The Wold Decomposition

Herman J. Bierens*

February 1, 2012

Abstract

In Chapter 7 in Bierens (2004) the Wold decomposition was motivated by claiming that every zero-mean covariance stationary process X_t can be written as $X_t = \sum_{j=1}^{\infty} \beta_j X_{t-j} + U_t$, where $E[U_t.X_{t-j}] = 0$ for all $j \geq 1$, and $\sum_{j=1}^{\infty} \beta_j X_{t-j}$ is the projection of X_t on its past. However, in general this claim is incorrect. In this note I will give a more general (and hopefully correct) proof of the Wold decomposition.

1 Projections on spaces spanned by a sequence

The fundamental projection theorem states that:

Theorem 1. *Given a sub-Hilbert space $\mathcal{S}$ of a Hilbert space $\mathcal{H}$ and an element $y \in \mathcal{H}$, there exists a unique element $\widehat{y} \in \mathcal{S}$ such that $||y - \widehat{y}|| = \inf_{z \in \mathcal{S}} ||y - z||$. Moreover the residual $u = y - \widehat{y}$ is orthogonal to any $z \in \mathcal{S}$: $\langle u, z \rangle = 0$.*

Proof: See for example Bierens (2004, Th. 7.A.3, p. 202).

This result is the basis for the famous Wold (1938) decomposition for covariance stationary time series, which in its turn is the basis for time series analysis.

*Thanks to Peter Boswijk (University of Amsterdam) for pointing out an error in a previous version of this note. Moreover, the queries of the students in my graduate time series courses have led to substantial improvements of the proof of the Wold decomposition.

The proof of the Wold decomposition in Anderson (1994) is more transparent than the original proof by Wold (1938). However, rather than following Anderson's proof, I will in this note derive first a general Wold decomposition for a regular sequence[1] in a general Hilbert space, and then specialize this result to the Wold decomposition for covariance stationary time series.

First, we need to define sub-Hilbert spaces spanned by a sequence in a Hilbert space, as follows.

Let $\{x_k\}_{k=1}^{\infty}$ be a sequence of elements of a Hilbert space $\mathcal{H}$, and let

$$\mathcal{M}_m = \text{span}(\{x_j\}_{j=1}^m)$$

be the space spanned by $x_1, ..., x_m$, i.e., $\mathcal{M}_m$ consists of all linear combinations of $x_1, ..., x_m$. Then

Lemma 1. *$\mathcal{M}_m$ is a Hilbert space.*

Proof: Without loss of generality we may assume that the $m \times m$ matrix Σ_m with elements $\langle x_i, x_j \rangle$, $i, j = 1, ..., m$, is non-singular, as otherwise we can remove one or more x_j's from the list $\{x_j\}_{j=1}^m$ and still span the same space. For example, suppose that $\text{rank}(\Sigma_m) = m - 1$, and let $c = (c_1, ..., c_m)'$ be the eigenvector corresponding to the zero eigenvalue. Then $\left\|\sum_{j=1}^m c_j x_j\right\|^2 = c'\Sigma_m c = 0$, hence $\sum_{j=1}^m c_j x_j = 0$ (the latter being the zero element of $\mathcal{M}_m$). Since at least one component of c is non-zero, for example c_i, we can write

$$x_i = \begin{cases} -\sum_{j=2}^m (c_j/c_1)x_j & \text{if} \quad i = 1, \\ -\sum_{j=1}^{m-1}(c_j/c_m)x_j & \text{if} \quad i = m, \\ -\sum_{j=1}^{i-1}(c_j/c_i)x_j - \sum_{j=i+1}^m (c_j/c_i)x_j & \text{if} \quad 1 < i < m, \end{cases}$$

so that

$$\mathcal{M}_m = \begin{cases} \text{span}(\{x_j\}_{j=2}^m) & \text{if} \quad i = 1, \\ \text{span}(\{x_j\}_{j=1}^{m-1}) & \text{if} \quad i = m, \\ \text{span}(x_1, ..., x_{i-1}, x_{i+1}, ..., x_m) & \text{if} \quad 1 < i < m. \end{cases}$$

Now let $z_n = \sum_{j=1}^m \beta_{j,n} x_j$ be a Cauchy sequence in $\mathcal{M}_m$, and denote $\beta_n = (\beta_{1,n}, ..., \beta_{m,n})'$. Then for each j, $\beta_{j,n}$ is a Cauchy sequence in $\mathbb{R}$ because

$$0 = \lim_{\min(n_1,n_2)\to\infty} ||z_{n_1} - z_{n_2}||^2 = \lim_{\min(n_1,n_2)\to\infty} \left\|\sum_{j=1}^m (\beta_{j,n_1} - \beta_{j,n_2}) x_j\right\|^2$$

[1] See Definition 4 below.

$$
\begin{aligned}
&= \lim_{\min(n_1,n_2)\to\infty} (\beta_{n_1}-\beta_{n_2})' \Sigma_m (\beta_{n_1}-\beta_{n_2}) \\
&\geq \lambda_{\min}(\Sigma_m). \lim_{\min(n_1,n_2)\to\infty} ||\beta_{n_1}-\beta_{n_2}||^2,
\end{aligned}
$$

where $\lambda_{\min}(\Sigma_m) > 0$ is the smallest eigenvalue of Σ_m. Consequently, $\lim_{n\to\infty} ||z - z_n|| = 0$, where $z = \sum_{j=1}^{m} \beta_j x_j \in \mathcal{M}_m$ with $\beta_j = \lim_{n\to\infty} \beta_{j,n}$. Q.E.D.

Definition 1. *The space* $\mathcal{M}_\infty = \overline{\cup_{n=1}^\infty \mathcal{M}_n}$ *(which is the closure of* $\cup_{n=1}^\infty \mathcal{M}_n$*) is called the space spanned by* $\{x_j\}_{j=1}^\infty$*, and is also denoted by* $\text{span}(\{x_j\}_{j=1}^\infty)$.

Lemma 2. $\mathcal{M}_\infty$ *is a Hilbert space.*

Proof: Let z_m be a Cauchy sequence in $\mathcal{M}_\infty$, with limit $\overline{z} \in \mathcal{H}$. If $\overline{z} \notin \mathcal{M}_\infty$ then, because $\mathcal{M}_\infty$ is closed, there exists an $\varepsilon > 0$ such that the set $\{z \in \mathcal{H} : ||z - \overline{z}|| < \varepsilon\}$ is completely outside $\mathcal{M}_\infty$: $\{z \in \mathcal{H} : ||z - \overline{z}|| < \varepsilon\} \cap \mathcal{M}_\infty = \oslash$. But then there exists an m such that $z_m \notin \mathcal{M}_\infty$. Since this is impossible, $\overline{z} \in \mathcal{M}_\infty$. Q.E.D.

Next, let us focus on projections on a space spanned by a sequence in a Hilbert space.

Lemma 3. *For* $z \in \mathcal{M}_\infty$ *let* $\widehat{z}_n$ *be the projection of* z *on* $\mathcal{M}_n$*. Then* $\lim_{n\to\infty} ||z - \widehat{z}_n|| = 0$.

Proof: If $z \in \cup_{n=1}^\infty \mathcal{M}_n$ then there exists an n_0 such that $z \in \mathcal{M}_{n_0}$, hence for $n \geq n_0$, $\widehat{z}_n = z$ and thus $\lim_{n\to\infty} ||z - \widehat{z}_n|| = 0$. Now let $z \in \mathcal{M}_\infty \backslash (\cup_{n=1}^\infty \mathcal{M}_n)$. Since $\mathcal{M}_\infty = \overline{\cup_{n=1}^\infty \mathcal{M}_n}$ is closed and $\mathcal{M}_n \subset \mathcal{M}_{n+1}$, for each n there exists an $z_n \in \mathcal{M}_n$ such that $\lim_{n\to\infty} ||z - z_n||^2 = 0$, hence for $n \to \infty$, $||z - \widehat{z}_n||^2 \leq ||z - z_n||^2 \to 0$. Q.E.D.

More generally we have:

Theorem 2. *For* $z \in \mathcal{H}$*, let* $\widehat{z}$ *be the projection of* z *on* $\mathcal{M}_\infty = \text{span}(\{x_j\}_{j=1}^\infty)$ *and let* $\widehat{z}_n$ *be the projection of* z *on* $\mathcal{M}_n = \text{span}(\{x_j\}_{j=1}^n)$*. Then* $\lim_{n\to\infty} ||\widehat{z} - \widehat{z}_n|| = 0$.

Proof: We may without loss of generality assume that $\widehat{z} \in \mathcal{M}_\infty \backslash (\cup_{n=1}^\infty \mathcal{M}_n)$, as otherwise the result of Theorem 2 holds trivially. Since $\mathcal{M}_\infty$ is closed this assumption implies that for each n we can select a $z_n \in \mathcal{M}_n$ such that

$$\lim_{n\to\infty} ||\widehat{z} - z_n|| = 0. \tag{1}$$

Let $||z - \widehat{z}|| = \delta$ and $||z - \widehat{z}_n|| = \delta_n$, and note that $\delta_n \geq \delta$. Since

$$\begin{aligned} \delta_n^2 &= ||z - \widehat{z}_n||^2 \leq ||z - z_n||^2 = ||z - \widehat{z} + \widehat{z} - z_n||^2 \\ &= ||z - \widehat{z}||^2 + ||\widehat{z} - z_n||^2 + 2\langle z - \widehat{z}, \widehat{z} - z_n\rangle \\ &= \delta^2 + ||\widehat{z} - z_n||^2 \end{aligned}$$

it follows from (1) that

$$\lim_{n\to\infty} \delta_n = \delta. \tag{2}$$

Recall that $z = \widehat{z} + u$, where $\langle u, x\rangle = 0$ for all $x \in \mathcal{M}_\infty$. Hence

$$\begin{aligned} ||\widehat{z} - \widehat{z}_n||^2 &= ||z - \widehat{z}_n - u||^2 = ||z - \widehat{z}_n||^2 + ||u||^2 - 2\langle z - \widehat{z}_n, u\rangle \\ &= ||z - \widehat{z}_n||^2 + ||u||^2 - 2\langle z, u\rangle = \delta_n^2 - \delta^2 \end{aligned} \tag{3}$$

where the last equality follows from

$$\langle z, u\rangle - \langle u, u\rangle = \langle \widehat{z}, u\rangle = 0 \tag{4}$$

and $\langle u, u\rangle = ||u||^2 = \delta^2$. The theorem now follows from (2) and (3). Q.E.D.

Remark 1. Although each projection $\widehat{z}_n$ is a linear combination of $x_1, ..., x_n$, in general the result of Theorem 2 does **not** imply that there exists a sequence $\{\theta_j\}_{j=1}^{\infty}$ such that $\widehat{z} = \sum_{j=1}^{\infty} \theta_j x_j$.

As a counter example, consider the Hilbert space $\mathcal{R}_0$ of zero-mean random variables with finite second moments, endowed with the inner product $\langle X, Y\rangle = E[X.Y]$ and associated norm and metric. Let

$$X_t = V_t - V_{t-1},$$

where the V_t's are independent $N(0,1)$ distributed. This is clearly a zero-mean covariance stationary process, with covariance function $\gamma(0) = 2,\ \gamma(1) = -1,\ \gamma(m) = 0$ for $m \geq 2$. Hence $X_t \in \mathcal{R}_0$ for all t.

For given t, let

$$\mathcal{M}_{-\infty}^{t-1} = \text{span}\left(\{X_{t-m}\}_{m=1}^{\infty}\right),\ \mathcal{M}_{t-n}^{t-1} = \text{span}\left(X_{t-1}, ..., X_{t-n}\right).$$

The projection $\widehat{X}_{t,n}$ of X_t on $\mathcal{M}_{t-n}^{t-1}$ takes the form

$$\widehat{X}_{t,n} = \sum_{j=1}^{n} \theta_{n,j} X_{t-j}$$

where the coefficients $\theta_{n,j}$ are the solutions of the normal equations

$$\gamma(m) = \sum_{k=1}^{n} \gamma(|k-m|)\theta_{n,k}, \ \ m = 1, ..., n.$$

hence for $n \geq 3$,

$$\begin{aligned}
-1 &= 2.\theta_{n,1} - \theta_{n,2} \\
0 &= -\theta_{n,1} + 2\theta_{n,2} - \theta_{n,3} \\
0 &= -\theta_{n,2} + 2\theta_{n,3} - \theta_{n,4} \\
&\vdots \\
0 &= -\theta_{n,n-2} + 2\theta_{n,n-1} - \theta_{n,n} \\
0 &= -\theta_{n,n-1} + 2\theta_{n,n}
\end{aligned}$$

The solutions of these normal equations are

$$\theta_{n,j} = \frac{j}{n+1} - 1, \ \ j = 1, ..., n,$$

hence

$$\widehat{X}_{t,n} = \sum_{j=1}^{n} \left(\frac{j}{n+1} - 1 \right) X_{t-j} \tag{5}$$

Next, let $\widehat{X}_t$ be the projection of X_t on $\mathcal{M}_{-\infty}^{t-1}$, and suppose that there exists a sequence $\{\theta_j\}_{j=1}^{\infty}$ such that $\widehat{X}_t = \sum_{j=1}^{\infty} \theta_j X_{t-j}$. Note that the latter is merely a short-hand notation for

$$\lim_{n\to\infty} \left\| \widehat{X}_t - \sum_{j=1}^{n} \theta_j X_{t-j} \right\|^2 = \lim_{n\to\infty} E\left[\left(\widehat{X}_t - \sum_{j=1}^{n} \theta_j X_{t-j} \right)^2 \right] = 0 \tag{6}$$

If so, it follows from Theorem 2 and (5) that

$$0 \ = \ \lim_{n\to\infty} E\left[\left(\sum_{j=1}^{n} \theta_j X_{t-j} - \sum_{j=1}^{n} \left(\frac{j}{n+1} - 1 \right) X_{t-j} \right)^2 \right]$$

$$= \lim_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] \tag{7}$$

But

$$\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j} = \sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)(V_{t-j}-V_{t-j-1})$$

$$= -\left(\frac{n}{n+1}+\theta_1\right)V_{t-1} - \sum_{j=1}^{n-1}\left(\theta_{j+1}-\theta_j-\frac{1}{n+1}\right)V_{t-j-1}$$

$$+\left(\frac{1}{n+1}+\theta_n\right)V_{t-n-1}$$

hence

$$E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] = \left(\frac{n}{n+1}+\theta_1\right)^2$$

$$+\sum_{j=1}^{n-1}\left(\theta_{j+1}-\theta_j-\frac{1}{n+1}\right)^2 + \left(\frac{1}{n+1}+\theta_n\right)^2 \tag{8}$$

This equality implies that for arbitrary integers $m \geq 1$,

$$\liminf_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right]$$

$$\geq \liminf_{n\to\infty}\left(\frac{n}{n+1}+\theta_1\right)^2 + \liminf_{n\to\infty}\left(\theta_{m+1}-\theta_m-\frac{1}{n+1}\right)^2$$

$$= (\theta_1+1)^2 + (\theta_{m+1}-\theta_m)^2 .$$

Therefore, a necessary condition for (7) is that $\theta_m = -1$ for $m = 1, 2, 3, \ldots..$ But then it follows from (8) that

$$\lim_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] = \lim_{n\to\infty}\left(\frac{1}{n+1}-1\right)^2 = 1$$

which contradicts (7). Thus, in this case there does **not** exist a sequence $\{\theta_j\}_{j=1}^{\infty}$ such that (6) holds.

2 Projections on the span of an orthonormal sequence

On the other hand,

Theorem 3. *If a sequence* $\{x_j\}_{j=1}^{\infty}$ *in a Hilbert space* $\mathcal{H}$ *is orthonormal, i.e.,*

$$\langle x_i, x_j \rangle = \begin{cases} 1 & if \quad i = j, \\ 0 & if \quad i \neq j, \end{cases} \tag{9}$$

then any projection $\widehat{z}$ *of* $z \in \mathcal{H}$ *on* $\mathrm{span}(\{x_j\}_{j=1}^{\infty})$ *takes the form* $\widehat{z} = \sum_{j=1}^{\infty} \theta_j x_j$ *(in the sense that* $\lim_{n\to\infty} ||\widehat{z} - \sum_{j=1}^{n} \theta_j x_j|| = 0$*), where* $\theta_j = \langle z, x_j \rangle$ and $\sum_{j=1}^{\infty} \theta_j^2 < \infty$.

Proof: Let $\widehat{z}_n$ be the projection of z on $\mathrm{span}(\{x_j\}_{j=1}^{n})$. Then

$$\begin{aligned} ||z - \widehat{z}_n||^2 &= \min_{c_1,\ldots,c_n} \left\| z - \sum_{j=1}^{n} c_j x_j \right\|^2 \\ &= \min_{c_1,\ldots,c_n} \left\{ ||z||^2 - 2\sum_{j=1}^{n} c_j \langle z, x_j \rangle + \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \langle x_i, x_j \rangle \right\} \\ &= \min_{c_1,\ldots,c_n} \left\{ ||z||^2 - 2\sum_{j=1}^{n} c_j \langle z, x_j \rangle + \sum_{j=1}^{n} c_j^2 \right\}, \end{aligned}$$

hence,

$$\widehat{z}_n = \sum_{j=1}^{n} \theta_j x_j, \text{ where } \theta_j = \langle z, x_j \rangle . \tag{10}$$

Moreover, denoting $u_n = z - \widehat{z}_n$, it follows from (9) and (10) that

$$\begin{aligned} ||u_n||^2 &= \left\| z - \sum_{j=1}^{n} \theta_j x_j \right\|^2 = ||z||^2 - 2\sum_{j=1}^{n} \theta_j \langle z, x_j \rangle + \sum_{j=1}^{n}\sum_{i=1}^{n} \theta_j \theta_i \langle x_j, x_i \rangle \\ &= ||z||^2 - \sum_{j=1}^{n} \theta_j^2 \geq 0 \end{aligned} \tag{11}$$

so that $\sum_{j=1}^{n}\theta_j^2 \leq ||z||^2$ for all n and thus $\sum_{j=1}^{\infty}\theta_j^2 < \infty$. Finally, it follows from Theorem 2 that

$$\lim_{n\to\infty}\left\|\widehat{z}-\sum_{j=1}^{n}\theta_j x_j\right\| = \lim_{n\to\infty}||\widehat{z}-\widehat{z}_n|| = 0.$$

Q.E.D.

3 The general Wold decomposition

3.1 Preliminary definitions and results

Let $\mathcal{S}_1, \mathcal{S}_2$ be a pair of subspaces of a Hilbert space $\mathcal{H}$. We say that:

Definition 2. *$\mathcal{S}_1$ and $\mathcal{S}_2$ are orthogonal, denoted by $\mathcal{S}_1 \perp \mathcal{S}_2$, if for each $x_1 \in \mathcal{S}_1$ and each $x_2 \in \mathcal{S}_2$, $\langle x_1, x_2\rangle = 0$.*

Lemma 4. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be sub-Hilbert spaces satisfying $\mathcal{S}_1 \perp \mathcal{S}_2$. Then*

$$\text{span}(\mathcal{S}_1, \mathcal{S}_2) = \{y = x_1 + x_2 : x_1 \in \mathcal{S}_1,\ x_2 \in \mathcal{S}_2\}$$

is a Hilbert space.

Proof: Let y_n be a Cauchy sequence in $\text{span}(\mathcal{S}_1, \mathcal{S}_2)$. Then $y_n = x_{1,n} + x_{2,n}$, where $x_{1,n} \in \mathcal{S}_1$ and $x_{2,n} \in \mathcal{S}_2$. Since $x_{1,n} - x_{1,m} \in \mathcal{S}_1$ and $x_{2,n} - x_{2,m} \in \mathcal{S}_2$ it follows from the orthogonality condition $\mathcal{S}_1 \perp \mathcal{S}_2$ that

$$\begin{aligned}
||y_n - y_m||^2 &= ||(x_{1,n} - x_{1,m}) + (x_{2,n} - x_{2,m})||^2 \\
&= ||x_{1,n} - x_{1,m}||^2 + ||x_{2,n} - x_{2,m}||^2 \\
&\quad +2\langle x_{1,n} - x_{1,m}, x_{2,n} - x_{2,m}\rangle \\
&= ||x_{1,n} - x_{1,m}||^2 + ||x_{2,n} - x_{2,m}||^2,
\end{aligned}$$

hence $\lim_{\min(n,m)\to\infty}||y_n - y_m|| = 0$ implies that $\lim_{\min(n,m)\to\infty}||x_{1,n} - x_{1,m}|| = 0$ and $\lim_{\min(n,m)\to\infty}||x_{2,n} - x_{2,m}|| = 0$. Because $\mathcal{S}_1$ and $\mathcal{S}_2$ are Hilbert spaces there exist an $x_1 \in \mathcal{S}_1$ and an $x_2 \in \mathcal{S}_2$ such that $\lim_{n\to\infty}||x_{1,n} - x_1|| = 0$ and $\lim_{n\to\infty}||x_{2,n} - x_2|| = 0$, hence $\lim_{n\to\infty}||y_n - y|| = 0$, where $y = x_1 + x_2 \in \text{span}(\mathcal{S}_1, \mathcal{S}_2)$. Q.E.D.

We also need the definitions of orthogonal complement and regularity:

Definition 3. *The orthogonal complement of a subspace* $\mathcal{S}$ *of a Hilbert space* $\mathcal{H}$*, denoted by* $\mathcal{S}^{\perp}$*, is the subset of* $\mathcal{H}$ *defined by*

$$\mathcal{S}^{\perp} = \{y \in \mathcal{H} : \langle x, y \rangle = 0 \text{ } for \text{ } all \text{ } x \in \mathcal{S}\}.$$

Lemma 5. *Orthogonal complements are Hilbert spaces.*

Proof: Let $x \in \mathcal{S}$ be arbitrary and let y_n be a Cauchy sequence in $\mathcal{S}^{\perp}$. Then there exists an $y \in \mathcal{H}$ such that $\lim_{n\to\infty} ||y - y_n|| = 0$. Since $\langle x, y_n \rangle = 0$ we have $\langle x, y \rangle = \langle x, y - y_n \rangle$. It follows now from the Cauchy-Schwarz inequality that $| \langle x, y \rangle | = | \langle x, y - y_n \rangle | \leq ||x||.||y - y_n|| \to 0$. Hence $y \in \mathcal{S}^{\perp}$. Q.E.D.

Note that the result of Lemma 5 does not require that $\mathcal{S}$ is a Hilbert space, although in the application below it is.

Definition 4. *Let* $\{x_k\}_{k=1}^{\infty}$ *be a sequence in a Hilbert space* $\mathcal{H}$*. Let* $\widehat{x}_k$ *be the projection of* x_k *on span*$(\{x_m\}_{m=k+1}^{\infty})$*, and denote* $u_k = x_k - \widehat{x}_k$*. The sequence* $\{x_k\}_{k=1}^{\infty}$ *is called regular if* $||u_k|| > 0$ *for all* $k \geq 1$*.*

Note that the regularity concept is related to the concept of linear independence in Euclidean spaces.

3.2 The Wold decomposition for regular sequences in a Hilbert space

We can now formulate the following general version of the Wold decomposition:

Theorem 4. *Given a regular sequence* $\{x_k\}_{k=1}^{\infty}$ *in a Hilbert space, every* $x \in \mathcal{M}_{\infty} = \mathrm{span}(\{x_k\}_{k=1}^{\infty})$ *can be written as*

$$x = \sum_{k=1}^{\infty} \alpha_k e_k + w, \tag{12}$$

in the sense that $\lim_{n\to\infty}\|x-w-\sum_{k=1}^{n}\alpha_k e_k\| = 0$, *where* $\{e_k\}_{k=1}^{\infty}$ *is an orthonormal sequence in* $\mathcal{M}_\infty$, $\alpha_k = \langle x, e_k\rangle$, $\sum_{k=1}^{\infty}\alpha_k^2 < \infty$, *and*

$$w \in \mathcal{S}_\infty \cap \mathcal{U}_\infty^\perp, \tag{13}$$

with $\mathcal{S}_\infty = \cap_{n=1}^{\infty}\text{span}(\{x_k\}_{k=n}^{\infty})$ *and* $\mathcal{U}_\infty^\perp$ *the orthogonal complement of* $\mathcal{U}_\infty = \text{span}(\{e_k\}_{k=1}^{\infty})$. *Note that* (13) *implies that* w *is orthogonal to all the* e_k*'s*: $\langle e_k, w\rangle = 0$ *for* $k = 1, 2, 3, ...$

Proof: Denote

$$\mathcal{S}_n = \text{span}(\{x_k\}_{k=n}^{\infty}).$$

Note that $\mathcal{M}_\infty = \mathcal{S}_1$. Project each x_k on $\mathcal{S}_{k+1}$, so that $x_k = \widehat{x}_k + u_k$ with projection $\widehat{x}_k \in \mathcal{S}_{k+1}$ and residual u_k. Recall that by the regularity condition, $||u_k|| > 0$, hence $e_k = u_k/||u_k||$ is well defined. It is not hard to verify that the residuals u_k are orthogonal, so that the e_k's are orthonormal, and that $\mathcal{U}_\infty \subset \mathcal{M}_\infty$. It follows now from Theorem 3 that (12) holds with $\alpha_k = \langle x, e_k\rangle$, $\sum_{k=1}^{\infty}\alpha_k^2 < \infty$, and $w \in \mathcal{U}_\infty^\perp$, where the latter follows from the fact that w the residual of the projection of x on $\mathcal{U}_\infty$. Therefore, the actual contents of Theorem 4 is that $w \in \mathcal{S}_\infty$.

The theorem under review will be proved in six steps:

Step 1. As before, let $\mathcal{M}_n = \text{span}(\{x_k\}_{k=1}^{n})$. I will show first that

$$\mathcal{M}_n \subset \ \text{span}\left(\mathcal{U}_n, \mathcal{U}_n^\perp \cap \mathcal{S}_2\right) \tag{14}$$

[c.f. Lemma 4], where $\mathcal{U}_n = \text{span}(e_1, ..., e_n) = \text{span}(u_1, ..., u_n)$ and $\mathcal{U}_n^\perp$ is the orthogonal complement of $\mathcal{U}_n$.

Proof. Let $z \in \mathcal{M}_n$ be arbitrary. Recall that z takes the form $z = \sum_{k=1}^{n} c_k x_k$. Substituting $x_k = \widehat{x}_k + u_k = \widehat{x}_k + ||u_k||e_k$ we can write z as

$$\begin{aligned} z &= \sum_{k=1}^{n} c_k\left(\widehat{x}_k + u_k\right) = \sum_{k=1}^{n} c_k u_k + \sum_{k=1}^{n} c_k \widehat{x}_k \\ &= \sum_{k=1}^{n} c_k ||u_k|| e_k + \sum_{k=1}^{n} c_k \widehat{x}_k \\ &= \sum_{k=1}^{n} c_k ||u_k|| e_k + z_2 \end{aligned}$$

where

$$z_2 = \sum_{k=1}^{n} c_k \widehat{x}_k$$

Note that

$$z_2 = \sum_{k=1}^{n} c_k \widehat{x}_k \in \mathcal{S}_2 \tag{15}$$

because $\widehat{x}_k \in \mathcal{S}_{k+1} \subset \mathcal{S}_2$ for $k = 1, 2, ..., n$.

Next, project z_2 on $\mathcal{U}_n$. This projection takes the form

$$\widehat{p}_n = \sum_{k=1}^{n} d_k e_k, \text{ where } d_k = \langle z_2, e_k \rangle,$$

with residual

$$w_{n+1} \in \mathcal{U}_n^{\perp}. \tag{16}$$

However, e_1 is orthogonal to any element of $\mathcal{S}_2$, and $z_2 \in \mathcal{S}_2$. Therefore, $d_1 = \langle z_2, e_1 \rangle = 0$ and thus

$$\widehat{p}_n = \sum_{k=2}^{n} d_k e_k \in \text{ span}\left(\{e_k\}_{k=2}^{n}\right) \subset \mathcal{S}_2,$$

where the latter follows from $e_k \in \mathcal{S}_k \subset \mathcal{S}_2$ for $k = 2, 3, ..., n$. Because $w_{n+1} = z_2 - \widehat{p}_n$ where both terms are elements of $\mathcal{S}_2$, it follows that

$$w_{n+1} \in \mathcal{S}_2. \tag{17}$$

Combining (16) and (17) now yields

$$w_{n+1} \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_2.$$

Thus, denoting $\alpha_1 = c_1||u_1||$, $\alpha_k = c_k||u_k|| + d_k$ for $k = 2, 3, ..., n$, we can write

$$z = \sum_{k=1}^{n} \alpha_k e_k + w_{n+1}, \text{ where } w_{n+1} \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_2,$$

hence $z \in \text{span}(\mathcal{U}_n, \mathcal{U}_n^{\perp} \cap \mathcal{S}_2)$. This proves (14).

Step 2. Next, it will be shown that

$$\text{span}\left(\mathcal{U}_n, \mathcal{U}_n^{\perp} \cap \mathcal{S}_2\right) = \text{ span}\left(\mathcal{U}_n, \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}\right), \tag{18}$$

so that by (14),

$$\mathcal{M}_n \subset \text{ span}\left(\mathcal{U}_n, \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}\right). \tag{19}$$

Proof. Denote for $k < m$,

$$\mathcal{S}_{k,m} = \text{ span}(\{x_j\}_{j=k}^{m}).$$

I will show first that for $m > n$,

$$\mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m} \subset \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m}, \tag{20}$$

as follows. Let $z \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}$ be arbitrary. Since $z \in \mathcal{S}_{2,m}$ there exists constants c_k such that

$$\begin{aligned} z &= \sum_{k=2}^{m} c_k x_k = \sum_{k=2}^{n} c_k \left(\widehat{x}_k + u_k\right) + \sum_{k=n+1}^{m} c_k x_k \\ &= \sum_{k=2}^{n} c_k \|u_k\| e_k + \sum_{k=2}^{n} c_k \widehat{x}_k + \sum_{k=n+1}^{m} c_k x_k. \end{aligned}$$

Moreover, since $z \in \mathcal{U}_n^{\perp}$ it follows that $\langle z, e_k \rangle = 0$ for $k = 1, ..., n$. In particular,

$$0 = \langle z, e_2 \rangle = c_2 \|u_2\| + \sum_{k=2}^{n} c_k \langle \widehat{x}_k, e_2 \rangle + \sum_{k=n+1}^{m} c_k \langle x_k, e_2 \rangle = c_2 \|u_2\|$$

because $\widehat{x}_k \in S_{k+1} \subset S_3$ for $k = 2, ..., n$, $x_k \in S_k \subset S_{n+1}$ for $k = n+1, ..., m$, and e_2 is orthogonal to S_3 and S_{n+1}. Hence $c_2 = 0$ and thus

$$z = \sum_{k=3}^{n} c_k \|u_k\| e_k + \sum_{k=3}^{n} c_k \widehat{x}_k + \sum_{k=n+1}^{m} c_k x_k.$$

It follows now similarly that $c_k = 0$ for $k = 3, ..., n$, hence

$$z = \sum_{k=n+1}^{m} c_k x_k \in \mathcal{S}_{n+1,m}.$$

Because $z \in \mathcal{U}_n^{\perp}$ as well, it follows now that

$$z \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m},$$

which implies (20).

However, $\mathcal{S}_{n+1,m} \subset \mathcal{S}_{2,m}$ and therefore

$$\mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m} \subset \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}. \tag{21}$$

Combining (20) and (21) now yields

$$\mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m} = \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m} \text{ for } m > n,$$

which in its turn implies that

$$\mathcal{U}_n^{\perp} \cap \left(\overline{\cup_{m=n+1}^{\infty} \mathcal{S}_{2,m}}\right) = \mathcal{U}_n^{\perp} \cap \left(\overline{\cup_{m=n+1}^{\infty} \mathcal{S}_{n+1,m}}\right). \tag{22}$$

Finally, note that $\mathcal{S}_2 = \overline{\cup_{m=n+1}^{\infty} \mathcal{S}_{2,m}}$ and $\mathcal{S}_{n+1} = \overline{\cup_{m=n+1}^{\infty} \mathcal{S}_{n+1,m}}$, hence it follows from (22) that (18) holds.

Step 3. Denote $\mathcal{R}_n = \text{span}\left(\mathcal{U}_n, \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}\right)$. Then

$$\mathcal{M}_{\infty} = \overline{\cup_{n=1}^{\infty} \mathcal{R}_n}. \tag{23}$$

Proof. It follows from (19) that $\mathcal{M}_n \subset \mathcal{R}_n$, hence

$$\mathcal{M}_{\infty} = \overline{\cup_{n=1}^{\infty} \mathcal{M}_n} \subset \overline{\cup_{n=1}^{\infty} \mathcal{R}_n}. \tag{24}$$

However, we also have $\mathcal{R}_n \subset \mathcal{M}_{\infty}$, as is not hard to verify, hence

$$\overline{\cup_{n=1}^{\infty} \mathcal{R}_n} \subset \mathcal{M}_{\infty}. \tag{25}$$

Thus, the result (23) follows from (24) and (25).

Step 4. For an $x \in \mathcal{M}_{\infty}$, let $\widehat{x}_n$ be the projection of x on $\mathcal{R}_n$. Then

$$\widehat{x}_n = \sum_{j=1}^{n} \alpha_j e_j + w_{n+1} \tag{26}$$

where $\alpha_j = \langle x, e_j \rangle$ and w_{n+1} is the projection of x on $\mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}$. Moreover,

$$\sum_{j=1}^{\infty} \alpha_j^2 < \infty. \tag{27}$$

Furthermore,

$$\lim_{n\to\infty} \left\| x - \sum_{j=1}^{n} \alpha_j e_j - w_{n+1} \right\| = 0. \tag{28}$$

Proof. By the definition of $\mathcal{R}_n$ and Lemma 4, $\widehat{x}_n = \sum_{j=1}^{n} \theta_j e_j + w$ for some constants θ_j and a $w \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}$. To determine the θ_j's and w, note that

$$\begin{aligned} \left\| x - \sum_{j=1}^{n} \theta_j e_j - w \right\|^2 &= ||x-w||^2 - 2\sum_{j=1}^{n} \theta_j \langle e_j, x\rangle + 2\sum_{j=1}^{n} \theta_j \langle e_j, w\rangle \\ &\quad + \left\| \sum_{j=1}^{n} \theta_j e_j \right\|^2 \\ &= ||x-w||^2 - 2\sum_{j=1}^{n} \theta_j \langle e_j, x\rangle + \sum_{j=1}^{n} \theta_j^2 \end{aligned}$$

because $w \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1} \subset \mathcal{U}_n^{\perp}$ implies $\langle e_j, w\rangle = 0$ and

$$\left\| \sum_{j=1}^{n} \theta_j e_j \right\|^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} \theta_j \theta_i \langle e_j, e_i\rangle = \sum_{j=1}^{n} \theta_j^2 \langle e_j, e_j\rangle = \sum_{j=1}^{n} \theta_j^2.$$

Thus

$$\begin{aligned} ||x - \widehat{x}_n||^2 &= \inf_{\theta_1,\ldots,\theta_n, w\in\mathcal{U}_n^{\perp}\cap\mathcal{S}_{n+1}} \left\| x - \sum_{j=1}^{n} \theta_j e_j - w \right\|^2 \\ &= \inf_{\theta_1,\ldots,\theta_n, w\in\mathcal{U}_n^{\perp}\cap\mathcal{S}_{n+1}} \left(||x-w||^2 - 2\sum_{j=1}^{n} \theta_j \langle e_j, x\rangle + \sum_{j=1}^{n} \theta_j^2 \right) \\ &= \inf_{w\in\mathcal{U}_n^{\perp}\cap\mathcal{S}_{n+1}} ||x-w||^2 - \sum_{j=1}^{n} \alpha_j^2 \\ &= ||x - w_{n+1}||^2 - \sum_{j=1}^{n} \alpha_j^2 \end{aligned} \tag{29}$$

where $\alpha_j = \langle x, e_j\rangle$ and w_{n+1} is the projection of x on $\mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}$.

This result implies that for all n,

$$\sum_{j=1}^{n} \alpha_j^2 \le ||x - w_{n+1}||^2 \le ||x||^2 \tag{30}$$

so that (27) holds.

Finally, to prove (28), let $\widehat{x}$ be the projection of x on $\overline{\cup_{n=1}^{\infty}\mathcal{R}_n}$. Then it follows from Theorem 2 that $\lim_{n\to\infty} ||\widehat{x}_n - \widehat{x}|| = 0$. But (23) implies $\widehat{x} \in \mathcal{M}_\infty$, hence $x = \widehat{x}$, so that $\lim_{n\to\infty} ||\widehat{x}_n - x|| = 0$.

Step 5. Let $z_n = \sum_{j=1}^{n} \alpha_j e_j$. Then

$$\lim_{n\to\infty} \|z - z_n\| = 0, \text{ where } z \in \mathcal{U}_\infty. \tag{31}$$

Proof. This follows from the fact that z_n is a Cauchy sequence in $\mathcal{U}_\infty$ because

$$\begin{aligned} \|z_n - z_m\|^2 &= \left\| \sum_{j=\min(m,n)+1}^{\max(m,n)} \alpha_j e_j \right\|^2 = \sum_{j=\min(m,n)+1}^{\max(m,n)} \alpha_j^2 \\ &\le \sum_{j=\min(m,n)+1}^{\infty} \alpha_j^2 \to 0 \end{aligned}$$

as $\min(m,n) \to \infty$, where the latter is due to $\sum_{j=1}^{\infty} \alpha_j^2 < \infty$.

Step 6. There exists a $w \in \mathcal{U}_\infty^\perp \cap \mathrm{S}_\infty$ such that

$$\lim_{n\to\infty} ||w_{n+1} - w|| = 0. \tag{32}$$

Proof. Recall from Step 4 that $w_{n+1} \in \mathcal{U}_n^\perp \cap \mathcal{S}_{n+1}$. Moreover, it is easy to verify that $\mathcal{U}_{n+1}^\perp \subset \mathcal{U}_n^\perp$ whereas it is trivial that $\mathcal{S}_{n+2} \subset \mathcal{S}_{n+1}$, hence $\mathcal{U}_{n+1}^\perp \cap \mathcal{S}_{n+2} \subset \mathcal{U}_n^\perp \cap S_{n+1}$. Thus for an arbitrary $k \ge 1$ and all $n \ge k$,

$$w_{n+1} \in \mathcal{U}_k^\perp \cap \mathcal{S}_{k+1}$$

Furthermore for $n \ge k$, w_{n+1} is a Cauchy sequence in $\mathcal{U}_k^\perp \cap \mathcal{S}_{k+1}$ because

$$\begin{aligned} ||w_{n+1} - w_{m+1}|| &= ||\widehat{x}_n - z_n - \widehat{x}_m + z_m|| \\ &\le ||\widehat{x}_n - \widehat{x}_m|| + ||z_n - z_m|| \\ &\le ||\widehat{x}_n - x|| + ||\widehat{x}_m - x|| + ||z_n - z_m|| \\ &\to 0 \end{aligned}$$

as $\min(m,n) \to \infty$. Therefore, there exists a $w \in \mathcal{U}_k^\perp \cap \mathcal{S}_{k+1}$ such that (32) holds. Since k was arbitrary we now have $w \in \cap_{k=1}^\infty \mathcal{U}_k^\perp = \mathcal{U}_\infty^\perp$ and $w \in \cap_{k=1}^\infty \mathcal{S}_{k+1} = \mathcal{S}_\infty$, hence

$$w \in \mathcal{U}_\infty^\perp \cap \mathcal{S}_\infty.$$

This completes the proof of Step 6.

The theorem now follows from (27), (31), (32) and the fact that $w \in \mathcal{U}_\infty^\perp \cap S_\infty \subset \mathcal{U}_\infty^\perp$,which implies that $\langle w, e_k \rangle = 0$ for $k = 1, 2, 3, ...$ Q.E.D.

4 The Wold decomposition for covariance stationary time series

In the case of the Hilbert space $\mathcal{R}_0$ of zero-mean random variables with finite second moments, with inner product $\langle X, Y \rangle = E[X.Y]$ and associated norm and metric, the results of Theorem 4 translate as follows:

Theorem 5. *Let X_t be a regular univariate zero-mean covariance stationary time series process. Then X_t can be written as*

$$X_t = \sum_{j=0}^{\infty} \alpha_j U_{t-j} + W_t \ a.s., \tag{33}$$

where the U_t is a zero-mean uncorrelated process with variance 1,

$$\alpha_j = E[X_t U_{t-j}], \ \sum_{j=0}^{\infty} \alpha_j^2 < \infty, \tag{34}$$

and W_t is a zero-mean covariance stationary process satisfying

$$W_t \in \mathcal{U}_t^\perp \cap \mathcal{S}_{-\infty}, \tag{35}$$

where $\mathcal{S}_{-\infty} = \cap_n \text{span}(\{X_{n-k}\}_{k=1}^\infty)$ and $\mathcal{U}_t^\perp$ is the orthogonal complement of $\mathcal{U}_t = \text{span}(\{U_{t-k}\}_{k=0}^\infty)$. The result (35) *implies that*

$$W_t \in \ \text{span}\left(\{W_{t-m}\}_{m=1}^\infty\right), \tag{36}$$

which in its turn implies that W_t *is perfectly predictable from its past values* $W_{t-1}, W_{t-2}, W_{t-3}, \ldots$*. In other words,* W_t *is a deterministic process. Moreover,* (35) *implies that*

$$E[W_t U_{t-m}] = 0 \tag{37}$$

for all leads and lags m*.*

Proof: Recall that $U_t = \widetilde{U}_t / \sqrt{E\left[\widetilde{U}_t^2\right]}$, where $\widetilde{U}_t = X_t - \widehat{X}_t$ with $\widehat{X}_t$ the projection of X_t on span$(\{X_{t-j}\}_{j=1}^{\infty})$. The uncorrelatedness of the $\widetilde{U}_t$'s follows from Theorem 4, but we still need to show that $E[\widetilde{U}_t] = 0$ and $E[\widetilde{U}_t^2] = \sigma^2$ for all t.

Proof of $E[\widetilde{U}_t] = 0$

Let $\widehat{X}_{t,n}$ be the projection of X_t on span$(\{X_{t-j}\}_{j=1}^{n})$. Then $\widehat{X}_{t,n}$ takes the form

$$\widehat{X}_{t,n} = \sum_{j=1}^{n} \beta_{j,n} X_{t-j},$$

where the $\beta_{j,n}$'s do not depend on t. The latter follows from the fact that the $\beta_{j,n}$'s are the solutions of the normal equations

$$\sum_{j=1}^{n} \beta_{j,n} \gamma(i-j) = \gamma(i), \ \ i = 1, 2, ..., n,$$

where $\gamma(i) = E\left[X_t X_{t-i}\right]$ is the covariance function of X_t. Hence $E[\widehat{X}_{t,n}] = 0$.

It follows from Theorem 2 that

$$\lim_{n\to\infty} \left\|\widehat{X}_{t,n} - \widehat{X}_t\right\|^2 = \lim_{n\to\infty} E\left[\left(\widehat{X}_{t,n} - \widehat{X}_t\right)^2\right] = 0 \tag{38}$$

so that by Liapounov's inequality and $E[\widehat{X}_{t,n}] = 0$,

$$\begin{aligned} \lim_{n\to\infty} \left|E[\widehat{X}_t]\right| &= \lim_{n\to\infty} \left|E[\widehat{X}_t - \widehat{X}_{t,n}]\right| \leq \lim_{n\to\infty} E\left[\left|\widehat{X}_t - \widehat{X}_{t,n}\right|\right] \\ &\leq \sqrt{\lim_{n\to\infty} E\left[\left(\widehat{X}_{t,n} - \widehat{X}_t\right)^2\right]} = 0. \end{aligned}$$

Thus $E[\widehat{X}_t] = 0$ and therefore $E[\widetilde{U}_t] = E[X_t - \widehat{X}_t] = 0$.

Proof of $E[\widetilde{U}_t^2] = \sigma^2$
Let $\widetilde{U}_{t,n} = X_t - \widehat{X}_{t,n}$. It follows from (38) that

$$\lim_{n\to\infty} E\left[\left(\widetilde{U}_t - \widetilde{U}_{t,n}\right)^2\right] = \lim_{n\to\infty} E\left[\left(\widehat{X}_{t,n} - \widehat{X}_t\right)^2\right] = 0 \tag{39}$$

Moreover,

$$\begin{aligned} E\left[\widetilde{U}_{t,n}^2\right] &= \left\|X_t - \widehat{X}_{t,n}\right\|^2 = E\left[\left(X_t - \sum_{j=1}^{n} \beta_{j,n} X_{t-j}\right)^2\right] \\ &= \gamma(0) - 2\sum_{j=1}^{n} \beta_{j,n}\gamma(j) + \sum_{j=1}^{n}\sum_{i=1}^{n} \beta_{j,n}\beta_{i,n}\gamma(i-j) \\ &= \sigma_n^2 \end{aligned}$$

say, which does not depend on t. Furthermore, note that σ_n^2 is non-increasing in n, so that

$$\lim_{n\to\infty} \sigma_n^2 = \sigma^2$$

exists, and that

$$\begin{aligned} E\left[\left(\widetilde{U}_t - \widetilde{U}_{t,n}\right)^2\right] &= \left\|\widehat{X}_{t,n} - \widehat{X}_t\right\|^2 = \left\|\widehat{X}_{t,n} - X_t + \widetilde{U}_t\right\|^2 \\ &= \left\|\widehat{X}_{t,n} - X_t\right\|^2 + 2\left\langle \widehat{X}_{t,n} - X_t, \widetilde{U}_t\right\rangle + ||\widetilde{U}_t||^2 \\ &= \left\|\widetilde{U}_{t,n}\right\|^2 - 2\left\langle X_t, \widetilde{U}_t\right\rangle + ||\widetilde{U}_t||^2 \\ &= \left\|\widetilde{U}_{t,n}\right\|^2 - 2\left\langle \widehat{X}_t + \widetilde{U}_t, \widetilde{U}_t\right\rangle + ||\widetilde{U}_t||^2 \\ &= \left\|\widetilde{U}_{t,n}\right\|^2 - 2\left\langle \widetilde{U}_t, \widetilde{U}_t\right\rangle + ||\widetilde{U}_t||^2 \\ &= \left\|\widetilde{U}_{t,n}\right\|^2 - ||\widetilde{U}_t||^2 \\ &= E\left[\widetilde{U}_{t,n}^2\right] - E\left[\widetilde{U}_t^2\right]. \end{aligned}$$

Thus,

$$E\left[\widetilde{U}_t^2\right] = \sigma_n^2 - E\left[\left(\widetilde{U}_t - \widetilde{U}_{t,n}\right)^2\right] \to \sigma^2.$$

Proof of (34), (35) and (37)
The result of Theorem 4 can now be translated as

$$\lim_{n\to\infty}\left\|X_t-\sum_{j=0}^{n}\alpha_jU_{t-j}-W_t\right\|=0, \tag{40}$$

where U_t is a zero-mean uncorrelated covariance stationary process with unit variance, and $\alpha_k=\langle X_t,U_{t-k}\rangle=E\left[X_tU_{t-k}\right]$ with $\sum_{k=0}^{\infty}\alpha_k^2<\infty$.

We still need to prove that the α_k's do not depend on t, as follows. Recall from the proof of $E[\widetilde{U}_t^2]=\sigma^2$ that $\widetilde{U}_{t,n}=X_t-\sum_{j=1}^{n}\beta_{j,n}X_{t-j}$, so that

$$E\left[X_{t+k}\widetilde{U}_{t,n}\right]=\gamma\left(k\right)-\sum_{j=1}^{n}\beta_{j,n}\gamma(k+j),$$

which does not depend on t. Moreover, by the Cauchy-Schwarz inequality and (39),

$$\lim_{n\to\infty}\left|E\left[X_{t+k}\left(\widetilde{U}_{t,n}-\widetilde{U}_t\right)\right]\right|^2\leq\gamma\left(0\right)\lim_{n\to\infty}E\left[\left(\widetilde{U}_{t,n}-\widetilde{U}_t\right)^2\right]=0.$$

Thus $E\left[X_{t+k}\widetilde{U}_t\right]=\lim_{n\to\infty}E\left[X_{t+k}\widetilde{U}_{t,n}\right]$. Since the latter does not depend on t, neither does $\alpha_k=E\left[X_{t+k}U_t\right]=E\left[X_{t+k}\widetilde{U}_t/||\widetilde{U}_t||\right]$.

The results (35) and (37) follow straightforwardly from Theorem 4.

Proof of (33)
The result (40) implies, by Chebyshev's inequality, that

$$X_t=p\lim_{n\to\infty}\sum_{j=0}^{n}\alpha_jU_{t-j}+W_t. \tag{41}$$

Recall that convergence in probability for $n\to\infty$ is equivalent to a.s. convergence along a further subsequence k_m of an arbitrary subsequence of n. See for example Bierens (2004, Theorem 6.B.3, p. 168). Thus for such a subsequence k_m,

$$\sum_{j=0}^{k_m}\alpha_jU_{t-j}\to X_t-W_t\text{ a.s.} \tag{42}$$

as $m \to \infty$, and the same holds for any further subsequence of k_m.

Without loss of generality we may choose $k_0 = 0$. Then for each $n > 0$ we can find an m_n such that

$$k_{m_{n-1}} < n \leq k_{m_n}. \tag{43}$$

Moreover, (42) implies that

$$\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} \to X_t - W_t \text{ a.s. as } n \to \infty. \tag{44}$$

Due to (43),

$$\sum_{n=1}^{\infty} E\left[\left(\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} - \sum_{j=0}^{n} \alpha_j U_{t-j}\right)^2\right] = \sum_{n=1}^{\infty} E\left[\left(\sum_{j=n+1}^{k_{m_n}} \alpha_j U_{t-j}\right)^2\right]$$
$$\leq \sum_{n=1}^{\infty} \sum_{j=k_{m_{n-1}}+1}^{k_{m_n}} \alpha_j^2 \leq \sum_{j=0}^{\infty} \alpha_j^2 < \infty,$$

so that by Chebyshev's inequality, for arbitrary $\varepsilon > 0$,

$$\sum_{n=0}^{\infty} P\left[\left|\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} - \sum_{j=0}^{n} \alpha_j U_{t-j}\right| > \varepsilon\right] < \infty.$$

This result implies, by the Borel-Cantelli lemma,[2] that

$$\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} - \sum_{j=0}^{n} \alpha_j U_{t-j} \to 0 \text{ a.s. as } n \to \infty. \tag{45}$$

Combining (44) and (45) it follow now that

$$\sum_{j=0}^{n} \alpha_j U_{t-j} \to X_t - W_t \text{ a.s. as } n \to \infty. \tag{46}$$

Since $\sum_{j=0}^{\infty} \alpha_j U_{t-j}$ is defined as $\lim_{n\to\infty} \sum_{j=0}^{n} \alpha_j U_{t-j}$, (33) is equivalent to (46).

[2]See for example Bierens (2004, Theorem 6.B.2, p. 168).

The zero-mean covariance stationarity of W_t

It follows now trivially from (33) that $E[W_t] = 0$. Moreover, W_t is covariance stationary because for $m \geq 0$,

$$\begin{aligned} E[W_t W_{t-m}] &= E\left[\left(X_t - \sum_{j=0}^{\infty} \alpha_j U_{t-j}\right)\left(X_{t-m} - \sum_{j=0}^{\infty} \alpha_j U_{t-m-j}\right)\right] \\ &= E[X_t X_{t-m}] - \sum_{j=m}^{\infty} \alpha_j E[U_{t-j} X_{t-m}] \\ &\quad - \sum_{j=0}^{\infty} \alpha_j E[U_{t-m-j} X_t] + \sum_{j=0}^{\infty} \alpha_{j+m}\alpha_j \\ &= \gamma(m) - \sum_{j=0}^{\infty} \alpha_{j+m}\alpha_j. \end{aligned}$$

Proof of (36)

Finally, $W_t \in \cap_n \text{span}(\{X_{n-j}\}_{j=0}^{\infty})$ implies that $W_t \in \text{span}(\{X_{t-j}\}_{j=1}^{\infty})$, hence the projection of W_t on $\text{span}(\{X_{t-j}\}_{j=1}^{\infty})$ is W_t itself. Since by (33),

$$\text{span}(\{X_{t-j}\}_{j=1}^{\infty}) = \text{span}\left(\text{span}(\{U_{t-j}\}_{j=1}^{\infty}), \text{span}(\{W_{t-j}\}_{j=1}^{\infty})\right)$$

and the projection of W_t on $\text{span}(\{U_{t-j}\}_{j=1}^{\infty})$ is zero, it follows that the projection of W_t on $\text{span}(\{W_{t-j}\}_{j=1}^{\infty})$ is W_t itself, which proves (36). Q.E.D.

Remark 2. The condition $\text{var}(U_t) = 1$ is not essential as long as X_t is regular. Without loss of generality we may then replace U_t with $\widetilde{U}_t = \sigma U_t$, $\sigma > 0$, and α_k with $\widetilde{\alpha}_k/\sigma$, where σ can be pinned down by normalizing $\widetilde{\alpha}_0 = 1$.

5 Further analysis of the deterministic term

5.1 An example

The conclusion that the term W_t in (33) is deterministic does *not* imply that W_t is non-random. For example, consider the following sequence of random variables:

$$W_t = U\cos(t) + V\sin(t),$$

where U and V are independent standard normal random variables which do not depend on t. Suppose that for a given t, W_{t-1} and W_{t-2} are observed. Then

$$\begin{pmatrix} W_{t-1} \\ W_{t-2} \end{pmatrix} = \begin{pmatrix} \cos(t-1) & \sin(t-1) \\ \cos(t-2) & \sin(t-2) \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix}$$

hence

$$\begin{aligned} \begin{pmatrix} U \\ V \end{pmatrix} &= \frac{1}{\sin(t-2)\cos(t-1) - \cos(t-2)\sin(t-1)} \\ &\quad \times \begin{pmatrix} \sin(t-2) & -\sin(t-1) \\ -\cos(t-2) & \cos(t-1) \end{pmatrix} \begin{pmatrix} W_{t-1} \\ W_{t-2} \end{pmatrix} \\ &= \frac{-1}{\sin(1)} \begin{pmatrix} \sin(t-2) & -\sin(t-1) \\ -\cos(t-2) & \cos(t-1) \end{pmatrix} \begin{pmatrix} W_{t-1} \\ W_{t-2} \end{pmatrix}, \end{aligned}$$

so that

$$\begin{aligned} W_t &= \frac{-1}{\sin(1)} (\cos(t), \sin(t)) \begin{pmatrix} \sin(t-2) & -\sin(t-1) \\ -\cos(t-2) & \cos(t-1) \end{pmatrix} \begin{pmatrix} W_{t-1} \\ W_{t-2} \end{pmatrix} \\ &= \frac{-1}{\sin(1)} \left(\sin(t-2)\cos(t) - \cos(t-2)\sin(t)\right) W_{t-1} \\ &\quad - \frac{1}{\sin(1)} \left(\sin(t)\cos(t-1) - \cos(t)\sin(t-1)\right) W_{t-2} \\ &= \frac{\sin(2)}{\sin(1)} . W_{t-1} - W_{t-2}. \end{aligned}$$

Moreover, it is trivial that $E[W_t] = 0$ and

$$E[W_t W_{t-m}] = \cos(t)\cos(t-m) + \sin(t).\sin(t-m) = \cos(m),$$

hence W_t is a zero-mean covariance stationary process.

5.2 Measurability

Again, the crux of Theorem 5 is that

$$W_t \in \mathcal{S}_{-\infty} = \cap_n \text{span}(\{X_{n-j}\}_{j=0}^{\infty}), \tag{47}$$

as the other conclusions of Theorem 5 follow straightforwardly from Theorem 3. The question I will now address is how (47) translates in terms of σ-algebras generated by sequences of the type $\{X_{n-j}\}_{j=0}^{\infty}$.

Denote for natural numbers n and m,

$$\begin{aligned}\mathcal{F}_{t-n-m}^{t-n} &= \sigma\left(X_{t-n}, X_{t-n-1}, ..., X_{t-n-m}\right), \\ \mathcal{S}_{t-n-m}^{t-n} &= \text{span}\left(X_{t-n}, X_{t-n-1}, ..., X_{t-n-m}\right)\end{aligned}$$

where $\mathcal{F}_{t-n-m}^{t-n}$ is the σ-algebra generated by $X_{t-n}, X_{t-n-1}, ..., X_{t-n-m}$. Recall that

$$\mathcal{F}_{-\infty}^{t-n} = \sigma\left(\cup_{m=1}^{\infty}\mathcal{F}_{t-n-m}^{t-n}\right)$$

is the smallest σ-algebra containing $\cup_{m=1}^{\infty}\mathcal{F}_{t-n-m}^{t-n}$, which is the σ-algebra generated by the sequence $\{X_{t-n-j}\}_{j=0}^{\infty}$, and that

$$\mathcal{F}_{-\infty} = \cap_{n=1}^{\infty}\mathcal{F}_{-\infty}^{t-n} = \cap_t\mathcal{F}_{-\infty}^{t}$$

is the remote σ-algebra of the sequence X_t, whereas

$$\begin{aligned}\mathcal{S}_{-\infty}^{t-n} &= \overline{\cup_{m=1}^{\infty}\mathcal{S}_{t-n-m}^{t-n}} = \text{span}\left(\{X_{t-n-j}\}_{j=0}^{\infty}\right), \\ \mathcal{S}_{-\infty} &= \cap_n\mathcal{S}_{-\infty}^{t-n}.\end{aligned}$$

The similarity between $\mathcal{S}_{-\infty}$ and $\mathcal{F}_{-\infty}$ suggests that, possibly, W_t is measurable with respect to (w.r.t.) the remote σ-algebra $\mathcal{F}_{-\infty}$. In Bierens (2004, p. 182) I have stated the latter as a fact, but in hindsight this is not trivial.

To prove this proposition, we need the following lemma:

Lemma 6. *Let $\{Z_n\}_{n=1}^{\infty}$ be a sequence of random variables [vectors] defined on a common probability space $\{\Omega, \mathcal{F}, P\}$ such that for a set $A \in \mathcal{F}$ with $P(A) = 1$, $\lim_{n\to\infty} Z_n(\omega) = Z(\omega)$ pointwise in $\omega \in A$. Then Z is a random variable [vector][3] defined on $\{\Omega, \mathcal{F}, P\}$, i.e., Z is measurable with respect to $\mathcal{F}$.*

Proof: See for example Theorem 2.13 in Bierens (2004, p. 47).

Let $\widehat{W}_{t-n,m}$ be the projection of W_t on $\mathcal{S}_{t-n-m}^{t-n}$, and let $\widehat{W}_{t-n}$ be the projection of W_t on $\mathcal{S}_{-\infty}^{t-n}$. Then it follows from Theorem 2 that

$$\lim_{m\to\infty} ||\widehat{W}_{t-n,m} - \widehat{W}_{t-n}||^2 = \lim_{m\to\infty} E\left[\left(\widehat{W}_{t-n,m} - \widehat{W}_{t-n}\right)^2\right] = 0,$$

[3] Possibly extended arbitrarily for $\omega \in \Omega\backslash A$.

hence by Chebyshev inequality, $\text{plim}_{m\to\infty}\widehat{W}_{t-n,m} = \widehat{W}_{t-n}$. Recall[4] that the latter implies that along a subsequence m_k,

$$\widehat{W}_{t-n,m_k} \overset{a.s.}{\to} \widehat{W}_{t-n} \text{ as } k \to \infty. \tag{48}$$

Moreover, recall that $\widehat{W}_{t-n,m_k} \in \mathcal{S}_{t-n-m_k}^{t-n}$ takes the form of a linear combination of $X_{t-n}, X_{t-n-1}, ..., X_{t-n-m_k}$, hence $\widehat{W}_{t-n,m_k}$ is measurable w.r.t. $\mathcal{F}_{t-n-m_k}^{t-n}$ and therefore is also measurable w.r.t. $\mathcal{F}_{-\infty}^{t-n}$. It follows now from (48) and Lemma 6 that $\widehat{W}_{t-n}$ is measurable w.r.t. $\mathcal{F}_{-\infty}^{t-n}$.

Because $W_t \in \mathcal{S}_{-\infty} \subset \mathcal{S}_{-\infty}^{t-n}$ and therefore $\widehat{W}_{t-n} = W_t$ a.s., it follows now that for arbitrary n, W_t is measurable w.r.t. $\mathcal{F}_{-\infty}^{t-n}$ and therefore W_t is also measurable w.r.t. $\cap_{n=1}^{\infty}\mathcal{F}_{-\infty}^{t-n}$. Thus,

Theorem 6. *The deterministic term W_t in the Wold decomposition* (33) *is measurable with respect to the remote σ-algebra $\mathcal{F}_{-\infty}$ of the time series X_t.*

5.3 When is the deterministic term equal to zero?

Finally, I will address the question: under what condition(s) is $W_t = 0$ a.s.. For this we need the concept of *vanishing memory* introduced in Bierens (2004, Definition 7.3, p.183):

Definition 5. *A time series process X_t is said to have a vanishing memory if all the sets in its remote σ-algebra $\mathcal{F}_{-\infty}$ have either probability zero or one.*

In this case any random variable W that is measurable w.r.t. to such a σ-algebra $\mathcal{F}_{-\infty}$ satisfies

$$E[W|\mathcal{F}_{-\infty}] = E[W] \text{ a.s.}$$

To see this, let $Z = E[W|\mathcal{F}_{-\infty}]$ and $A = \{\omega \in \Omega : Z(\omega) - E[W] > 0\}$. Recall from the definition of conditional expectation relative to a σ-algebra that Z is measurable w.r.t. $\mathcal{F}_{-\infty}$, so that $A \in \mathcal{F}_{-\infty}$, and that $\int_A Z\mathrm{d}P = \int_A W\mathrm{d}P$. Now suppose that $P(A) = 1$. Then

$$\begin{aligned}\int_A E[W]\mathrm{d}P &= E[W].P(A) = E[W] = \int W\mathrm{d}P \\ &= \int_A W\mathrm{d}P + \int_{\Omega\backslash A} W\mathrm{d}P = \int_A W\mathrm{d}P\end{aligned}$$

[4] See for example Bierens (2004, Theorem 6.B.3, p. 168).

hence $\int_A (Z - E[W])\mathrm{d}P = 0$. But this implies that $P(A) = 0$,[5] so that $P(A) = 1$ is not possible, and therefore the only alternative is that $P(A) = 0$. The same applies to $A = \{\omega \in \Omega : Z(\omega) - E[W] < 0\}$. Hence, $Z = E[W]$ a.s..

Moreover, because W is measurable w.r.t. to $\mathcal{F}_{-\infty}$ we also have[6]

$$E[W|\mathcal{F}_{-\infty}] = W \text{ a.s.}$$

Because $E[W_t] = 0$ it follows now that

Theorem 7. *If the time series X_t in Theorem 5 has a vanishing memory then the deterministic term W_t in the Wold decomposition* (33) *is a.s. zero.*

6 The multivariate Wold decomposition

To prove the multivariate version of the Wold decomposition for a k-variate covariance stationary process X_t, consider the Hilbert space $\mathcal{R}_k$ of zero mean random vectors in $\mathbb{R}^k$ with finite second moment matrices, endowed with the inner product $\langle X, Y\rangle = E[X'Y]$ and associated norm and metric. Let $\widehat{X}_t$ be the projection of X_t on $\text{span}(\{X_{t-j}\}_{j=1}^{\infty})$, with residual vector $V_t = X_t - \widehat{X}_t$, and let $\Sigma = E[V_t V_t']$. In this case we need to extend the notion of regularity by requiring that Σ is positive definite rather than only $||V_t||^2 = E[V_t' V_t] > 0$, so that we can define $U_t = \Sigma^{-1/2} V_t$. Then the projection $\widetilde{X}_t$ of X_t on $\text{span}(\{U_{t-j}\}_{j=0}^{n})$ takes the form $\widetilde{X}_t = \sum_{j=1}^{n} A_j U_{t-j}$, where $A_j = E\left[X_t U_{t-j}'\right]$. It follows now straightforwardly from the proofs of Theorems 4 and 5 that

$$X_t = \sum_{j=1}^{\infty} A_j U_{t-j} + W_t \text{ a.s.,}$$

where the process U_t is uncorrelated with zero expectation vector and variance matrix I_k, and $W_t \in \mathcal{U}_t^{\perp} \cap \mathcal{S}_{-\infty}$, with $\mathcal{U}_t^{\perp}$ and $\mathcal{S}_{-\infty}$ defined in Theorem 5.

7 References

Anderson, T. W. (1994), *The Statistical Analysis of Time Series*, Wiley

[5] See for example Bierens (2004, Lemma 3.1, p.71).
[6] See for example Bierens (2004, Theorem 3.4, p. 73).

Bierens, H. J. (2004). *Introduction to the Mathematical and Statistical Foundations of Econometrics.* Cambridge University Press.
Wold, H. (1938), *A Study in the Analysis of Stationary Time Series.* Almqvist and Wiksell, Sweden.