered and stability and optimality of the minimum variance controller were proved. It was shown that for the stability of the adaptive controller is not necessary to modify the gain matrix. The tracking problem when the noise is non-Gaussian and when, also, unmodeled dynamics are present was considered in (Filipovic, 1996). Global convergence for a robust adaptive one-step ahead predictor is proved in (Filipovic, 2001).

In this paper we will investigate an adaptive minimum-variance controller for a system which is described by a multivariable ARMAX model. It is assumed that the system is minimum phase and that the relevant vectors of signals have different dimensions (rectangular systems). For parameter estimation, a stochastic approximation algorithm is used. Using the concept of the stochastic Lyapunov function, stability and optimality of the feedback system are established.

2. Problem Formulation

Let the system under consideration be described by a linear multi-input, multi-output ARMAX model with m-dimensional output and l-dimensional input,

$$\boldsymbol{A}(z)\boldsymbol{y}_{n+1} = \boldsymbol{B}(z)u_n + \boldsymbol{C}(z)\boldsymbol{w}_{n+1}, \quad n \geq 0, \qquad (1)$$

$$\boldsymbol{y}_n = \boldsymbol{w}_n = 0, \quad \boldsymbol{u}_n = 0, \quad n < 0,$$

where $\boldsymbol{A}(z)$, $\boldsymbol{B}(z)$ and $\boldsymbol{C}(z)$ are matrix polynomials in the shift-back operator $z\boldsymbol{y}_n = \boldsymbol{y}_{n-1}$ of orders p, q and r, respectively, i.e.,

$$\boldsymbol{A}(z) = \boldsymbol{I} + \boldsymbol{A}_1 z + \cdots + \boldsymbol{A}_p z^p, \qquad p \geq 0, \quad (2)$$

$$\boldsymbol{B}(z) = \boldsymbol{B}_1 + \boldsymbol{B}_2 z + \cdots + \boldsymbol{B}_q z^{q-1}, \quad q \geq 1, \quad (3)$$

$$\boldsymbol{C}(z) = \boldsymbol{I} + \boldsymbol{C}_1 z + \cdots + \boldsymbol{C}_r z^r, \qquad r \geq 0. \quad (4)$$

The noise $\{\boldsymbol{w}_n\}$ is assumed to be a martingale-difference sequence with respect to a nondecreasing family of σ-algebras $\{F_n\}$.

The unknown matrix coefficients are

$$\boldsymbol{\theta}^M = \left[-\boldsymbol{A}_1 \cdots -\boldsymbol{A}_p \boldsymbol{B}_1 \cdots \boldsymbol{B}_q \boldsymbol{C}_1 \cdots \boldsymbol{C}_r\right]^T. \qquad (5)$$

Model (1) can then be rewritten in the form

$$\boldsymbol{y}_{n+1} = \left(\boldsymbol{\theta}^M\right)^T \boldsymbol{\phi}_n^0 + \boldsymbol{w}_{n+1}, \qquad (6)$$

where

$$\left(\boldsymbol{\phi}_n^0\right)^T = \left[\boldsymbol{y}_n^T \cdots \boldsymbol{y}_{n-p+1}^T \boldsymbol{u}_n^T \cdots \boldsymbol{u}_{n-q+1}^T \boldsymbol{w}_n^T \cdots \boldsymbol{w}_{n-r+1}^T\right]. \qquad (7)$$

Let us introduce

$$\boldsymbol{X}_n^0 = \begin{bmatrix} \left(\boldsymbol{\phi}_n^0\right)^T & & & 0 \\ & \left(\boldsymbol{\phi}_n^0\right)^T & & \\ & & \ddots & \\ 0 & & & \left(\boldsymbol{\phi}_n^0\right)^T \end{bmatrix} = \boldsymbol{I} \otimes \left(\boldsymbol{\phi}_n^0\right)^T, \qquad (8)$$

where $\otimes$ stands for the Kronecker product. Also, a new vector $\boldsymbol{\theta}$ is constructed by stacking the columns of the $\boldsymbol{\theta}^M$ matrix. The relation (6) now has the form

$$\boldsymbol{y}_{n+1} = \boldsymbol{X}_n^0 \boldsymbol{\theta} + \boldsymbol{w}_{n+1}. \qquad (9)$$

In this paper we will consider a direct adaptive minimum-variance controller. The algorithm for estimating the unknown parameters can be reduced to the minimization of the functional (Filipovic and Kovacevic, 1994):

$$J(\theta) = E\left\{\Phi\left(\boldsymbol{\varepsilon}_{n+1}\right)\right\}, \quad \Phi : \mathbb{R}^m \to \mathbb{R}^1, \qquad (10)$$

where $\boldsymbol{\varepsilon}_{n+1}$ is the prediction error, i.e., $\boldsymbol{\varepsilon}_{n+1} = \boldsymbol{y}_{n+1} - \hat{\boldsymbol{y}}_{n+1}$, in which $\hat{\boldsymbol{y}}_{n+1}$ is the prediction of $\boldsymbol{y}_{n+1}$. The algorithm will have a stochastic approximation form (Kushner and Yin, 2003).

The functional $J(\boldsymbol{\theta})$ depends on the probability of observations, which is, in general, non-Gaussian. From identification theory it is known that

$$\Phi(\boldsymbol{x}) = -\log \boldsymbol{p}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^m, \qquad (11)$$

where $\boldsymbol{p}(\cdot)$ is probability density. As a result of applying the methodology of (Filipovic and Kovacevic, 1994) to (8) and (9), we get

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n + \frac{a\boldsymbol{X}_n^T}{r_n}\boldsymbol{\Psi}\left(\boldsymbol{y}_{n+1} - \boldsymbol{X}_n\boldsymbol{\theta}_n\right), \quad 0 < a \leq 1, \qquad (12)$$

$$r_n = r_{n-1} + \operatorname{tr} \boldsymbol{X}_n^T \boldsymbol{M} X_n, \qquad (13)$$

$$\boldsymbol{\Psi}(\boldsymbol{x}) = -\nabla_x \log \boldsymbol{p}(\boldsymbol{x}), \qquad (14)$$

$$\left(\boldsymbol{\phi}_n\right)^T = \Big[\boldsymbol{y}_n^T \cdots \boldsymbol{y}_{n-p+1}^T \boldsymbol{u}_n^T \cdots \boldsymbol{u}_{n-q+1}^T \boldsymbol{y}_n^T - \left(\boldsymbol{X}_{n-1}\boldsymbol{\theta}_{n-1}\right)^T \cdots \boldsymbol{y}_{n-r}^T - \left(\boldsymbol{X}_{n-r+1}\boldsymbol{\theta}_{n-r+1}\right)^T\Big], \qquad (15)$$

$$\boldsymbol{X}_n = \boldsymbol{I} \otimes \boldsymbol{\phi}_n^T, \qquad (16)$$

$$\boldsymbol{M} = E\left\{\nabla_x \boldsymbol{\Psi}(\boldsymbol{x})\right\}, \quad \boldsymbol{x} \in \mathbb{R}^m, \qquad (17)$$

$$\boldsymbol{X}_n \boldsymbol{\theta}_n = \boldsymbol{y}_{n+1}^*. \qquad (18)$$

The controller (18) is a minimum-variance one, where $\{\boldsymbol{y}_{n+1}^*\}$ is a sequence of bounded deterministic signals.

Remark 1. If we can use an *a priori* assumption that the distribution of the real noise lies in a specified class of distributions F which is convex and vaguely compact (Huber, 2003), it is possible to construct a robust real-time procedure in a min-max sense. The members of F are symmetric and F contains the standard normal distribution N. Two important classes are:

(a) *the gross error model:*

$$F_{1\varepsilon} = \big\{F : F = (1-\varepsilon)\, N + \varepsilon G, \quad G \text{ is symmetric}\big\}, \quad \varepsilon \in [0,1),$$

b) *the Kolmogorov model:*

$$F_{2\varepsilon} = \big\{F : F \text{ is symmetric and } \sup_x |F(x) - N(x)| \le \varepsilon\big\}, \quad \exists \varepsilon > 0.$$

Recent applications of robust statistics in engineering are presented in (Hubert *et al.*, 2004). A statistical analysis for atypical observations in economic and financial time series is made in (Lucas *et al.*, 2005).

3. Analysis of the Adaptive Algorithm

In this part of the paper global stability of the control system and a self-optimizing property of the adaptive controller are established. These results are formulated in the form of the following theorem:

Theorem 1. *Assume that for the model (1) and the algorithm (12)–(18) the following conditions are satisfied:*

C1: $\boldsymbol{B}_1$ *has full rank and* $\boldsymbol{B}_1^+\boldsymbol{B}(z)$ *is an asymptotically stable matrix polynomial, where* $\boldsymbol{B}_1^+$ *denotes the pseudoinverse of* $\boldsymbol{B}_1$.

C2: *All zeros of the polynomial* $\det \boldsymbol{C}(z)$ *lie inside the unit circle.*

C3: *Upper bounds for* p, q *and* r *are known.*

C4: *All finite-dimensional distributions of* $\{\boldsymbol{x}_0, \boldsymbol{w}\}$ *are absolutely continuous with respect to the Lebesque measure, where* $\boldsymbol{x} = \{\boldsymbol{y}_0, \dots, \boldsymbol{y}_{1-n}; \boldsymbol{u}_0, \dots, \boldsymbol{u}_{1-n}; \boldsymbol{w}_0, \dots, \boldsymbol{w}_{1-n}\}$, $n = \max\{p, q, r\}$.

C5: *The reference signal* $\{\boldsymbol{y}_n^*\}$ *is uniformly bounded.*

C6: $\{\boldsymbol{w}_n, F_n\}$ *is a martingale-difference sequence having a symmetric probability distribution function* $P(\cdot)$ *and satisfying*

$$E\left\{\boldsymbol{w}_{n+1}\boldsymbol{w}_{n+1}^T \mid F_{n-1}\right\} = \boldsymbol{R} \text{ a.s.},$$

$$E\left\{\|\boldsymbol{w}_{n+1}\|^4 \mid F_{n-1}\right\} \le C_w < \infty \text{ a.s.}$$

C7: *The functions* $\psi_i(\cdot)$, $i = 1, \dots, m$ *are odd and continuous everywhere.*

C8: *The functions* $\psi_i(\cdot)$, $i = 1, \dots, m$ *are uniformly bounded.*

C9: $\lambda_{\min}\{\boldsymbol{M}\} > 0$.

C10: *There exists a passive operator* $\boldsymbol{H}$ *such that for every* $n \ge 1$ *we have*

$$\boldsymbol{H}\boldsymbol{Z}_{1n} = \boldsymbol{\Phi}_1\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n}\right) - \frac{1}{2}\boldsymbol{\Phi}_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n}\right)\boldsymbol{Z}_{1n},$$

$$\boldsymbol{Z}_{1n} = -\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n, \quad \tilde{\boldsymbol{\theta}}_n = \boldsymbol{\theta}_n - \boldsymbol{\theta},$$

$$\boldsymbol{\Phi}_1\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n}\right) = E\left\{\boldsymbol{\psi}\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n} - \boldsymbol{w}_{n+1}\right)|F_n\right\},$$

$$\boldsymbol{\Phi}_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n}\right) = E\left\{\boldsymbol{\psi}'\left(\boldsymbol{C}^{-1}(z)\boldsymbol{Z}_{1n} - \boldsymbol{w}_{n+1}\right)|F_n\right\}.$$

C11: $0 < \|\boldsymbol{\Phi}_2(x)\| < \infty$, $\forall x$.

Then the adaptive controller is stable and optimal in the following sense:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\left(\|\boldsymbol{y}_{i+1}\|^2 + \|\boldsymbol{u}_{i+1}\|^2\right) < \infty \text{ a.s.},$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\left(\boldsymbol{y}_{i+1} - \boldsymbol{y}_{i+1}^*\right)\left(\boldsymbol{y}_{i+1} - \boldsymbol{y}_{i+1}^*\right)^T = \boldsymbol{R} \text{ a.s.}$$

Proof. Introducing the stochastic Lyapunov function

$$V_{n+1} = \tilde{\boldsymbol{\theta}}_{n+1}^T\tilde{\boldsymbol{\theta}}_{n+1}, \quad \tilde{\boldsymbol{\theta}}_{n+1} = \boldsymbol{\theta}_{n+1} - \boldsymbol{\theta}, \tag{19}$$

and using (12), one can get

$$\begin{aligned} V_{n+1} &= \left[\tilde{\boldsymbol{\theta}}_n + \frac{a\boldsymbol{X}_n^T\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n}\right]^T\left[\tilde{\boldsymbol{\theta}}_n + \frac{a\boldsymbol{X}_n^T\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n}\right] \\ &= \tilde{\boldsymbol{\theta}}_n^T\tilde{\boldsymbol{\theta}}_n + \frac{a\tilde{\boldsymbol{\theta}}_n^T\boldsymbol{X}_n^T\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n} + \frac{a\boldsymbol{\Psi}^T(\boldsymbol{\varepsilon}_{n+1})\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n}{r_n} \\ &\quad + \frac{a^2\boldsymbol{\Psi}^T(\boldsymbol{\varepsilon}_{n+1})\boldsymbol{X}_n\boldsymbol{X}_n^T\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n^2} \\ &= V_n + \frac{2a\left(\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\right)^T\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n} \\ &\quad + \frac{a^2\boldsymbol{\Psi}^T(\boldsymbol{\varepsilon}_{n+1})\boldsymbol{X}_n^T\boldsymbol{X}_n\boldsymbol{\Psi}(\boldsymbol{\varepsilon}_{n+1})}{r_n^2}. \end{aligned} \tag{20}$$

The prediction error $\boldsymbol{\varepsilon}_{n+1}$ has the form (Duflo, 1997):

$$\boldsymbol{\varepsilon}_{n+1} = -\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n + \boldsymbol{w}_{n+1}. \tag{21}$$

Using Condition C8, we conclude that

$$\|\boldsymbol{\Psi}(\cdot)\| \le k, \quad \in (0, \infty), \tag{22}$$

where $\boldsymbol{\Psi}(\cdot) = [\psi_1(\cdot) \cdots \psi_m(\cdot)]^T$.

Taking conditional expectations and having Condition C7 in mind from (20)–(22) we obtain

$$
\begin{aligned}
&E\{V_{n+1}|F_n\}\\
&\leq V_n-\frac{2a(\boldsymbol{X}_n\boldsymbol{\theta}_n)^T}{r_n}E\Big\{\boldsymbol{\Psi}\Big(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n-\boldsymbol{w}_{n+1}\Big)\mid F_n\Big\}\\
&\quad+\frac{a^2k\|\boldsymbol{X}_n\|^2}{r_n^2}\\
&=V_n-\frac{2a(\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n)^T}{r_n}\boldsymbol{\Phi}_1\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\right)\\
&\quad+\frac{a(\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n)^T(\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n)}{r_n}\boldsymbol{\Phi}_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\right)\\
&\quad-\frac{a\|\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\|^2}{r_n}\boldsymbol{\Phi}_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\right)+\frac{a^2k\,\|\boldsymbol{X}_n\|^2}{r_n^2}\\
&=V_n-\frac{2a(\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n)^T}{r_n}\\
&\quad\times\Big[\boldsymbol{\Phi}_1\Big(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\Big)-\frac{1}{2}\boldsymbol{\Phi}_2\Big(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\Big)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\Big]\\
&\quad-\frac{a\|\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\|^2}{r_n}\boldsymbol{\Phi}_2\Big(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\Big)+\frac{a^2k\,\|\boldsymbol{X}_n\|^2}{r_n^2}. \qquad (23)
\end{aligned}
$$

Define

$$
\begin{aligned}
S_n=2a\sum_{i=1}^{n}\left(\boldsymbol{X}_i\tilde{\boldsymbol{\theta}}_i\right)^T\Big[\boldsymbol{\Phi}_1\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_i\tilde{\boldsymbol{\theta}}_i\right)\\
-\frac{1}{2}\boldsymbol{\Phi}_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_i\tilde{\boldsymbol{\theta}}_i\right)\boldsymbol{X}_i\tilde{\boldsymbol{\theta}}_i\Big]+k_0. \qquad (24)
\end{aligned}
$$

From Condition C10 it follows that

$$S_n\geq 0,\quad \exists k_0>0.$$

Now we will define the nonnegative random variable:

$$T_{n+1}=V_{n+1}+\frac{S_n}{r_n}. \qquad (25)$$

From the definitions of S_n and r_n, we obtain

$$E\{T_{n+1}|F_n\}=E\{V_{n+1}|F_n\}+\frac{S_n}{r_n},\quad r_{n-1}\leq r_n. \qquad (26)$$

Using the relations (23)–(26), we have

$$
\begin{aligned}
&E\{T_{n+1}\mid F_n\}\\
&\leq V_n+\frac{S_{n-1}}{r_{n-1}}\\
&\quad-\frac{a\|\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\|^2}{r_n}\Phi_2\left(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\right)+\frac{a^2k\,\|\boldsymbol{X}_n\|^2}{r_n^2}\\
&=T_n-\frac{a\|\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\|^2}{r_n}\Phi_2\Big(\boldsymbol{C}^{-1}(z)\boldsymbol{X}_n\tilde{\boldsymbol{\theta}}_n\Big)\\
&\quad+\frac{a^2k\,\|\boldsymbol{X}_n\|^2}{r_n^2}. \qquad (27)
\end{aligned}
$$

For the last term in (27) we can write

$$
\begin{aligned}
&\sum_{i=0}^{\infty}\frac{a^2k\|\boldsymbol{X}_i\|^2}{r_i^2}\leq a^2k\sum_{i=0}^{\infty}\frac{\|\boldsymbol{X}_i\|^2}{1+\lambda_{\min}\{M\}\sum_{k=1}^{i}\|\boldsymbol{X}_n\|^2}\\
&=\frac{a^2k}{\lambda_{\min}\{M\}}\sum_{i=0}^{\infty}\frac{\lambda_{\min}\{\boldsymbol{M}\}\|\boldsymbol{X}_i\|^2}{1+\lambda_{\min}\{\boldsymbol{M}\}\sum_{k=1}^{i}\|\boldsymbol{X}_n\|^2}<\infty\text{ a.s.}
\end{aligned}
\qquad (28)
$$

The last result is a consequence of the Abel-Deany theorem (Rudin, 1964).

Using Condition C11, the Robbins-Siegmund martingale convergence theorem (Robins and Siegmund, 1971), (27) and (28), we get

$$\sum_{i=0}^{\infty}\frac{\|\boldsymbol{X}_i\tilde{\boldsymbol{\theta}}_i\|^2}{r_i}=O(1). \qquad (29)$$

Now we will prove that, under Condition C6, we have $r_n\to\infty$ as $n\to\infty$. From C6 we obtain

$$E\left\{\left(\boldsymbol{w}_{n+1}\boldsymbol{w}_{n+1}^T-\boldsymbol{R}\right)|F_n\right\}=0. \qquad (30)$$

Further, we can write (having in mind Condition C6):

$$
\begin{aligned}
&\Big\|\sum_{i=0}^{\infty}\frac{1}{(i+1)^2}E\Big\{\left(\boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T+\boldsymbol{R}\right)^T\\
&\qquad\times\left(\boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T-\boldsymbol{R}\right)\mid F_i\Big\}\Big\|\\
&\leq\Big\|\sum_{i=0}^{\infty}\frac{1}{(i+1)^2}E\Big\{\left(\boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T+\boldsymbol{R}\right)^T\\
&\qquad\times\left(\boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T+\boldsymbol{R}\right)\mid F_i\Big\}\Big\|\\
&\leq\sum_{i=0}^{\infty}\frac{1}{(i+1)^2}\Big\|E\Big\{\left(\|\boldsymbol{w}_{i+1}\|^2+\|\boldsymbol{R}\|\right)^2\mid F_i\Big\}\Big\|
\end{aligned}
$$

$$= \sum_{i=0}^{\infty} \frac{E\left\{\|\boldsymbol{w}_{i+1}\|^4 \mid F_i\right\}}{(i+1)^2} + 2\|\boldsymbol{R}\| \sum_{i=0}^{\infty} \frac{E\left\{\|\boldsymbol{w}_{i+1}\|^2 \mid F_{i-1}\right\}}{(i+1)^2} + \|\boldsymbol{R}\|^2 \sum_{i=0}^{\infty} \frac{1}{(i+1)^2}$$

$$\leq \left(C_w + 2\|\boldsymbol{R}\| \operatorname{tr} \boldsymbol{R} + \|\boldsymbol{R}\|^2\right) \sum_{i=0}^{\infty} \frac{1}{(i+1)^2} < \infty \quad \text{a.s.} \tag{31}$$

Using Theorem 2.18 of (Hall and Heyde, 1980, p. 35), we thus obtain

$$\sum_{i=0}^{\infty} \frac{\boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T - \boldsymbol{R}}{i+1} < \infty \quad \text{a.s.}, \tag{32}$$

and then using Kronecker's lemma (Shiryayev, 2004), we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \boldsymbol{w}_{i+1}\boldsymbol{w}_{i+1}^T = \boldsymbol{R} > 0 \quad \text{a.s.} \tag{33}$$

From the last relation we get

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{w}_{i+1}\|^2 > 0 \quad \text{a.s.} \tag{34}$$

Using (1) and Condition C2, we have

$$\sum_{i=0}^{n} \|\boldsymbol{w}_{i+1}\|^2 \leq k_1 \sum_{i=0}^{n} \left(\|\boldsymbol{y}_{i+1}\|^2 + \|\boldsymbol{u}_i\|^2\right), \quad k_1 \in (0,\infty). \tag{35}$$

Assume that

$$\lim_{n\to\infty} r_n < \infty \quad \text{a.s.} \tag{36}$$

From the definition of r_n, cf. (13), and (35) it follows that

$$r_n \geq 1 + \lambda_{\min}\{\boldsymbol{M}\} \|\boldsymbol{X}_n\|^2 \geq 1 + k_2 \left(\sum_{i=0}^{n} \|\boldsymbol{w}_{i+1}\|^2 + \sum_{i=0}^{n} \|\boldsymbol{y}_{i+1}\|^2 + \sum_{k=1}^{i} \|\boldsymbol{u}_i\|^2\right) \geq 1 + k_3 \sum_{i=0}^{n} \|\boldsymbol{w}_{i+1}\|^2, \quad k_2 \in (0,\infty), \quad k_3 \in (0,\infty). \tag{37}$$

Using (35) and (37), we obtain

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{w}_{i+1}\|^2 = 0 \quad \text{a.s.} \tag{38}$$

The last relation contradicts (34) so that we always have

$$\lim_{n\to\infty} r_n = \infty \quad \text{a.s.} \tag{39}$$

Using Kronecker's lemma (Shiryayev, 2004), from (29) and (39) we get

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_{i=0}^{n} \left\|\boldsymbol{X}_i \tilde{\boldsymbol{\theta}}_i\right\|^2 = 0 \quad \text{a.s.} \tag{40}$$

Condition C1, taken in conjunction with (1) and (34), yields

$$\frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{u}_i\|^2 \leq \frac{k_4}{n} \sum_{i=0}^{n} \|\boldsymbol{y}_{i+1}\|^2 + k_5, ; \quad (k_4, k_5) \in (0,\infty). \tag{41}$$

Also, for r_n we can write (having in mind C11):

$$r_n \leq 1 + \lambda_{\min}\{\boldsymbol{M}\} \sum_{i=0}^{n} \|\boldsymbol{X}_n\|^2. \tag{42}$$

Using (41) and (42), we have

$$\frac{r_n}{n} \leq \frac{k_6}{n} \sum_{i=0}^{n} \|\boldsymbol{y}_{i+1}\|^2 + k_7, \quad (k_6, k_7) \in (0,\infty). \tag{43}$$

Similary to (Goodwin *et al.*, 1981), from (41) and (43) it follows that

$$\liminf_{n\to\infty} \frac{n}{r_n} > 0 \quad \text{a.s.} \tag{44}$$

Now from (40) and (44) we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \left\|\boldsymbol{X}_i \tilde{\boldsymbol{\theta}}_i\right\|^2 = 0 \quad \text{a.s.} \tag{45}$$

Since (44) is equivalent to

$$\limsup_{n\to\infty} \frac{r_n}{n} < \infty \quad \text{a.s.}, \tag{46}$$

from the definition of r_n it follows that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{y}_{i+1}\|^2 < \infty \quad \text{a.s.}, \tag{47}$$

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{u}_i\|^2 < \infty \quad \text{a.s.} \tag{48}$$

So the stability of the adaptive controller has been established.

The next goal is to prove system optimality. Write

$$\boldsymbol{\zeta}_{n+1} = \boldsymbol{C}^{-1}(z) \boldsymbol{X}_n \tilde{\boldsymbol{\theta}}_n. \tag{49}$$

Using Condition C2 and (45), we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \|\boldsymbol{\zeta}_{i+1}\|^2 = 0 \quad \text{a.s.} \tag{50}$$

From (21) and (49) it follows that

$$\boldsymbol{y}_{n+1} - \boldsymbol{y}^*_{n+1} = -\boldsymbol{\zeta}_{n+1} + \boldsymbol{w}_{n+1}. \tag{51}$$

Using the last relation and the Cauchy-Schwarz inequality (Rudin, 1964), we can writte

$$\begin{aligned}
&\left\| \frac{1}{n}\sum_{i=0}^{n} \left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)\left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)^T - \frac{1}{n}\sum_{i=0}^{n} \boldsymbol{w}_{i+1}\boldsymbol{w}^T_{i+1} \right\| \\
&= \left\| \frac{1}{n}\sum_{i=0}^{n} \left(-\boldsymbol{\zeta}_{i+1} + \boldsymbol{w}_{i+1}\right)\left(-\boldsymbol{\zeta}_{i+1} + \boldsymbol{w}_{i+1}\right)^T - \frac{1}{n}\sum_{i=0}^{n} \boldsymbol{w}_{i+1}\boldsymbol{w}^T_{i+1} \right\| \\
&= \left\| \frac{1}{n}\sum_{i=0}^{n} \left(\boldsymbol{\zeta}_{i+1}\boldsymbol{\zeta}^T_{i+1} - 2\boldsymbol{\zeta}_{i+1}\boldsymbol{w}^T_{i+1} + \boldsymbol{w}_{i+1}\boldsymbol{w}^T_{i+1}\right) - \frac{1}{n}\sum_{i=0}^{n} \boldsymbol{w}_{i+1}\boldsymbol{w}^T_{i+1} \right\| \\
&= \left\| \frac{1}{n}\sum_{i=0}^{n} \boldsymbol{\zeta}_{i+1}\boldsymbol{\zeta}^T_{i+1} - \frac{2}{n}\sum_{i=0}^{n} \boldsymbol{\zeta}_{i+1}\boldsymbol{w}^T_{i+1} \right\| \\
&\le \frac{1}{n}\sum_{i=0}^{n} \left\|\boldsymbol{\zeta}_{i+1}\right\|^2 + 2\left(\frac{1}{n}\sum_{i=0}^{n} \left\|\boldsymbol{\zeta}_{i+1}\right\|^2\right)^{\frac{1}{2}} \left(\frac{1}{n}\sum_{i=0}^{n} \left\|\boldsymbol{w}_{i+1}\right\|^2\right)^{\frac{1}{2}}.
\end{aligned} \tag{52}$$

Having (34), (50) and (51), in mind from the last relation we get

$$\left\| \frac{1}{n}\sum_{i=0}^{n} \left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)\left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)^T - \frac{1}{n}\sum_{i=0}^{n} \boldsymbol{w}_{i+1}\boldsymbol{w}^T_{i+1} \right\| \underset{n\to\infty}{\longrightarrow} 0. \tag{53}$$

Finally, (30) and (50) yield

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} \left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)\left(\boldsymbol{y}_{i+1} - \boldsymbol{y}^*_{i+1}\right)^T = \boldsymbol{R}. \tag{54}$$

Our theorem is thus proven. ■

Remark 2. Condition C4 in Theorem 1 can be replaced by some modifications of B_{1n} (Åström and Wittenmark, 1973; Bercu, 1995; Chen and Guo, 1991).

Remark 3. In order to justify Condition C10, we need some concepts of passive systems. Assume that a system is described by an operator. Let $\mathbb{Z}$ denote the integers, $\mathbb{Z}_+$ the positive integers (i.e., those greater then or equal to zero) and l_2 the Hilbert space with the inner product $\langle\cdot,\cdot\rangle$ defined by

$$\langle a, b\rangle = \sum_{k=0}^{\infty} a^T(k)\, b(k).$$

We adopt some results from (Desoer and Vidyasagar, 1975) for discrete-time systems:

D1. Let $f(k) : \mathbb{Z}_+ \to \mathbb{Z}$. Then for each $k \in \mathbb{Z}_+$, the function $f_n(k) : \mathbb{Z}_+ \to \mathbb{Z}$ is defined by

$$f_n(k) = \begin{cases} f(k), & 0 \le k \le n, \\ 0, & n \le k, \end{cases}$$

and is called the trunction of $f(k)$ to the interval $[0, n]$.

D2. The set l_{2e} consists of all measurable functions $f(k) : \mathbb{Z}_+ \to \mathbb{Z}$ with the property that $f_n(k) \in l_2$ for all finite n. It is called the extension of l_2 or the extended l_2 space.

D3. An operator $G : l_{2e} \to l_{2e}$ is said to be passive if

$$\langle x, Gx\rangle_n \ge 0, \quad \forall n \ge 0, \quad \forall x \in l_{2e}.$$

When the operator G is nonlinear, Definition D3 implies Condition C10, and hence (24).

When the stochastic disturbance w_n is a Gaussian process, Condition C10 has the form

$$\sum_{n=1}^{n} \boldsymbol{Z}^T_{1k}\left(\boldsymbol{C}^{-1}(z) - \frac{1}{2}\boldsymbol{I}\right)\boldsymbol{Z}_{1k} > 0.$$

The above relation is correct when

$$\mathrm{Re}\left\{\boldsymbol{C}^{-1}(z) - \frac{1}{2}\boldsymbol{I}\right\} > 0.$$

That is a well-known condition from the theory of linear recursive algorithms.

4. Conclusions

In this paper we have presented a methodology for adaptive control of discrete-time dynamic stochastic MIMO systems when the disturbance has a non-Gaussian distribution. Using the Huber min-max approach, the methodology is extended to the case when *a-priori* information exists about the class of distributions to which the real disturbance belongs. The main contribution of the paper is the proof of global stability and optimality of the adaptive control system considered.

There are a number of interesting directions for future research in this area. First, it is interesting to consider robust recursive algorithms with the matrix gain which would possess a higher speed of convergence. Also, the problem of global stability and optimality of the minimum-variance controller for systems with time-delays and non-Gaussian disturbances could be considered.

References

Åström K.J. and Wittenmark B. (1973): *On self-tuning regulators*. — Automatica, Vol. 9, No. 2, pp. 185–199.

Åström K.J. and Wittenmark B. (1989): *Adaptive Control*. — New York: Addison Wesley.

Åström K.J. and Wittenmark B. (1995): *Adaptive Control*. — New York: Addison Wesley.

Becker J.A.H, Kumar P.R. and Wei C.Z. (1985): *Adaptive control with the stochastic approximation algorithm*. — IEEE Trans. Automat. Contr., Vol. 30, pp. 330–338.

Bercu B. (1995): *Weighted estimation and tracking for ARMAX models*. — SIAM J. Contr. Optim., Vol. 33, No. 1, pp. 89–106.

Caines P. (1988): *Linear Stochastic Systems*. — New York: Wiley.

Chen H.F. and Guo L. (1991): *Identification and Stochastic Adaptive Control*. — Basel: Birkhäuser.

Desoer C.A. and Vidyasagar M. (1975): *Feedback Systems: Input-Output Properties*. — New York: Academic Press.

Duflo M. (1997): *Random Iterative Models*. — New York: Springer.

Filipovic V. and Kovacevic B. (1994): *On robust AML identification algorithms*. — Automatica, Vol. 30, No. 12, pp. 1775–1778.

Filipovic V. (1996): *Robustness of adaptive tracking for stochastic multivariable minimum variance controller*. — Proc. 13th *IFAC World Congress*, San Francisco, USA, Vol. K, pp. 391–396.

Filipovic V. (1999): *Convergence and optimality of stochastic adaptive control scheme when the disturbance is non-Gaussian*. — Proc. 14th *IFAC World Congress*, Beijing, China, pp. 875–880.

Filipovic V. (2001): *Robust adaptive one-step ahead predictor*. — IMA J. Math. Contr. Inf. Vol. 18, pp. 491–500.

Goodwin G.C. and Sin K.S. (1984): *Adaptive Filtering, Prediction and Control*. — New Jersey: Prentice-Hall.

Goodwin G.C., Ramadge P. and Caines P. (1981): *Discrete time stochastic adaptive control*. — SIAM J. Contr. Optim., Vol. 19, No. 6, pp. 829–853.

Hall P. and Heyde C.C. (1980): *Martingale Limit Theory and Its Applications*. — New York: Academic Press.

Huber P. (2003): *Robust Statistics*. — New York: Wiley.

Hubert M., Pison G., Strouf A. and Van Aelst S. (Eds.) (2004): *Theory and Applications of Recent Robust Methods*. — Basel: Birkhauser.

Ioannou P.A. and Sun J. (1996): *Robust Adaptive Control*. — New Jersey: Prentice Hall.

Kumar P.R. and Varaija P. (1986): *Stochastic Systems: Estimation, Identification, and Adaptive Control*. — New Jersey: Prentice Hall.

Kumar P.R. and Praly L. (1987): *Self-tuning tracker*. — SIAM J. Contr. Optim., Vol. 25, No. 4, pp. 1053–1071.

Kushner J.H. and Yin G.G. (2003): *Stochastic Approximation. Algorithms and Applications*. — New York: Springer.

Landau I.D., Lozano R. and M'Saad M. (1998): *Adaptive Control*. — New York: Springer.

Lai T.L. and Wei C.Z. (1986): *Extended least squares and their applications to adaptive control and prediction in linear systems*. — IEEE Trans. Automat. Contr., Vol. 31, No. 6, pp. 898–906.

Lin W., Kumar P.R. and Seidman T. (1985): *Will the self-tuning approach work for general cost criteria?* — Syst. Contr. Lett., Vol. 6, No. 1, pp. 77–85.

Lucas A., Frances P.H. and Van Dijk D. (2005): *Outlier Robust Analysis of Economic Time Series*. — Oxford: Oxford University Press.

Praly L., Lin S.F. and Kumar P.R. (1989): *A robust adaptive minimum variance controller*. — SIAM J. Contr. Optim., Vol. 27, No. 2, pp. 235–266.

Radenkovic M.S. and Michel A.N. (1992): *Robust adaptive systems and self-stabilization*. — IEEE Trans. Automat. Contr., Vol. 37, No. 9, pp. 1355–1369.

Robins H. and Siegmund D. (1971): *A convergence theory for nonnegative almost supermartingale and same applications*, In: Optimization Methods in Statistics (J.S. Rustagi, Ed.). — New York: Academic Press, p. 233–257.

Rudin V. (1964): *Principles of Mathematical Analysis*. — New York: Mc Graw Hill.

Sastry S. and Bodson M. (1989): *Adaptive Control: Stability, Convergence and Robustness*. — New Jersey: Prentice Hall.

Shiryayev A.N. (2004): *Probability, Vols. 1 and 2*. — Moscow: MCNMO, (in Russian).

Received: 25 March 2005
Revised: 28 June 2005