ME 233 Advanced Control II

Lecture 21

Parameter Convergence in Least Squares Estimation and Persistence of Excitation

Estimation of ARMA model

$$A(q^{-1})y(k) = q^{-\mathrm{d}}B(q^{-1})u(k)$$

Where

- $u(k)$ known ***bounded*** input

- $y(k)$ measured output

Estimation of ARMA model

$$A(q^{-1})y(k) = q^{-\mathrm{d}}B(q^{-1})u(k)$$

Where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n}$$ **(anti-Schur)**

$$B(q^{-1}) = b_o + b_1 q^{-1} + \cdots + b_m q^{-m}$$

- Orders n and m are <u>known</u>
- Relative degree d is <u>known</u>
- a's and b's are <u>unknown</u> but <u>constant</u> coefficients

ARMA Model

$$y(k) = \phi^T(k-1)\,\theta$$

Unknown parameter vector:

$$\theta = \begin{bmatrix} \left.\begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix}\right\} n \\ \left.\begin{matrix} b_o \\ \vdots \\ b_m \end{matrix}\right\} m+1 \end{bmatrix}$$

Known regressor vector:

$$\phi(k-1) = \left.\begin{bmatrix} -y(k-1) \\ \vdots \\ -y(k-n) \\ u(k-\mathrm{d}) \\ \vdots \\ u(k-\mathrm{d}-m) \end{bmatrix}\right\} n+m+1$$

ARMA series-parallel estimation

- A-priori output

$$\widehat{y}^{o}(\underline{k}) = \phi^{T}(k-1)\,\widehat{\theta}(\underline{k-1})$$

$$\widehat{\theta}(k) = \left[\begin{array}{cccccc} \widehat{a}_1(k) & \cdots & \widehat{a}_n(k) & \widehat{b}_o(k) \cdots & \widehat{b}_m(k) \end{array}\right]^T$$

- A-priori error

$$e^{o}(k) = y(k) - \widehat{y}^{o}(k)$$

ARMA series-parallel estimation

- A-priori error

$$e^o(k) \;=\; y(k) - \widehat{y}^o(k)$$

$$e^o(k) \;=\; \phi^T(k-1)\,\tilde{\theta}(k-1)$$

- Parameter error

$$\tilde{\theta}(k) = \theta - \widehat{\theta}(k)$$

RLS Estimation Algorithm

$$e^o(k+1) = y(k+1) - \phi^T(k)\widehat{\theta}(k)$$

$$e(k+1) = \frac{\lambda_1(k)}{\lambda_1(k) + \phi^T(k)F(k)\phi(k)} e^o(k+1)$$

$$\widehat{\theta}(k+1) = \widehat{\theta}(k) + \frac{1}{\lambda_1(k)} F(k)\phi(k)e(k+1)$$

$$F(k+1) = \frac{1}{\lambda_1(k)} \left[F(k) - \lambda_2(k) \frac{F(k)\phi(k)\phi^T(k)F(k)}{\lambda_1(k) + \lambda_2(k)\phi^T(k)F(k)\phi(k)} \right]$$

Overview

- In Lecture 21 we learned how to analyze the stability of adaptive systems and proved:
  - Convergence of the output error

$$e^o(k) \rightarrow 0 \qquad e(k) \rightarrow 0$$

- Today we will provide conditions on the input sequence $u(k)$ that guarantee that

$$\tilde{\theta}(k) = \theta - \hat{\theta}(k)$$

also converges to zero.

Parameter error convergence

- Remember that $e^{o}(k) \rightarrow 0$

It can be shown that the n+m+1 parameter error also converges:

$$\lim_{k \rightarrow \infty} \tilde{\theta}(k) = \bar{\theta}$$

$$\lim_{k \rightarrow \infty} \tilde{\theta}(k) = \bar{\theta} = \begin{bmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_n \\ \\ \bar{b}_o \\ \vdots \\ \bar{b}_m \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} \bar{a}_1 \\ \vdots \\ \bar{a}_n \end{matrix}}\right\} n \\ \\ \left.\vphantom{\begin{matrix} \bar{b}_o \\ \vdots \\ \bar{b}_m \end{matrix}}\right\} m+1 \end{array}$$

Parameter error convergence

The steady-state parameter error satisfies

$$\boxed{E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} \;=\; 0}$$

Regressor

$$\phi(k) = \begin{array}{c} \\ \left[\begin{array}{c} -y(k) \\ \vdots \\ -y(k-n+1) \\ u(k) \\ \vdots \\ u(k-m) \end{array}\right] \end{array} \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} n \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} m+1 \end{array}$$

Parameter error convergence

The steady-state parameter error satisfies

$$\boxed{E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} = 0}$$

Where the regressor correlation $E\left\{\phi(k)\phi^T(k)\right\}$ is:

$$E\{\phi(k)\phi(k)^T\} = \lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{j=-N}^{j=N}\phi(k+j)\phi^T(k+j)\right\}$$

Parameter error convergence

Since the steady-state parameter error satisfies

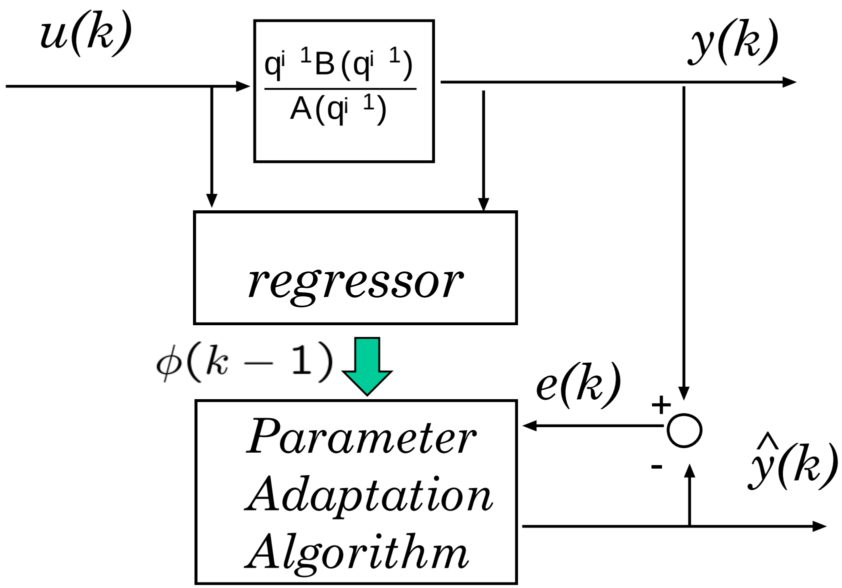
$$E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} \;=\; 0$$

$$E\left\{\phi(k)\phi^T(k)\right\} \succ 0 \quad \Rightarrow \quad \lim_{k\to\infty}\tilde{\theta}(k) = \bar{\theta} = 0$$

The regressor vector $\phi(k)$ is <u>persistently exciting</u> if

$$E\left\{\phi(k)\phi^T(k)\right\} \succ 0$$

Persistence of Excitation



$$\phi(k) = \begin{array}{c} \left[\begin{array}{c} -y(k) \\ \vdots \\ -y(k-n+1) \\ \\ u(k) \\ \vdots \\ u(k-m) \end{array} \right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c} -y(k) \\ \vdots \\ -y(k-n+1) \end{array}}\right\} n \\ \\ \left.\vphantom{\begin{array}{c} u(k) \\ \vdots \\ u(k-m) \end{array}}\right\} m+1 \end{array}$$

We need to find the conditions that the input sequence $u(k)$ must satisfy to guarantee that $\phi(k)$ is persistently exciting.

$$E\left\{\phi(k)\phi^T(k)\right\} \succ 0$$

Excitation matrix

Given and input sequence $u(k) \in \mathcal{R}$

Define the u-regressor of order n:

$$\phi_{u_n}(k) = \begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-n+1) \end{bmatrix} \in \mathcal{R}^n$$

only present and past values of u(k) are used

Excitation matrix

Given and input sequence $u(k) \in \mathcal{R}$

$$\phi_{u_n}(k) = \begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-n+1) \end{bmatrix} \in \mathcal{R}^n$$

Define the n x n excitation matrix:

$$C_n = \underbrace{\lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} \phi_{u_n}(k)\phi_{u_n}^T(k) \right\}}_{\text{Time average of } \phi_{u_n}(k)\phi_{u_n}^T(k)}$$

Persistence of Excitation (PE)

The input sequence $u(k)$

is **persistently exciting** of order n if
the $n \times n$ excitation matrix is **positive definite**

$$C_n \succ 0$$

$$\phi_{u_n}(k) = \begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-n+1) \end{bmatrix} \in \mathcal{R}^n$$

PE inputs in FIR models

Theorem:

$u(k)$ is persistently exciting (PE) of order $\boldsymbol{n}$ iff

the following holds **for all** nonzero polynomials

$A(q^{-1})$ of order at most n-1

$$U = \lim_{N \to \infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} w^2(k) \right\} > 0$$

where $w(k) = A(q^{-1})u(k)$

$w(k)$ is PE of order 1

PE inputs in FIR models

Alternate statement of Theorem:

The following are equivalent:

- $u(k)$ is PE of order n
- $A(q^{-1})u(k)$ is PE of order 1 for all nonzero polynomials $A(q^{-1})$ of degree at most n-1

PE inputs in FIR models

Proof: Let

$$A(q^{-1}) = a_o + a_1 q^{-1} + \cdots + a_{n-1}\, q^{n-1}$$

Then

$$A(q^{-1})\, u(k) = \underbrace{\left[\begin{array}{cccc} a_o & a_1 & \cdots & a_{n-1} \end{array}\right]}_{a^T} \underbrace{\left[\begin{array}{c} u(k) \\ u(k-1) \\ \vdots \\ u(k-n-1) \end{array}\right]}_{\phi(k)}$$

$$\boxed{w(k) = A(q^{-1})u(k) = a^T \phi(k) = \phi^T(k) a}$$

PE inputs in FIR models

Proof (cont'd):

$$U = \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} w^2(k) \right\}$$

$$= \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} a^T \phi(k) \phi^T(k) a \right\}$$

$$= a^T \left[\lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} \phi(k) \phi^T(k) \right\} \right] a$$

$$= a^T C_n a$$

PE inputs in FIR models

Proof (cont'd):

Since $U = a^T C_n a$, we see that $U > 0$, $\forall a \neq 0$ if and only if $C_n \succ 0$.

Therefore, $U > 0$ for all nonzero polynomials $A(q^{-1})$ of order at most n-1 <u>if and only if</u> $C_n \succ 0$

■

PE inputs in FIR models

To determine the PE order of a sequence $u(k)$

1. Find a nonzero polynomial $A(q^{-1})$ of order n such that $A(q^{-1})u(k)$ is not PE of order 1

   this means that $u(k)$ is PE of order <u>at most</u> n

2. Compute the excitation matrix C_n and verify that it is positive definite.

Conditions for PE

Examples: Constant input

$$u(k) = 1, \quad \forall k$$

$$(1 - q^{-1})u(k) = 0 \quad \Longrightarrow$$

u(k) is not PE of order 2

$$C_1 = \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{j=-N}^{j=N} u^2(k) \right\} = 1 > 0$$

$\Longrightarrow$ *u(k)* is PE of order 1

Conditions for PE in FIR Models

Examples: Sinusoid input

Consider the pure sinusoid input

$$u(k) = \sin(\omega\, k)\,. \qquad\qquad 0 < \omega < \pi$$

$$[1 - 2\cos(\omega)q^{-1} + q^{-2}]u(k) = 0$$

$\Longrightarrow$ *u(k)* is not PE of order 3

Conditions for PE in FIR Models

Examples: Sinusoid input

Let $\phi(k) = \begin{bmatrix} u(k) & u(k-1) \end{bmatrix}^T$. $\qquad u(k) = \sin(\omega\, k)$.

$$C_2 = \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{k=-N}^{N} \phi(k)\phi^T(k) \right\}$$

$$= \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \begin{bmatrix} \sum_{k=-N}^{N} u^2(k) & \sum_{k=-N}^{N} u(k)u(k-1) \\ \sum_{k=-N}^{N} u(k)u(k-1) & \sum_{k=-N}^{N} u^2(k-1) \end{bmatrix} \right\}$$

$$C_2 = \frac{1}{2} \begin{bmatrix} 1 & \cos(\omega) \\ \cos(\omega) & 1 \end{bmatrix} \succ 0 \qquad\qquad 0 < \omega < \pi$$

Conditions for PE in FIR Models

Examples: Sinusoid input

$$[1 - 2\cos(\omega)q^{-1} + q^{-2}]u(k) = 0 \quad \Longrightarrow$$

u(k) is not PE of order 3

$$C_2 = \frac{1}{2}\begin{bmatrix} 1 & \cos(\omega) \\ \cos(\omega) & 1 \end{bmatrix} \succ 0 \quad \Longrightarrow$$

u(k) is PE of order 2

Conditions for PE in FIR Models

Examples: Sum of Sinusoids

Consider an input that is a sum of m sinusoids, with m distinct frequencies

$$u(k) = \sum_{i=1}^{m} \sin(\omega_i\, k)\,.$$

$$0 < \omega_i < \pi$$
$$\omega_i \neq w_j$$

$$\boxed{u(k) \text{ is PE of order } n = 2m.}$$

Conditions for PE in FIR Models

Examples: Random process

Consider a colored random process

$$u(k) = G(q)\, w(k)$$

where $w(k)$ is white noise and G(q) is nonzero.

$$\boxed{u(k) \text{ is PE of any order.}}$$

PE in Filtered Signals

Filtered signals:

Let

- $u(k)$ be PE of order n
- *v*(*k*) be the output of the model

$$v(k) = A(q^{-1})u(k)$$

- $A(q^{-1})$ is a nonzero polynomial of degree $m < n$

$v(k)$ is PE of order r

for some r satisfying.

$$n - m \leq r \leq n$$

PE in Filtered Signals

Filtered signals:

- $u(k)$ be PE of order n
- Let $$\boxed{v(k) = \frac{1}{A(q^{-1})}u(k)}$$
- $A(q^{-1})$ is an anti-Schur polynomial

$$\boxed{v(k) \text{ is also PE of order } n.}$$

PE in Filtered Signals

Theorem

Let $v(k)$ be the output of the model

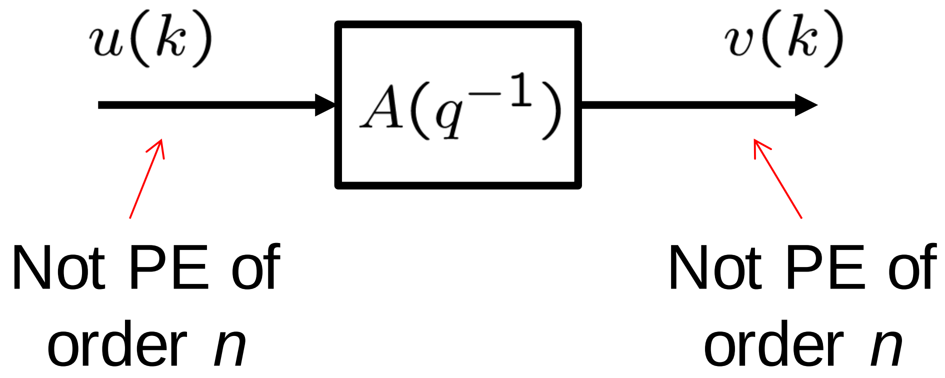
$$v(k) = A(q^{-1})u(k)$$

where $A(q^{-1})$ is a nonzero polynomial

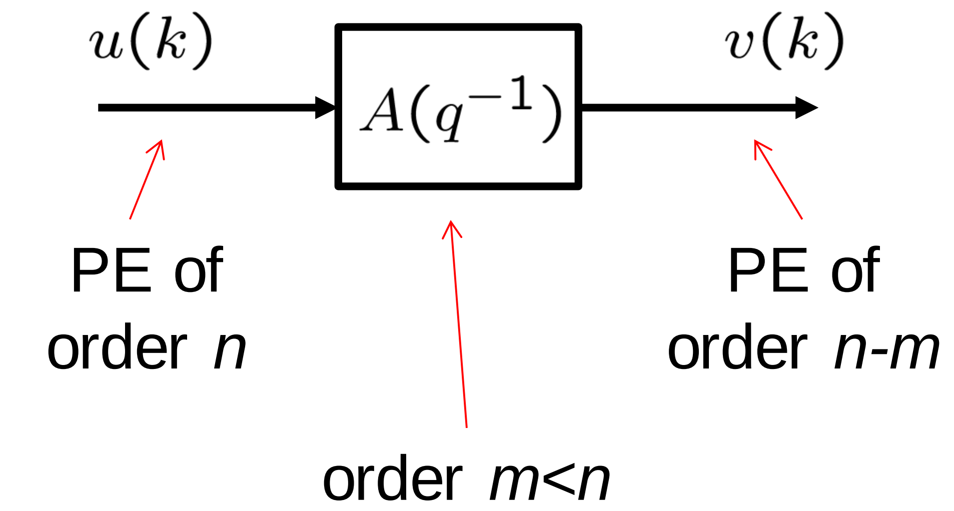
1. If $u(k)$ is <u>not</u> PE of order n, then $v(k)$ is <u>not</u> PE of order n
2. If u(k) is PE of order n and $A(q^{-1})$ has degree $m < n$, then $v(k)$ is PE of order n-m
3. If $A(q^{-1})$ is anti-Schur, then u(k) is PE of order n if and only if $v(k)$ is PE of order n

Interpretation of Theorem

1.

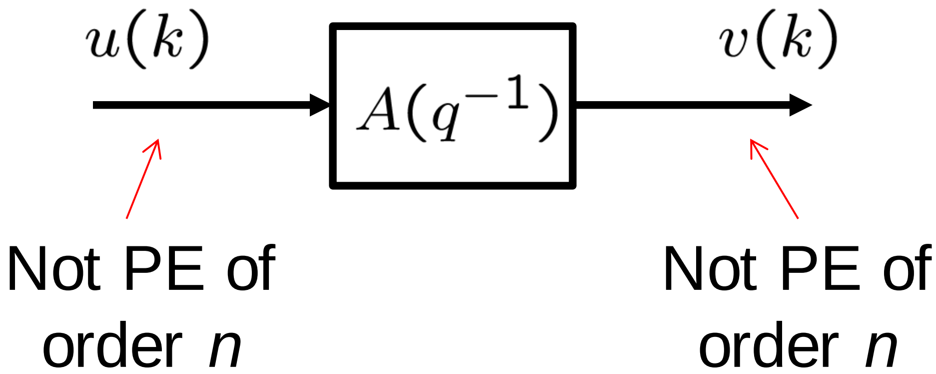


2.

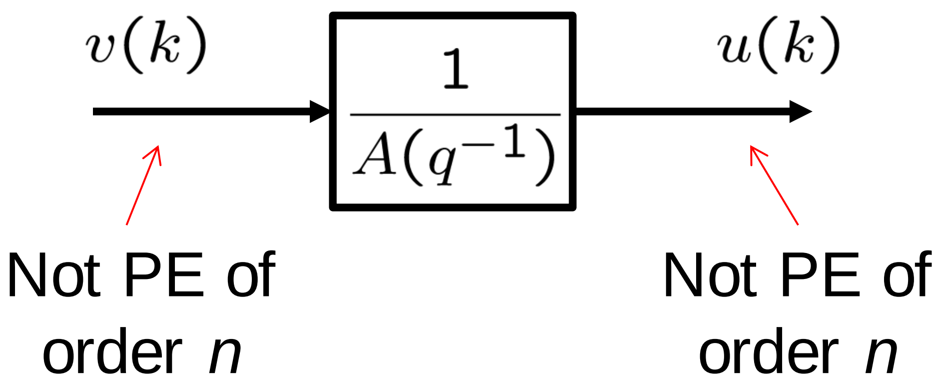


Interpretation of Theorem

3. When $A(q^{-1})$ anti-Schur



(this is redundant with part 1 of the theorem)



PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Preliminary result 1:

If $u(k)$ is not PE of order 1, then $v(k)$ is not PE of order 1

Proof:

Let $A(q^{-1}) = a_0 + a_1 q^{-1} + \cdots + a_{n-1} q^{-n+1}$

$$\Rightarrow A(q^{-1})\, u(k) = \underbrace{\begin{bmatrix} a_o & a_1 & \cdots & a_{n-1} \end{bmatrix}}_{a^T} \underbrace{\begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-n-1) \end{bmatrix}}_{\phi(k)}$$

PE in Filtered Signals

Proof of preliminary result 1 (continued):

$$v(k) = A(q^{-1})u(k) = a^T\phi(k)$$

$$U = \lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{k=-N}^{N} v^2(k)\right\}$$

$$= a^T\left(\lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{k=-N}^{N}\phi(k)\phi^T(k)\right\}\right)a \quad = a^T C_n a$$

Since $u(k)$ is not PE of order 1, $C_1 = 0$

$\Longrightarrow$ The diagonal elements of C_n are zero

Since $C_n \succeq 0$, this implies that $C_n = 0$

$\Longrightarrow$ $U = 0$, which implies that $v(k)$ is not PE of order 1

∎

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Preliminary result 2:

If $A(q^{-1})$ is anti-Schur and $v(k)$ is not PE of order 1,

then $\dfrac{1}{A(q^{-1})}v(k)$ is not PE of order 1

The proof is based on frequency domain techniques for deterministic signals that are analogous to power spectral density techniques for wide sense stationary random signals

(see the additional material at the end of this lecture)

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Proof of (1):

Let $u(k)$ not be PE of order n

Choose nonzero $B(q^{-1})$ of degree at most n-1 such that $w(k) = B(q^{-1})\, u(k)$ is not PE of order 1

$$B(q^{-1})v(k) = A(q^{-1})B(q^{-1})u(k) \quad = A(q^{-1})w(k)$$

By the preliminary result, $A(q^{-1})\, w(k)$ is not PE of order 1, which implies that $B(q^{-1})\, v(k)$ is not PE of order 1

$\Rightarrow$ $v(k)$ is not PE of order n

$\blacksquare$

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Proof of (2):

Let u(k) be PE of order n and $A(q^{-1})$ have degree $m < n$

Suppose $B(q^{-1})v(k)$ is not PE of order 1 where $B(q^{-1})$ has order at most n-m-1

$\Rightarrow$ $B(q^{-1})A(q^{-1})u(k)$ is not PE of order 1

Since $B(q^{-1})A(q^{-1})$ has order at most n-1 and u(k) is PE of order n, $B(q^{-1})A(q^{-1})$ is the zero polynomial

Since $A(q^{-1})$ is a nonzero polynomial, $B(q^{-1})$ is the zero polynomial

$\Rightarrow$ $v(k)$ is PE of order n-m $\blacksquare$

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Proof of (3):

By statement (1) of the theorem, if $u(k)$ is not PE of order n, then $v(k)$ is not PE of order n

It only remains to show that if $v(k)$ is not PE of order n, then $u(k)$ is not PE of order n

Let $v(k)$ not be PE of order n and choose nonzero $B(q^{-1})$ of order at most n-1 such that w(k)= $B(q^{-1})v(k)$ is not PE of order 1

This implies that $A(q^{-1})B(q^{-1})u(k)$ is not PE of order 1

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Proof of (3), continued:

$A(q^{-1})B(q^{-1})u(k)$ is not PE of order 1

Since $A(q^{-1})$ is anti-Schur, we use preliminary result 2 to see that $B(q^{-1})u(k)$ is not PE of order 1

Since $B(q^{-1})$ is a nonzero polynomial of order at most n-1

$\Longrightarrow$ $u(k)$ is not PE of order n

$\blacksquare$

ARMA Model (review)

$$A(q^{-1})y(k) = q^{-\mathrm{d}}B(q^{-1})u(k)$$

Where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n}$$ **(anti-Schur)**

$$B(q^{-1}) = b_o + b_1 q^{-1} + \cdots + b_m q^{-m}$$

- Orders n and m are <u>known</u>
- Relative degree d is <u>known</u>
- a's and b's are <u>unknown</u> but <u>constant</u> coefficients

ARMA Model (review)

$$y(k) = \phi^T(k-1)\,\theta$$

Unknown parameter vector:

$$\theta = \begin{bmatrix} \left.\begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix}\right\} n \\ \left.\begin{matrix} b_o \\ \vdots \\ b_m \end{matrix}\right\} m+1 \end{bmatrix}$$

Known regressor vector:

$$\phi(k-1) = \left.\begin{bmatrix} -y(k-1) \\ \vdots \\ -y(k-n) \\ u(k-\mathrm{d}) \\ \vdots \\ u(k-\mathrm{d}-m) \end{bmatrix}\right\} n+m+1$$

PE in ARMA models

Theorem:

Consider the parameter estimation of the ARMA system using the LS estimation algorithm. If

- $A(q^{-1})$ is anti-Schur
- $A(q^{-1})$ and $B(q^{-1})$ are co-prime
- $\boxed{u(k) \text{ is PE of order } n+m+1}$

Parameter estimates convergence to the true values

PE in ARMA models - Proof

Simplifying assumption: the parameter error converges

$$\bar{\theta} = \lim_{k \to \infty} \tilde{\theta}(k) \;\; = \begin{bmatrix} \bar{a}_1 & \cdots & \bar{a}_n & \bar{b}_0 & \cdots & \bar{b}_m \end{bmatrix}^T$$

Define: the LS output estimation error by

$$e(k) = \phi(k-1)^T \bar{\theta}$$

We know that

$$e(k) \to 0$$

PE in ARMA models - Proof

Notice that,

$$0 = \lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{k=-N}^{N} e^2(k)\right\}$$

$$= \bar{\theta}^T \lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{k=-N}^{N} \phi(k-1)\phi^T(k-1)\right\}\bar{\theta}$$

$$= \bar{\theta}^T C_{n+m+1}\bar{\theta}$$

Therefore, if we can show that $C_{n+m+1} \succ 0$, we will be able to conclude that $\bar{\theta} = 0$

PE in ARMA models - Proof

Notice that

$$e(k) = q^{-\mathsf{d}}\,\bar{B}(q^{-1})\,u(k) - \bar{A}(q^{-1})y(k)$$

where

$$\begin{aligned}\bar{A}(q^{-1}) &= A(q^{-1}) - \hat{A}(q^{-1}) \\ &= \bar{a}_1 q^{-1} + \cdots + \bar{a}_n q^{-n}\end{aligned}$$

$$\begin{aligned}\bar{B}(q^{-1}) &= B(q^{-1}) - \hat{B}(q^{-1}) \\ &= \bar{b}_0 + \cdots + \bar{b}_m q^{-m}\end{aligned}$$

PE in ARMA models - Proof

From

$$e(k) = q^{-\mathsf{d}}\,\bar{B}(q^{-1})\,u(k) - \bar{A}(q^{-1})y(k)$$

$$y(k) = \frac{q^{-\mathsf{d}}B(q^{-1})}{A(q^{-1})}\,u(k)$$

We obtain

$$e(k) = q^{-\mathsf{d}}\,\bar{B}(q^{-1})\,u(k) - \bar{A}(q^{-1})\frac{q^{-\mathsf{d}}\,B(q^{-1})}{A(q^{-1})}u(k)$$

$$= q^{-\mathsf{d}}\,\left[\bar{B}(q^{-1})A(q^{-1}) - \bar{A}(q^{-1})B(q^{-1})\right]\,\frac{1}{A(q^{-1})}\,u(k)\,.$$

PE in ARMA models - Proof

$$e(k) = q^{-\mathrm{d}} \underbrace{\left[\bar{B}(q^{-1})\, A(q^{-1}) - \bar{A}(q^{-1}) B(q^{-1})\right]}_{\text{Polynomial of order } \boldsymbol{n+m}} \underbrace{\frac{1}{A(q^{-1})}\, u(k)}_{v(k)} \; .$$

Notice that since $A(q^{-1})$ is anti-Schur and

$$v(k) = \frac{1}{A(q^{-1})} u(k)$$

$u(k)$ is PE of order $n+m+1$

$\Updownarrow$

$v(k)$ is PE of order $n+m+1$

PE in ARMA models - Proof

$$e(k) = q^{-\mathrm{d}} \underbrace{\left[\bar{B}(q^{-1})\,A(q^{-1}) - \bar{A}(q^{-1})B(q^{-1})\right]}_{\text{Polynomial of order } n+m} \underbrace{\frac{1}{A(q^{-1})}\,u(k)}_{v(k)}\,.$$

- $v(k)$ is PE of order $n+m+1$
- $e(k)$ is PE of order 1 **unless**

$$\left[\bar{B}(q^{-1})\,A(q^{-1}) - \bar{A}(q^{-1})\,B(q^{-1})\right] = 0$$

Since *e(k)* = 0, it cannot be PE of order 1

Therefore, $\left[\bar{B}(q^{-1})\,A(q^{-1}) - \bar{A}(q^{-1})\,B(q^{-1})\right] = 0$

PE in ARMA models - Proof

So far, we know that if

$$u(k) \text{ is PE of order } n+m+1,$$

then

$$\left[\bar{B}(q^{-1})\,A(q^{-1}) - \bar{A}(q^{-1})\,B(q^{-1})\right] = 0$$

where $A(q^{-1})$ and $B(q^{-1})$ are co-prime

$$\bar{A}(q^{-1}) = A(q^{-1}) - \widehat{A}(q^{-1})$$

$$\bar{B}(q^{-1}) = B(q^{-1}) - \widehat{B}(q^{-1})$$

PE in ARMA models - Proof

$$\left[\bar{B}(q^{-1})\,A(q^{-1}) - \bar{A}(q^{-1})\,B(q^{-1})\right] = 0$$

This equation can be written as follows:

$$D\,\bar{\theta}^{*} = 0$$

where

$$\bar{\theta}^{*} = \begin{bmatrix} \bar{b}_o \cdots & \bar{b}_m & -\bar{a}_1 & \cdots & -\bar{a}_n \end{bmatrix}^T \in \mathcal{R}^{n+m+1}$$

and:

$$\bar{a}_i = a_i - \widehat{a}_i$$

$$\bar{b}_i = b_i - \widehat{b}_i$$

PE in ARMA models - Proof

$$D = \left[\underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_1 & 1 & \ddots & \vdots \\ a_2 & a_1 & \ddots & 0 \\ \vdots & a_2 & \ddots & 1 \\ a_{n-1} & \vdots & \ddots & a_1 \\ a_n & a_{n-1} & \ddots & a_2 \\ 0 & a_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1} \\ 0 & \cdots & 0 & a_n \end{bmatrix}}_{m+1} \; \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ b_0 & 0 & \cdots & 0 & 0 \\ b_1 & b_0 & \ddots & \vdots & \vdots \\ \vdots & b_1 & \ddots & 0 & 0 \\ b_{m-1} & \vdots & \ddots & b_0 & 0 \\ b_m & b_{m-1} & \ddots & b_1 & b_0 \\ 0 & b_m & \ddots & \vdots & b_1 \\ 0 & 0 & \ddots & b_{m-1} & \vdots \\ \vdots & \vdots & \ddots & b_m & b_{m-1} \\ 0 & 0 & \ddots & 0 & b_m \end{bmatrix}}_{n} \right]$$

PE in ARMA models - Proof

$$D\,\bar{\theta}^* = 0$$

$A(q^{-1})$ and $B(q^{-1})$ are co-prime

$\Downarrow$

D is nonsingular and $\bar{\theta}^* = 0$

Therefore, when $u(k)$ is PE of order $n+m+1$

Parameter estimates convergence to the true values

Example

- Plant:

$$y(k) = \frac{q^{-1}\,0.1(1+0.5q^{-1})}{(1+0.9q^{-1})(1+0.8q^{-1})}\,u(k)$$

$$y(k+1) \;=\; \theta^T\,\phi(k)$$

$$\theta = \begin{bmatrix} 1.7 \\ 0.72 \\ 0.1 \\ 0.05 \end{bmatrix} \in \mathcal{R}^4 \qquad \phi(k) = \begin{bmatrix} -y(k) \\ -y(k-1) \\ u(k) \\ u(k-1) \end{bmatrix} \in \mathcal{R}^4$$

- We need *u(k)* to be a PE sequence of order 4 to guarantee parameter convergence

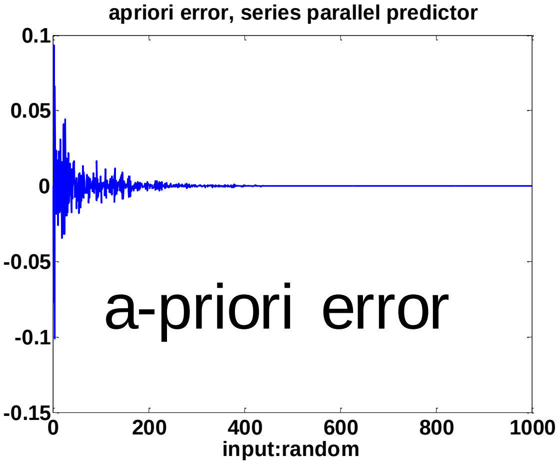
Example: Input Random Noise

u(k) : zero mean uniform white noise between [-1,1]

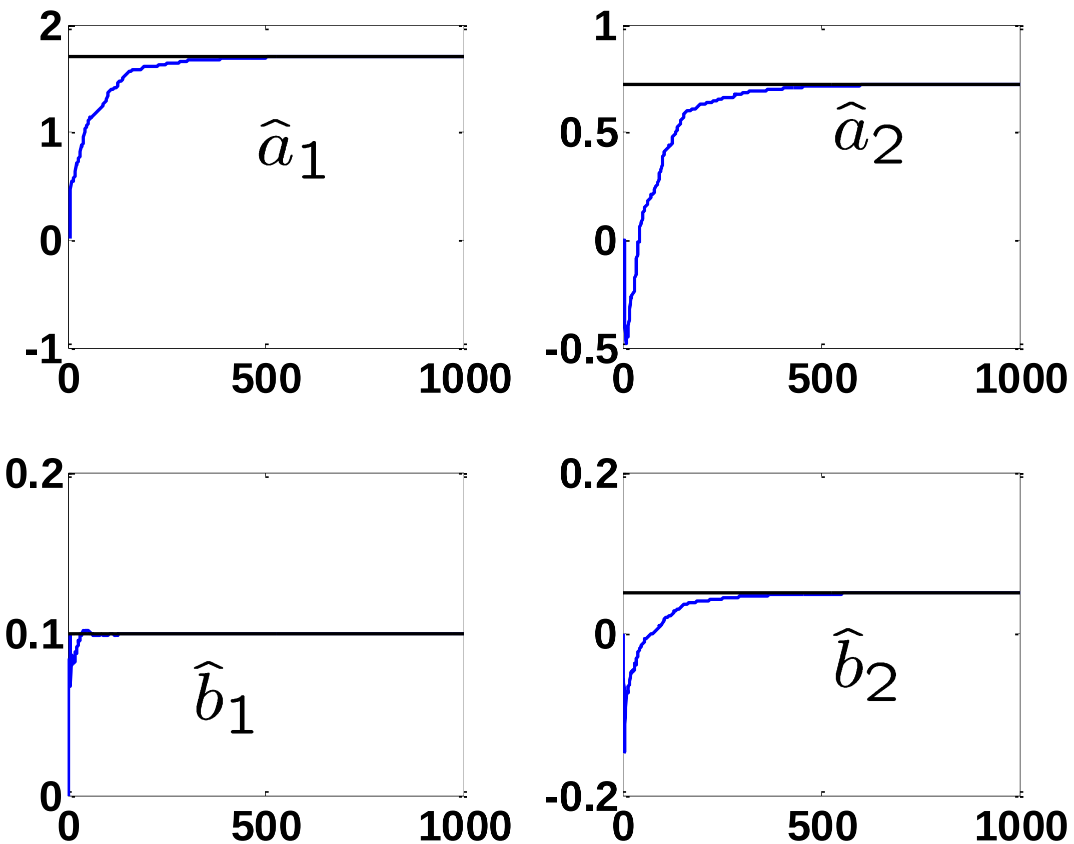
$\boxed{u(k) \text{ is PE of any order.}}$

$$F(0) = 100 * I_4$$

$$\lambda_1 = 0.99 \qquad \lambda_2 = 1$$



parameter convergence



Example: Step Input

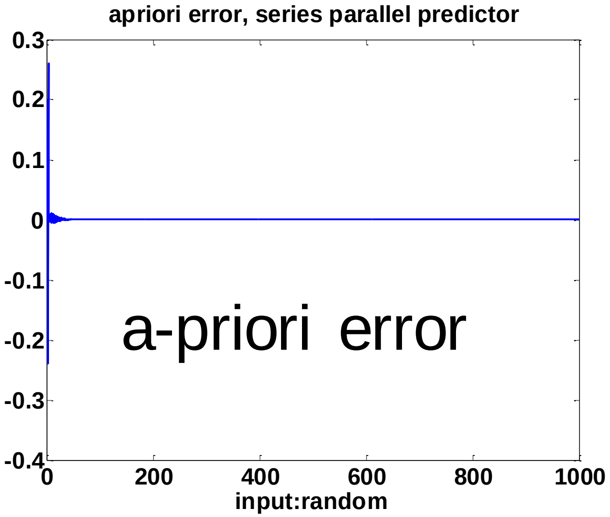
u(k) **= 2*1(t)**

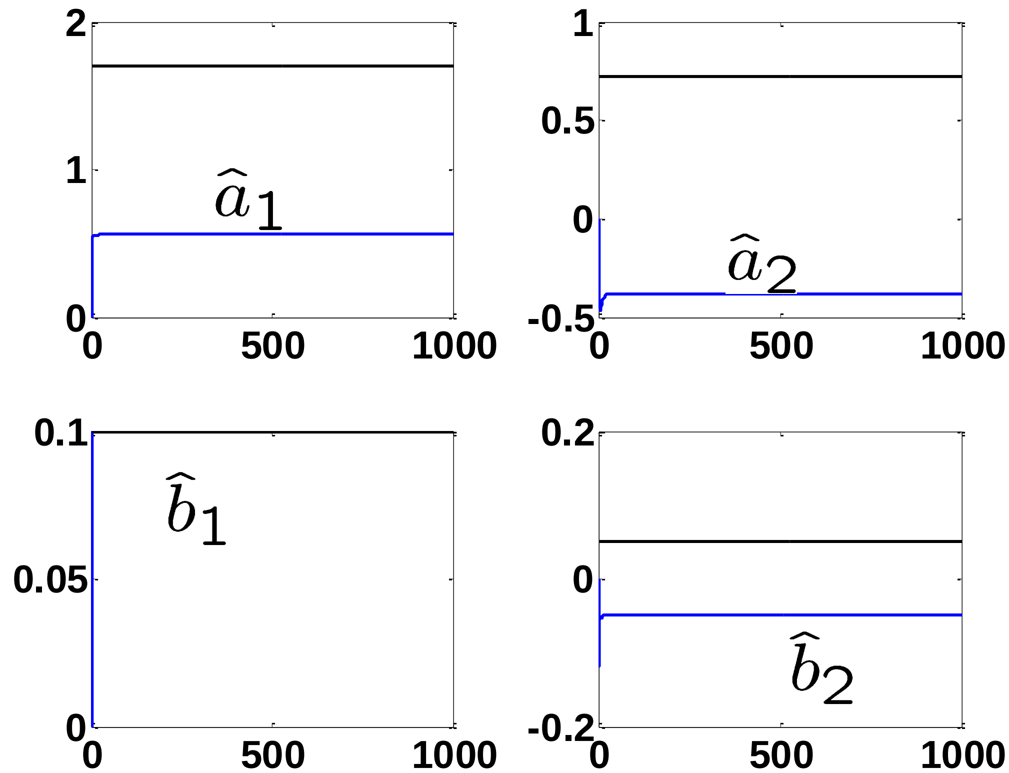
$$\boxed{u(k) \text{ is PE of order 1.}}$$

NO parameter convergence

$$F(0) = 100 * I_4$$

$$\lambda_1 = 0.99 \qquad \lambda_2 = 1$$



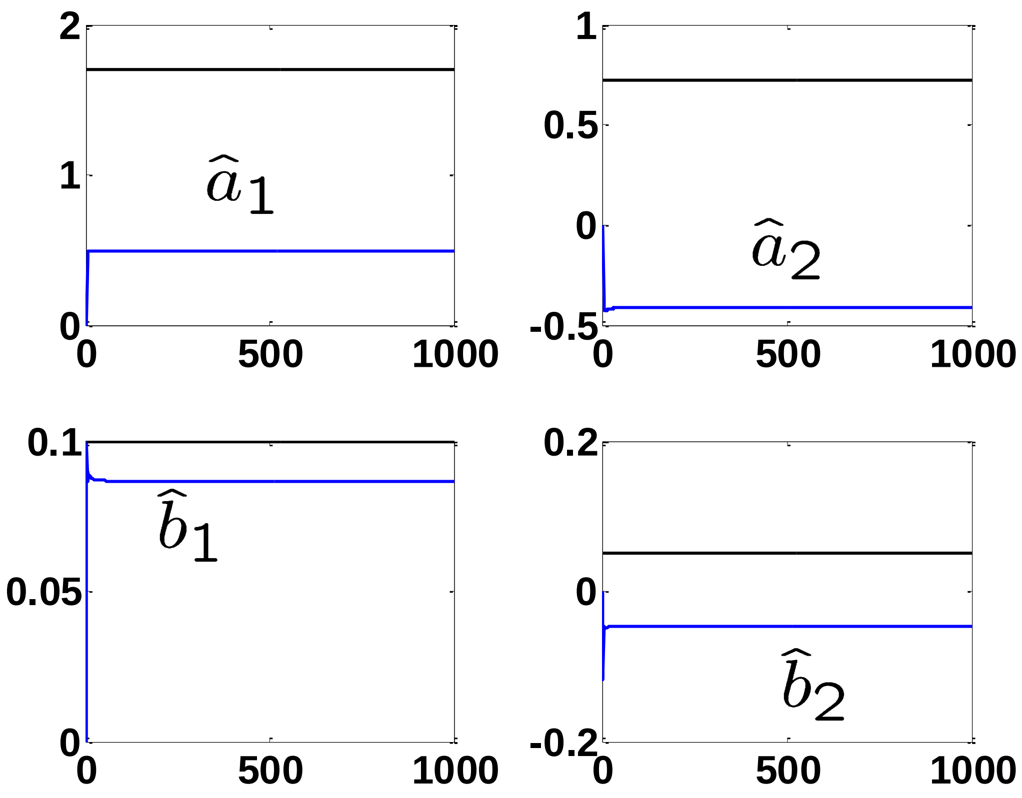


Example: Sinusoidal input – 1 frequency

$u(k)$ = 2*sin(t)

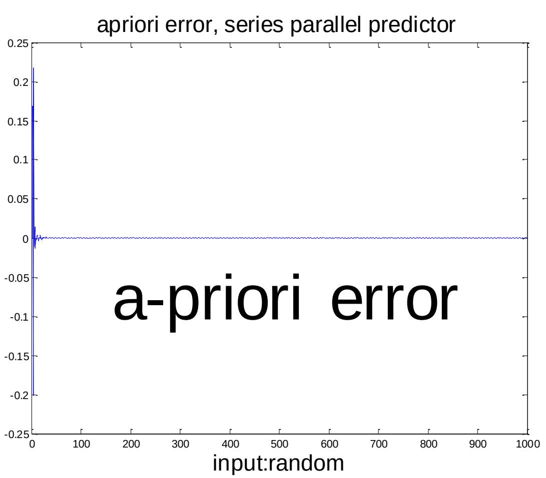
$$\boxed{u(k) \text{ is PE of order 2.}}$$

<u>NO</u> parameter convergence



$$F(0) = 100 * I_4$$

$$\lambda_1 = 0.99 \qquad \lambda_2 = 1$$

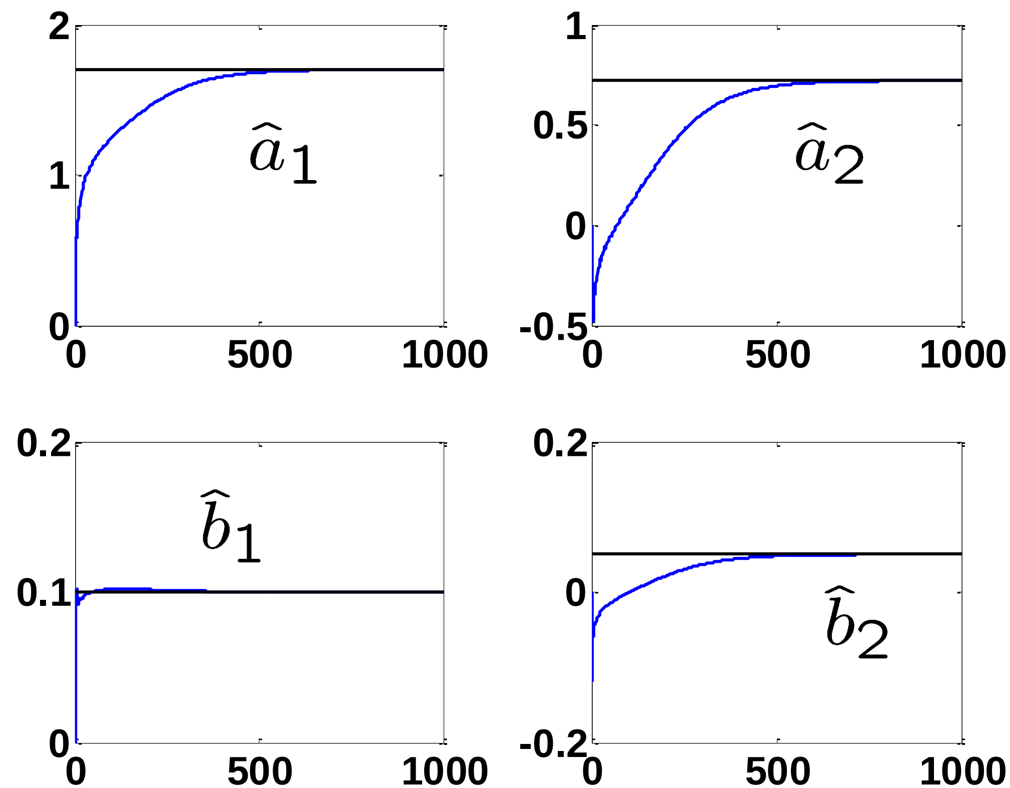


Example: Sinusoidal input – 2 frequencies

u(k) **= 2*sin(t)+2cos(2*t)**

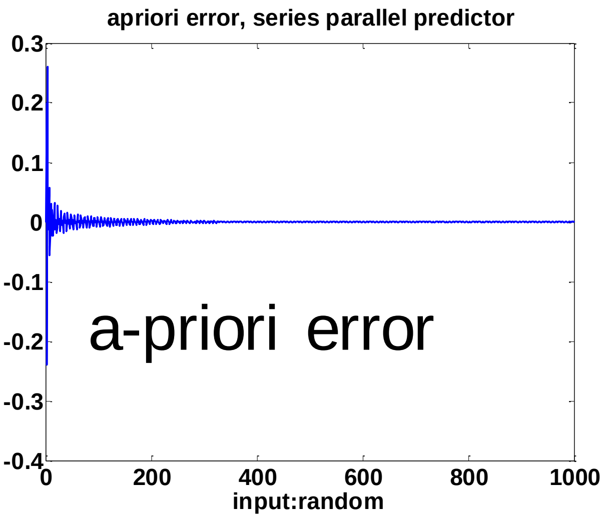
$\boxed{u(k) \text{ is PE of order 4.}}$

parameter convergence



$$F(0) = 100 * I_4$$

$$\lambda_1 = 0.99 \qquad \lambda_2 = 1$$



Additional Material
(you are not responsible for this)

- Proof of preliminary result 2

PE in Filtered Signals

$$v(k) = A(q^{-1})u(k)$$

Preliminary result 2:

If $A(q^{-1})$ is anti-Schur and $v(k)$ is not PE of order 1,

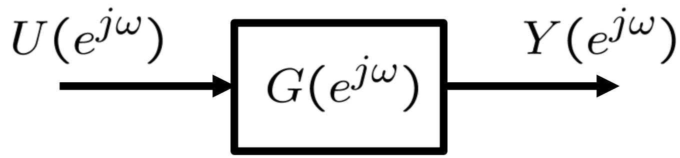
then $\dfrac{1}{A(q^{-1})}v(k)$ is not PE of order 1

The proof is based on frequency domain techniques for deterministic signals that are analogous to power spectral density techniques for wide sense stationary random signals

Stochastic and Deterministic Signals

WSS zero-mean random signals, *X*(*k*) and *Y*(*k*)	Deterministic signals, *x*(*k*) and *y*(*k*)
$\Lambda_{XY}(j) = E\left\{X(k+j)Y^T(k)\right\}$	$\Gamma_{xy}(j) = \underbrace{\lim_{N\to\infty}\left(\frac{1}{2N+1}\sum_{k=-N}^{N} x(k+j)y^T(k)\right)}_{\text{Average value of } x(k+j)y^T(k) \text{ over } k}$
$\Phi_{XX}(\omega) = \mathcal{F}\left\{\Lambda_{XX}(\cdot)\right\}$	$\Psi_{xx}(\omega) = \mathcal{F}\left\{\Gamma_{xx}(\cdot)\right\}$
$\Lambda_{XX}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{XX}(\omega)d\omega$	$\Gamma_{xx}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Psi_{xx}(\omega)d\omega$

Stochastic and Deterministic Signals



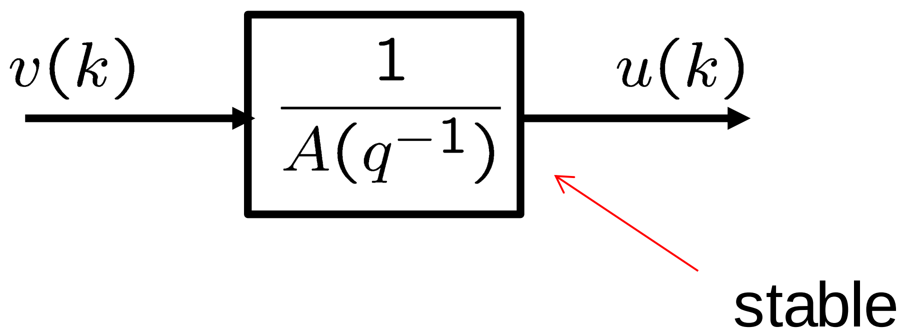
where $G(z)$ is stable

WSS zero-mean random signals, $U(k)$ and $Y(k)$	Deterministic signals, $u(k)$ and $y(k)$
$\Phi_{YY}(\omega) = G(e^{j\omega})\Phi_{UU}(\omega)G^*(e^{j\omega})$	$\Psi_{YY}(\omega) = G(e^{j\omega})\Psi_{UU}(\omega)G^*(e^{j\omega})$
↓ Scalar $U(k)$ and $Y(k)$	↓ Scalar $u(k)$ and $y(k)$
$\Phi_{YY}(\omega) = \lvert G(e^{j\omega})\rvert^2\Phi_{UU}(\omega)$	$\Psi_{YY}(\omega) = \lvert G(e^{j\omega})\rvert^2\Psi_{UU}(\omega)$

Proof of Preliminary Result 2

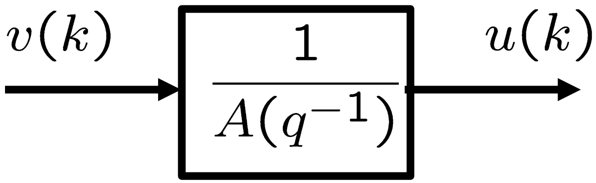
Let

- $A(q^{-1})$ be anti-Schur
- $v(k)$ not be PE of order 1
- u(k) be generated by



Choose M such that $\left|\frac{1}{A(e^{-j\omega})}\right|^2 \leq M, \quad \forall\omega \in [0, 2\pi]$

Proof of Preliminary Result 2



$$\left|\frac{1}{A(e^{-j\omega})}\right|^2 \leq M, \quad \forall \omega \in [0, 2\pi]$$

$$\Gamma_{uu}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_{uu}(\omega) d\omega \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\left|\frac{1}{A(e^{-j\omega})}\right|^2 \Psi_{vv}(\omega)\right] d\omega$$

$$\leq M \frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_{vv}(\omega) d\omega \;=\; M\Gamma_{vv}(0)$$

Therefore, we have

$$0 \leq \Gamma_{uu}(0) \leq M\Gamma_{vv}(0)$$

Proof of Preliminary Result 2

$$0 \leq \Gamma_{uu}(0) \leq M\Gamma_{vv}(0)$$

v(*k*) not PE of order 1 $\Longrightarrow$ $\Gamma_{vv}(0) = 0$

$\Longrightarrow$ $0 \leq \Gamma_{uu}(0) \leq 0$

$\Longrightarrow$ $\Gamma_{uu}(0) = 0$

$\Longrightarrow$ *u*(*k*) not PE of order 1

$\blacksquare$

PE inputs

To determine the PE order of a sequence $u(k)$

1. Find an annihilating polynomial $A_n(q^{-1})$ of order n such

$$A_n(q^{-1})u(k) = 0$$

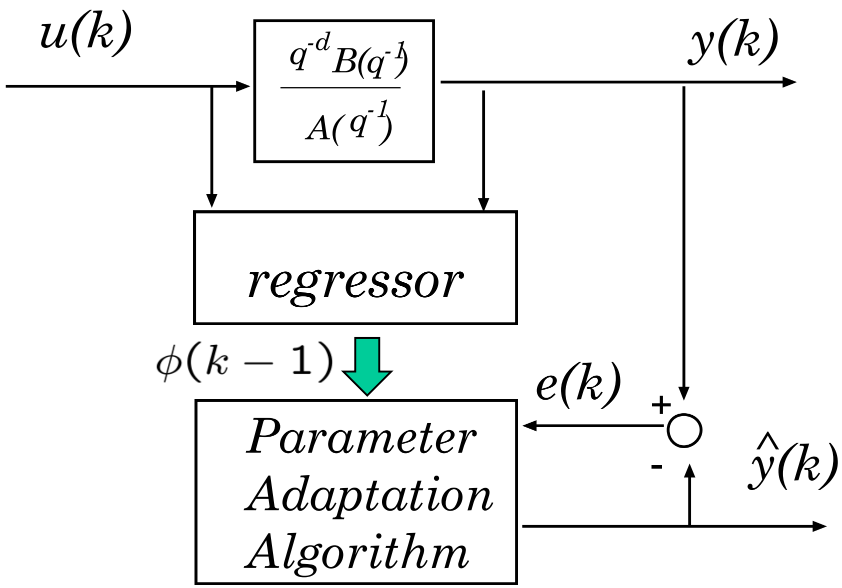
this means that $u(k)$ <u>is at most PE of order n</u>

2. Compute the excitation matrix

$$C_n = E\{\phi_{u_n}(k)\phi_{u_n}^T(k)\} \succ 0$$

and verify that it is positive definite.

Persistence of excitation for ARMA model identification



$$\phi(k) = \begin{array}{c} \left[\begin{array}{c} -y(k) \\ \vdots \\ -y(k-n+1) \\ \\ u(k-d) \\ \vdots \\ u(k-m-d) \end{array} \right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c} a \\ b \\ c \end{array}}\right\} n \\ \\ \left.\vphantom{\begin{array}{c} a \\ b \\ c \end{array}}\right\} m+1 \end{array}$$

We need to find what conditions must the input sequence *u(k)* satisfy so that $\phi(k)$ is persistently exciting.

$$E\left\{\phi(k)\phi^T(k)\right\} \succ 0$$

PE in ARMA models

Given:

$$y(k) = \frac{q^{-\mathsf{d}} B(q^{-1})}{A(q^{-1})} u(k) \qquad \phi(k) = \begin{array}{l} \left[\begin{array}{c} -y(k) \\ \vdots \\ -y(k-n+1) \\ u(k-d) \\ \vdots \\ u(k-m-d) \end{array} \right] \begin{array}{l} \left.\vphantom{\begin{array}{c} a \\ b \\ c \end{array}}\right\} n \\ \left.\vphantom{\begin{array}{c} a \\ b \\ c \end{array}}\right\} m+1 \end{array} \end{array}$$

- $u(k)$ is bounded
- $A(q^{-1})$ is Schur
- $A(q^{-1})$ and $B(q^{-1})$ are co-prime

$$\boxed{\begin{array}{c} u(k) \text{ is PE of order } n+m+1 \\ \Updownarrow \\ E\left\{\phi(k)\phi^T(k)\right\} \succ 0 \end{array}}$$

Derivation of Results

1. Determine conditions on the input sequence

$$u(k) \in \mathcal{R}$$

- For the parameter convergence of a Moving Average (MA) model

$$y(k) \;=\; q^{-\mathsf{d}}\, B(q^{-1})\, u(k)$$

- For the parameter convergence of an ARMA model

$$A(q^{-1})\, y(k) \;=\; q^{-\mathsf{d}}\, B(q^{-1})\, u(k)$$

Statistical Interpretation of LS Estimation

Stochastic Model

$$\boxed{y(k) = \phi^T(k-1)\,\theta + \epsilon(k)}$$

Where

- $y(k)$ observed output
- $\epsilon(k)$ **zero-mean noise**
- $\phi(k) = \begin{bmatrix} \phi_1(k) & \cdots & \phi_n(k) \end{bmatrix}^T$ regressor
- $\theta = \begin{bmatrix} \theta_1 & \cdots & \theta_n \end{bmatrix}^T$ unknown parameter vector

Statistical Interpretation of LS Estimation

Assumptions:

- $E\{\epsilon(k)\} = 0$ zero-mean
- Independence or orthogonality:

$$E\{\phi(k)\epsilon(k)\} = E\{\phi(k)\}E\{\epsilon(k)\} = 0$$

- Ergodicity

$$E\{\phi(k)\phi(k)^T\} = \lim_{N\to\infty}\left\{\frac{1}{2N+1}\sum_{j=-N}^{j=N}\phi(k+j)\phi^T(k+j)\right\}$$

Statistical Interpretation of LS Estimation

Collect data for k observations:

$$y(k) = \phi^T(k-1)\,\theta + \epsilon(k)$$

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(k) \end{bmatrix}}_{Y(k)} = \underbrace{\begin{bmatrix} \phi_1(0) & \cdots & \phi_n(0) \\ \phi_1(1) & \cdots & \phi_n(1) \\ . & \cdots & . \\ \phi_1(k-1) & \cdots & \phi_n(k-1) \end{bmatrix}}_{\Phi^T(k-1)} \underbrace{\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}}_{\theta} + \underbrace{\begin{bmatrix} \epsilon(1) \\ \epsilon(2) \\ \vdots \\ \epsilon(k) \end{bmatrix}}_{\mathcal{E}(k)}$$

LS Statistical Interpretation

Collect data for k observations:

$$Y(k) = \Phi^T(k-1)\,\theta + \mathcal{E}(k)$$

Where

- $$Y(k) \;=\; \left[\begin{array}{ccc} y(1) & \cdots & y(k) \end{array}\right]^T \in \mathcal{R}^k$$
- $$\Phi(k-1) \;=\; \left[\begin{array}{c} \phi(0)\cdots\phi(k-1) \end{array}\right] \in \mathcal{R}^{n\times k}$$
- $$\mathcal{E}(k) \;=\; \left[\begin{array}{ccc} \epsilon(1) & \cdots & \epsilon(k) \end{array}\right]^T \in \mathcal{R}^k$$
- $$\theta = \left[\begin{array}{ccc} \theta_1 & \cdots & \theta_n \end{array}\right]^T \in \mathcal{R}^n$$

LS Statistical Interpretation

$$\Phi(k-1) = \left[\ \phi(0)\cdots\phi(k-1)\ \right] \in \mathcal{R}^{n\times k}$$

$$= \begin{bmatrix} \phi_1(0) & \cdots & \phi_1(k-1) \\ \phi_2(0) & \cdots & \phi_2(k-1) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \phi_n(0) & \cdots & \phi_n(k-1) \end{bmatrix}$$

Deterministic Least Squares Estimation

Parameter estimate after k observations: $\widehat{\theta}(k)$

$$y(1), \cdots, y(k)$$

$$\phi(0), \cdots, \phi(k-1)$$

Which minimizes the following cost functional:

$$V(\widehat{\theta}(k)) = \frac{1}{2} \sum_{j=1}^{k} \left[y(j) - \phi^T(j-1)\, \widehat{\theta}(k) \right]^2$$

Notice that $\widehat{\theta}(k)$ is kept constant in the summation

Deterministic Least Squares Estimation

$\widehat{\theta}(k)$: Parameter estimate which minimizes

$$V(\widehat{\theta}(k))$$

Is given by the **Normal Equation**:

$$\boxed{\Phi(k-1)\Phi(k-1)^T\,\widehat{\theta}(k) = \Phi(k-1)\,Y(k)}$$

LS Statistical Interpretation

Normal equation:

$$\boxed{\Phi(k-1)\Phi(k-1)^T\,\widehat{\theta}(k) = \Phi(k-1)\,Y(k)}$$

Stochastic model:

$$\boxed{Y(k) = \Phi^T(k-1)\,\theta + \mathcal{E}(k)}$$

Parameter error vector:

$$\boxed{\widetilde{\theta}(k) = \theta - \widehat{\theta}(k)}$$

LS Statistical Interpretation

Substitute the stochastic model

$$Y(k) = \Phi^T(k-1)\,\theta + \mathcal{E}(k)$$

Into the normal equation:

$$\Phi(k-1)\Phi(k-1)^T\,\widehat{\theta}(k) = \Phi(k-1)\,Y(k)$$

To obtain:

$$\boxed{\Phi(k-1)\Phi^T(k-1)\,\tilde{\theta}(k) = -\Phi(k-1)\mathcal{E}(k)\,.}$$

LS Statistical Interpretation

$$\Phi(k-1)\Phi^T(k-1)\,\tilde{\theta}(k) = -\Phi(k-1)\mathcal{E}(k)\,.$$

Notice that

$$\Phi(k-1) \;=\; \left[\; \phi(0)\cdots\phi(k-1) \;\right]$$

$$\mathcal{E}(k) \;=\; \left[\; \epsilon(1) \quad \cdots \quad \epsilon(k) \;\right]^T$$

Therefore,

$$\boxed{\left\{\sum_{j=0}^{k-1}\phi(j)\phi^T(j)\right\}\tilde{\theta}(k) = -\sum_{j=1}^{k}\phi(j-1)\epsilon(j)}$$

LS Statistical Interpretation

Assume now that the parameter error converges:

$$\boxed{\bar{\theta} = \lim_{k \to \infty} \tilde{\theta}(k)}$$

Multiply by $1/k$ and take limits as $k \longrightarrow \infty$

$$\lim_{k \to \infty} \left\{ \frac{1}{k} \sum_{j=0}^{k-1} \phi(j)\phi^T(j) \right\} \tilde{\theta}(k) = - \lim_{k \to \infty} \left\{ \frac{1}{k} \sum_{j=1}^{k} \phi(j-1)\epsilon(j) \right\}$$

$$\boxed{\lim_{k \to \infty} \left\{ \frac{1}{k} \sum_{j=0}^{k-1} \phi(j)\phi^T(j) \right\} \bar{\theta} = - \lim_{k \to \infty} \left\{ \frac{1}{k} \sum_{j=1}^{k} \phi(j-1)\epsilon(j) \right\}}$$

LS Statistical Interpretation

$$\lim_{k\to\infty}\left\{\frac{1}{k}\sum_{j=0}^{k-1}\phi(j)\phi^T(j)\right\}\bar{\theta} = -\lim_{k\to\infty}\left\{\frac{1}{k}\sum_{j=1}^{k}\phi(j-1)\epsilon(j)\right\}$$

By Ergodicity,

$$\boxed{E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} \;=\; -E\left\{\phi(k)\epsilon(k+1)\right\}}$$

LS Statistical Interpretation

$$E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} = -E\left\{\phi(k)\epsilon(k+1)\right\}$$

If $\phi(k)$ and $\epsilon(k)$ are independent or orthogonal,

$$\begin{aligned} E\left\{\phi(k)\epsilon(k+1)\right\} &= -E\left\{\phi(k)\right\} E\left\{\epsilon(k+1)\right\} \\ &= 0 \end{aligned}$$

Since, $E\left\{\epsilon(k)\right\} = 0$

LS Statistical Interpretation

The parameter error vector satisfies:

$$\boxed{E\left\{\phi(k)\phi^T(k)\right\}\bar{\theta} \;=\; 0}$$

Thus, a sufficient condition for $\bar{\theta} \;=\; 0$ is that

$$E\left\{\phi(k)\phi^T(k)\right\} > 0 \qquad \text{(positive definite)}$$

LS Statistical Interpretation

We now define the Excitation matrix $C_n \in \mathcal{R}^{n\times n}$

$$C_n = E\left\{\phi(k)\phi^T(k)\right\} \qquad \phi(k) \in \mathcal{R}^n$$

$$= \lim_{k\to\infty}\left\{\frac{1}{k}\sum_{j=0}^{k-1}\phi(j)\phi^T(j)\right\}$$

$$C_n = C_n^T \qquad C_n \geq 0$$

LS Statistical Interpretation

Theorem:

$$y(k) = \phi^T(k-1)\,\theta + \epsilon(k)$$

Under the conditions:

- $E\{\epsilon(k)\} = 0$
- $E\{\phi(k-1)\epsilon(k)\} = E\{\phi(k-1)\}\,E\{\epsilon(k)\} = 0 = 0$

If the excitation matrix C_n is positive definite,

the parameter error vector of the least square algorithm converges to zero.

$$\bar{\theta} = \lim_{k \longrightarrow \infty} \tilde{\theta}(k) = 0$$

Persistence of Excitation (PE)

Persistently exciting regressor: $\phi(k) \in \mathcal{R}^n$

There exist finite constants:

- $0 < m$
- $0 < \rho_1 < \rho_2 < \infty$

For all k

$$\boxed{\rho_2 \, I_n \geq \sum_{j=k}^{k+m} \phi(j)\phi(j)^T \geq \rho_1 \, I_n}$$

Persistence of Excitation (PE)

Persistently exciting regressor: $\phi(k) \in \mathcal{R}^n$

$$\boxed{\rho_2 I_n \geq \sum_{j=k}^{k+m} \phi(j)\phi(j)^T \geq \rho_1 I_n}$$

$$0 < \rho_1 < \lambda_{min} \left\{ \sum_{j=k}^{k+m} \phi(j)\phi(j)^T \right\}$$

$$\infty > \rho_2 > \lambda_{max} \left\{ \sum_{j=k}^{k+m} \phi(j)\phi(j)^T \right\}$$

for all k

and a fixed m

PE in Moving Average (MA) models

Finite Impulse Response (FIR) model:

$$
\begin{aligned}
y(k+1) &= B(q^{-1})\, u(k) \\
&= b_o\, u(k) + \cdots + b_{n-1}\, u(k-n+1) \\
&= \theta^T \phi(k)
\end{aligned}
$$

where

$$\theta = \begin{bmatrix} b_o & b_1 \cdots & b_{n-1} \end{bmatrix}^T \in \mathcal{R}^n$$

$$\phi(k) = \begin{bmatrix} u(k) & u(k-1) \cdots & u(k-n+1) \end{bmatrix}^T \in \mathcal{R}^n$$

Conditions for PE in FIR Models

Persistently exciting input sequence:

$u(k)$ Is persistently exciting (PE) of order n

if the regressor vector

$$\phi(k) = \left[\begin{array}{cccc} u(k) & u(k-1) \cdots & u(k-n+1) \end{array} \right]^T \in \mathcal{R}^n$$

is persistently exciting

Conditions for PE in FIR Models

For a persistently exciting input sequence $u(k)$ with regressor

$$\phi(k) = \begin{bmatrix} u(k) & u(k-1) \cdots & u(k-n+1) \end{bmatrix}^T \in \mathcal{R}^n$$

The excitation matrix C_n is a Positive Definite Toeplitz matrix

$$C_n = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ c_{1n} & \cdots & c_{nn} \end{bmatrix} \qquad \begin{aligned} c_{ij} &= c_{ji} \\ &= E\{u(k)\,u(k+i-j)\} \\ &= R_{uu}(i-j) \end{aligned}$$