

Lecture Notes 1. Linear Space Theory

Much of time series analysis is based on the construction of linear decompositions of the processes under study. This approach, in turn, is based on various results from the mathematics of linear spaces. Linear spaces are sometimes called vector spaces. These lecture notes describe some basic definitions and results. The results are all standard. I follow Royden (1988) and Simmons (1963) closely in terms of presentation. Ash (1976) chapter 3 is also useful.

1. Basic definitions

All of the spaces we work with will represent refinements of the concept of a linear space. Following Simmons (1963 p. 80)

Definition 1.1.1 Linear space

Let Γ denote a non-empty set. For each pair of elements γ_i and γ_j contained in Γ , assume that there exists an addition operation, $+$ which combines pairs of elements to produce an element, $\gamma_i + \gamma_j$, which is also contained in Γ . Suppose that for the combination of the set Γ and the addition operation, the following hold:

i. $\gamma_i + \gamma_j = \gamma_j + \gamma_i$

ii. $\gamma_i + (\gamma_j + \gamma_k) = (\gamma_i + \gamma_j) + \gamma_k$

iii. Γ contains a unique element, 0, such that $\gamma + 0 = \gamma \forall \gamma \in \Gamma$

iv. For each $\gamma \in \Gamma \exists$ an element $-\gamma$ such that $\gamma + (-\gamma) = 0$

Further, for any scalar element c of the set of complex numbers C , assume that c may be combined with any $\gamma \in \Gamma$ to produce an element of Γ $c\gamma$. This is the multiplication operation. Suppose that for the combination of the set and the multiplication operation, the following hold:

v. $c(\gamma_i + \gamma_j) = c\gamma_i + c\gamma_j$

vi. $(c + d)\gamma = c\gamma + d\gamma$

vii. $(cd)\gamma = c(d\gamma)$

viii. $1\gamma = \gamma$

Then, the set Γ , addition operator and multiplication operator together define a linear space.

Two comments. First, one can replace the set of complex numbers with the set of real numbers R and define a linear space analogously. We will use real and complex weights in different contexts. Second, notice that the addition and multiplication operators are defined to according to how they operate on elements of Γ . Thus the definition of a linear space is quite general, we will work with linear spaces of number, functions, etc.

In essence, the definition of a linear space requires that the space is closed with respect to addition of any pair of elements of the space and that multiplying any element by a scalar produces another element of the space. This definition does not require the space to be complete. A space is complete when the limit of every sequence of elements in the space is also a element of the

space. One reason why it is useful to work with a complete space concerns optimization. In order to rigorously discuss whether the space is complete, it is necessary to equip the linear space with a norm. A norm is a function that defines distances between elements of a space. This definition is required to fulfill certain properties.

Definition 1.1.2. Norm

For a linear space Γ , a norm, denoted as $\|\cdot\|$, is a mapping from Γ to R such that

$$i. \|\gamma\| \geq 0; \|\gamma\| = 0 \rightarrow \gamma = 0$$

$$ii. \|c\gamma\| = |c|\|\gamma\| \quad \forall c \in C$$

$$iii. \|\gamma_i + \gamma_j\| \leq \|\gamma_i\| + \|\gamma_j\|$$

The first condition requires the norm to measure something interpretable as a length. The second condition requires that the norm is linear. The third condition is the triangle inequality. This is a hint that intuitions about 3-dimensional Euclidean space that we have from elementary geometry will apply to more general spaces. The next definition will come as no surprise.

Definition 1.1.3. Normed linear space

A normed linear space is a linear space equipped with a norm.

When a space is endowed with a norm, one can discuss the convergence properties of various sequences of elements in the space. An important type of sequence is a Cauchy sequence.

Definition 1.1.4. Cauchy sequence

The sequence γ_i is a Cauchy sequence if, for any $\varepsilon > 0$, there exists an I (which may depend on ε) such that for all $i, j > I$ $\|\gamma_i - \gamma_j\| < \varepsilon$.

The notion of a Cauchy sequence naturally leads to the notion of a limit point for the sequence.

Definition 1.1.5. Limit of a Cauchy sequence

The element $\bar{\gamma}$ is the limit to Cauchy sequence γ_i if, for any $\varepsilon > 0$, there exists an I (which may depend on ε) such that for all $i > I$, $\|\gamma_i - \bar{\gamma}\| < \varepsilon$

This definition naturally leads to

Definition 1.1.6. Completeness

A linear space Γ is complete if the limit to every Cauchy sequence in the space is an element of the space.

It is not necessarily true that the limit point of Cauchy sequence is an element of a normed. Normed linear spaces that are complete form an important class of spaces, known as Banach spaces.

Definition 1.1.7. Banach space

A Banach space is a complete normed linear space.

A noted, the norm function provides a way of measuring the magnitude of an element γ_i ; it also provides a measure of the distance between two elements,

$\|\gamma_i - \gamma_j\|$. It is important to also be able to characterize the relationship between pairs of elements relative to one another. The key notion for this is the inner product of two elements of a space. For the complex number c , \bar{c} denotes its complex conjugate.

Definition 1.1.8. Inner product

An inner product function $\langle \cdot, \cdot \rangle$ is a mapping from $\Gamma \times \Gamma$ to \mathbb{C} such that

$$i. \langle a\gamma_i + b\gamma_j, \gamma_k \rangle = a\langle \gamma_i, \gamma_k \rangle + b\langle \gamma_j, \gamma_k \rangle$$

$$ii. \langle \gamma_i, \gamma_j \rangle = \overline{\langle \gamma_j, \gamma_i \rangle}$$

$$iii. \langle \gamma_i, \gamma_i \rangle \geq 0$$

$$iv. \langle \gamma_i, \gamma_i \rangle = 0 \rightarrow \gamma_i = 0$$

Our final definition describes Banach spaces where the inner product defines the norm.

Definition 1.1.9. Hilbert space

A Hilbert space is a Banach space with an inner product such that $\langle \gamma_i, \gamma_i \rangle = \|\gamma_i\|^2$.

Hilbert spaces provide the foundation for the theory of time series.

2. Properties of Hilbert spaces

This section discusses a few properties of the general structure of Hilbert space. These properties depend critically on the notion of orthogonality between two elements of a Hilbert space.

Definition 1.2.1 Orthogonality

Two elements γ_i and γ_j are orthogonal if $\langle \gamma_i, \gamma_j \rangle = 0$

i. Projections

Orthogonality allows one to construct various decompositions of elements of Hilbert spaces and hence of the spaces themselves. To see this, take any two elements of a space, γ_i and γ_j and consider the element

$$\gamma_i - \frac{\langle \gamma_i, \gamma_j \rangle}{\langle \gamma_j, \gamma_j \rangle} \gamma_j \quad (1.1)$$

This element is also a member of the Hilbert space since the space is linear.

Further, this element is orthogonal to γ_j since

$$\langle \gamma_i - \frac{\langle \gamma_i, \gamma_j \rangle}{\langle \gamma_j, \gamma_j \rangle} \gamma_j, \gamma_j \rangle = \langle \gamma_i, \gamma_j \rangle - \langle \gamma_i, \gamma_j \rangle = 0 \quad (1.2)$$

This construction is suggestive of a fundamental property of Hilbert spaces, namely that there exist ways to decompose these spaces into mutually orthogonal subspaces. Ash, Royden and Simmons all provide proofs of the following.

Theorem 1.2.1. Decomposability of Hilbert space into orthogonal subspaces

Suppose that Γ and Γ_1 are Hilbert spaces such that $\Gamma_1 \subseteq \Gamma$. Then there exists a unique Hilbert space $\Gamma_2 \subseteq \Gamma$ such that

$$i. \Gamma = \Gamma_1 \oplus \Gamma_2$$

$$ii. \Gamma_1 \perp \Gamma_2$$

The operator “ \oplus ” is called the direct sum. The direct sum of two spaces produces a third space whose elements consist of all linear combinations of elements of the original spaces and the limits of all such combinations. The following corollary is nothing more than the restriction of Theorem 1.2.1 to consideration of a single element of Γ .

Corollary 1.2.1. Decomposition of an element of a Hilbert space into orthogonal components

Suppose that Γ and Γ_1 are Hilbert spaces such that $\Gamma_1 \subseteq \Gamma$. For a given element $\gamma \in \Gamma$, there exists a unique pair of elements $\gamma_1 \in \Gamma_1$ and $\gamma_2 \in \Gamma_2$, (where Γ_2 is defined in Theorem 1.2.1) such that

$$i. \gamma = \gamma_1 + \gamma_2$$

$$ii. \gamma_1 \perp \gamma_2$$

In this corollary, γ_1 is called the *projection* of γ onto Γ_1 . This projection has an interpretation in terms of an optimization problem. Suppose one want to solve the problem

$$\min_{\xi \in \Gamma_1} \|\gamma - \xi\|^2 \tag{1.3}$$

In other words, one wants to find that element of the subspace Γ_1 that is closest to an element of Γ . The solution to this problem is $\xi = \gamma_1$. To see this, note that

$$\begin{aligned} \|\gamma - \xi\|^2 &= \|\gamma_1 + \gamma_2 - \xi\|^2 = \\ &\|\gamma_1 - \xi\|^2 + \|\gamma_2\|^2 + \langle \gamma_1 - \xi, \gamma_2 \rangle + \langle \gamma_2, \gamma_1 - \xi \rangle = \\ &\|\gamma_1 - \xi\|^2 + \|\gamma_2\|^2 \end{aligned} \tag{1.4}$$

where the last equality follows from the orthogonality of Γ_1 and Γ_2 . Since ξ only affects the value of $\|\gamma - \xi\|^2$ through $\|\gamma_1 - \xi\|^2$, the solution is immediate

ii. Orthonormal bases for Hilbert space

For R^n , it is possible to think of each element as a linear combination of vectors $(1,0,\dots)$, $(0,1,\dots)$, etc. These unit length vectors are, of course the axes of the space. This property holds for general Hilbert spaces.

We start with two definitions.

Definition 1.2.3. Orthonormal set

A subset S of a given Hilbert space is called an orthonormal set if

- i.* each element of S is orthogonal to every other element
- ii.* the norm of every element equals 1

Definition 1.2.4. Orthonormal basis

An orthonormal set S is an orthonormal basis with respect to a given Hilbert space H if it is not a proper subset of any other orthonormal set in the space.

If a set of vectors span a space, this means that if one takes all linear combinations and limits of the linear combination of elements of the set, one reproduces the space. Orthonormal bases have the important property that they span the space. A set of elements of a space span it if, starting with the set, linear combinations of the elements and their associated limits together constitute the space.

Theorem 1.2.2 Relationship between orthonormal basis and a given Hilbert space

An orthonormal basis of a given Hilbert space Γ spans the space.

The next two theorems establish the existence of orthonormal bases. Particularly clear proofs are found in Ash section 3.2 and Royden section 10.8.

Theorem 1.2.3. Existence of an orthonormal basis

Every Hilbert space contains an orthonormal basis.

Theorem 1.2.4. Existence of a countable orthonormal basis

A Hilbert space contains a countable orthonormal basis iff the space is separable.¹

Countable orthonormal bases are of interest because they provide the “axes” for a Hilbert space that are analogous to the orthogonal unit vectors one uses to construct R^k . When S is a countable orthonormal basis for Γ , a typical element γ can either be represented as

$$\gamma = \sum_{i=1}^I a_i s_i \quad (1.5)$$

or can be arbitrarily well approximated this way. (The difference is that since the space is complete, it contains limits of all such sums as well as the sums themselves.) The coefficients are implicitly determined by the projection of γ onto each of the 1-dimensional subspaces spanned by the s_i 's.² Using the construction of orthogonal decompositions developed above, it is clear that $a_i = \langle \gamma, s_i \rangle$. (Recall that $\langle s_i, s_i \rangle = 1$ by assumption. These are also known as Fourier coefficients.

3. Applications

To provide an application of the abstract ideas we have discussed considering Hilbert space, consider R^k , k -dimensional Euclidean space. Letting x and y denote vectors in this space, the inner product of the vectors is $\sum_i x_i y_i$ with associated norm $(\sum_i x_i x_i)^{1/2}$. To see how the Hilbert space decomposition theorem applies, consider a Hilbert space generated around the vector $(1,0,0\dots 0)$. (Recall that generating a Hilbert space around some set of elements means adding all linear combinations and limits of linear combinations to elements to the set.) This will be a subspace of R^k . What is the orthogonal complement to this space? It is the Hilbert space spanned by the vectors $(0,1,0,0\dots 0)$, $(0,0,1,0\dots 0)$, ..., $(0,0,0,0\dots 1)$.

¹A space is separable if it contains a countable dense subset. The spaces we study will always be separable.

²Each s_i spans a distinct 1-dimensional subspace by the fact that the s_i 's form an orthonormal basis.

Similarly, a number of simple properties of Hilbert spaces are generalizations of results that one finds in a range of mathematical contexts.

Example 1.3.1. Pythagorean Theorem

Suppose that some element of a Hilbert space, call it x , is the sum of two orthogonal elements y and z . In this case, the square of the norm of x obeys

$$\langle x, x \rangle = \langle y + z, y + z \rangle = \langle y, y \rangle + \langle z, z \rangle \quad (1.6)$$

since by orthogonality, $\langle y, z \rangle = \langle z, y \rangle = 0$. But this is nothing more than the Pythagorean theorem for arbitrary Hilbert spaces, as opposed to R^2 .

Example 1.3.2. Cauchy-Schwarz Inequality

For any a , the inner product of $x + ay$ with can be written as

$$\langle x + ay, x + ay \rangle = \langle x, x \rangle + a\langle y, x \rangle + \bar{a}\langle x, y \rangle + |a|^2 \langle y, y \rangle \geq 0 \quad (1.7)$$

Suppose that $a = -\frac{\langle x, y \rangle}{\langle y, y \rangle}$. Substituting into the above,

$$\begin{aligned} \langle x + ay, x + ay \rangle &= \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle} \langle x, y \rangle + \left| \frac{\langle x, y \rangle}{\langle y, y \rangle} \right|^2 \langle y, y \rangle = \\ &\langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} \geq 0. \end{aligned} \quad (1.8)$$

Rearranging terms, this expression implies

$$\langle x, x \rangle \langle y, y \rangle \geq |\langle x, y \rangle|^2 \quad (1.9)$$

with equality iff $x = ay$ (The if and only if claim is something you should verify.)

This is the Cauchy-Schwarz inequality for an arbitrary Hilbert space.

References

Ash, R., (1976), *Real Analysis and Probability*, New York: Academic Press.

Royden, H., (1988), *Real Analysis*, Third Edition, New York: MacMillan.

Simmons, G., (1963), *Introduction to Topology and Modern Analysis*, New York: Academic Press.