Stochastic Signals and Systems

Lecture 9.

Unstable AR and MA processes
Multivariate systems

3 December 2020
REMINDER
Unstable MA processes, I.

Consider now an MA process s.t. \(C(z^{-1})\) not necessarily stable.

\[ y = C'(q^{-1})e' \quad (1) \]

where \(C'(q^{-1})\) is a polynomial of \(q^{-1}\) and \((e'_n)\) is a w.s.st. orthogonal process.

Assuming \(\sigma^2(e') = 1\) the spectral density of \((y_n)\) is given by

\[ f(\omega) = |C'(e^{-i\omega})|^2. \]

This can be obtained by restricting the complex-function

\[ g(z) = C'(z^{-1})C'(z) \]

to \(z = e^{-i\omega} \).
Unstable MA processes, II.

Proposition

Let \( C'(z^{-1}) \) be a polynomial of \( z^{-1} \) such that \( C'(z^{-1}) \neq 0 \) for \( |z| = 1 \).

Then \( \exists \) a stable polynomial \( C(z^{-1}) \) with \( \deg C = \deg C' \) such that

\[
C'(z^{-1})C'(z) = C(z^{-1})C(z). \tag{2}
\]

The stable factor is unique if the leading coefficient is fixed to 1.
The spectral density of a MA process can be written as

\[ f(\omega) = C(e^{-i\omega})C(e^{i\omega}) = |C(e^{-i\omega})|^2. \]

where \( C(z^{-1}) \) is a stable polynomial. This is called spectral factorization, and \( C(e^{-i\omega}) \) is called a stable spectral factor.

Define a new process \( e \) by its spectral representation process as

\[ d\zeta^e(\omega) = \frac{1}{C(e^{-i\omega})} d\zeta^y(\omega). \tag{3} \]

Write the spectral representation process of \( e \) as

\[ d\zeta^e(\omega) = \frac{1}{C(e^{-i\omega})} d\zeta^y(\omega) = \frac{C'(e^{-i\omega})}{C(e^{-i\omega})} d\zeta'^e(\omega). \tag{4} \]

The definition of \( e \), given in (4), implies

\[ y = C(q^{-1})e, \]

and since \( C(z^{-1}) \) stable, \( e \) is the innovation process of \( y \).
Proposition

Let \( y = (y_n) \) be an MA process given by \( y = C'(q^{-1})e' \) (1). Assume that \( C'(z^{-1}) \neq 0 \) for \( |z| = 1 \). Let \( C(e^{-i\omega}) \) be the stable spectral factor of \( f^y(\omega) \), and define \( e = (e_n) \) by (3). Then \( e \) is a w.s.st. orthogonal process,

\[
y = C(q^{-1})e,
\]

and \( e \) is the innovation process of \( y \).
ARMA processes, I.

The combination of AR and MA processes is called an ARMA process.

**Definition**

A w.s.st. process $y = (y_n)$ is called an ARMA process if it satisfies the dynamics

$$A(q^{-1}) y = C(q^{-1}) e,$$

where $(e_n)$ is a w.s.st. orthogonal process, and $A(q^{-1})$ and $C(q^{-1})$ are polynomials of the backward shift operator $q^{-1}$.

We use the following notations in defining an ARMA$(p, r)$ process:

$$A(q^{-1}) = \sum_{k=0}^{p} a_k q^{-k}, \quad C(q^{-1}) = \sum_{k=0}^{r} c_k q^{-k},$$

assuming that $a_0 = c_0 = 1$, and $a_p \neq 0$ and $c_r \neq 0$. 
ARMA processes, II.

**Proposition**

Consider the ARMA dynamics (5). Assume that \( A(z^{-1}) \neq 0 \) for \( |z| = 1 \). Then there is a unique w.s.st. process \( y = (y_n) \) satisfying (5). The process \( (y_n) \) has a spectral density equal to

\[
f^y(\omega) = \sigma^2(e) \frac{|C(e^{-i\omega})|^2}{|A(e^{-i\omega})|^2}.
\]

**Remark.** We assume that \( A(z^{-1}) \) and \( C(z^{-1}) \) have no common factor.
Let us consider an ARMA($p$, $r$) process $y = (y_n)$ defined by

$$A(q^{-1}) y = C(q^{-1}) e \quad \text{with} \quad \deg A = p, \ \deg C = r. \quad (6)$$

Here $e = (e_n)$ is a w.s.st. orthogonal process. We can ask ourselves: under what conditions $e$ is the innovation process of $y$.

**Proposition**

Assume that both $A(z^{-1})$ and $C(z^{-1})$ are stable, i.e. $A(z^{-1}) \neq 0$ and $C(z^{-1}) \neq 0$ for $|z| \geq 1$. Then $e = (e_n)$ is the innovation process of $y = (y_n)$.

The idea of the proof: expand both $\frac{C(e^{-i\omega})}{A(e^{-i\omega})}$ and $\frac{A(e^{-i\omega})}{C(e^{-i\omega})}$ into a power series of $e^{-i\omega}$, to infer both $H^y_n \subset H^e_n$ and $H^e_n \subset H^y_n$ $\forall n$. 
Stability and inverse stability, II.

A simple extension of the lemma on the expansion of $1/A(e^{-i\omega})$:

**Lemma**

If both $A(z^{-1})$ and $C(z^{-1})$ are stable, $a_0 = c_0 = 1$, then

$$\frac{C(e^{-i\omega})}{A(e^{-i\omega})} = \sum_{k=0}^{\infty} h_k e^{-ik\omega}, \text{ and } \frac{A(e^{-i\omega})}{C(e^{-i\omega})} = \sum_{k=0}^{\infty} g_k e^{-ik\omega},$$

with $h_0 = g_0 = 1$, where convergence on the r.h.s. is uniform in $\omega$.

It follows that the r.h.s. of $\frac{C(e^{-i\omega})}{A(e^{-i\omega})}$ converges in $L_2^c (dF^e)$, hence the random measure $\frac{C(e^{-i\omega})}{A(e^{-i\omega})} d\zeta^e(\omega)$ is well-defined. Similarly, the r.h.s. of $\frac{A(e^{-i\omega})}{C(e^{-i\omega})}$ converges in $L_2^y (dF^y)$, hence the random measure $\frac{A(e^{-i\omega})}{C(e^{-i\omega})} d\zeta^y(\omega)$ is well-defined.

© L. Gerencsér, Zs. Vágó and B. Gerencsér
End of REMINDER
Unstable ARMA processes, I.

The analysis of unstable MA or AR processes can be extended. Let $y = (y_n)$ be a w.s.st. ARMA process defined by

$$A'(q^{-1}) y = C'(q^{-1}) e'$$

where the polynomials $A'(z^{-1})$ and $C'(z^{-1})$ are not necessarily stable. The process $e'$ is a w.s.st. orthogonal process with $\sigma^2(e') = 1$. Assume $A'(z^{-1}) \neq 0$ and $C'(z^{-1}) \neq 0$ for $|z| = 1$. The spectral density of $y$:

$$f(\omega) = \left| \frac{C'(e^{-i\omega})}{A'(e^{-i\omega})} \right|^2.$$

Let $A(e^{-i\omega})$ and $C(e^{-i\omega})$ be the stable spectral factors of the denominator and the numerator, resp.
Unstable ARMA processes, II.

Then

\[ f_y(\omega) = \left| \frac{C'(e^{-i\omega})}{A'(e^{-i\omega})} \right|^2 = \left| \frac{C(e^{-i\omega})}{A(e^{-i\omega})} \right|^2. \]

The rational function \( \frac{C(e^{-i\omega})}{A(e^{-i\omega})} \) is called a stable spectral factor of \( f_y \).

Now define the w.s.st. process \( e \) by

\[ d\zeta^e(\omega) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} d\zeta^y(\omega), \]

Equivalently, define the process \( e \) by the inverse dynamics

\[ C(q^{-1}) e = A(q^{-1}) y \quad (8) \]

Let us spell out the definition of \( e \):

\[ d\zeta^e(\omega) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} d\zeta^y(\omega) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} \cdot \frac{C'(e^{-i\omega})}{A'(e^{-i\omega})} d\zeta^e'(\omega). \]
Unstable ARMA processes, III.

\[ d\zeta^e(\omega) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} \, d\zeta^y(\omega) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} \cdot \frac{C'(e^{-i\omega})}{A'(e^{-i\omega})} \, d\zeta^e'(\omega). \]

Note that the transfer function

\[ G(e^{-i\omega}) = \frac{A(e^{-i\omega})}{C(e^{-i\omega})} \cdot \frac{C'(e^{-i\omega})}{A'(e^{-i\omega})} \]

is such that

\[ G(e^{-i\omega})G(e^{i\omega}) = |G(e^{-i\omega})|^2 = 1 \quad \forall \omega. \quad (9) \]

We say that the transfer function \( G(e^{-i\omega}) \) is **all-pass**: all frequencies are passed through the filter \( G \) with unchanged energy.

It is readily seen that the process \( e = (e_n) \) is a w.s.st. orthogonal process:
Proposition

Let \( y = (y_n) \) be a w.s.st. ARMA process defined by

\[
A'(q^{-1}) y = C'(q^{-1}) e'
\]  \hspace{1cm} (10)

where \( e' \) is a w.s.st. orthogonal process with \( \sigma^2(e') = 1 \).

Assume \( A'(z^{-1}) \neq 0 \) and \( C'(z^{-1}) \neq 0 \) for \( |z| = 1 \). Let \( C(e^{-i\omega})/A(e^{-i\omega}) \) denote the stable spectral factor of the spectral density of \( y \):

\[
f(\omega) = \left| \frac{C(e^{-i\omega})}{A'(e^{-i\omega})} \right|^2 = \left| \frac{C(e^{-i\omega})}{A(e^{-i\omega})} \right|^2.
\]

Then the innovation process of \( y \) is obtained from the equation below:

\[
C(q^{-1}) e = A(q^{-1}) y, \hspace{1cm} (11)
\]
VECTOR VALUED PROCESSES
Vector valued w.s.st. processes, I.

Motivation: simultaneous price movements of several commodities. Interaction between individual prices: get better predictions.

Definition

The $\mathbb{R}^s$-valued stochastic process $y = (y_n)$, $-\infty < n < +\infty$ is called wide sense stationary if $\mathbb{E}|y_n|^2 < \infty$ for all $n$, $\mathbb{E}y_n = 0$ for all $n$, and the covariance matrix

$$R(\tau) = R^y(\tau) = \mathbb{E}(y_{n+\tau}y_n^\top)$$

is independent of $n$.

The condition $\mathbb{E}y_n = 0$ can be replaced by the condition that $\mathbb{E}y_n = m$ with some fixed vector $m \in \mathbb{R}^s$ for all $n$.

The matrix-valued function $R(.)$ is called the auto-covariance function of $(y_n)$. Obviously, we have

$$R(-\tau) = R(\tau)^\top.$$
Vector valued w.s.st. processes, II.

The definition extends to $\mathbb{C}^s$-valued (complex-valued) processes requiring

$$R(\tau) = \mathbb{E}(y_{n+\tau} \overline{y}_n)$$

to be independent of $n$. Obviously, $R(-\tau) = \overline{R}(\tau)^\top$.

An eminent role is played by vector-valued w.s.st. orthogonal processes.

**Definition**

A $\mathbb{R}^s$-valued w.s.st. stochastic process $(e_n)$ is called a w.s.st orthogonal process if

$$\mathbb{E}e_{n+\tau}e_n^\top = 0 \quad \text{for} \quad \tau \neq 0 \quad \text{and} \quad \mathbb{E}e_ne_n^\top = \Sigma \quad \text{for all} \quad n,$$

where $\Sigma$ is a fixed, $s \times s$ symmetric positive semi-definite matrix.

Note that $\Sigma$ is not assumed to be the identity $I$, it may be even singular.
Results for vector-valued w.s.st. processes

Let \((y_n)\) be a \(\mathbb{R}^s\)-valued w.s.st. process. Define a new process:

\[
    z_n = \sum_{k=1}^{p} a_k^T y_{n-k}, \quad \text{with} \quad a_1, \ldots, a_p \in \mathbb{R}^s.
\]

Exercise. (HW) Show that \((z_n)\) is an \(\mathbb{R}\)-valued w.s.st. process, and

\[
    \mathbb{E} z_n^2 = \sum_{k=1}^{p} \sum_{j=1}^{p} a_k^T R(k-j) a_j \geq 0.
\]

Corollary. The block matrix \(R\), which is defined by the blocks\(^{(12)}\)

\[
    R_{k,j} = R(k-j) \in \mathbb{R}^{s \times s}, \quad k, j = 1, \ldots, p
\]

is positive semi-definite. The size of \(R\) is \((ps) \times (ps)\).
Prediction and innovation, I.

Let \((y_n)\) be a \(\mathbb{R}^s\)-valued w.s.st. process.

Any component of \(y_n\) or a linear combination of components of \(y_n\) will be predicted by taking a set of vectors \(a_1, \ldots, a_p\) in \(\mathbb{R}^s\), and defining

\[
    z_n = \sum_{k=1}^{p} a_k^\top y_{n-k}.
\]

**Exercise.** Show that \((z_n)\) is an \(\mathbb{R}\)-valued w.s.st. process and we have

\[
    \mathbb{E}z_n^2 = \sum_{k=1}^{p} \sum_{l=1}^{p} a_k^\top R(k - l)a_l. \tag{13}
\]

Thus the block matrix \(R = (R_{k,l})\) defined by the blocks

\[
    R_{k,l} = R(k - l), \quad k, l = 1, \ldots, p \tag{14}
\]

is positive semi-definite. The size of \(R\) is \(ps \times ps\).
Block Toeplitz matrices

Recall: $R_{k,j} = R(k - j) = \mathbb{E}(y(n - j)y^T(n - k))$, $k, j = 1, \ldots, p$, and

$$R = \begin{pmatrix}
R(0) & R(1) & \cdots & \cdots & R(p - 1) \\
R(-1) & R(0) & R(1) & \cdots & R(p - 2) \\
R(-2) & R(-1) & R(0) & \cdots & R(p - 3) \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
R(1 - p) & R(2 - p) & \cdots & \cdots & R(0)
\end{pmatrix}$$

Definition

A $ps \times ps$ matrix $R$ consisting of $m \times m$ blocks is called a **block-Toeplitz matrix**.

Note. A block-Toeplitz matrix is typically not Toeplitz in the usual sense.
The history of \((y_n)\)

Let \((y_n)\) be an \(\mathbb{R}^s\)-valued w.s.st. process.

Consider first the linear space of \(\mathbb{R}\)-valued (!) random variables:

\[
\mathcal{L}_{n-1}^y = \left\{ \sum_{k=1}^{p} a_k^T y_{n-k}, \text{ for some } p, \text{ and } a_k \in \mathbb{R}^s \right\} \subset L_2(\Omega, \mathcal{F}, P).
\]

We define \(H_{n-1}^y\) as the closure of \(\mathcal{L}_{n-1}^y\) in \(L_2(\Omega, \mathcal{F}, P)\).

Remark. \(H_{n-1}^y\) is a Hilbert space of real-valued r.v.-s.
Projection in $L_2^s(\Omega, \mathcal{F}, P)$

We define $L_2^s(\Omega, \mathcal{F}, P) = \{x \in \mathbb{R}^s \text{ is a r.v.} : \mathbb{E}|x|^2 < \infty\}$.

The projection of $x \in L_2^s(\Omega, \mathcal{F}, P)$ onto $H_{n-1}^y$ is defined componentwise:

$$\hat{x} = (x|H_{n-1}^y) = ((x_1|H_{n-1}^y), \ldots, (x_s|H_{n-1}^y))^T.$$

For the error vector $\tilde{x} = x - \hat{x}$ we have

$$\tilde{x} \perp H_{n-1}^y \quad \text{i.e.} \quad \tilde{x}_k \perp H_{n-1}^y \quad \forall \ k.$$
The innovation process

Define the innovation process of \( y = (y_n) \), denoted by \( e = (e_n) \), via

\[
e_n = y_n - (y_n | H_{n-1}^y).
\]

A vector valued process \( (y_n) \) is singular, if its innovation process identically 0, or equivalently if

\[
H_n^y = H_{n-1}^y \text{ for all } n.
\] (15)

A novel phenomenon is that \( \Lambda = \mathbb{E}e_ne_n^T \) may be singular, although \( \Lambda \neq 0 \).
Spectral theory, I.

Question: can we extend Herglotz’s theorem on the representation of the auto-covariance sequence to vector valued w.s.st. processes?

Assume first that the auto-covariance function of \((y_n)\) satisfies

\[
\sum_{\tau = -\infty}^{\infty} \| R(\tau) \|^2 < \infty,
\]

(16)

where \(\| R \|\) denotes the operator norm of the matrix \(R\), i.e.

\[
\| R \| = \max_{x \neq 0} |Rx| / |x|.
\]
Spectral theory, II.

**Theorem**

Let \( y = (y_n) \) be an \( \mathbb{R}^s \)-valued w.s.st. process, and let \( R(.) \) be its auto-covariance function. Assume that \( R(.) \) satisfies (16). Then we have

\[
R(\tau) = \frac{1}{2\pi} \int_{0}^{2\pi} e^{i\tau \omega} f(\omega) d\omega, \quad (17)
\]

where \( f(.) \) is a symmetric, positive semi-definite function in \( L_2^{s \times s}(d\omega) \).

The function \( f(.) \) is called the spectral density of \( y = (y_n) \).

**Exercise.**  **HW** Show that for an \( \mathbb{R}^s \)-valued orthogonal wide sense stationary process \( (e_n) \) with covariance matrix \( \Lambda = \mathbb{E} e_n e_n^T \) we have

\[
f(\omega) = \Lambda \quad \forall \omega.
\]
Spectral theory, III.

Proof (Outline.) Let $\alpha \in \mathbb{R}^s$ and consider the scalar process

$$z_{\alpha,n} = \alpha^\top y_n.$$ 

The auto-covariance function of $z_{\alpha,n}$ is

$$r^\alpha(\tau) = \alpha^\top R(\tau)\alpha.$$ 

It is obvious by (16) that $\sum_{\tau = -\infty}^{\infty} (r^\alpha(\tau))^2 < +\infty$. By the special case of the scalar Herglotz’s theorem we have

$$r^\alpha(\tau) = \frac{1}{2\pi} \int_{0}^{2\pi} e^{i\omega \tau} f^\alpha(\omega) d\omega,$$

where $f^\alpha(\omega) \geq 0$ is the spectral density corresponding to $r^\alpha(\tau)$. 

© L. Gerencsér, Zs. Vágó and B. Gerencsér
Spectral theory, IV.

We know how to get $f^\alpha(.)$ from $r^\alpha(.)$ explicitly:

$$f^\alpha(\omega) = \sum_{\tau=-\infty}^{\infty} r^\alpha(\tau) e^{-i\tau\omega}.$$  

Here convergence on the r.h.s. is meant in $L^2_c(d\omega) = L^2([0,2\pi], d\omega)$.

Substituting $r^\alpha(\tau) = \alpha^\top R(\tau)\alpha$ we get

$$f^\alpha(\omega) = \sum_{\tau=-\infty}^{\infty} \alpha^\top R(\tau)\alpha e^{-i\tau\omega}. \quad (18)$$

Taking finite truncations of the r.h.s. we get that

$$\sum_{\tau=-N}^{N} \alpha^\top R(\tau)\alpha e^{-i\tau\omega} = \alpha^\top \left( \sum_{\tau=-N}^{N} R(\tau) e^{-i\tau\omega} \right)\alpha \quad (19)$$

converges in $L_2(d\omega)$ for any $\alpha$. 
Spectral theory, V.

From here we would like to conclude that

\[ f_N(\omega) = \sum_{\tau=-N}^{N} R(\tau)e^{-i\tau \omega} \]

itself converges in \( L^{s \times s}_2(d\omega) \). To see this we need what follows:

**Exercise.** Prove that a quadratic form \( \alpha^\top F \alpha \), with \( F \) symmetric, determines the bilinear form corresponding to \( F \) uniquely as

\[ \beta^\top F \gamma = \frac{1}{4} \left( (\beta + \gamma)^\top F(\beta + \gamma) - (\beta - \gamma)^\top F(\beta - \gamma) \right). \quad (20) \]
Spectral theory, VI.

Take $F = f_N$, and any pair of unit vectors in $\mathbb{R}^s$, say, $\beta = e_k, \gamma = e_l$.

We conclude that infinite series below converges in $L_2(d\omega) \forall k, l$:

$$ e_k \top \sum_{\tau = -\infty}^{\infty} R(\tau)e^{-i\tau \omega}e_l = \sum_{\tau = -\infty}^{\infty} R_{k,l}(\tau)e^{-i\tau \omega} =: f_{k,l}(\omega). $$

Thus the infinite series below converges in $L_{2}^{s \times s}(d\omega)$:

$$ \sum_{\tau = -\infty}^{\infty} R(\tau)e^{-i\tau \omega} =: f(\omega) $$

It is readily seen that

$$ f^\alpha(\omega) = \alpha \top f(\omega)\alpha \quad \forall \alpha \in \mathbb{R}^s. $$

Obviously $f(\omega)$ is symmetric, and $f^\alpha(\omega) \geq 0$

that $f(\omega)$ is positive semi-definite. This concludes the proof.
Spectral theory: the general case

Theorem

Let $y = (y_n)$ be an $\mathbb{R}^s$-valued w.s.st. process, and let $R(.)$ be its auto-covariance function. Then we have

$$R(\tau) = \frac{1}{2\pi} \int_0^{2\pi} e^{i\tau \omega} dF(\omega),$$

where $F(.)$ is a matrix-valued function such that $F(0) = 0$ and $F(2\pi)$ is finite, and its increments of $F(\cdot)$ are symmetric positive semi-definite matrices.

The matrix-valued $F(\omega)$ is the spectral distribution function of $(y_n)$