The space spanned by a countable infinite sequence in an Hilbert space, with application to the Wold decomposition

Herman J. Bierens*

January 30, 2009

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

Let $\mathcal{H}$ be a Hilbert space and let $\{x_k\}_{k=1}^{\infty}$ be a sequence of elements of $\mathcal{H}$. Let $\mathcal{M}_n$ be the linear manifold spanned by $x_1, ..., x_n$, i.e., $\mathcal{M}_n$ consists of all linear combinations of $x_1, ..., x_n$. Then

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

Proof: Let $z_m = \sum_{j=1}^{n}\beta_{j,m}x_j$ be a Cauchy sequence in $\mathcal{M}_n$. Then there exists a $z \in \mathcal{H}$ such that $\lim_{m\to\infty}||z - z_m|| = 0$. Let $\widehat{z}_n$ be the projection of

*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 Spring 2009 graduate time series course have led to substantial improvements of the proof of the Wold decomposition.

z on $\mathcal{M}_n$, i.e.,

$$\begin{aligned}||z-\widehat{z}_n||^2 &= \min_{\theta_1,\ldots,\theta_n}\left\|z-\sum_{j=1}^{n}\theta_j x_j\right\|^2 \\ &= \min_{\theta_1,\ldots,\theta_n}\left\{||z||^2-2\sum_{j=1}^{n}\theta_j\left\langle x_j,z\right\rangle+\sum_{i=1}^{n}\sum_{j=1}^{n}\theta_i\theta_j\left\langle x_i,x_j\right\rangle\right\}.\end{aligned}$$

The optimal θ_j's are the solutions of the normal equations

$$\langle z,x_j\rangle=\sum_{i=1}^{n}\theta_{n,i}\left\langle x_i,x_j\right\rangle,\ j=1,...,n, \tag{1}$$

hence $\widehat{z}_n=\sum_{j=1}^{n}\theta_{n,j}x_j$, and $||z-\widehat{z}_n||^2\leq||z-z_m||^2\to 0$, so that $z=\sum_{j=1}^{n}\theta_{n,j}x_j\in\mathcal{M}_n$. 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* $\mathrm{span}(\{x_j\}_{j=1}^{\infty})$.

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.

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

Proof: Adopting the notation in the previous lemmas, 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 1 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

$$\lim_{n\to\infty} ||\widehat{z} - z_n|| = 0. \tag{2}$$

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 (2) that

$$\lim_{n\to\infty} \delta_n = \delta. \tag{3}$$

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

$$\begin{aligned} ||\widehat{z} - \widehat{z}_n||^2 &= ||z - \widehat{z}_n - u||^2 = ||z - \widehat{z}_n||^2 + ||u||^2 - 2\langle z - \widehat{z}_n, u\rangle \\ &= ||z - \widehat{z}_n||^2 + ||u||^2 - 2\langle z, u\rangle = \delta_n^2 - \delta^2 \end{aligned} \tag{4}$$

where the last equality follows from

$$\langle z, u\rangle - \langle u, u\rangle = \langle \widehat{z}, u\rangle = 0 \tag{5}$$

and $\langle u, u\rangle = ||u||^2 = \delta^2$. The theorem now follows from (3) and (4). Q.E.D.

Remark. Although each projection $\widehat{z}_n$ is a linear combination of $x_1, ..., x_n$, in general the result of Theorem 1 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 an example of such a case, 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 V_t is distributed i.i.d. $N(0,1)$. 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

$$\widehat{X}_{t,n} = \sum_{j=1}^{n} \left(\frac{j}{n+1} - 1\right) X_{t-j} \tag{6}$$

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

$$\lim_{n\to\infty} \left\| \widehat{X}_t - \sum_{j=1}^{n} \theta_j X_{t-j} \right\|^2 = \lim_{n\to\infty} E\left[\left(\widehat{X}_t - \sum_{j=1}^{n} \theta_j X_{t-j}\right)^2\right] = 0 \tag{7}$$

If so, it follows from Theorem 1 and (6) that

$$\begin{aligned} 0 &= \lim_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\theta_j X_{t-j} - \sum_{j=1}^{n}\left(\frac{j}{n+1}-1\right)X_{t-j}\right)^2\right] \qquad (8) \\ &= \lim_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] \end{aligned}$$

But

$$\begin{aligned} &\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j} = \sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)(V_{t-j}-V_{t-j-1}) \\ &= -\left(\frac{n}{n+1}+\theta_1\right)V_{t-1} - \sum_{j=1}^{n-1}\left(\theta_{j+1}-\theta_j-\frac{1}{n+1}\right)V_{t-j-1} \\ &+ \left(\frac{1}{n+1}+\theta_n\right)V_{t-n-1} \end{aligned}$$

hence

$$\begin{aligned} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] &= \left(\frac{n}{n+1}+\theta_1\right)^2 \\ &+ \sum_{j=1}^{n-1}\left(\theta_{j+1}-\theta_j-\frac{1}{n+1}\right)^2 + \left(\frac{1}{n+1}+\theta_n\right)^2 \qquad (9) \end{aligned}$$

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

$$\begin{aligned} &\liminf_{n\to\infty} E\left[\left(\sum_{j=1}^{n}\left(\frac{j}{n+1}-1-\theta_j\right)X_{t-j}\right)^2\right] \\ &\geq \liminf_{n\to\infty}\left(\frac{n}{n+1}+\theta_1\right)^2 + \liminf_{n\to\infty}\left(\theta_{m+1}-\theta_m-\frac{1}{n+1}\right)^2 \\ &= (\theta_1+1)^2 + (\theta_{m+1}-\theta_m)^2 . \end{aligned}$$

Therefore, a necessary condition for (8) is that $\theta_m = -1$ for $m = 1, 2, 3, \ldots..$ But then it follows from (9) 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 (8). Thus, in this case there does **not** exist a sequence $\{\theta_j\}_{j=1}^{\infty}$ such that (7) holds.

2 Projections on the span of an orthonormal sequence

On the other hand,

Theorem 2. *If a sequence $\{x_j\}_{j=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is orthonormal, i.e.,*

$$\langle x_i, x_j \rangle = \begin{cases} 0 & for \quad i \neq j, \\ 1 & for \quad i = 1, \end{cases} \tag{10}$$

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

Proof: Observe from the normal equations (1) that the projection $\widehat{z}_n$ of z on $\text{span}(\{x_j\}_{j=1}^{n})$ takes the form

$$\widehat{z}_n = \sum_{j=1}^{n} \theta_j x_j, \text{ where } \theta_j = \langle z, x_j \rangle . \tag{11}$$

Moreover, denoting $u_n = z - \widehat{z}_n$, it follows from (10) and (11) that

$$\begin{aligned} ||u_n||^2 &= \left\| z - \sum_{j=1}^{n} \theta_j x_j \right\|^2 = ||z||^2 - 2\sum_{j=1}^{n} \theta_j \langle z, x_j \rangle + \sum_{j=1}^{n} \sum_{i=1}^{n} \theta_j \theta_i \langle x_j, x_i \rangle \\ &= ||z||^2 - \sum_{j=1}^{n} \theta_j^2 \geq 0 \end{aligned} \tag{12}$$

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

$$\lim_{n\to\infty} \left\| \widehat{z} - \sum_{j=1}^{n} \theta_j x_j \right\|^2 = \lim_{n\to\infty} ||\widehat{z} - \widehat{z}_n||^2 = 0$$

so that we can write $\widehat{z} = \sum_{j=1}^{\infty} \theta_j x_j$. Q.E.D.

3 The Wold decomposition

Let $\mathcal{S}_1, ..., \mathcal{S}_n$ be subspaces of a Hilbert space $\mathcal{H}$. Then similar to Definition 1,

Definition 2. $\text{Span}(\mathcal{S}_1, ..., \mathcal{S}_n)$ *is the closure of the space of all linear combinations* $\sum_{j=1}^{n} c_j x_j$, *where* $x_j \in \mathcal{S}_j$.

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* $\mathcal{H}$ *such that for each* $x \in S$ *and* $y \in \mathcal{S}^{\perp}$, $\langle x, y \rangle = 0$.

Lemma 4. *Orthogonal complements are subspaces.*

Proof: Let x be an arbitrary element of a subspace $\mathcal{S}$ of an Hilbert space $\mathcal{H}$ 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[1] 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.

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.

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

Theorem 3. *Given a regular sequence* $\{x_k\}_{k=1}^{\infty}$ *in a Hilbert space, every* $x \in \mathcal{S} = \text{span}(\{x_k\}_{k=1}^{\infty})$ *can be written as* $x = \sum_{k=1}^{\infty} \alpha_k e_k + w$, *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* $\mathcal{S}$, $\alpha_k = \langle x, e_k \rangle$, $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$, *and*

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

[1] See for example Bierens (2007a).

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

$$\mathcal{U}_n = \text{ span}\left(e_1, ..., e_n\right) = \text{ span}\left(u_1, ..., u_n\right),$$

and let $\mathcal{U}_n^\perp$ be the orthogonal complement of $\mathcal{U}_n$. Moreover, as before, let $\mathcal{M}_n = \text{span}(\{x_k\}_{k=1}^{n})$.

A typical element $z \in \mathcal{M}_n$ takes the form $z = \sum_{k=1}^{n} c_k x_k$. Hence, z can be written as

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

Note that $\sum_{k=1}^{n} c_k \widehat{x}_k \in \mathcal{S}_2$ because $\widehat{x}_k \in \mathcal{S}_{k+1}$.

Next, project $\sum_{k=1}^{n} c_k \widehat{x}_k$ on $\mathcal{U}_n$. This projection takes the form $\sum_{k=1}^{n} d_k e_k$ with residual $w_{n+1} \in \mathcal{S}_2$. But we also have $\langle e_k, w_{n+1}\rangle = 0$ for $k = 1, ..., n$, hence $w_{n+1} \in \mathcal{U}_n^\perp$. Thus, denoting $\alpha_k = c_k||u_k|| + d_k$, we can write every $z \in \mathcal{M}_n$ as

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

Therefore

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

I will now show that

$$\text{span}\left(\mathcal{U}_n, \mathcal{U}_n^\perp \cap \mathcal{S}_2\right) = \text{ span}\left(\mathcal{U}_n, \mathcal{U}_n^\perp \cap \mathcal{S}_{n+1}\right), \tag{15}$$

as follows. Denote $\mathcal{S}_{k,m} = \text{span}(\{x_j\}_{j=k}^{m})$ for $m \geq k$ and let $z \in \mathcal{U}_n^\perp \cap \mathcal{S}_{2,m}$ for some $m \geq 2$. Consider first the case $m > n$. Since $z \in \mathcal{S}_{2,m}$ there exists constants c_k such that

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

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

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

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

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

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

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

Thus, $\mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m} \subset \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m}$. However, $\mathcal{S}_{n+1,m} \subset \mathcal{S}_{2,m}$ and therefore $\mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1,m} \subset \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}$,so that

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

This result implies that

$$\mathcal{U}_n^{\perp} \cap \left(\cup_{m=n+1}^{\infty} \mathcal{S}_{2,m}\right) = \mathcal{U}_n^{\perp} \cap \left(\cup_{m=n+1}^{\infty} \mathcal{S}_{n+1,m}\right) \tag{16}$$

In the case $m \leq n$, $z \in \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}$ implies that $z = 0$, as can be straightforwardly verified from the above argument, so that $\mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m} = \{0\}$ for $m = 2, 3, ..., n$. Since Hilbert spaces are vector spaces and therefore always contain the null element it follows that

$$\begin{aligned} \cup_{m=2}^{\infty} \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m} &= \{0\} \cup \left(\cup_{m=n+1}^{\infty} \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}\right) \\ &= \cup_{m=n+1}^{\infty} \mathcal{U}_n^{\perp} \cap \mathcal{S}_{2,m}, \end{aligned}$$

hence

$$\mathcal{U}_n^{\perp} \cap \left(\cup_{m=2}^{\infty} \mathcal{S}_{2,m}\right) = \mathcal{U}_n^{\perp} \cap \left(\cup_{m=n+1}^{\infty} \mathcal{S}_{2,m}\right). \tag{17}$$

It follows now from (17) that

$$\begin{aligned}
\mathcal{U}_n^{\perp} \cap \mathcal{S}_2 &= \mathcal{U}_n^{\perp} \cap \overline{\cup_{m=2}^{\infty} \mathcal{S}_{2,m}} \\
&= \mathcal{U}_n^{\perp} \cap \overline{\cup_{m=n+1}^{\infty} \mathcal{S}_{n+1,m}} = \mathcal{U}_n^{\perp} \cap \mathcal{S}_{n+1}
\end{aligned}$$

which implies that (15) holds.

Combining (14) and (15) yields

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

say. However, we also have $\mathcal{R}_n \subset \mathcal{S}_1$, as is not hard to verify. Consequently,

$$\mathcal{S}_1 = \overline{\cup_{n=1}^{\infty} \mathcal{M}_n} = \overline{\cup_{n=1}^{\infty} \mathcal{R}_n}.$$

Let $x \in \mathcal{S}_1$ and let $\widehat{x}_n$ be the projection of x on $\mathcal{R}_n$. Then

$$\begin{aligned}
||x - \widehat{x}_n||^2 &= \inf_{\theta_1, \ldots, \theta_n, w \in \mathcal{U}_n^{\perp} \cap S_{n+1}} \left\| x - \sum_{j=1}^{n} \theta_j u_j - w \right\|^2 \\
&= \inf_{w \in \mathcal{U}_n^{\perp} \cap S_{n+1}} \left\| x - \sum_{j=1}^{n} \langle \widehat{x}, e_j \rangle u_j - w \right\|^2 \\
&= \inf_{w \in \mathcal{U}_n^{\perp} \cap S_{n+1}} ||x - w||^2 - \sum_{j=1}^{n} \alpha_j^2 \\
&= ||x - w_{n+1}||^2 - \sum_{j=1}^{n} \alpha_j^2
\end{aligned} \tag{18}$$

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

$$\sum_{j=1}^{n} \alpha_j^2 \le ||x - w_{n+1}||^2 \le ||\widehat{x}||^2 \tag{19}$$

so that

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

Therefore, $z = \sum_{j=1}^{\infty} \alpha_j e_j$ is well defined, as $\lim_{n\to\infty} ||z - \sum_{j=1}^{n} \alpha_j e_j|| = 0$, and so is

$$w = x - \sum_{j=1}^{\infty} \alpha_j e_j$$

which is defined as $\lim_{n\to\infty} ||x - \sum_{j=1}^{n} \alpha_j e_j - w|| = 0$.

Note that w is the limit of the Cauchy sequence $w_{n+1} = x - \sum_{j=1}^{n} \alpha_j e_j$. That w_{n+1} is a Cauchy sequence follows from

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

as $\min(m,n) \to \infty$, where the latter is due to $\sum_{j=1}^{\infty} \alpha_j^2 < \infty$. Since for arbitrary fixed $m \geq 1$ and all $n \geq m$, $w_{n+1} \in \mathcal{S}_{m+1}$, it follows that $w \in \mathcal{S}_{m+1}$ for all $m \geq 1$, hence

$$w \in \cap_{n=1}^{\infty} \mathcal{S}_n.$$

Similarly, $w_{n+1} \in \mathcal{U}_n^{\perp} \subset \cup_{k=1}^{\infty} \mathcal{U}_k^{\perp} \subset \overline{\cup_{k=1}^{\infty} \mathcal{U}_k^{\perp}} = \mathcal{U}_{\infty}^{\perp}$, say, where $\mathcal{U}_{\infty}^{\perp}$ is the orthogonal complement of $\mathcal{U}_{\infty} = \text{span}(\{e_k\}_{k=1}^{\infty})$. Hence

$$w \in \mathcal{U}_{\infty}^{\perp},$$

which implies that $\langle w, e_k \rangle = 0$ for $k = 1, 2, 3, \ldots\ldots..$ Q.E.D.

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 3 translate as follows:

Theorem 4. *Let X_t be a regular univariate zero-mean covariance stationary time series process. Then X_t can be written as*

$$X_t = \sum_{j=0}^{\infty} \alpha_j U_{t-j} + W_t \ a.s., \tag{20}$$

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

$$\alpha_j = E[X_t U_{t-j}],\ \sum_{j=0}^{\infty} \alpha_j^2 < \infty, \tag{21}$$

and W_t *is a zero-mean covariance stationary process satisfying*

$$W_t \in \mathcal{U}_t^{\perp} \cap \mathcal{S}_{-\infty}, \tag{22}$$

where $\mathcal{S}_{-\infty} = \cap_n \mathrm{span}(\{X_{n-k}\}_{k=1}^{\infty})$ *and* $\mathcal{U}_t^{\perp}$ *is the orthogonal complement of* $\mathcal{U}_t = \mathrm{span}(\{U_{t-k}\}_{k=0}^{\infty})$. *The result* (22) *implies that*

$$W_t \in \ \mathrm{span}\left(\{W_{t-m}\}_{m=1}^{\infty}\right), \tag{23}$$

which in its turn implies that W_t *is perfectly predictable from the past values* $W_{t-1}, W_{t-2}, W_{t-3}, ...$ *Moreover,* (22) *implies that*

$$E[W_t U_{t-m}] = 0 \tag{24}$$

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 $\mathrm{span}(\{X_{t-j}\}_{j=1}^{\infty})$. The uncorrelatedness of the $\widetilde{U}_t$'s follows from Theorem 3, 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 $\mathrm{span}(\{X_{t-j}\}_{j=1}^{n})$. 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 $E\left[\widehat{X}_{t,n}\right] = 0$.

It follows from Theorem 1 that

$$\lim_{n\to\infty}\left\|\widehat{X}_{t,n} - \widehat{X}_t\right\|^2 = \lim_{n\to\infty} E\left[\left(\widehat{X}_{t,n} - \widehat{X}_t\right)^2\right] = 0 \tag{25}$$

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

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

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 (25) that

$$\lim_{n\to\infty} E\left[\left(\widetilde{U}_t - \widetilde{U}_{t,n}\right)^2\right] = \lim_{n\to\infty} E\left[\left(\widehat{X}_{t,n} - \widehat{X}_t\right)^2\right] = 0 \tag{26}$$

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

$$E\left[\left(\widetilde{U}_t - \widetilde{U}_{t,n}\right)^2\right] = \left\|\widehat{X}_{t,n} - \widehat{X}_t\right\|^2 = \left\|\widehat{X}_{t,n} - X_t + \widetilde{U}_t\right\|^2$$

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

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 (21), (22) and (24)

The result of Theorem 3 can now be translated as

$$\lim_{n\to\infty}\left\|X_t - \sum_{j=0}^{n}\alpha_j U_{t-j} - W_t\right\| = 0, \tag{27}$$

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

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

which does not depend on t. Moreover, by the Cauchy-Schwarz inequality and (26),

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

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 (22) and (24) follow straightforwardly from Theorem 3.

Proof of (20)
The result (27) implies, by Chebyshev's inequality, that

$$X_t = p \lim_{n\to\infty} \sum_{j=0}^{n} \alpha_j U_{t-j} + W_t. \tag{28}$$

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,

$$\sum_{j=0}^{k_m} \alpha_j U_{t-j} \to X_t - W_t \text{ a.s.} \tag{29}$$

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

$$k_{m_{n-1}} < n \leq k_{m_n}. \tag{30}$$

Moreover, (29) implies that

$$\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} \to X_t - W_t \text{ a.s. as } n \to \infty. \tag{31}$$

Due to (30),

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

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

$$\sum_{j=0}^{k_{m_n}} \alpha_j U_{t-j} - \sum_{j=0}^{n} \alpha_j U_{t-j} \to 0 \text{ a.s. as } n \to \infty. \tag{32}$$

Combining (31) and (32) it follow now that

$$\sum_{j=0}^{n} \alpha_j U_{t-j} \to X_t - W_t \text{ a.s. as } n \to \infty. \tag{33}$$

Since $\sum_{j=0}^{\infty} \alpha_j U_{t-j}$ is defined as $\lim_{n\to\infty} \sum_{j=0}^{n} \alpha_j U_{t-j}$, (20) is equivalent to (33).

The zero-mean covariance stationarity of W_t
It follows now trivially from (20) that $E[W_t] = 0$. Moreover, it is left as an exercise to show that for $m \geq 0$,

$$E[W_t W_{t-m}] = \gamma(m) - \sum_{j=0}^{\infty} \alpha_{j+m}\alpha_j. \tag{34}$$

Proof of (23)
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 (20),

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

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 (23). Q.E.D.

Remarks

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 $\widetilde{\alpha}_0 = 1$.

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

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

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 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 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 3 and 4 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 4.

Finally, for further reading on Hilbert spaces, see for example Bierens (2007, 2008, 2009) and Young (1988). For more on the Wold decomposition, see Anderson (1994).

4 References

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

Bierens, H. J. (2004). *Introduction to the Mathematical and Statistical Foundations of Econometrics.* Cambridge University Press.

Bierens, H. J. (2007), *Introduction to Hilbert Spaces*, Lecture notes (http:// econ.la.psu.edu/~hbierens/HILBERT.PDF)

Bierens, H. J. (2008), *Orthonormal Polynomials, Related Orthonormal Functions and the Hilbert Spaces They Span*, Lecture notes (http://econ.la.psu.edu /~hbierens/ORTHONORMAL_POL.PDF)

Bierens, H. J. (2009), *Hilbert Space Theory and Its Applications in Econometrics,* Lecture notes (http://econ.la.psu.edu/~hbierens/HILBERTSPACE.PDF)

Young, N. (1988), *An Introduction to Hilbert Space.* Cambridge University Press.