Further results for ARX models in adaptive tracking

Bernard Bercu and Victor Vazquez

Universities of Bordeaux and Puebla, France, Mexico

47th IEEE Conference on Decision and Control

Cancun, Mexico, December 11, 2008

Outline

1. Introduction
   - The ARX Model
   - Matrix Polynomials and Causality
2. Strong Controllability
   - Preliminars
   - Strong Controllability
   - Schur Complement
3. Estimation and Adaptive Control
   - Estimation
   - Adaptative Control
4. Main Results
   - Almost sure convergence
   - Central Limit Theorem
   - Law of Iterated Logarithm
5. Simulations

Outline

1. Introduction
   - The ARX Model
   - Matrix Polynomials and Causality
2. Strong Controllability
   - Preliminars
   - Strong Controllability
   - Schur Complement
3. Estimation and Adaptive Control
   - Estimation
   - Adaptative Control
4. Main Results
   - Almost sure convergence
   - Central Limit Theorem
   - Law of Iterated Logarithm
5. Simulations

Introduction

Consider the d-dimensional $\text{ARX}_d(p, q)$ model given by

$$A(R)X_{n+1} = B(R)U_n + \varepsilon_{n+1}$$

where

1. R the shift-back operator,
2. X_n the system output,
3. U_n the system input,
4. ε_n the driven noise.

Introduction

Consider the d-dimensional $\text{ARX}_d(p, q)$ model given by

$$A(R)X_{n+1} = B(R)U_n + \varepsilon_{n+1}$$

where

1. R the shift-back operator,
2. X_n the system output,
3. U_n the system input,
4. ε_n the driven noise.

Introduction

Consider the d-dimensional $\text{ARX}_d(p, q)$ model given by

$$A(R)X_{n+1} = B(R)U_n + \varepsilon_{n+1}$$

where

1. R the shift-back operator,
2. X_n the system output,
3. U_n the system input,
4. ε_n the driven noise.

Introduction

Consider the d-dimensional ARX$_d(p, q)$ model given by

$$A(R)X_{n+1} = B(R)U_n + \varepsilon_{n+1}$$

where

1. R the shift-back operator,
2. X_n the system output,
3. U_n the system input,
4. ε_n the driven noise.

Matrix Polynomials Causality

The polynomials A and B are given for all $z \in \mathbb{C}$ by

$$\begin{aligned} A(z) &= I_d - A_1 z - \cdots - A_p z^p, \\ B(z) &= I_d + B_1 z + \cdots + B_q z^q, \end{aligned}$$

where A_i and B_j are unknown square matrices of order d and I_d is the identity matrix.

Definition

The matrix polynomial B is **causal** if for all $z \in \mathbb{C}$ with $|z| \leq 1$

$$\det(B(z)) \neq 0.$$

The Unknown Parameter

Denote by θ the unknown parameter of the model

$$\theta^t = (A_1, \ldots, A_p, B_1, \ldots, B_q).$$

The ARX model can be rewritten as

$$\boldsymbol{X_{n+1} = \theta^t \Phi_n + U_n + \varepsilon_{n+1}}$$

where the regression vector

$$\Phi_n = \left(X_n^t, \ldots, X_{n-p+1}^t, U_{n-1}^t, \ldots, U_{n-q}^t\right)^t.$$

About the noise

We assume that (ε_n) is a **martingale difference sequence** adapted to $\mathbb{F} = (\mathcal{F}_n)$ such that for all $n \geq 0$,

$$\mathbb{E}[\varepsilon_{n+1}\varepsilon_{n+1}^t|\mathcal{F}_n] = \Gamma \qquad \text{a.s.}$$

where Γ is a positive definite covariance matrix. Moreover, we assume that (ε_n) has finite conditional moment of order > 2 so

$$\Gamma_n = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k \varepsilon_k^t \to \Gamma \qquad \text{a.s.}$$

Outline

1 Introduction
- The ARX Model
- Matrix Polynomials and Causality

2 Strong Controllability
- Preliminars
- Strong Controllability
- Schur Complement

3 Estimation and Adaptive Control
- Estimation
- Adaptative Control

4 Main Results
- Almost sure convergence
- Central Limit Theorem
- Law of Iterated Logarithm

5 Simulations

Preliminars

For all $z \in \mathbb{C}$ such that $|z| \leq r$, where $r > 1$ is strictly less than the smallest modulus of the zeros of $\det(B(z))$, we denote

$$P(z) = B^{-1}(z)(A(z) - I_d) = \sum_{k=1}^{\infty} P_k z^k.$$

All the matrices P_k may be explicitly calculated as functions of the matrices A_i and B_j. For exemple, if $p = 2$, $q = 2$,

$$P_1 = -A_1,$$
$$P_2 = B_1 A_1 - A_2,$$
$$P_3 = (B_2 - B_1^2)A_1 + B_1 A_2.$$

Preliminars

For all $z \in \mathbb{C}$ such that $|z| \leq r$, where $r > 1$ is strictly less than the smallest modulus of the zeros of $\det(B(z))$, we denote

$$P(z) = B^{-1}(z)(A(z) - I_d) = \sum_{k=1}^{\infty} P_k z^k.$$

All the matrices P_k may be explicitly calculated as functions of the matrices A_i and B_j. For exemple, if $p = 2$, $q = 2$,

$$P_1 = -A_1,$$
$$P_2 = B_1 A_1 - A_2,$$
$$P_3 = (B_2 - B_1^2) A_1 + B_1 A_2.$$

Preliminars

For all $z \in \mathbb{C}$ such that $|z| \leq r$, where $r > 1$ is strictly less than the smallest modulus of the zeros of $\det(B(z))$, we denote

$$P(z) = B^{-1}(z)(A(z) - I_d) = \sum_{k=1}^{\infty} P_k z^k.$$

All the matrices P_k may be explicitly calculated as functions of the matrices A_i and B_j. For exemple, if $p = 2$, $q = 2$,

$$\begin{aligned} P_1 &= -A_1, \\ P_2 &= B_1 A_1 - A_2, \\ P_3 &= (B_2 - B_1^2) A_1 + B_1 A_2. \end{aligned}$$

Consider the square matrix of order dq given, if $p \geq q$, by

$$\Pi = \begin{pmatrix} P_p & P_{p+1} & \cdots & P_{p+q-2} & P_{p+q-1} \\ P_{p-1} & P_p & P_{p+1} & \cdots & P_{p+q-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{p-q+2} & \cdots & P_{p-1} & P_p & P_{p+1} \\ P_{p-q+1} & P_{p-q+2} & \cdots & P_{p-1} & P_p \end{pmatrix}$$

while, if $p \leq q$, by

$$\Pi = \begin{pmatrix} P_p & P_{p+1} & \cdots & \cdots & P_{p+q-2} & P_{p+q-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_1 & P_2 & \cdots & \cdots & P_{q-1} & P_q \\ 0 & P_1 & P_2 & \cdots & P_{q-2} & P_{q-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & P_1 & \cdots & P_p \end{pmatrix}.$$

Strong Controllability

Definition

The $\text{ARX}_d(p,q)$ process is said to be **strongly controllable** if B is causal and Π is invertible,

$$\mathbf{det(\Pi) \neq 0.}$$

Remark. The concept of strong controllability is not restrictive. For exemple, if $p = q = 2$, it holds as soon as

$$\det \begin{pmatrix} A_1 & A_2 - B_1 A_1 \\ A_2 - B_1 A_1 & -B_1 A_2 + (B_1^2 - B_2) A_1 \end{pmatrix} \neq 0.$$

For all $1 \leq i \leq q$, denote by H_i be the square matrix of order d

$$H_i = \sum_{k=i}^{\infty} P_k \Gamma P_{k-i+1}^t.$$

In addition, let H be the symmetric square matrix of order dq

$$H = \begin{pmatrix} H_1 & H_2 & \cdots & H_{q-1} & H_q \\ H_2^t & H_1 & H_2 & \cdots & H_{q-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ H_{q-1}^t & \cdots & H_2^t & H_1 & H_2 \\ H_q^t & H_{q-1}^t & \cdots & H_2^t & H_1 \end{pmatrix}.$$

For all $1 \leq i \leq p$, let $K_i = P_i \Gamma$ and, if $q \leq p$,

$$K = \begin{pmatrix} 0 & K_1 & K_2 & \cdots & \cdots & K_{p-2} & K_{p-1} \\ 0 & 0 & K_1 & \cdots & \cdots & K_{p-3} & K_{p-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & K_1 & K_2 & \cdots & K_{p-q+1} \\ 0 & \cdots & \cdots & 0 & K_1 & \cdots & K_{p-q} \end{pmatrix}$$

while, if $p \leq q$,

$$K = \begin{pmatrix} 0 & K_1 & \cdots & K_{p-2} & K_{p-1} \\ 0 & 0 & K_1 & \cdots & K_{p-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & K_1 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Denote by L the block diagonal matrix of order dp

$$L = \text{diag}\,(\Gamma, \cdots, \Gamma)\,.$$

Let Λ be the symmetric square matrix of order $\delta = d(p+q)$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{L} & \boldsymbol{K}^t \\ \boldsymbol{K} & \boldsymbol{H} \end{pmatrix}.$$

Schur Complement

Lemma

Let S be the Schur complement of L in Λ

$$\boldsymbol{S = H - KL^{-1}K^t.}$$

If the $ARX_d(p,q)$ process is **strongly controllable**, *then S and Λ are invertible and Λ^{-1} is given by*

$$\Lambda^{-1} = \begin{pmatrix} L^{-1} + L^{-1}K^tS^{-1}KL^{-1} & -L^{-1}K^tS^{-1} \\ -S^{-1}KL^{-1} & S^{-1} \end{pmatrix}.$$

Outline

1. Introduction
   - The ARX Model
   - Matrix Polynomials and Causality
2. Strong Controllability
   - Preliminars
   - Strong Controllability
   - Schur Complement
3. Estimation and Adaptive Control
   - Estimation
   - Adaptative Control
4. Main Results
   - Almost sure convergence
   - Central Limit Theorem
   - Law of Iterated Logarithm
5. Simulations

Weighted least squares

The weighted least squares estimator $\widehat{\theta}_n$ of θ satisfies

$$\widehat{\theta}_{n+1} = \widehat{\theta}_n + a_n S_n^{-1}(a)\Phi_n \left(X_{n+1} - U_n - \widehat{\theta}_n^t \Phi_n \right)^t$$

$$S_n(a) = \sum_{k=0}^{n} a_k \Phi_k \Phi_k^t + I_\delta$$

The **standard least squares estimator** is given by

$$a_n = 1.$$

Weighted least squares

The **weighted least squares estimator** is given for $\gamma > 0$ by

$$a_n = \left(\frac{1}{\log s_n}\right)^{1+\gamma} \quad \text{where} \quad s_n = \sum_{k=0}^{n} \| \Phi_k \|^2 .$$

We always have the decomposition

$$\widehat{\theta}_n - \theta = S_{n-1}^{-1}(a) M_n(a)$$

$$M_n(a) = \sum_{k=0}^{n-1} a_k \Phi_k \varepsilon_{k+1}.$$

Adaptive Control

The role played by U_n is to force X_n to track step by step a given trajectory (x_n). We make use of the **adaptive tracking control**

$$U_n = x_{n+1} - \widehat{\theta}_n^t \Phi_n.$$

Then, we obtain the closed-loop system

$$X_{n+1} - x_{n+1} = \pi_n + \varepsilon_{n+1}$$

where the prediction error

$$\pi_n = (\theta - \widehat{\theta}_n)^t \Phi_n.$$

We assume that the reference trajectory (x_n) satisfies

$$\sum_{k=1}^{n} \| x_k \|^2 = o(n) \qquad \text{a.s.}$$

Let (C_n) be the average cost matrix sequence defined by

$$C_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - x_k)(X_k - x_k)^t.$$

Definition

The tracking is said to be **optimal** if C_n converges a.s. to Γ.

Outline

1 Introduction
- The ARX Model
- Matrix Polynomials and Causality

2 Strong Controllability
- Preliminars
- Strong Controllability
- Schur Complement

3 Estimation and Adaptive Control
- Estimation
- Adaptative Control

4 Main Results
- Almost sure convergence
- Central Limit Theorem
- Law of Iterated Logarithm

5 Simulations

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable** *and that (ε_n) has finite conditional moment of order > 2. Then, for the* **LS estimator**, *we have*

$$\lim_{n\to\infty} \frac{S_n}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to θ*

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable** *and that (ε_n) has finite conditional moment of order > 2. Then, for the* **LS estimator**, *we have*

$$\lim_{n\to\infty} \frac{S_n}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to θ*

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable** *and that (ε_n) has finite conditional moment of order > 2. Then, for the* **LS estimator**, *we have*

$$\lim_{n\to\infty} \frac{S_n}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to θ*

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{\log n}{n}\right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable**. *Suppose that (ε_n) is a white noise or (ε_n) has finite conditional moment of order > 2. Then, for the* **WLS estimator**, *we have*

$$\lim_{n\to\infty} (\log n)^{1+\gamma} \frac{S_n(a)}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = o\left(\frac{(\log n)^{1+\gamma}}{n}\right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to* θ

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{(\log n)^{1+\gamma}}{n}\right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable**. *Suppose that (ε_n) is a white noise or (ε_n) has finite conditional moment of order > 2. Then, for the* **WLS estimator**, *we have*

$$\lim_{n \to \infty} (\log n)^{1+\gamma} \frac{S_n(a)}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = o\left(\frac{(\log n)^{1+\gamma}}{n} \right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to θ*

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{(\log n)^{1+\gamma}}{n} \right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable**. *Suppose that (ε_n) is a white noise or (ε_n) has finite conditional moment of order > 2. Then, for the* **WLS estimator**, *we have*

$$\lim_{n\to\infty} (\log n)^{1+\gamma} \frac{S_n(a)}{n} = \Lambda \quad a.s.$$

In addition, the **tracking is optimal**

$$\| C_n - \Gamma_n \| = o\left(\frac{(\log n)^{1+\gamma}}{n}\right) \quad a.s.$$

Finally, $\widehat{\theta}_n$ **converges almost surely** *to θ*

$$\| \widehat{\theta}_n - \theta \|^2 = \mathcal{O}\left(\frac{(\log n)^{1+\gamma}}{n}\right) \quad a.s.$$

Theorem

Assume that the $ARX_d(p, q)$ process is **strongly controllable** *and that (ε_n) has finite conditional moment of order $\alpha > 2$. In addition, suppose that (x_n) has the same regularity in norm as (ε_n) which means that for all $2 < \beta < \alpha$*

$$\sum_{k=1}^{n} \| x_k \|^{\beta} = \mathcal{O}(n) \quad a.s.$$

Then, the **LS and WLS estimators** *share the same* **CLT**

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Lambda^{-1} \otimes \Gamma)$$

Theorem

In addition, the **LS and WLS estimators** *share the same* **LIL** *which means that for any vectors $u \in \mathbb{R}^d$ and $v \in \mathbb{R}^\delta$,*

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} v^t(\widehat{\theta}_n - \theta)u = -\liminf_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/2} v^t(\widehat{\theta}_n - \theta)u = \left(v^t\Lambda^{-1}v\right)^{1/2}\left(u^t\Gamma u\right)^{1/2} \quad a.s.$$

Outline

1. Introduction
   - The ARX Model
   - Matrix Polynomials and Causality
2. Strong Controllability
   - Preliminars
   - Strong Controllability
   - Schur Complement
3. Estimation and Adaptive Control
   - Estimation
   - Adaptative Control
4. Main Results
   - Almost sure convergence
   - Central Limit Theorem
   - Law of Iterated Logarithm
5. Simulations

Consider the $ARX_2(1,1)$ process

$$X_{n+1} = AX_n + U_n + BU_{n-1} + \varepsilon_{n+1}$$

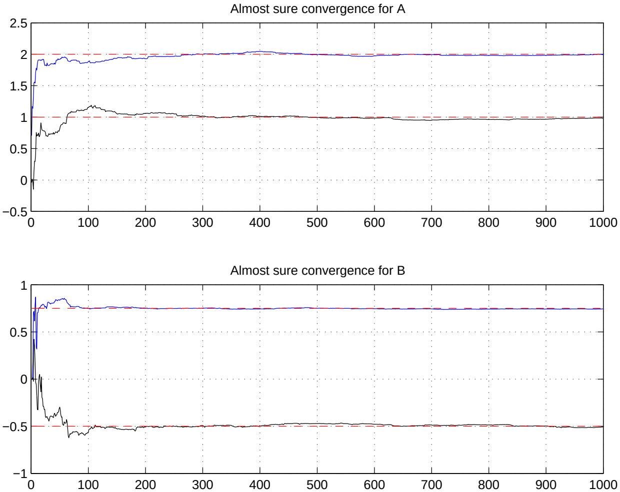
where

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \frac{1}{4}\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}.$$

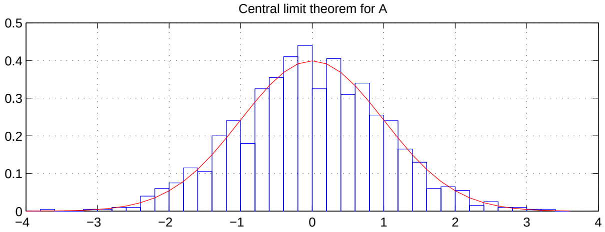
It is **strongly controllable** with limiting matrix

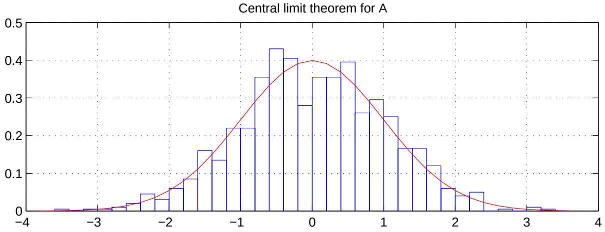
$$\Lambda = \frac{1}{21}\begin{pmatrix} 21 & 0 & 0 & 0 \\ 0 & 21 & 0 & 0 \\ 0 & 0 & 192 & 0 \\ 0 & 0 & 0 & 28 \end{pmatrix}$$

Almost sure convergence

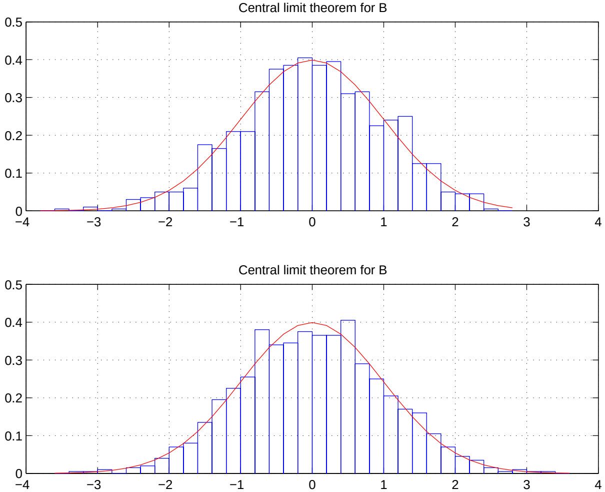


Central Limit Theorem





Central Limit Theorem



Consider the $ARX_2(1,1)$ process

$$X_{n+1} = AX_n + U_n + BU_{n-1} + \varepsilon_{n+1}$$

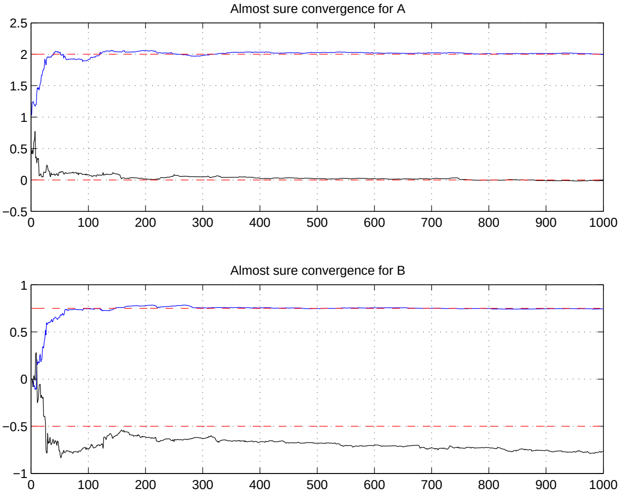
where

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \frac{1}{4}\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}.$$

It is **not strongly controllable** because the limiting matrix

$$\Lambda = \frac{1}{21}\begin{pmatrix} 21 & 0 & 0 & 0 \\ 0 & 21 & 0 & 0 \\ 0 & 0 & 192 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Almost sure convergence



Central Limit Theorem