Lesson 11: The Wold Decomposition Theorem

Umberto Triacca

Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica
Università dell'Aquila,
umberto.triacca@univaq.it

The Wold Decomposition Theorem

In this lesson we present the Wold decomposition theorem.

The Wold decomposition theorem states that any covariance stationary process can be decomposed into two mutually uncorrelated component processes,

- one a linear combination of lags of a white noise process
- and the other a process, future values of which can be predicted exactly by some linear function of past observations.

As we will see, one reason for the popularity of the ARMA models derives from Wold's Theorem.

The Wold Decomposition Theorem

We start with two definitions.

Definition. Let $\{x_t; t \in \mathbb{Z}\}$ be a covariance-stationary process. The random variable

$$P\left[x_{t+h}|x_{t-1},...,x_{t-N}\right] = \alpha_0^{(N)} + \alpha_1^{(N)}x_{t-1} + ... + \alpha_N^{(N)}x_{t-N}$$

where the coefficients $\alpha_0^{(N)}, \alpha_1^{(N)}, ..., \alpha_N^{(N)}$ are such that

$$S(\alpha_0^{(N)}, \alpha_1^{(N)}, ..., \alpha_N^{(N)}) = E(x_{t+h} - \alpha_0^{(N)} - \alpha_1^{(N)}x_{t-1} - ... - \alpha_N^{(N)}x_{t-N})^2$$

is minimum, is called the **orthogonal projection** of x_{t+h} on $x_{t-1}, ..., x_{t-N}$

The orthogonal projection of x_{t+h} on $x_{t-1}, x_{t-2}...,$ denoted $P\left[x_{t+h}\,|x_{t-1}, x_{t-2}...,\right]$, is defined by

$$P[x_{t+h}\,|x_{t-1}, x_{t-2}...,] = \lim_{N\to\infty} P\left[x_{t+h}\,|x_{t-1}, ..., x_{t-N}\right]$$

Deterministic processes

Definition. A covariance-stationary process, $\{x_t; t \in \mathbb{Z}\}$, is called (linearly) deterministic if

$$P\left[x_t \left| x_{t-1}, x_{t-2} \ldots, \right.\right] = x_t$$

The Wold Decomposition Theorem

We have that a stationary process $\{x_t; t \in \mathbb{Z}\}$ is deterministic if x_t can be predicted correctly (with zero error) using the entire past $x_{t-1}, x_{t-2}, \ldots$.

For a deterministic process the one-step prediction error is zero.

The Wold Decomposition Theorem

An example. Let $\{x_t; t \in \mathbb{Z}\}$ be a stochastic process defined by

$$x_t = A\cos(t) + B\sin(t)$$

where A and B are independent standard normal random variables. **This process is deterministic**. In fact it is possible to show that

$$x_t = \frac{\sin(2)}{\sin(1)} x_{t-1} - x_{t-2}.$$

and hence

$$P\left[x_t | x_{t-1}, x_{t-2} \ldots, \right] = \frac{\sin(2)}{\sin(1)} x_{t-1} - x_{t-2} = x_t$$

It is important to note that deterministic does not mean that x_t is non-random.

We can now introduce the Wold decomposition theorem.

The Wold Decomposition Theorem

Theorem (Wold's Decomposition Theorem) Any zero-mean nondeterministic covariance-stationary process $\{x_t; t \in \mathbb{Z}\}$ can be decomposed as

$$x_t = \sum_{j=0}^{\infty} \psi_j u_{t-j} + d_t$$

where

1. $\psi_0 = 1, \; \sum_{j=1}^{\infty} \psi_j^2 < \infty,$
2. $u_t \sim WN(0, \sigma_u^2),$
3. $\{\psi_j\}$ and $\{u_t\}$ are unique,
4. $\{d_t; t \in \mathbb{Z}\}$ is deterministic,
5. u_t is the limit of linear combinations of $x_s, \; s \leq t$
6. $E(d_t u_s) = 0 \;\; \forall t, s$

The Wold Decomposition Theorem

The Wold representation is the **unique** linear representation where the innovations are linear forecast errors.

Purely nondeterministic processes

Definition. A zero-mean nondeterministic covariance-stationary process, $\{x_t; t \in \mathbb{Z}\}$, is called purely nondeterministic (or regular) if $d_t = 0$.

The Wold Decomposition Theorem

Thus if the process $\{x_t; t \in \mathbb{Z}\}$ is purely nondeterministic then

$$x_t = \sum_{j=0}^{\infty} \psi_j u_{t-j}$$

where

1. $\psi_0 = 1, \ \sum_{j=1}^{\infty} \psi_j^2 < \infty$,
2. $u_t \sim WN(0, \sigma_u^2)$,
3. u_t is the limit of linear combinations of $x_s, \ s \leq t$

The Wold Decomposition Theorem

The Wold theorem plays a central role in time series analysis.

It implies that the dynamic of any purely nondeterministic covariance-stationary process can be arbitrarily well approximated by an ARMA process.

The Wold Decomposition Theorem

In fact, by Wold's Decomposition Theorem, we have that any purely nondeterministic covariance-stationary process can be written as a linear combination of lagged values of a white noise process (MA(∞)) representation), that is

$$x_t = \sum_{j=0}^{\infty} \psi_j u_{t-j}$$

The Wold Decomposition Theorem

Now, we note that, under general conditions, the infinite lag polynomial of the Wold decomposition can be approximated by the ratio of two finite-lag polynomials:

$$\psi(L) \approx \frac{\theta(L)}{\phi(L)}.$$

Therefore x_t can be accurately approximated by a ARMA process

$$x_t^* = \frac{\theta(L)}{\phi(L)} u_t.$$

The Wold Decomposition Theorem

Any purely nondeterministic covariance-stationary process has an ARMA representation!

This means that the stationary $ARMA(p, q)$ models are a class of linear stochastic processes that are general enough.

The Wold Decomposition Theorem

Are the covariance-stationary ARMA processes purely nondeterministic processes?

The Wold Decomposition Theorem

Consider a covariance-stationary ARMA(p,q) process defined by

$$\phi(L)x_t = \theta(L)u_t \quad u_t \sim WN(0, \sigma^2)$$

where

$$\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$

and

$$\theta(L) = 1 + \theta_1 L ... + \theta_q L^q$$

Suppose that this representation is causal and invertible.

The Wold Decomposition Theorem

The causality assumption implies that there exists constants $\psi_0, \psi_1, ...$ such that

$$\sum_{j=0}^{\infty} |\psi_j| < \infty, \text{ with } \psi_0 = 1$$

and

$$x_t = \sum_{j=0}^{\infty} \psi_j u_{t-j} \ \ \forall t.$$

$$\sum_{j=0}^{\infty} |\psi_j| < \infty \ \ \Rightarrow \sum_{j=1}^{\infty} \psi_j^2 < \infty$$

$u_t \sim WN(0, \sigma_u^2)$ and the invertibility condition implies that u_t is the limit of linear combinations of $x_s, \ s \leq t$

We can conclude that the covariance-stationary ARMA(p,q) process,x_t, is a purely nondeterministic process.

The Wold Decomposition Theorem

Now, suppose the representation

$$\phi(L)x_t = \theta(L)u_t \quad u_t \sim WN(0, \sigma_u^2)$$

where

$$\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$

and

$$\theta(L) = 1 + \theta_1 L ... + \theta_q L^q$$

is not causal and not invertible.

The Wold Decomposition Theorem

It is possible to show that if $\theta(z) \neq 0$ when $|z| = 1$, then it is always possible to find polynomials $\phi^*(L)$ and $\theta^*(L)$ and a white noise $v_t \sim WN(0, \sigma_v^2)$ such that the representation

$$\phi^*(L)x_t = \theta^*(L)v_t \quad v_t \sim WN(0, \sigma_v^2)$$

is causal and invertible.

The Wold Decomposition Theorem

Thus a covariance-stationary ARMA(p,q) process defined by

$$\phi(L)x_t = \theta(L)u_t \quad u_t \sim WN(0, \sigma^2)$$

where

$$\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$

and

$$\theta(L) = 1 + \theta_1 L ... + \theta_q L^q$$

with $\theta(z) \neq 0$ if $|z| = 1$ is a purely nondeterministic process.

The Wold Decomposition Theorem

A covariance-stationary ARMA process, with $\theta(z) \neq 0$ if $|z| = 1$, is a purely nondeterministic process